# A Comparison of Machine Learning Techniques Applied to Landsat-5 TM Spectral Data for Biomass Estimation

**Pablito M. López-Serrano[1], Carlos A. López-Sánchez[2], Juan G. Álvarez-González[3], and Jorge García-Gutiérrez[4,*]**

*[1]Ciencias Agropecuarias y Forestales, Universidad Juárez del Estado de Durango, Negrete 800, Centro, 34000 Durango, Dgo., Mexico*

*[2]Instituto de Silvicultura e Industria de la Madera, Universidad Juárez del Estado de Durango, Universidad Juárez del Estado de Durango, Negrete 800, Centro, 34000 Durango, Dgo., Mexico [3]Departamento de Ingeniería Agroforestal, Universidad de Santiago de Compostela, Avenida Dr.Ángel Echeverri, s/n. Campus Vida, 15782 Santiago de Compostela, C, Spain*

*[4]Departamento de Lenguajes y Sistemas Informáticos, Universidad de Sevilla, Reina Mercedes s/n., Sevilla 41012, Spain*

**Abstract. Machine learning combines inductive and automated techniques for recognizing patterns. These techniques can be used with remote sensing datasets to map aboveground biomass (AGB) with an acceptable degree of accuracy for evaluation and management of forest ecosystems. Unfortunately, statistically rigorous comparisons of machine learning algorithms are scarce. The aim of this study was to compare the performance of the 3 most common nonparametric machine learning techniques reported in the literature, vis., Support Vector Machine (SVM), k-nearest neighbor (kNN) and Random Forest (RF), with that of the parametric multiple linear regression (MLR) for estimating AGB from Landsat-5 Thematic Mapper (TM) spectral reflectance data, texture features derived from the Normalized Difference Vegetation Index (NDVI), and topographical features derived from a digital elevation model (DEM). The results obtained for 99 permanent sites (for calibration/validation of the models) established during the winter of 2011 by systematic sampling in the state of Durango (Mexico), showed that SVM performed best once the parameterization had been optimized. Otherwise, SVM could be outperformed by RF. However, the kNN yielded the best overall results in relation to the goodness-of-fit measures. The findings confirm that nonparametric machine learning algorithms are powerful tools for estimating AGB with datasets derived from sensors with medium spatial resolution.**

**Résumé. L'apprentissage automatique combine des techniques inductives et automatisées pour la reconnaissance des formes. Ces techniques peuvent être utilisées avec des ensembles de données de télédétection pour cartographier la biomasse aérienne « aboveground biomass » (AGB) avec un degré de précision acceptable pour l'évaluation et la gestion des écosystèmes forestiers. Malheureusement, des comparaisons statistiquement rigoureuses des algorithmes d'apprentissage automatique sont rares. Le but de cette étude était de comparer les performances des 3 méthodes d'apprentissage automatique non paramétriques les plus fréquemment rapportées dans la littérature, vis., les machines à vecteurs de support « Support Vector Machine » (SVM), les *k* plus proches voisins « k-nearest neighbor » (kNN) et les forêts aléatoires « Random Forest » (RF), avec celle de la régression linéaire multiple paramétrique (MLR) pour l'estimation de l'AGB provenant des données de réflectance spectrale de Landsat-5 Thematic Mapper (TM), des caractéristiques de texture dérivées de l'indice de végétation par différence normalisée « Normalized Difference Vegetation Index » (NDVI) et des caractéristiques topographiques dérivées d'un modèle numérique de terrain « digital elevation model » (DEM).Les résultats obtenus pour 99 sites permanents (pour la calibration/validation des modèles) établis au cours de l'hiver 2011 par l'échantillonnage systématique dans l'État de Durango (Mexique), ont montré que les SVM montrent leurs meilleures performances une fois que le paramétrage a été optimisé. Par ailleurs, les SVM pourraient être surpassées par les RF. Cependant, les kNN ont donné les meilleurs résultats globaux par rapport aux mesures d'ajustement. Les résultats confirment que les algorithmes d'apprentissage automatique non paramétriques sont des outils puissants pour l'estimation de l'AGB avec des ensembles de données provenant de capteurs avec une résolution spatiale moyenne.**

