Peer-reviewed version available at *Forests* **2017**, *8*, 269; doi:10.3390/f8080269

Article

Allometric Equations for Estimation of Biomass and Carbon Stocks in Temperate Forests of North-Western Mexico

Benedicto Vargas-Larreta [1,*], Carlos Antonio López-Sánchez [2], José Javier Corral-Rivas [2], Jorge Omar López-Martínez [3], Cristóbal Gerardo Aguirre-Calderón [1] and Juan Gabriel Álvarez-González [4]

[1] Instituto Tecnológico de El Salto. Mesa del Tecnológico s/n. 34942 El Salto, Durango, México; g_aguirremx@yahoo.com
[2] Instituto de Silvicultura e Industria de la Madera. Universidad Juárez del Estado de Durango. Boulevard del Guadiana 501. Ciudad Universitaria. Torre de Investigación. 34120 Durango, Dgo. México; calopez@ujed.mx (C.A.L.-S.); jcorral@ujed.mx (J.J.C.-R.)
[3] El Colegio de la Frontera Sur. Av. Centenario km 5.5, 77014 Chetumal, Quintana Roo, México; lmjorgeomar@gmail.com
[4] Departamento de Ingeniería Agroforestal. Universidad de Santiago de Compostela. Escuela Politécnica Superior, Lugo, España; juangabriel.alvarez@usc.es

* Correspondence: bvargas@itelsalto.edu.mx; Tel.: +52-675-876-5533

Abstract: This paper presents new above-ground biomass (AGB) and biomass components equations for seventeen forest species in the temperate forests of northwestern Mexico. A data set corresponding to 1336 destructively sampled oak and pine trees was used to fit the models. Generalized method of moments was used to simultaneously fit systems of equations for biomass components and AGB, to ensure additivity. Additionally, the carbon content of each tree component was calculated by the dry combustion method, in a TOC analyser. The fitted equations accounted for on average 91, 83, 84 and 78% of the observed variance in stem wood and stem bark, branch and foliage biomass, respectively, whereas the total AGB equations explained on average 93% of the total observed variance in AGB. The inclusion of h or d^2h as additional predictor in the d-only based equations systems slightly improved estimates of stem wood, stem bark and total above-ground biomass, and greatly improved the estimates produced by the branch and foliage biomass equations. The fitted equations were used to estimate AGB stocks at stand level from a database on growing stock from 429 permanent sampling plots. Three machine-learning techniques were used to model the estimated stand level AGB and carbon contents; the selected models were applied to map the AGB and carbon distributions in the study area, which yielded mean values of 129.84 Mg ha^{-1} and 63.80 Mg ha^{-1}, respectively.

Keywords: aboveground biomass; GMM; allometry; biomass allocation; machine learning technique

1. Introduction

Better knowledge of carbon stocks and fluxes is needed to understand the current state of the carbon cycle and how it might evolve with changing land use and climatic conditions [1]. This has led to an increased interest in estimating forest biomass for both practical forestry purposes and scientific purposes. Tree biomass is an important component of the carbon pool in forests, and it can be estimated by using biomass expansion factors [2] or by relating biomass functions to tree-level data obtained in forest inventories [3]. In both cases, generic biomass functions are used to quantify the carbon in forests [4] because they improve the accuracy in carbon accounting systems and thus allow accurate planning of whole tree and residual biomass utilization for bioenergy production.

Temperate forests occupy 32 330 508 ha in Mexico, which represents 17 per cent of the country. These are the richest ecosystems in Mexico with some 7000 plant species [5], of which ~150 are species of pines and 170 of oaks, and these represent over 50 per cent of all known pine and oak species [6]. Durango is the most important forestry state in Mexico, with 10.5 million ha of forest cover. Timber production in the state is 1.9 million m^3 per year, which represents 32.8% of the national forest production [7].

Mexico is one of the largest emitters of CO_2 by deforestation and contributes with 1.6 per cent of global emissions [8], mainly from temperate and tropical forests (12.9 and 54.1 Mt C $year^{-1}$, respectively) [9]. Thus, most of the states in Mexico are implementing action plans for mitigating the effects of climate change and for accessing economic incentives favoring carbon sequestration in forests. The role of temperate forests in Mexico is important because of their potential to accumulate C and emit large amounts of CO_2 into the atmosphere; however, the lack of sets of species-level biomass equations for oak and pine species growing in these forests, in addition to equations that incorporate aspects of forest structure that vary significantly at regional scales is required [10], so that species and site-specific biomass equations must therefore be developed. Initiatives such as REDD+ (Reducing Emissions from Deforestation and Forest Degradation and enhancement of carbon stocks) are important efforts aimed at combating climate change; however, for effective implementation of such mechanisms accurate estimation and monitoring AGB and associated carbon stocks in forests is first required [11].

The objectives of the present study were: (i) to develop species-specific systems of additive equations for predicting total above-ground biomass; and (ii) to model the forest biomass and carbon in the temperate forests in north-western Mexico by using remote sensing Landsat-5 TM imagery, terrain parameters and data from permanent research plots.

2. Materials and Methods

2.1 Study area

The study was conducted in the temperate uneven-aged and multi-species forest of Durango (22°20′49″ to 26°46′33″ N; 103°46′38″ to 107°11′36″ W), which occupies about 23% of the Sierra Madre Occidental ecosystem (Figure 1). The elevation above sea level varies between 363 and 3200 m (average 2264 m). Precipitation ranges from 443 to 1452 mm, with an annual average of 917 mm, whereas the mean annual temperature varies from 8.2 to 26.2 °C, with an annual average of 13.3 °C. The predominant forest types are pine and uneven-aged pine-oak, with one or two pine species (usually *P. cooperi* and *P. durangensis*) dominating the overstorey and *Q. sideroxyla* dominating the understorey [12].

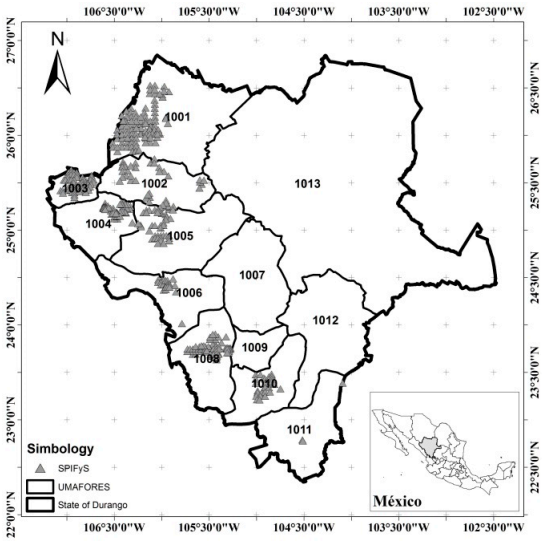

Figure 1. Location of the study area. Triangles indicate the location of the 429 permanent sampling plots.

2.2 Biomass data

Biomass estimates for stem wood, stem bark, branches and foliage were obtained from 1336 destructively sampled trees of 17 tree species: *Pinus cooperi*, *P. durangensis*, *P. engelmannii*, *P. leiophylla*, *P. herrerae*, *P. teocote*, *P. lumholtzii*, *P. strobiformis*, *P. oocarpa*, *P. douglasiana*, *P. michoacana*, *Juniperus depeanna*, *Arbutus bicolor*, *Quercus sideroxyla*, *Q. rugosa*, *Q. durifolia* and *Q. crassifolia*. The *P. douglasiana* and *P. michoacana* trees were considered together in order to improve the model fit. A general equation for each biomass component for all pines and all oaks trees was also developed.

The fieldwork for aboveground biomass measurement included tree selection, carrying out standing measurements, felling trees, collecting dimensional data, cutting and separating the tree components and weighing fresh components of each biomass component on site. Trees were sampled in 5 cm diameter classes, from 5 cm until the maximum diameter found in the area. The number of trees sampled varied from a minimum of 30 trees in the case of *P. douglasiana* to a maximum of 130 trees of *P. durangensis* and *Q. crassifolia*. Diameter at breast height (*d*) and total height (*h*) were measured in each sample tree.

The following biomass components were considered: stem wood, stem bark, branches (including both wood and bark) and foliage (needles/leaves). For each felled tree, stem diameter outside bark was measured at 0.3, 0.6, 1.3 and thereafter every 2.5 m along the stem to take into account variations in moisture content along the stem. The green weight of stem and branches was determined by weighing the logs and branches in the field by placing them on a 1000 kg balance (precision 100 g). Foliage was totally separated from the trunk and weighed on an analytical balance (precision 1 g). A disk of about 5 cm thick was cut from each log, and representative samples of branches and needles/leaves were weighed in the field (fresh weight) before being transported to the laboratory where they were oven-dried at 70-85 °C to constant weight (dry weight, measured to the nearest 0.1 g). On the basis of the ratio of dry biomass to fresh biomass, the biomass of each tree component was calculated and then summed to produce the total AGB of each tree sampled. The carbon content of each tree component was determined by the dry combustion method, in a TOC analyzer. The number of observations used for tree biomass estimations and the basic description of the tree biomass components data for each species are summarized in Table 1.

Table 1. Summary statistics of the felled trees and the main aboveground biomass components used to develop the biomass equations.