## INTRODUCTION

Forest biomass plays an important role in the global climate system because forest ecosystems absorb approximately 1/12 of Earth's atmospheric carbon stocks every year (Malhi et al. 2002), and much of this carbon is stored as aboveground biomass (AGB). The importance of forest biomass has been underlined by the United Nations Framework Convention on Climate Change (UNFCCC), which has identified AGB as an Essential Climate Variable (GCOS 2010). Moreover, quantification of AGB and modeling of the associated dynamics are important

*Corresponding author e-mail: jorgarcia@us.es.

to support decision-making models in different fields, including energy and materials provision for human use (FAO 2001, 2006), forest fragmentation (e.g., Malhi and Phillips 2004), and biodiversity conservation (e.g., Bunker et al. 2005). Accurate monitoring of forest biomass and how it changes at local to global scales is, therefore, of critical importance toward a better understanding of these processes (Lu 2006; Hartig et al. 2012; Le Toan and Quegan 2015).

The most accurate method of estimating forest biomass is based on field measurements; however, estimating biomass in large areas is not an easy task and is hindered by the high costs (both time and money) associated with fieldwork (Lu et al. 2016).

Remote sensing has been shown to be a practical option that helps to overcome these limitations because it enables obtaining forest information in large areas with reasonable effort. This is now the primary data source for large-scale biomass estimation (e.g., Andersen et al. 2011; Lu et al. 2016). Over the past few decades, the so-called passive sensors (i.e., sensors that use the solar radiation reflected or emitted by the objects detected at Earth's surface) have been used to estimate AGB (e.g., Lu et al. 2012; Frazier et al. 2014). Considering the advantages and limitations of different remote sensing images, the medium-resolution (pixel size, 30 m) Landsat-5 TM sensor is one of the most widely used for biomass estimation (e.g., Agarwal et al. 2014; Pflugmacher et al. 2014; Dube and Mutanga 2015; Zhu and Liu 2015). The advantages of using the Landsat-5 TM sensor over high-resolution sensors, particularly for analysis of large áreas, are that numerous historical spatiotemporal archives are available (images since 1972) and the Landsat data is free of cost for users. For a review of Landsat imagery-based AGB estimations, see Wu et al. (2016).

Regardless of thept type of sensor used, model accuracy and error estimation vary in relation to a series of factors such as the structure of the field data and the statistical techniques used (Ghosh et al. 2014). The most common model used in estimating forest biomass from remote sensing data is the regression-based model (e.g., Tian et al. 2012; Lu et al. 2012; Næsset et al. 2013); however, the accuracy of estimates obtained with small numbers of sample plots or when there is a weak linear relationship between variables and biomass is rather low (Lu et al. 2016). Nonparametric modeling approaches, which make no assumptions about the statistical distributions of the original data and relationships between predictor and response variables, have also been used to relate AGB and remotely sensed features. Various recent studies have explored the use of nonparametric approaches for estimating AGB with remote sensing data (e.g., Breidenbach et al. 2012; Mutanga et al. 2012; Jung et al. 2013; Fassnacht et al. 2014).

Machine learning involves different techniques (mainly nonparametric) that focus on automated and inductive learning to recognize patterns (Cracknell and Reading 2014) in data (e.g., patterns in remote sensing data related to AGB in a set of located plots); once the pattern is learned, it can be applied to yield a prediction or classification in areas where it is not possible to carry out fieldwork to quantify an objective variable (e.g., AGB). In the last decade, various machine learning techniques such as Support Vector Machine (SVM), k-nearest neighbor (kNN) and Random Forest (RF) have been used to develop predictive models of AGB in large areas. Thus, Shataee (2013) showed that kNN performed better than SVMs, RF, and Artificial Neural Networks (ANN) for estimating biophysical variables such as basal area. More recently, Garcia-Gutierrez et al. (2015) showed that SVM models performed best for estimating forest variables from Light Detection and Ranging (LIDAR), while Wang et al. (2016) showed that RF outperformed SVM and ANN for estimating wheat biomass from remote sensing data. For a more complete review of research being carried out to retrieve vegetation biomass from remote sensing data, using machine learning methods, see Ali et al. (2016).

The goodness-of-fitpt of models derived from spectral data are usually evaluated by the coefficient of determination ($R^2$) and the root mean square error (RMSE). These measures report the performance of the model in predicting the data used to fit the model; however, because the quality of the fit does not necessarily reflect the quality of the prediction, assessment of their validity is often needed to ensure that the predictions represent the most likely outcome in the real world (Yang et al. 2004). The only method that can be regarded as "true" validation involves the use of a new independent dataset (Pretzsch et al. 2002; Yang et al. 2004); however, the scarcity of such data forces the use of alternative approaches, such as Cross Validation (CV), to enable evaluation of the quality of a particular fitting technique and minimize the risk of overfitting (Molinaro et al. 2005). Unfortunately, most studies involving estimation of AGB do not use CV as part of the model development.

For rigorous comparison of the performance of different machine learning techniques, the study should also be accompanied by statistical validation of the results within a statistical framework (i.e., not merely calculating statistics such as $R^2$ or RMSE). Although this is well known in the field of machine learning (García et al. 2010), this type of validation is not common in remote sensing, even though machine learning plays an important role in many biomass estimation studies. This fact might have led to some degree of discordance in the scientific literature, in which we can find examples of kNN, SVM, and RF outperforming each other (Shataee 2013; Garcia-Gutierrez et al. 2015; Wang et al. 2016).

The objective of this study was to analyze and statistically compare the performance of 3 nonparametric techniques (SVM, kNN, and RF) and the parametric Multiple Linear Regression (MLR) technique for estimating AGB. The techniques were tested with Landsat-5 TM surface spectral reflectance data, texture features derived from the Normalized Difference Vegetation Index (NDVI), and topographical features derived from a digital elevation model (DEM) in the Sierra Madre Occi-