Species (*n*)		Variables		Biomass components				
		d	*h*	W_w	W_b	W_{br}	W_f	W_t
Pinus cooperi (103)	Max	52.3	28.0	1586.5	101.0	317.8	28.8	1833.8
	Min	5.5	4.2	2.3	0.3	1.5	0.2	4.1
	Mean	28.9	17.1	397.5	30.6	97.2	12.5	537.8
	SD	10.9	5.0	346.0	23.7	81.4	7.4	444.3
P. durangensis (130)	Max	44.5	29.8	1027.5	89.2	312.0	38.4	1319.0
	Min	8.7	6.9	8.9	1.3	3.0	0.4	14.4
	Mean	25.9	16.6	269.2	37.2	106.6	11.8	424.8
	SD	8.5	4.4	210.9	23.4	67.5	8.2	294.5
P. engelmannii (89)	Max	51.8	27.8	1825.0	44.3	308.3	30.4	2024.7
	Min	6.5	3.8	6.0	0.3	0.5	0.2	6.9
	Mean	27.7	15.0	507.5	18.5	107.0	11.8	644.7
	SD	9.4	5.4	421.7	11.5	58.9	6.9	479.5
P. leiophylla (84)	Max	55.3	29.2	1129.8	89.7	392.0	21.8	1614.6
	Min	8.4	5.4	5.7	0.5	1.1	0.4	7.7
	Mean	29.6	16.6	329.5	23.2	149.6	7.4	509.7
	SD	12.0	5.4	278.6	20.8	111.3	5.7	402.6
P. herrerae (97)	Max	46.4	31.0	1056.9	72.8	139.4	31.8	1200.4
	Min	5.0	5.2	4.0	0.3	0.8	0.4	5.4
	Mean	27.8	16.3	354.3	20.3	49.4	13.6	437.6
	SD	9.2	4.9	273.0	13.9	34.0	8.6	319.2
P. teocote (81)	Max	45.0	24.7	789.3	47.9	161.3	34.8	975.0
	Min	10.0	4.5	3.2	0.3	1.1	0.6	5.2
	Mean	29.6	15.6	288.9	16.6	50.5	12.2	368.2
	SD	9.4	4.0	207.6	12.2	38.3	7.9	261.7
P. lumholtzii (35)	Max	42.0	24.9	832.2	37.6	86.9	37.0	981.5
	Min	5.0	3.6	2.9	0.2	0.4	0.5	4.0
	Mean	22.4	14.7	236.3	10.7	32.0	11.7	290.7
	SD	8.6	4.3	209.7	8.5	27.1	10.2	252.7
P. strobiformis (98)	Max	49.0	26.6	1240.1	37.6	211.3	40.4	1501.1
	Min	5.0	6.3	1.8	0.3	1.5	0.4	4.0
	Mean	27.1	16.2	292.9	17.9	73.1	17.2	401.3
	SD	9.8	4.8	277.2	9.8	56.2	11.2	354.4
P. oocarpa (37)	Max	35.7	18.7	448.9	49.5	173.1	38.0	647.5
	Min	7.5	3.2	6.0	0.6	3.0	2.0	11.5
	Mean	21.6	12.9	153.4	19.6	53.4	16.8	243.3
	SD	6.3	3.6	109.6	12.4	42.0	9.8	167.3
P. douglasiana (30)	Max	39.0	25.6	718.4	60.8	124.2	31.2	884.1
	Min	8.9	5.6	2.7	0.5	1.1	0.3	4.6
	Mean	25.0	17.5	261.9	21.8	40.5	12.4	336.6
	SD	7.2	4.4	174.9	15.2	29.5	9.2	224.9
P. michoacana (32)	Max	42.1	24.7	866.7	56.2	137.2	39.8	1073.5
	Min	12.9	13.5	48.8	4.0	12.9	4.2	69.9
	Mean	31.3	20.3	416.6	27.4	64.6	21.4	530.0
	SD	8.2	3.0	232.5	13.5	36.6	9.2	285.6
Juniperus deppeana (48)	Max	43.7	21.5	357.0	21.8	56.3	27.5	456.3
	Min	10.0	4.5	3.7	0.5	1.2	0.8	6.2
	Mean	32.9	11.4	183.7	11.4	25.8	13.7	234.7
	SD	7.9	3.1	92.7	4.8	15.8	6.9	117.7

Species (*n*)		Variables		Biomass components				
		d	*h*	W_w	W_b	W_{br}	W_f	W_t
Arbutus bicolor (49)	Max	44.8	2.5	236.1	9.6	231.1	14.5	375.6
	Min	7.9	2.4	3.8	0.2	3.1	0.2	7.2
	Mean	22.8	8.9	83.3	3.4	48.5	5.4	140.6
	SD	7.8	2.9	53.9	2.4	46.8	3.7	97.6
Quercus sideroxyla (123)	Max	57.0	24.8	1018.0	276.3	308.3	29.6	1559.0
	Min	11.0	6.3	12.3	3.3	0.4	0.4	16.4
	Mean	30.7	14.6	290.9	87.3	75.6	8.5	462.2
	SD	9.7	3.7	213.4	90.7	62.2	6.7	331.9
Q. rugosa (61)	Max	41.3	20.2	456.1	129.8	132.0	19.4	648.6
	Min	9.3	3.3	4.2	1.6	3.2	1.5	14.3
	Mean	22.5	11.2	107.2	40.6	51.1	19.2	218.0
	SD	8.1	3.7	95.8	31.4	33.4	13.0	166.2
Q. durifolia (131)	Max	45.5	22.1	930.9	118.4	421.9	63.9	1526.3
	Min	7.0	5.4	6.4	1.3	2.1	0.3	10.2
	Mean	27.2	12.4	344.4	32.5	125.6	25.1	527.6
	SD	8.8	3.5	226.6	26.9	109.6	18.3	371.7
Q. crassifolia (108)	Max	43.1	18.8	533.9	88.3	274.6	35.5	865.0
	Min	8.5	6.0	14.3	1.3	11.0	1.0	33.4
	Mean	25.5	11.2	218.5	25.6	93.3	14.2	351.5
	SD	7.7	2.6	134.7	18.7	62.0	8.2	212.5

were *n*= number of trees, *d*=diameter at breast height (cm), *h*= total height (m), W_w= wood biomass of stem (kg tree^{-1}), W_b= stem bark biomass (kg tree^{-1}), W_{br}= wood plus bark biomass of branches (kg tree^{-1}), W_f= foliage (leaves/needles) biomass (kg tree^{-1}), W_t= total above-ground biomass (W_w + W_b + W_{br} + W_f) (kg tree^{-1}).

2.3 Procedures for developing the species-specific biomass equations

2.3.1 Basic models

We used three basic model forms as starting points for model selection (eq. 1-3).

$$w = \alpha d^{\beta} + \varepsilon_i \quad (1)$$

$$w = \alpha d^{\beta} h^{\gamma} + \varepsilon_i \quad (2)$$

$$w = \alpha d^{2} h + \varepsilon_i \quad (3)$$

where α, β, and γ are the equation parameters, w can be total tree AGB or any of the tree biomass components considered in the study and ε_i is the model error.

A first regression procedure was used to select the definitive tree variables for each biomass component equation, over the linearized version of the models taking natural logarithms. The significance level for entering and maintaining variables in the model was restricted to 0.001 [2]. In this first step, the best model for each tree biomass component and species was chosen. A species-specific system of equations with cross-equation constraints on the structural parameters and cross-equation error correlation was then defined for additive prediction of tree component and above-ground biomass [13, 14].

The species-specific biomass equation system is formulated as follows:

$$W_i = \alpha_i X_j^{\beta_{ij}} + \varepsilon_i \quad (4)$$

$$W_t = \sum_{i=1}^{n} W_i + \varepsilon_t \quad (5)$$

where W_i represents the tree biomass for the i-th component, W_t is the total tree AGB (i.e. the sum of all the tree biomass components), X_j are tree variables ($j = d, h, d^2h$), α_i, β_{ij} are parameters to be estimated in the fitting process, and ε_i, ε_t are inter-correlated error terms [14].

2.3.2 Simultaneous fitting of tree biomass components and total AGB

Total tree AGB was formulated as the sum of the equations for each tree component, and the system of equations was fitted using the Generalized Method of Moments (GMM) in the PROC MODEL procedure of SAS/ETS® [15]. This method produces efficient parameter estimates under potentially heteroscedastic conditions, without specifying the nature of the heteroscedasticity [16], and thus avoids estimating the heteroscedastic error variance. Bi *et al*. [14] and Castedo *et al*. [2] state that the GMM procedure of SAS overcomes the above mentioned problem of the linear combination of the error terms of the tree biomass component equations by computing a generalized inverse of the variance covariance matrix -by setting part of the matrix for whole-tree biomass to zero- and that estimation of parameters by simultaneously fitting the biomass component equations guarantees that AGB will be the sum of the tree component estimates.

2.3.3 Comparison of equations

Statistical and graphical analyses were used to compare the performance of the equations. The goodness-of-fit of each biomass fraction model was evaluated using the root mean squared error (RMSE) and the coefficient of determination (R^2) (equations 6-7).

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}\left(Y_{ij} - \hat{Y}_{ij}\right)^2}{n-p}} \quad (6)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(Y_{ij} - \hat{Y}_{ij}\right)^2}{\sum_{i=1}^{n}\left(Y_{ij} - \bar{Y}_{ij}\right)^2} \quad (7)$$

where Y_{ij} and $\hat{Y}_{ij}$ are the j-th observed and predicted values of biomass for component i, $\bar{Y}_{ij}$ is the mean of n observed values for the same component and p is the number of parameters in the model.

2.4 Applying the above-ground biomass equations developed

A data set from a network of 429 permanent research sampling plots (Sitios Permanentes de Investigación Forestal y de Suelos (SPIFyS)), distributed across the Sierra Madre Occidental in Durango [17] was used to relate stand biomass and carbon stocks to variables obtained from remote sensors. The total AGB in each stand was calculated by applying the developed tree-level biomass equations (to each tree), converted into carbon content by using the carbon proportion estimated for each component and expressed per unit area (ha). The species or group species-specific AGB models reported by Rojas-García *et al*. [18] were used to estimate total AGB for the tree species present in the permanent plots and for which no biomass equations were developed in this study.