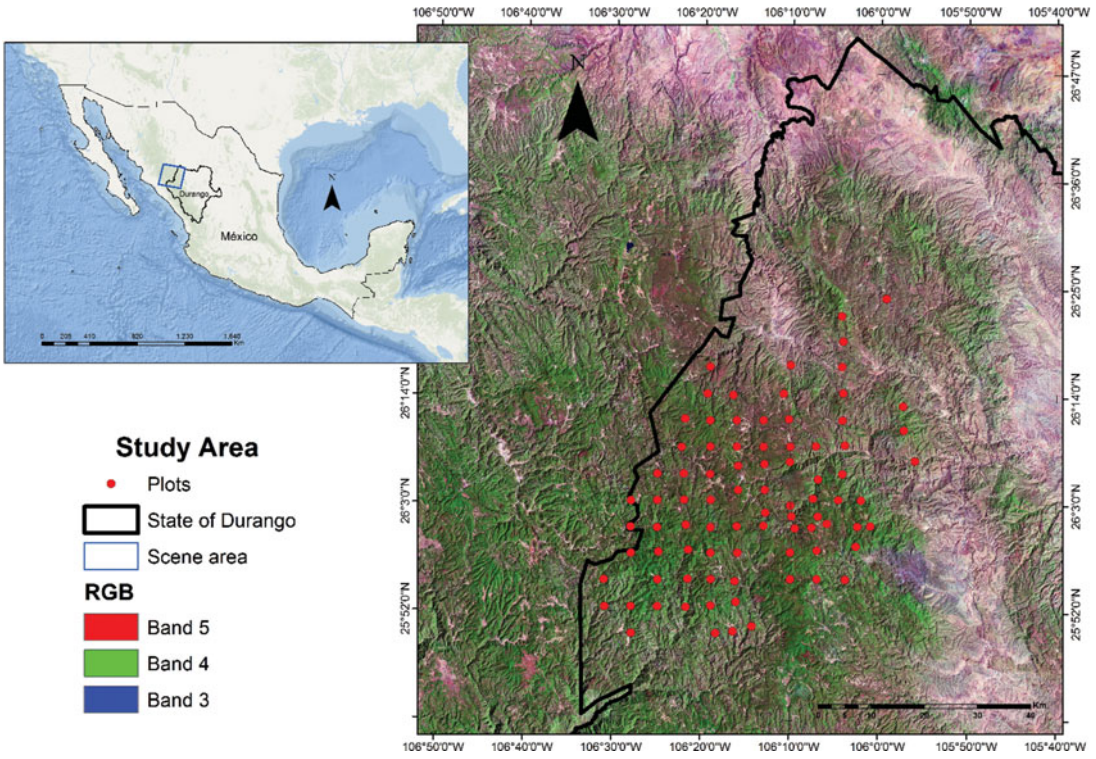


FIG. 1. Geographical location of the study site and sample plots used in the study.

dental (state of Durango, Mexico). The results obtained with each technique were compared after application of CV and posterior statistical validation of the mean rankings obtained for each.

## MATERIAL AND METHODS

### Study Area

The study site is located in the Sierra Madre Occidental, in the north of the state of Durango (Mexico), and covers an area of 1,142,916 ha (Figure 1). The climate is humid temperate, with rainfall in summer (relative humidity, 50.1%). The average temperature ranges from 8 °C to 20 °C, and the annual precipitation is from 400 mm to 1200 mm. The average altitude above sea level in this area is 1,900 m. The vegetation comprises pine, oak, Douglas fir, pine-oak, and oak-pine forest, according to the description in the Land Use and Vegetation Cover Chart, scale 1:250,000, Series V (INEGI 2012). The forests are basically mixed and uneven-aged pine-oak stands, with a canopy cover ranging from 32% to 100%. These forests have been subject to selective harvesting for almost a century to provide a mixture of services to local communities. This structure is the result of the management history, which has depended on land ownership and the economic and social changes that have taken place in the state, as well as natural conditions (Wehenkel et al. 2011).

### Dataset

*Field Data*

A network of 99 permanent sampling plots was established during the winter of 2011, following the method described by Corral-Rivas et al. (2009). The plots were located by systematic sampling (with some exceptions to avoid nonforested areas) of a grid of equidistant points separated by 3 km or 5 km, depending on the accessibility, which is limited by the rugged terrain of the study area. In each plot (squares of side 50 m), all species of trees were recorded and the diameters at breast height (cm) and total height (m) of all standing trees were measured.

Species-specific individual tree models developed by Vargas-Larreta (2013) were used to estimate the total AGB of field plots by tree value aggregation. The $R^2$ and the RMSE of the models used ranged from 0.87 kg–0.99 kg and 22.8 kg–95.2 kg, respectively. The mean, minimum, maximum and standard deviation of the AGB values per hectare of the sample plots are summarized in Table 1.

*Spectral Data*

The spectral data were derived from a satellite image Landsat-5 TM obtained in April 2011 (path 32, row 42) and covering the entire study área.[1] Landsat-5 TM data have a

[1]Available from the US Geological Service webpage, at http://glovis.usgs.gov/

TABLE 1
Total biomass statistics expressed in Mg ha$^{-1}$

| No. of Observations | Mean | Standard Deviation | Minimum Value | Maximum Value |
|---|---|---|---|---|
| 99 | 89.03 | 43.45 | 2.70 | 234.03 |

spatial resolution of 30 m with a revisit period of 16 days. Bands 1, 2, 3, 4, 5, and 7 (level L1T) of Landsat-5 TM were used in the present study; band 6 was not used because of its thermal characteristics, its coarse spatial resolution (120 m), and the low contrast in the forest area (NASA 2011). The satellite images were radiometrically, atmospherically, and topographically corrected by using the ATCOR3® module (Geosystems 2013), regarded as particularly suitable for mountainous zones. The ATCOR3® module first calculates the radiance at sensor level (W sr$^{-1}$ m$^{-2}$) from the image pixel. Several input parameters were required for this calculation and were retrieved from the image metadata (header file): date of acquisition, scale factors, geometry (solar zenith angle and solar azimuth), and other information about the sensor calibration file ("gain and bias"). Other parameters were adjusted by taking into account the characteristics of the input datasets and the conditions of the imagery dates, e.g., visibility (35 km), pixel size of the DEM (15 m), aerosol type (rural), among others. Because the image was cloudless and no suitable water vapor bands were available, dehazing/cloud removal and atmospheric water retrieval settings were kept as "default," which, in this case, is recommended by the *ATCOR3® User Manual* (Geosystems 2013). The corrections were implemented with the ERDAS® IMAGINE® 2013 software. (ERDAS Inc. 2014). A number of vegetation indices were computed from the atmospherically and topographically corrected image bands and included in the biomass estimation models for evaluation as possible regressor features (Table 2).

*Texture Features*

The texture features homogeneity, contrast, dissimilarity, mean, standard deviation, entropy, second-order angular moment, and correlation (Haralick et al. 1973) were calculated from the NDVI image based on grey level cooccurrence matrices, with the aim of including information combining the spatial and spectral domain of the remotely sensed imagery in the biomass estimation models. We used NDVI texture features rather than each spectral band of Landsat-5 TM to avoid saturating high biomass values (Mutanga and Skidmore 2004). Because it also becomes more difficult to obtain an optimal subset as the number of attributes increases, we therefore aimed for a compromise between quantity and quality. The features were calculated using PCI Geomatica2013® software,[2] and 3 different scales of operation were considered by using moving window sizes of 3 × 3 pixels, 5 × 5 pixels, and 7 × 7 pixels (Table 2).

TABLE 2
Features (independent variables) for biomass estimation in comparison of machine learning techniques

| Abbreviation | Variable | Reference |
|---|---|---|
| | Vegetation Index | |
| NDVI | Normalized Difference Vegetation Index | Rouse et al. (1974) |
| MSAVI2 | Modified Soil-Adjusted Vegetation Index | Qi et al. (1994) |
| SAVI | Adjusted Soil Vegetation Index | Huete (1988) |
| IAF | Leaf Area Index | Baret and Guyot (1991) |
| ALB | Albedo | Asrar (1989) |
| Fpar | Fraction of Photosynthetically Active Radiation | Asrar et al. (1984) |
| FSR | Flow Solar Radiation | Brutsaerts (1975) |
| | Texture (NDVI) | |
| HOL | Homogeneity | Haralick et al. (1973) |
| CO | Contrast | |
| DI | Dissimilarity | |
| ME | Mean | |
| STD | Standard Deviation | |
| EN | Entropy | |
| ASM | Angular Second Moment | |
| CR | Correlation | |
| | Terrain (DEM) | |
| Altitude | Altitude | |
| B | Slope | |
| TRASP | Transformed Aspect | Roberts and Cooper (1989) |
| TSI | Terrain Shape Index | McNab (1989) |
| WI | Wetness Index | Moore and Nieber (1989) |
| PC | Profile Curvature | Wilson and Gallant (2000) |
| PLC | Plan Curvature | |
| C | Curvature | |

*Terrain Features*

Terrain features are directly related to forest species composition, tree height growth, and other forest stand variables, en-

[2]PCI Geomatics Inc. 2013

abling these to be modeled (McNab 1989; Roberts and Cooper 1989). First- and second-order terrain features were, therefore, derived from the 5 × 5-pixel low pass filtered DEM of the study area with a spatial resolution of 15 m. The DEM was derived from LIDAR data and corresponds to an array of elevation data interpolated to 15 m resolution from the coordinates of the last return of the pulses emitted (INEGI 2014). The final set of features derived from Landsat-5 TM sensor and from the DEM, which were used as possible predictors (independent variables) for estimating AGB (which played the role of dependent variable), are shown in Table 2.

Finally, the sample plots were geopositioned with the aim of extracting the pixel value average with an associated buffer of 25 m for each described feature, to obtain a database with the mean biomass values and the associated features for each plot. The extraction was carried out using R statistical software (R Core Team 2014) and the "raster" package.