The spectral data were derived from Landsat TM5 (Thematic Mapper) satellite images obtained on the same dates that the SPIFyS were established (2007 to 2011) and available from the National Landsat Archive Processing System (NLAPS). The Landsat 4-5 Thematic Mapper product, level 1 of surface reflectance (radiometrically and atmospherically corrected) was processed using the Standard Landsat Product Generation System (LPGS) via the Landsat Ecosystem Disturbance

Adaptive Processing System (LEDAPS) algorithm (both available at https://espa.cr.usgs.gov). Landsat TM5 bands 1, 2, 3, 4, 5 and 7 were used; band 6 was not used, because of its thermal characteristics [19]. Some common Vegetation Indexes (VI) and other derived parameters were computed from the atmospherically corrected image bands: the Normalized Difference Vegetation Index (NDVI); the Soil Adjusted Vegetation Index (SAVI) and its modification (MSAVI); the Enhanced Vegetation Index (EVI); the Normalized Burn Ratio using bands 4 and 7 (NBR); the difference Normalized Burn Ratio 2 (NBR2) using bands 5 and 7; and the Normalized Difference Moisture Index (NDMI). These indices have been widely used as comprehensive indicators of the interaction between land cover and solar radiation in the visible and near-infrared regions of the electromagnetic spectrum [20-23].

The direct relationship between terrain variables and forest species composition, tree height growth and other stand variables enables these forest variables to be modelled [24, 25]. First and second order terrain parameters were thus derived from the 5×5 low pass filtered Digital Elevation Model (DEM) of the study area with a spatial resolution of 15 m [26]. The first order terrain parameters selected were elevation, slope, aspect, transformed aspect, profile curvature, plan curvature and curvature, while the second order terrain parameters were terrain shape index and wetness index. These parameters are potentially related to key features for forest stand development, such as overall climate characteristics, insolation, evapotranspiration, run-off, infiltration, wind exposure and site productivity [24, 27].

The sample plots were geolocated in order to extract the average pixel value with an associated buffer of 25 m for each potential predictor. The pixel data were extracted using R statistical software [28] and the "raster" package. Finally, a database was constructed with the mean biomass values for each plot; the corrected bands of the Landsat-5 TM sensor (6 bands: 1, 2, 3, 4, 5 and 7), the vegetation indexes (7 indexes) and the terrain variables derived from the DEM (9 variables).

2.5 Machine Learning Techniques (MLTs)

We compared the performance of three machine learning techniques for estimating the ABG and carbon at stand level: (i) the non-parametric Support Vector Machine for Regression (SVM) technique, (ii) Regression by Discretization based on Random Forest (RD-RF), and (iii) parametric multiple linear regression (MLR).

MLR is the technique most commonly used in this kind of study [29]; furthermore, this type of model is easy to understand and is widely used in most scientific disciplines. To select the best set of independent variables, the model was initially built on all descriptors, and descriptors with the smallest standardized regression coefficients were then removed step-wise from the model until no improvement was observed in the estimate of the average prediction error given by the Akaike information criterion [30]. On the other hand, Support Vector Machines for Regression (SVM), originating from statistical learning theory, have become a subject of intensive study [31], as they enable the user to deal with highly nonlinear problems [32] such as estimating complex forest structures. These models are developed by a set of vectors that minimize the mean error. SVM are robust in generalization, even when the training data are noisy, and are guaranteed to have a unique global solution that is not trapped in multiple local minima [33]. SVM have proven to be useful in remote sensing of forest environments [34, 35]. The Shevade *et al*. [36] modification of the Sequential Minimal Optimization (SMO) with a polynomial kernel and a trade-off parameter value of 1 was used for SVM ensemble. Finally, Regression by Discretization based on Random Forest (RD-RF) employs a classifier (random forest, in this case) on a copy of the data in which the property/activity value is discretized with equal width. The predicted value is the expected value of the mean class value for each discretized interval (based on the predicted probabilities for each interval). In this study, we used the random forest classification algorithm [37]. The success of this technique is based on the use of numerous trees developed with different independent variables that are randomly selected from the complete original set of variables. The number of bins for discretization was fixed at 10 and the number of trees fitted was established at 100. WEKA open source software [38] was used to implement all three techniques used.

Several approaches can be used to test the accuracy of supervised learning algorithms. We used the common method of *k*-fold cross validation. In this process, the data set is divided into *k* subsets. In each application, one of the *k* subsets is used as the test set and the other *k*-1 subsets form the training set. Error statistics are calculated across all *k* trials. This provides a good indication of how well the classifier will perform on unseen data. We used *k*=10 and compute several standard performance metrics to calculate the goodness-of-fit statistics for each technique.

We compared the performance of the MLTs by using the root mean squared error (RMSE), the coefficient of determination (R^2) and a paired T-test (corrected) based on Student's *t*-criterion. Finally, we used the selected MLT to generate above-ground biomass and carbon maps for the study area.

3. Results

3.1 Tree-level biomass equations

The parameter estimates and the goodness-of-fit statistics for the tree biomass component equations for each species are presented in Table 2. All the parameter estimates were significant at α=0.05. Stem wood biomass and AGB estimates were the most accurate, as indicated by the R^2 and RMSE values, whereas foliage and branch biomass estimates were the least accurate. The selected equations fitted the data well and generally explained between 70 and 98% of the observed total and per component biomass variation for all species. Nevertheless, as mentioned above, the explained variation was lower than 70% for two of the components: branch biomass in *P. engelmannii* (60%) and foliage biomass in *P. herrerae* (61%).

Table 2. Estimated parameters and goodness of fit statistics for the biomass equations for temperate forest species.

Tree biomass component equation	Goodness of fit		Tree biomass component equation	Goodness of fit	
	R^2	RMSE		R^2	RMSE
P. cooperi			*P. douglasiana* and *P. michoacana*		
$W_w = 0.031126d^{2.09355}h^{0.768845}$	0.98	36.47	$W_w = 0.11997d^{2.34448}$	0.90	67.17
$W_b = 0.011361d^{1.676006}h^{0.74627}$	0.82	9.91	$W_b = 0.03126d^{1.978201}$	0.78	6.75
$W_{br} = 0.007965d^{1.599044}h^{1.347388}$	0.90	25.52	$W_{br} = 0.014982d^{2.40887}$	0.90	11.01
$W_f = 0.049925d^{1.122846}h^{0.600293}$	0.75	3.69	$W_f = 0.01168d^{2.148631}$	0.84	4.09
$W_t = \Sigma W_i$	0.97	61.67	$W_t = \Sigma W_i$	0.92	74.28
P. durangensis			*Juniperus depeanna*		
$W_w = 0.01204d^{1.76074}h^{1.45047}$	0.97	39.67	$W_w = 0.01289d^2h$	0.91	28.34
$W_b = 0.01706d^{1.37067}h^{1.10922}$	0.87	8.40	$W_b = 0.000772d^2h$	0.61	2.99
$W_{br} = 0.14589d^{1.64608}h^{0.41109}$	0.71	36.56	$W_{br} = 0.00204d^2h$	0.84	5.36
$W_f = 0.00301d^{1.50863}h^{1.15013}$	0.83	3.35	$W_f = 0.00098d^2h$	0.71	3.69
$W_t = \Sigma W_i$	0.97	40.24	$W_t = \Sigma W_i$	0.90	36.89
P. engelmannii			*Arbutus bicolor*		
$W_w = 0.09798d^{1.67370}h^{1.02867}$	0.91	126.71	$W_w = 0.06438d^{1.66448}h^{0.86489}$	0.94	14.45
$W_b = 0.037974d^{1.11488}h^{0.88389}$	0.98	1.62	$W_b = 0.00328d^{1.51833}h^{0.95083}$	0.89	0.77
$W_{br} = 1.39092d^{1.25795}h^{0.05199}$	0.60	37.38	$W_{br} = 0.02125d^{1.81216}h^{0.8198}$	0.95	6.81
$W_f = 0.069316d^{1.24804}h^{0.34853}$	0.71	3.74	$W_f = 0.00698d^{1.50325}h^{0.84794}$	0.81	1.64
$W_t = \Sigma W_i$	0.92	136.82	$W_t = \Sigma W_i$	0.95	22.68
P. leiophylla			*Q. sideroxyla*		
$W_w = 0.015582d^2h$	0.94	67.23	$W_w = 0.033462d^{1.739234}h^{1.111268}$	0.91	65.24
$W_b = 0.001074d^2h$	0.90	6.49	$W_b = 0.03022d^{1.66927}h^{0.80238}$	0.79	27.73
$W_{br} = 0.007269d^2h$	0.65	65.53	$W_{br} = 0.005154d^{1.92033}h^{1.06618}$	0.81	27.16

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 May 2017 doi:10.20944/preprints201705.0178.v1

Peer-reviewed version available at Forests 2017, 8, 269; doi:10.3390/f8080269

Tree biomass component equation	Goodness of fit R^2	RMSE	Tree biomass component equation	Goodness of fit R^2	RMSE
$W_f = 0.000343d^2h$	0.81	2.51	$W_f = 0.00092d^{1.57684}h^{1.34236}$	0.74	3.43
$W_t = \Sigma W_i$	0.93	104.34	$W_t = \Sigma W_i$	0.92	94.71
P. herrerae			*Q. rugosa*		
$W_w = 0.06741d^{2.520356}$	0.83	111.87	$W_w = 0.01988d^{2.28684}h^{0.52175}$	0.91	28.33
$W_b = 0.02468d^{1.98837}$	0.72	7.35	$W_b = 0.05621d^{2.0764}$	0.78	14.84
$W_{br} = 0.05085d^{2.03574}$	0.72	17.88	$W_{br} = 0.11276d^{1.52164}h^{0.53343}$	0.87	12.28
$W_f = 0.05437d^{1.64335}$	0.61	5.36	$W_f = 0.0377d^{1.42193}h^{0.70675}$	0.84	5.15
$W_t = \Sigma W_i$	0.86	121.45	$W_t = \Sigma W_i$	0.94	40.78
P. teocote			*Q. durifolia*		
$W_w = 0.04428d^{1.8953}h^{0.84674}$	0.93	56.20	$W_w = 0.1254d^{1.7402}h^{0.8114}$	0.94	55.41
$W_b = 0.00341d^{1.96215}h^{0.66178}$	0.94	8.91	$W_b = 0.00235d^{2.30073}h^{0.68619}$	0.88	9.48
$W_{br} = 0.0051d^{1.78757}h^{1.12341}$	0.93	2.99	$W_{br} = 0.00355d^{2.66464}h^{0.56009}$	0.90	35.29
$W_f = 0.03041d^{1.78361}$	0.74	16.22	$W_f = 0.0073d^{2.05034}h^{0.48685}$	0.86	6.83
$W_t = \Sigma W_i$	0.95	4.03	$W_t = \Sigma W_i$	0.95	79.37
P. lumholtzii			*Q. crassifolia*		
$W_w = 0.078527d^{2.5043}$	0.91	63.91	$W_w = 0.02542d^2h$	0.83	55.92
$W_b = 0.022809d^{1.93859}$	0.77	16.73	$W_b = 0.00305d^2h$	0.82	7.97
$W_{br} = 0.03461d^{2.1666}$	0.84	4.10	$W_{br} = 0.01113d^2h$	0.83	25.48
$W_f = 0.011642d^{2.16799}$	0.83	4.07	$W_f = 0.00165d^2h$	0.66	4.81
$W_t = \Sigma W_i$	0.90	81.33	$W_t = \Sigma W_i$	0.91	63.54
P. strobiformis			*All pine species*		
$W_w = 0.00716d^{2.02253}h^{1.30938}$	0.90	85.96	$W_w = 0.0291d^{1.74165}h^{1.16614}$	0.92	70.42
$W_b = 0.03088d^{1.10021}h^{1.09925}$	0.88	6.31	$W_b = 0.02029d^{1.33299}h^{0.92887}$	0.67	11.02
$W_{br} = 0.01613d^{1.90578}h^{0.70112}$	0.93	14.99	$W_{br} = 0.02508d^{1.83773}h^{0.54626}$	0.66	29.07
$W_f = 0.03886d^{1.53515}h^{0.31776}$	0.76	5.16	$W_f = 0.05227d^{1.28231}h^{0.43275}$	0.57	5.68
$W_t = \Sigma W_i$	0.94	88.67	$W_t = \Sigma W_i$	0.93	81.12
P. oocarpa			*All oak species*		
$W_w = 0.01753d^{1.8261}h^{1.28397}$	0.94	25.93	$W_w = 0.11618d^{1.77395}h^{0.68708}$	0.79	93.13
$W_b = 0.02898d^{2.08978}$	0.89	4.11	$W_b = 0.00827d^{2.54589}$	0.66	27.19
$W_{br} = 0.00948d^{2.7493}$	0.88	14.84	$W_{br} = 0.0502d^{1.97638}h^{0.34229}$	0.58	53.03
$W_f = 0.04163d^{1.93601}$	0.84	3.89	$W_f = 0.08234d^{1.59396}$	0.34	11.53
$W_t = \Sigma W_i$	0.95	35.93	$W_t = \Sigma W_i$	0.82	133.68

The coefficients of determination were usually highest for stem wood biomass and AGB, and lowest for branch and foliage biomass components. The average variance explained by the biomass equations was as follows: stem wood (91% ± 5.3%), branches (84% ± 9.4%), stem bark (83% ± 9.2%) and foliage (78% ± 7.3%), and the variance explained by the AGB equations was 93% (± 3.6%). The RMSE values ranged between 14.5 kg and 126.7 kg (stem wood biomass), 0.77 kg and 27.7 kg (stem bark biomass), 2.99 kg and 48.1 kg (branches biomass), 1.64 kg and 16.2 kg (foliage biomass) and 4.03 kg and 136.8 kg (AGB). The average RMSE values for AGB of all pine and all oak species groups were 71.2 kg (± 39.5 kg) and 74.1 kg (± 23.2 kg), respectively.

The allometric equation (eq. 1) produced the best estimates of all biomass components of *P. herrerae*, *P. lumholtzii*, *P. leiophylla*, *P. douglasiana*, *P. michoacana* and *P. oocarpa* (except stem wood biomass), as well as for the branch and foliage biomass of *P. leiophylla*, foliage biomass of *P. teocote*, stem bark biomass of *Q. rugosa*, and stem bark and foliage biomass of all of the oak species.

A strong relationship between biomass components and total height was found for the groups of pine and oak species (Figure 2). The biomass equation including total height (*h*) as a predictor was the best equation for all biomass components of *P. cooperi*, *P. durangensis*, *P. engelmannii*, *P. teocote*

(except foliage biomass), *P. strobiformis*, *Arbutus bicolor*, *Q. sideroxyla*, *Q. rugosa* (except stem bark biomass) and *Q. durifolia*, as well as for all of the pine species and for stem wood and branch biomass of all of the oak species. AGB and all biomass components of *Juniperus depeanna*, *P. leiophylla* and *Q. crassifolia* were best estimated using the combination of diameter at breast height and total height (d^2h) (eq. 3).

The biomass equation including total height (h) as a predictor was the best equation for all biomass components of *P. cooperi*, *P. durangensis*, *P. engelmannii*, *P. teocote* (except foliage biomass), *P. strobiformis*, *Arbutus bicolor*, *Q. sideroxyla*, *Q. rugosa* (except stem bark biomass) and *Q. durifolia*, as well as for all of the pine species and for stem wood and branch biomass of all of the oak species. AGB and all biomass components of *Juniperus depeanna*, *P. leiophylla* and *Q. crassifolia* were best estimated using the combination of diameter at breast height and total height (d^2h) (eq. 3).

3.2 Above-ground biomass allocation

Biomass distribution varied among tree components and between species. For all species, the largest amount of biomass was located in the stem wood (Figure 2). Stem wood biomass of pine ranged from 63% (*P. oocarpa*) to 80.9% (*P. lumholtzii*), while for oak species the values varied between 59.5% (*Q. rugosa*) and 65.3% (*Q. durifolia*). General equations for all pine and all oak species produced stem wood biomass values of 75.5 and 62.9%, respectively.

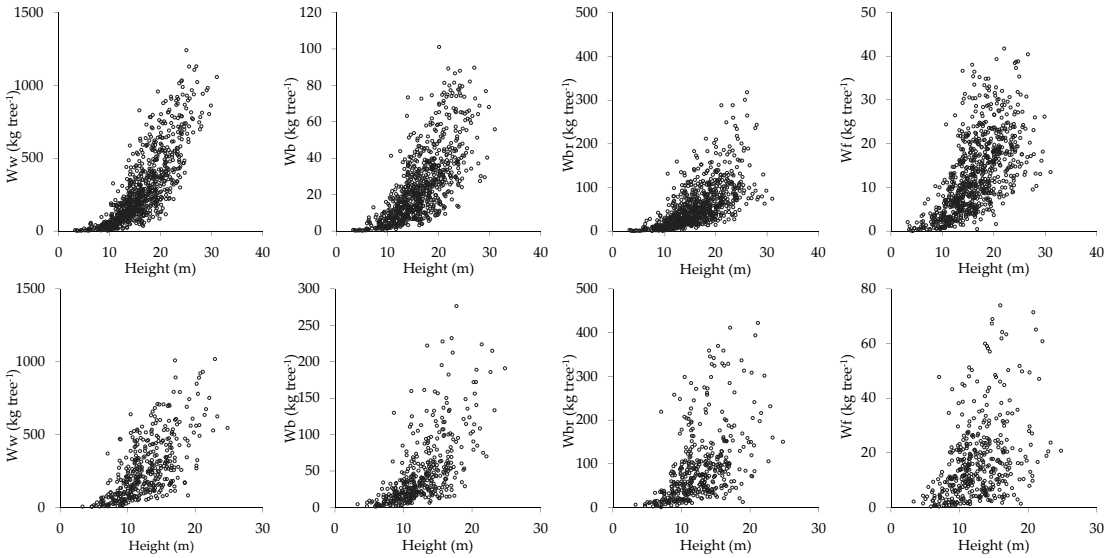

Figure 2. Scatter plots showing the relationship between tree biomass components (stem wood, stem bark, branches and foliage) and total height (h).

The lowest values for stem bark, branches and foliage biomass were respectively 2.4% (*J. depeanna*), 11.6% (*P. lumholtzii*), and 2.4% (*P. herrerae*), and the highest values of these components were 18.7% (*Q. rugosa*), 30.5% (*A. bicolor*), and 7.1% (*P. oocarpa*), respectively. For all pine species pooled together, the highest quantity of biomass was contained in stem wood (75.4%); followed by branches (14.9%), stem bark (6.4%) and foliage (3.3%). For all oak species the biomass values for the different components were respectively 62.9%, 22.2%, 10.9% and 4.0%.

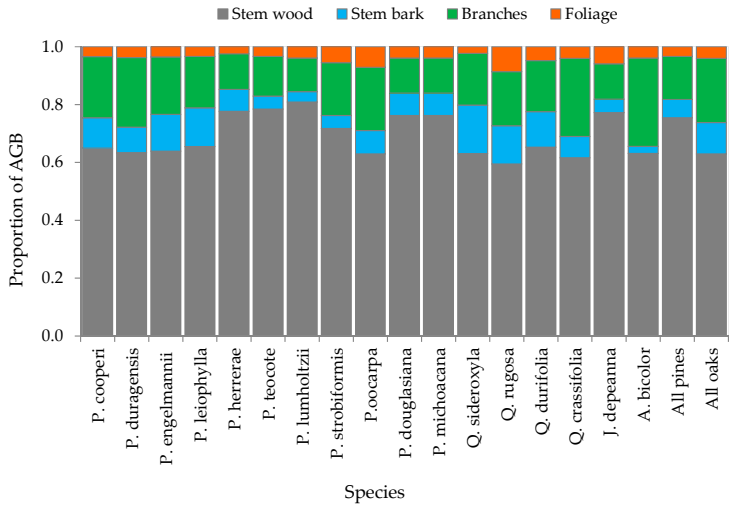

Figure 3. Estimated relative contributions of the biomass components to total tree above-ground biomass (AGB).