## Comparison Framework

*Machine Learning Techniques*

Three nonparametric machine learning techniques and one parametric technique were applied to data from the study area in order to compare their performance: (*i*) k-Nearest Neighbour (kNN), (*ii*) Support Vector Machine (SVM), (*iii*) Random Forest (RF), and (*iv*) Multiple Linear Regression (MLR). All these techniques were used to estimate AGB, using as possible predictors the variables included in Table 2.

The parametric MLR technique is the most commonly used in this kind of study (Fassnacht et al. 2014). Moreover, this type of model is easy to understand and is widely used in most scientific disciplines. However, unlike the nonparametric approaches, MLR relies on certain assumptions, such as the fundamental least squares assumption of independence and equal distribution of errors with zero mean and constant variance, which can be violated by factors such as nonnormality of variables, multicollinearity of variables, and heteroscedasticity of error variance.

Nearest neighbor (NN), a well-known machine learning technique used in remote sensing (Shataee 2013), makes a prediction by using the information about the neighbors of the instance to be regressed (Cover and Hart 1967). The NN depends on a parameter, usually called k, which determines the number of neighbors used by the algorithm. The technique is therefore usually called kNN when more than one neighbor is used. Although the idea behind this type of technique is quite intuitive, the resulting model is not easy to interpret because all results depend on a training set.

SVMs have been developed from artificial neural networks (Cortes and Vapnik 1995) and have been used in many scientific fields (e.g., Abedi et al. 2012; Bayoudh et al. 2015; Garcia-Gutierrez et al. 2015). SVM models are developed by a set of vectors (or hyperplanes if greater dimension is requested) that separate instances of different labels (classification) or minimize the mean error (regression). Kernel functions are used to overcome the limitations associated with linear separability in SVM models. Appropriate selection of the kernel function and the kernel regularization parameters is important in relation to the SVM model behavior, which can make this type of technique more difficult to implement for users. As with kNN, the models produced using SVM are more difficult to interpret than those of MLR.

RF is not exactly a classification or regression technique, but a combination of other techniques, mainly regression or classification trees (Breiman 2001). 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 features (e.g., Deschamps et al. 2012; Wang et al. 2016). The number of predictors used by trees and the number of trees are established by the users.

WEKA open source software (Hall et al. 2009) was used to implement all of the techniques compared. Thus, linear regression was used for MLR, IBk for kNN, SMOreg with polynomial and Gaussian kernels for SVM, and an adaptation of the RF implementation of WEKA for regression (using M5P as the basic regression technique for the development of this ensemble).

*Feature Selection, Parameterization and Validation*

In machine learning, spurious data features must be removed before a model is generated (Hall 1999). Thus, the variables that are potentially most important are selected. Some techniques (e.g., SVM and RF) carry out this selection, but others might be seriously affected by excessively large combinations of variables (e.g., the Hughes effect [Hughes 1968] in kNN and multicollinearity in MLR). This is a common situation in this type of analysis because of the large set of predictor variables that can be calculated from remote sensing data (Packalén et al. 2012). Moreover, correct functioning of different machine learning techniques depends on a proper parameterization (set-up of their parameters, i.e., variables that modify the behavior of the machine learning techniques). In this study, both of these steps (feature selection and parameterization) were carried out via a metaheuristic search (Samadzadegan et al. 2012). From the possible metaheuristic techniques (i.e., a method of optimization that provides a near-optimal solution in computationally affordable time), we selected an evolutionary algorithm, which is illustrated in Figure 2. The algorithm starts with a population of random solutions (Initial Population in Figure 2) called individuals and ranks them according to fitness of the individuals (Fitness Sorting in Figure 2). In the present study, the fitness was evaluated by the RMSE obtained with a training set. A new population of individuals is then created by mating parents (random selection of coefficients shown in Figure 2), selected with a probability proportional to their fitness, and later mutating the new individuals with a given probability (in this case, a value will be randomly selected and changed to a new random value, as can be seen in Figure 2).