The biomass allocation is highlighted per species group in Figure 4. The proportion of stem biomass in pine trees increased with *d* (exceeding 50% of AGB for all diameters); although the proportion of stem bark and foliage biomass decreased, it was fairly stable for branch biomass. In all of the oak species, stem wood and branch biomass increased from 0.25 and 0.14 for *d*=5 cm to 0.42 and 0.23 for *d*>50 cm, respectively. Stem bark of oaks decreased with *d* from 0.16 to about 0.10, while foliage biomass decreased from 0.13 to 0.03.

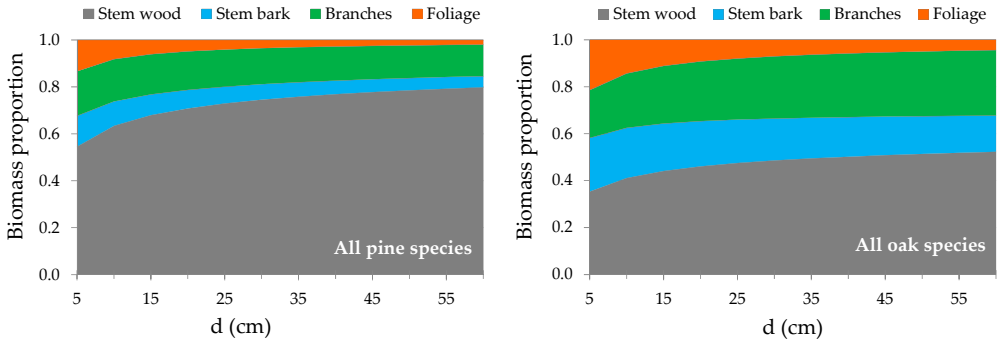

Figure 4. Tree-level biomass allocation patterns as a function of individual tree size (*d*), estimated by the proposed biomass equation systems.

3.3. Carbon fractions in different tree components

The specific carbon content (Mg of carbon per Mg of dry matter) in tree biomass components for the species evaluated is shown in Table 3. The mean values of carbon fraction of the pine species were in all cases close to the value provided by IPCC [39] (0.5), except for oak species, in which the average proportion of carbon was 0.45.

Table 3. Specific mean carbon contents of the main tree components (standard deviation is given in brackets).

Species	*n*	Carbon proportion			
		Wood	Bark	Leaves/needles	Total (s.d)
P. cooperi	27	0.485	0.511	0.471	0.489 (0.0020)
P. durangensis	27	0.489	0.531	0.487	0.505 (0.0030)
P. engelmannii	27	0.497	0.531	0.494	0.507 (0.0020)
P. leiophylla	27	0.495	0.540	0.515	0.516 (0.0023)
P. herrerae	27	0.479	0.524	0.480	0.512 (0.0025)
P. teocote	27	0.484	0.539	0.512	0.512 (0.0028)
P. lumholtzii	27	0.485	0.532	0.492	0.501 (0.0027)
P. strobiformis	27	0.494	0.533	0.492	0.506 (0.0023)
P. oocarpa	27	0.469	0.528	0.472	0.490 (0.0033)
P. douglasiana	27	0.483	0.525	0.494	0.500 (0.0021)
P. michoacana	27	0.48	0.508	0.486	0.491 (0.0015)
Q. sideroxyla	27	0.462	0.462	0.477	0.467 (0.0009)
Q. rugosa	27	0.455	0.466	0.459	0.438 (0.0034)
Q. durifolia	27	0.463	0.428	0.451	0.448 (0.0018)
Q. crassifolia	27	0.460	0.416	0.459	0.445 (0.0025)
J. depeanna	27	0.527	0.438	0.496	0.487 (0.0045)
A. bicolor	27	0.467	0.378	0.479	0.441 (0.0055)

3.4. Biomass and carbon estimates in the permanent research plots

The fitted equations were used to estimate the stand level total AGB and C contents in the 429 research plots. The total AGB ranged from 11.06 to 469.42 Mg ha^{-1} with a mean value of 129.84 Mg ha^{-1} and C contents ranged from 5.12 to 232.94 Mg ha^{-1} with a mean value of 63.80 Mg ha^{-1}. The goodness-of-fit statistics obtained by the different machine learning methods used to model AGB and carbon contents of the 429 sample plots are shown in Table 4.

Table 4. Summary of the goodness-of-fit statistics for each machine learning method using 10-fold cross-validations for AGB and carbon content. The best results are highlighted in bold type.

	AGB			**Carbon**		
	MLR	**DR-DF**	**SVM**	**MLR**	**RD-RF**	**SVM**
R^2	0.465	0.488	0.512	0.463	0.491	0.510
RMSE	51.044	49.971	49.771	25.282	24.630	24.727

The best AGB estimates were obtained by using Support Vector Machine for Regression (SVM), whereas the best estimates of carbon content were obtained using Regression by Discretization based on Random Forest (RD-RF). However, the results of the paired *t*-test (corrected) based on Student's *t*-criterion did not indicate any significance differences in the results ($\alpha = 0.05$) produced by the three approaches considered.

The variables that provided most information in the estimation of AGB and carbon content correspond to spectral bands 1 and 7; the vegetation indices NDMI, NBR2 and EVI; and the terrain variables aspect and wetness index (Figure 5). However, the results differ depending on the machine learning method used.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 May 2017 doi:10.20944/preprints201705.0178.v1

Peer-reviewed version available at Forests 2017, 8, 269; doi:10.3390/f8080269

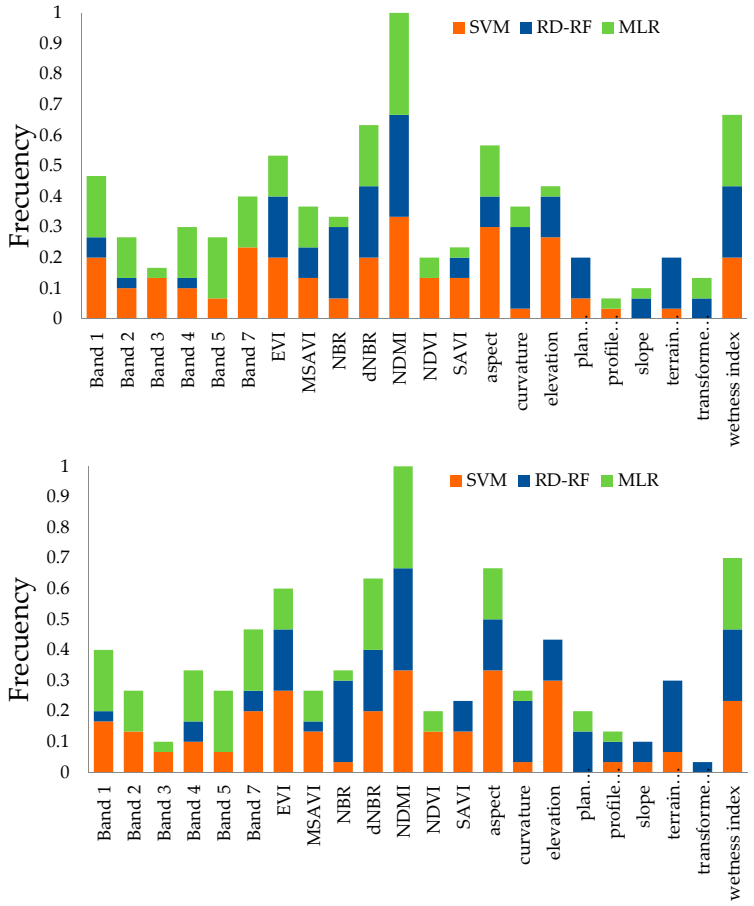

Figure 5. Frequency of appearance (importance) of each attribute (Wrapper feature selection) obtained by each machine learning technique using 10-fold cross validations for AGB (upper) and carbon content (lower).

The results obtained in terms of the absolute error of estimates for each of the 429 samples based on 10-fold cross validation were used for statistical comparison of the techniques. The comparison is summarized in histograms showing the relative positions reached (rankings) for each technique (Figure 6). In qualitative terms, the SVM technique yielded the best results.

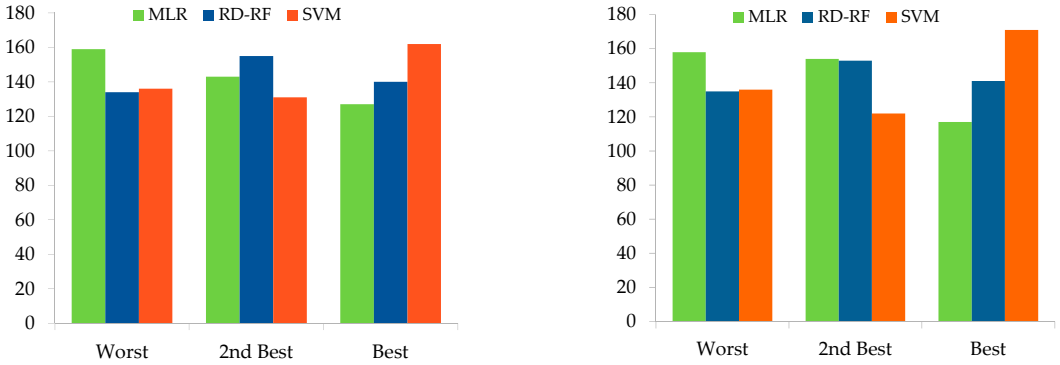

Figure 6. Percentage of sample plots achieved by each technique (ranking) for AGB (left) and carbon content (right) estimates based on 10-fold cross validation.

Finally, the spatial distributions of the estimates of total AGB and C content (Mg ha^{-1}) in the study, obtained by application of the SVM model, are shown in Figure 7.

Peer-reviewed version available at Forests 2017, 8, 269; doi:10.3390/f8080269

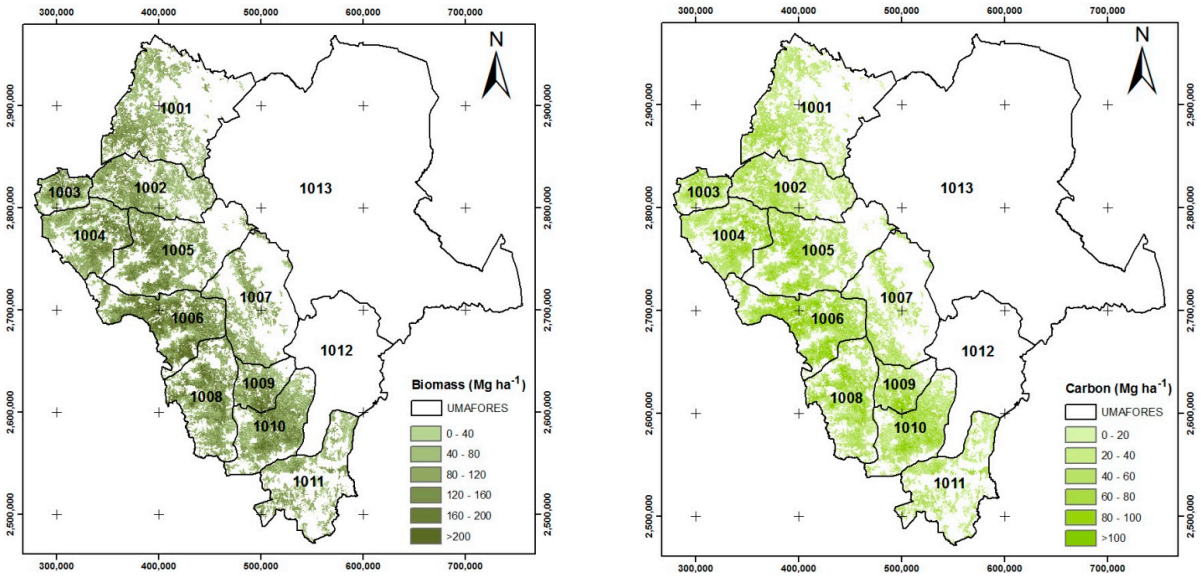

Figure 7. Spatial distribution of total AGB (left) and carbon content (right) in the study area estimated by the Support Vector Machine for Regression (SVM) approach.

4. Discussion

4.1 The allometric equations

The equations were mainly based on three model formulations (eq. 1-3) used in previous studies [2, 14, 40-42]. Diameter at breast height was the main explanatory tree variable used to estimate the tree biomass components for all species. Several authors have noted that inclusion of total height does not usually lead to a substantial increase in predictive ability precision of diameter-based biomass regressions, and they also assume that *d* is sufficient to obtain a reliable tree biomass prediction [43-45]. However, other authors have found a significant improvement in model fits [46, 47] or an increase in the accuracy of the biomass estimates [42, 48] when *h* is also used as a predictor. In this study, *d* by itself was a good predictor of biomass, but the addition of *h* as a second variable improved the predictions for several species. A strong relationship between biomass components and total height was found for the groups of pine and oak species, and it was a significant predictor of all tree biomass components and AGB of 12 out of 17 species and for the group of all pine species, as well as for stem wood and branches biomass of the oak species group. The *d* and *h* based system of equations yielded poor fits for branch and foliage biomass of *P. oocarpa* and *P. teocote*. Lambert *et al*. [49], Bi *et al*. [50] and Zhao *et al*. [51] also found that inclusion of tree height improved the accuracy of predicting the stem biomass but not the crown (needles and branches) biomass components.

Feldpausch *et al*. [52] reported that the mean relative error in biomass estimates when *h* was included was 50% less that when *h* was excluded. Similarly, we found that inclusion of total height in the biomass equations increased the accuracy of the biomass estimations (measured as RMSE) by between 10.4% (*P. teocote*) and 53.9% (*P. durangensis*) for stem wood biomass, by between 2.9% (*A. bicolor*) and 28.1% (*P. strobiformis*) for stem bark biomass, by between 13.1% (*Q. durifolia*) and 20.6% (*P. strobiformis*) for branch biomass, by between 1.1% (*P. engelmannii*) and 19.7% (*Q. sideroxyla*) for foliage biomass and by between 26.7% (*Q. rugosa*) and 55.9% (*P. cooperi*) for total tree AGB. These results are consistent with those of António *et al*. [40] who found that the use of height implied a

Peer-reviewed version available at Forests 2017, 8, 269; doi:10.3390/f8080269

decrease in the sum of squares of residuals by 72%, 8%, 12% and 10%, for stem wood, stem bark, leaves and branches respectively. They are also consistent with the findings of Li and Zhao [53] who reported that the height improved model performance of the fitted equation especially for total above-ground biomass, stem wood biomass and stem bark biomass.

Branch and foliage biomass components are always difficult to estimate with the same accuracy as bole or total AGB. This is illustrated by the fact that inclusion of height in the model only improved the accuracy of branch and foliage biomass estimates by up to 21%. We assume, as Ter-Mikaelian and Parker [54] found, that the influence on height of stand density (i.e. competition), stand structure and site quality may account for some of this variation, and that these variables represent a large proportion of the small contribution of h to explaining the variance in these biomass components for some of the species evaluated. Nevertheless, this was not tested in the present study because the data does not enable discrimination among the effects of these factors or the different silvicultural treatments on the crown structure.

The combined variable d^2h is usually used in biomass equations [55-57], and it has found that the accuracy of biomass estimation increased significantly (measured as R^2) when h or d^2h was also included, in addition to d. In fact, tree biomass is closely correlated with d^2h as shown by Parresol [58] and Carvalho and Parresol [59]. In the models developed by Parresol [58], height was a good predictor of stem wood but not of stem bark biomass; whereas Carvalho and Parresol [59] obtained the best estimates for stem, crown and total tree biomass of Pyrenean oak including the variable d^2h as the sole independent variable in the equation; Bi *et al*. [50] reached a similar conclusion, reporting that d^2h performed better for predicting stem and bark components than diameter alone but not for branch and leaf components. In the present study, d^2h yielded the best estimates of all biomass components and AGB for *P. leiophylla*, *J. depeanna* and *Q. crassifolia*. The results reported here suggest that the best equations for biomass estimation for most species are based on d and h; it is therefore possible to use the biomass equations systems developed for a specific species across different sites in the temperate forests of Durango, considering that total height is included in the models and that the addition of this variable may take into account different levels of competition induced by different stand density conditions [40].

Differences in biomass estimates between the equations developed in this study and previously published biomass functions for the same species in northwest Mexico [60] were found. The differences with respect to the biomass estimates reported by Návar [60] may be explained by the fact that these equations only use d as an explanatory variable.

Finally, we fitted the equations for the four components simultaneously with the total AGB, in order to take into account the correlation between the errors of the models and to guarantee additivity. This restriction has not been taken into account by several authors [60] with the consequent inconsistency in total biomass estimates.

4.2 Contribution of components to total AGB

For all species evaluated, stem wood contributed most to the total AGB, followed by branches, stem bark and foliage. The fact that the highest proportion of the AGB is allocated in the stem has been widely documented, with the proportions ranging from 50 to 92% for different species [4, 61, 62]. Our results are within the same range, varying between 59.5% (*Q. rugosa*) and 80.9% (*P. lumholtzii*). The biomass allocation patterns observed are consistent with those reported by Blujdea *et al*. [41] for broadleaf trees in Rumania, Wirt *et al*. [4] for Norway spruce in Central Europe,

Correia *et al*. [63] for *P. pinea* L. in Portugal, and by Pajtík *et al*. [64] for *Picea abies* [L.] Karst in Slovakia.

As expected, the proportion of biomass allocated to stem wood increased with increasing tree size, whereas the relative contribution of stem bark and foliage to AGB decreased. This is consistent with findings reported by Blujdea *et al*. [41] and Johansson [65]. The pattern was similar for pines and oaks, except for branch biomass. In oaks, branch biomass increased with *d*, whereas for pines it remained quite constant (around 15%) over the entire diameter range. The decrease in foliage biomass with increasing *d* can be attributed to competition for light among the trees; small branches and a smaller quantity of leaves occur in larger trees or those trees with a dominant position [66]. The foliage biomass of oaks decreased to below 3.5%, while for pines it decreased to about 2%.

4.3 Machine Learning Techniques (MLT)

The variables that were most important for estimating AGB and C contents by the MLT evaluated correspond to band 1 and band 7, the vegetation indices NDMI, NBR2 and EVI, and the terrain variables aspect and wetness index. Several studies have demonstrated that such variables are usually good predictors for estimating AGB [23, 67, 68].

The SVM technique yielded the best fits, thus confirming that this type of technique is of great potential for improving biomass estimation, independently of the type of sensor to which it is applied, as demonstrated in recent studies [69]. However, no significance differences were found between this method and the other two approaches used.

The values of goodness of fit statistics for the SVM technique obtained in the present study were slightly lower than those reported by Guo *et al*. [70] who estimated AGB in *Picea crassifolia* forests in China using Landsat TM data and two non-parametric methods (*k*-Nearest Neighbour *kNN*: R^2=0.54 and RMSE=26.62 Mg ha^{-1} and SVM: R^2=0.51 and RMSE=27.45 Mg ha^{-1}). They were also lower than those reported by Tian *et al*. [71] (*kNN*: R^2=0.59) and those obtained by López-Serrano *et al*. [23] for mixed and uneven-aged forests in the Sierra Madre Occidental of Mexico (R^2=0.73 and RMSE=22.59 Mg ha^{-1}).

5. Conclusions

This study describes a set of simultaneous species-specific allometric equation systems for seventeen temperate species in north-western Mexico. Breast height diameter-only based equations proved best for predicting all biomass components of only four species. Including *h* or d^2h as additional predictor in the equation systems only slightly improves stem wood, stem bark and above-ground biomass estimates, but greatly improves the predictions of branch and foliage biomass.

In all species, most of the biomass was allocated in stem wood, followed by branches, stem bark and foliage. The biomass allocation differed between tree components and among species; however, further study is required to clarify which factors affect the allocation patterns.

The developed biomass equations can be applied to tree-level data in forest inventories and may also improve the quality of biomass estimates and verify carbon stocks changes in the temperate uneven-age multi-species forests of the study area. They are also essential tools for accurate estimation of forest residues in the development of bioenergy projects.

The results of this study regarding the use of MLT to estimate AGB and C content from remote sensors indicate that moderate resolution sensors, such as the Landsat TM5, are sufficiently reliable

and accurate for monitoring these important variables in real time and at a low cost, because the spectral data are available free of charge.

Acknowledgments: This study was financially supported by the Mexican National Council for Science and Technology (CONACyT) and by the State of Durango Government (Project FOMIX- DGO-2011-C01-165681).

Author Contributions: B.V.L conceived the research design, conducted the field data collection, analysis, and wrote the manuscript. C.A.L. conducted the analysis of satellite data. J.J.C. and J.G.A. analyzed the data and revised the manuscript. C.G.A. and J.O.L. assisted to result analysis and revised the manuscript.

Conflicts of Interest: The authors declare no conflict of interest

References

1. Hollinger, D. Defining a Landscape-Scale Monitoring Tier for the North American Carbon Program. In: *Field Measurements for Forest Carbon Monitoring*; Hoover, C.M., Ed.; Springer: Netherlands, 2008, pp. 3-17. doi: 10.1007/978-1-4020-8506-2_1.
2. Castedo-Dorado, F.; Gómez-García, E.; Diéguez-Aranda, U.; Barrio-Anta, M.; Crecente-Campo, F. Aboveground stand-level biomass estimation: a comparison of two methods for major forest species in northwest Spain. *Ann For Sci* **2012**, *69*(6), 735-46. doi: 10.1007/s13595-012-0191-6.
3. Hall, G.M.J.; Wiser, S.K.; Allen, R.B.; Beets, P.N.; Goulding, C.J. Strategies to estimate national forest carbon stocks from inventory data: the 1990 New Zealand baseline. *Glob Chang Biol* **2001**, *7*(4), 389-403. doi: 10.1046/j.1365-2486.2001.00419.x.
4. Wirth. C.; Schumacher, J.; Schulze, E.D. Generic biomass functions for Norway spruce in Central Europe—a meta-analysis approach toward prediction and uncertainty estimation. *Tree Physiol* **2004**, *24*, 121-39. doi: 10.1093/treephys/24.2.121.
5. Challenger, A. *Utilización de los ecosistemas terrestres de México. Pasado, presente y futuro*. Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO), México, 1998.
6. CONABIO. *Estrategia Nacional sobre Biodiversidad de México*. Comisión Nacional para el Conocimiento y Uso de la Biodiversidad, México, 2000.
7. SEMARNAT. Anuario Estadístico de la Producción Forestal 2013. Secretaría de Medio Ambiente y Recursos Naturales, México, 2014; 236 p.
8. García-Romero, A.; Orozco, O.; Galicia, L. Land-use systems and resilience of tropical rainforests in the Tehuantepec Isthmus, Mexico. *Environ Manage* **2004**, *34*, 768–785. doi: 10.1007/s00267-004-0178-z.
9. Masera, O.; Ordóñez, A.; Dirzo, R. Carbon emissions from Mexican forests: current situation and long-term scenarios. *Clim Change* **1997**, *24*, 256–295. doi: 10.1023/A:1005309908420.
10. Návar-Cháidez, J.J. Biomass allometry for tree species of northwestern Mexico. *Trop Subtrop Agrosyst* **2010**, *12*, 507-19. Available online: https://www.cabdirect.org/cabdirect/abstract/20103282813 (accessed on 13 September 2015).
11. Salimon, C.I.; Putz, F.E.; Menezes-Filho, L.; Anderson, A.; Silveira, M.; Brown, I.F.; *et al*. Estimating state-wide biomass carbon stocks for a REDD plan in Acre, Brazil. *For Ecol Manage* **2011**, *262*, 555-60, doi:10.1016/j.foreco.2011.04.025.
12. INEGI. Instituto Nacional de Estadística Geográfica e Informática. *Uso del Suelo y Vegetación Escala 1: 250 000 Serie V*, Información vectorial, México. 2012.
13. Parresol, B.R. Additivity of nonlinear biomass equations. *Can J For Res* **2001**, *31*(5), 865-78, doi:10.1139/x00-202.
14. Bi, H.; Long, Y.; Turner, J.; Lei, Y.; Snowdon, P.; Li, Y.; *et al.* Additive prediction of aboveground biomass for *Pinus radiata* (D. Don) plantations. *For Ecol Manage* **2010**, *259*, 2301-14, doi:10.1016/j.foreco.2010.03.003.
15. Institute S. *SAS/STAT 9.2. User´s Guide Release*. SAS Institute Inc., Cary, NC, USA, 2009.
16. Greene, W.H. *Econometric Analysis*, 4th ed.; Prentice Hall, Upper Saddle River, N.J., USA, 1999.
17. Corral-Rivas, J.J.; Reyes, R.I.; Wehenkel, C.; Aguirre-Calderón, O.A.; Gadow, K.v. A Network of Forest Observational Studies in Durango (Mexico). In *Forest Observational Studies*; Zhao, X.H.; Zhang, C.Y.; Gadow, K.v., Eds.; Beijing Forestry University, Beijing, China, 2012; pp. 125-138.
18. Rojas-García, F.; De Jong, B.J.; Martínez-Zurimendí, P.; Paz-Pellat, F. Database of 478 allometric equations to estimate biomass for Mexican trees and forests. *Ann For Sci* **2015**, *72*(6), 835-64, doi:10.1007/s13595-015-0456-y.
19. NASA. *Landsat 7 Science Data User's Handbook*. NASA, USA, 2011.

20. Rouse, J.W.; Haas, R.H.; Deering, D.W. Monitoring vegetation systems in the Great Plains with ERTS. *Remote Sens* **1974**, *1*, 309-313. Available online: https://ntrs.nasa.gov/search.jsp?R=19740022614 (accessed on 13 September 2015).
21. Huete, A.R. A soil-adjusted vegetation index (SAVI). *Rem Sen Environ* **1988**, *25*, 295-309, doi:10.1016/0034-4257(88)90106-X.
22. Glenn, E.; Huete, A.; Nagler, P.; Nelson, S. Relationship between remotely-sensed Vegetation Indices, Canopy Attributes and Plant Physiological Processes: What Vegetation Indices Can and Cannot Tell Us About the Landscape. *Sensors* **2008**, *8*, 2136-60, doi:10.3390/s8042136.
23. López-Serrano, P.; López-Sánchez, C.; Díaz-Varela, R.; Corral-Rivas, J.; Solís-Moreno, R.; Vargas-Larreta, B.; *et al*. Estimating biomass of mixed and uneven-aged forests using spectral data and a hybrid model combining regression trees and linear models. *iForest - Biogeosciences and Forestry* (early view) **2015**, doi:10.3832/ifor1504-008 [online 2015-09-21].
24. McNab, W.H. Terrain Shape Index: Quantifying effect of minor landforms on tree height. *For Sci* **1989**, *35*(1), 91-104.
25. Moore, I.D.; Norton, T.W.; Williams, J.E. Modelling environmental heterogeneity in forested landscapes. *J Hydrol* **1993**, *150*(2-4), 717-47, doi:10.1016/0022-1694(93)90133-T.
26. INEGI. *Mexican continuous elevation 3.0 (CEM3.0)*. [Mexican continuous elevation 3.0 (CEM3.0)]. Instituto Nacional de Estadística Geográfica e Informática, México, 2004. Available online: http://www.inegi.org.mx/geo/contenidos/datosrelieve/continental/descarga.aspx (Accessed on 08 May 2015).
27. Wilson, J.P.; Gallant, J.C. Digital terrain analysis. In *Terrain Analysis: Principles and Applications,* Wilson, J.P.; Gallant, J.C.,Eds.; John Wiley and Sons, Inc., New York, USA, 2000; pp. 1-26.
28. R, C.T. *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria, 2014.
29. Fassnacht, F.E.; Hartig, F.; Latifi, H.; Berger, C.; Hernández, J.; Corvalán, P.; *et al*. Importance of sample size, data type and prediction method for remote sensing-based estimations of aboveground forest biomass. *Remote Sens Environ* **2014**, *154*, 102-14, doi:10.1016/j.rse.2014.07.028.
30. Akaike, H. A new look at the statistical model identification. *IEEE Trans Autom Control* **1974**, *19*(6), 716-723, doi:10.1109/TAC.1974.1100705.
31. Boser, B.E.; Guyon, I.M.; Vapnik, V.N. A training algorithm for optimal margin classifiers. In Proceedings of the Fifth Annual Workshop on Computational Learning Theory, Pittsburgh, USA, 1992; p. 144-152.
32. Vapnik, V. *The Nature of Statistical Learning Theory*. Springer, New York, USA, 1995.
33. Cristianini, N.; Shawe-Taylor, J. An Introduction to Support Vector Machines: And other Kernel-Based Learning Methods. Cambridge University Press, Cambridge, England, 2000.
34. Huang, D.; Knyazikhin, Y.; Wang, W.; Deering, D.W.; Stenberg, P.; Shabanov, N.; *et al*. Stochastic transport theory for investigating the three-dimensional canopy structure from space measurements. *Rem Sen Environ* **2008**, *11*(1), 35-50, doi:10.1016/j.rse.2006.05.026.
35. Kuemmerle, T.; Hostert, P.; St-Louis, V.; Radeloff, V.C. Using image texture to map farmland field size: a case study in Eastern Europe. *J Land Use Sci* **2009**, *4*(1-2), 85-107, doi:10.1080/17474230802648786.
36. Shevade, S.K.; Keerthi, S.S.; Bhattacharyya, C.; Murthy, K.R.K. Improvements to the SMO algorithm for SVM regression. Neural Networks, *IEEE Trans Neural Networks* **2000**, *11*(5), 1188-93, doi:10.1109/72.870050.
37. Breiman L. Random Forests. *Mach Learn* **2001**, *45*(1), 5-32, doi:10.1023/A:1010933404324.
38. Hall, M.; Frank, E.; Holmes, G.; Pfahringer, B.; Reutemann, P.; Witten, I.H. The WEKA data mining software: an update. SIGKDD *Explor* **2009**, *11*(1), 10-18, doi:10.1145/1656274.1656278.
39. IPCC. *Guidelines for National Greenhouse Gas Inventories*, National Greenhouse Gas Inventories Programme, Japan, 2006.
40. António, N.; Tomé, M.; Tomé, J.; Soares, P.; Fontes, L. Effect of tree, stand, and site variables on the allometry of Eucalyptus globulus tree biomass. *Can J For Res* **2007**, *37*(5), 895-906, doi:10.1139/X06-276.
41. Blujdea, V.N.B.; Pilli, R.; Dutca, I.; Ciuvat, L.; Abrudan, I.V. Allometric biomass equations for young broadleaved trees in plantations in Romania. *For Ecol Manage* **2012**, *264*, 172-184, doi:10.1016/j.foreco.2011.09.042.
42. González-García, M.; Hevia, A.; Majada, J.; Barrio-Anta, M. Above-ground biomass estimation at tree and stand level for short rotation plantations of *Eucalyptus nitens* (Deane & Maiden) Maiden in Northwest Spain. *Biomass Bioener* **2013**, *54*, 147-57, doi:10.1016/j.biombioe.2013.03.019.

43. Johansson, T. Biomass equations for determining fractions of pendula and pubescent birches growing on abandoned farmland and some practical implications. *Biomass Bioener* **1999**, *16*(3), 223-38, doi:10.1016/S0961-9534(98)00075-0.
44. Porté, A.; Trichet, P.; Bert, D.; Loustau, D. Allometric relationships for branch and tree woody biomass of Maritime pine (*Pinus pinaster* Ait.). *For Ecol Manage* **2002**, *158*(1-3), 71-83, doi:10.1016/S0378-1127(00)00673-3.
45. Jenkins, J.; Chojnacky, D.S.; Birdsey, R. National scale biomass estimators for United States tree species. *For Sci* **2003**, *49*(1), 12-35.
46. Reed, D.; Tomé, M. Total aboveground biomass and net dry matter accumulation by plant component in young *Eucalyptus globulus* in response to irrigation. *For Ecol Manage* **1998**, *103*, 21-32, doi:10.1016/S0378-1127(97)00174-6.
47. Bartelink, H.H. Allometric relationships on biomass and needle area of Douglas-fir. *For Ecol Manage* **1996**, *86*(1-3), 193-203, doi:10.1016/S0378-1127(96)03783-8.
48. Menéndez-Miguélez, M.; Canga, E.; Barrio-Anta, M.; Majada, J.; Álvarez-Álvarez, P. A three level system for estimating the biomass of *Castanea sativa* Mill. coppice stands in north-west Spain. *For Ecol Manage* **2013**, *291*, 417-26, doi:10.1016/j.foreco.2012.11.040.
49. Lambert, M.C.; Ung, C.H.; Raulier, F. Canadian national tree aboveground biomass equations. *Can J For Res* **2005**, *35*(8), 1996-2018, doi:10.1139/x05-112.
50. Bi, H.; Turner, J.; Lambert, M. Additive biomass equations for native eucalypt forest trees of temperate Australia. *Trees* **2004**, *18*(4), 467-79, doi:10.1007/s00468-004-0333-z.
51. Zhao, D.; Kane, M.; Markewitz, D.; Teskey, R.; Clutter, M. Additive tree biomass equations for midrotation loblolly pine plantations. *For Sci* **2015**, *61*(4), 613-23.
52. Feldpausch, T.R.; Lloyd, J.; Lewis, S.L.; Brienen, R.J.; Gloor, M.; Monteagudo, M.A.; *et al*. Tree height integrated into pantropical forest biomass estimates. *Biogeosciences* **2012**, *9*(8), 3381-3403, doi:10.5194/bg-9-3381-2012.
53. Li, H.; Zhao, P. Improving the accuracy of tree-level aboveground biomass equations with height classification at a large regional scale. *For Ecol Manage* **2013**, 289, 153-63, doi:10.1016/j.foreco.2012.10.002.
54. Ter-Mikaelian, M.; Parker, W. Estimating biomass of white spruce seedlings with vertical photo imagery. *New For* **2000**, *20*(2), 145-62, doi:10.1023/A:1006716406751.
55. Chave, J.; Andalo, C.; Brown, S.; Cairns, M.A.; Chambers, J.Q.; Eamus, D.; *et al*. Tree allometry and improved estimation of carbon stocks and balance in tropical forests. *Oecologia* **2005**, 145, 87-99, doi:10.1007/s00442-005-0100-x.
56. Schmidt, A.; Poulain, M.; Klein, D.; Krause, K.; Peña-Rojas, K.; Schmidt, H.; *et al*. Allometric above-belowground biomass equations for *Nothofagus pumilio* (Poepp. & Endl.) natural regeneration in the Chilean Patagonia. *Ann For Sci* **2009**, *66*(5), 513-513, doi:10.1051/forest/2009030.
57. Zeng, W.; Tang, S.; Huag, G.; Zhang, M. Population classification and sample structure on modeling of single tree biomass equations for national biomass estimation in China. *For Rest Manage* **2010**, *3*, 16-23. Available online: http://en.cnki.com.cn/Article_en/CJFDTOTAL-LYZY201003005.htm.
58. Parresol, B.R. Assessing tree and stand biomass: A Review with examples and critical comparisons. *For Sci* **1999**, *45*(4), 573-93.
59. Carvalho, J.P.; Parresol, B.R. Additivity in tree biomass components of Pyrenean oak (*Quercus pyrenaica* Willd.). *For Ecol Manage* **2003**, *179*(1-3), 269-76, doi:10.1016/S0378-1127(02)00549-2.
60. Návar, J. Allometric equations for tree species and carbon stocks for forests of northwestern Mexico. *For Ecol Manage* **2009**, *257*, 427-434, doi:10.1016/j.foreco.2008.09.028.
61. Fonseca, W.; Rey, B.J.M.; Alice, F.E. Carbon accumulation in the biomass and soil of different aged secondary forests in the humid tropics of Costa Rica. *For Ecol Manage* **2011**, *262*(8), 1400-8, doi:10.1016/j.foreco.2011.06.036.
62. Redondo-Brenes, A.; Montagnini, F. Growth, productivity, aboveground biomass, and carbon sequestration of pure and mixed native tree plantations in the Caribbean lowlands of Costa Rica. *For Ecol Manage* **2006**, *232*(1-3), 168-78, doi:10.1016/j.foreco.2006.05.067.
63. Correia, A.C.; Tomé, M.; Pacheco, C.A.; Faias, S.; Días, A.C.; Freire, J.; *et al*. Biomass allometry and carbon factors for a Mediterranean pine (*Pinus pinea* L.) in Portugal. *For Syst* **2010**, *19*(3), 418-33, doi:10.5424/fs/2010193-9082.

64. Pajtík, J.; Konôpka, B.; Lukac, M. Biomass functions and expansion factors in young Norway spruce (*Picea abies* [L.] Karst) trees. *For Ecol Manage* **2008**, *256*(5), 1096-1103, doi: 10.1016/j.foreco.2008.06.013.
65. Johansson, T. Biomass equations for hybrid larch growing on farmland. *Biomass Bioener* **2013**, *49*, 152-59, doi:10.1016/j.biombioe.2012.12.020.
66. Aquino-Ramírez, M.; Velázquez-Martínez, A.; Castellanos-Bolaños, J.F.; De los Santos-Posadas, H.M.; Etchevers-Barra, J.D. Partición de la biomasa aérea en tres especies arbóreas tropicales. *Agrociencia* **2015**, *49*(3), 299-314. Available online: http://www.redalyc.org/articulo.oa?id=30238027006 (Accessed on 17 May 2016).
67. Lu, D.; Chen, Q.; Wang, G.; Moran, E.; Batistella, M.; Zhang, M.; *et al*. Aboveground forest biomass estimation with Landsat and LiDAR data and uncertainty analysis of the estimates. *Int J For Res* **2012**, *2012*, 1-16, doi:10.1155/2012/436537.
68. Lu, D.; Chen, Q.; Wang, G.; Liu, L.; Li, G.; Moran, E. A survey of remote sensing-based aboveground biomass estimation methods in forest ecosystems. *International Journal of Digital Earth* **2014**, 1-43, doi:10.1080/17538947.2014.990526.
69. García-Gutiérrez, J.; Martínez-Álvarez, F.; Troncoso, A.; Riquelme, J.C. A comparison of machine learning regression techniques for LiDAR-derived estimation of forest variables. *Neurocomputing* **2015**, *167*, 24-31, doi:10.1016/j.neucom.2014.09.091.
70. Guo, Z.D.; Hu, H.F.; Pan, Y.D.; Birdsey, R.A.; Fang, J.Y. Increasing biomass carbon stocks in trees outside forests in China over the last three decades. *Biogeosciences* **2014**, *11*(15), 4115-22, doi:10.5194/bg-11-4115-2014.
71. Tian, X.; Li, Z.; Su, Z.; Chen, E.; van der Tol, C.; Li, X.; *et al*. Estimating montane forest above-ground biomass in the upper reaches of the Heihe River Basin using Landsat-TM data. *Int J Remote Sen* **2014**, *35*(21), 7339-62, doi:10.1080/01431161.2014.967888.