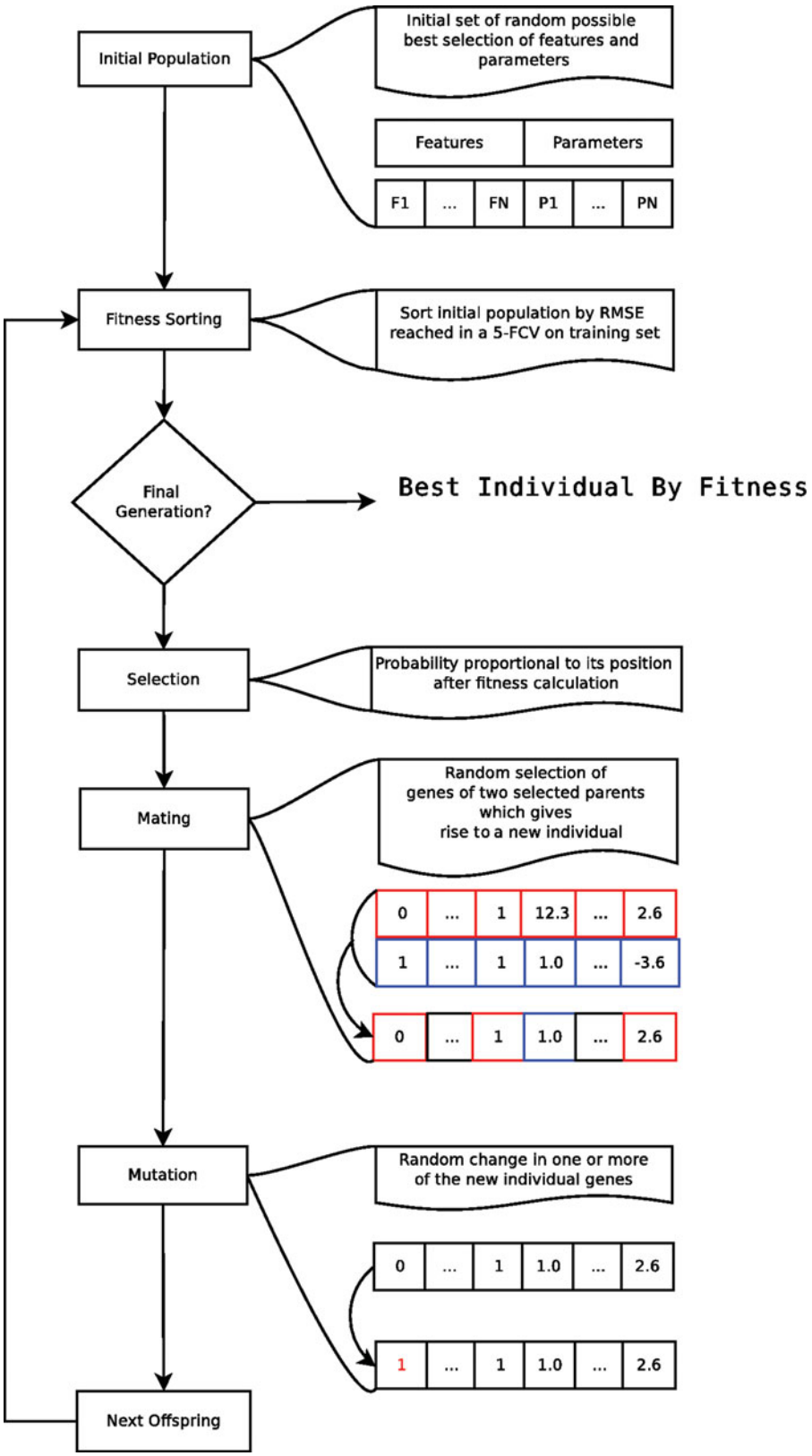


FIG. 2. Description of the evolutionary procedure used to determine the best methods for parameterization and feature selection.

TABLE 3
Intervals used by the evolutionary algorithm to search for the different optimal parameters*

| Technique | Name | Minimum | Maximum |
|---|---|---|---|
| kNN | k | 1 | 20 |
| SVM | GAMMA (Gaussian-kernel-only) | 0.01 | 2.0 |
| | EXP (Polynomial-kernel-only) | 1 | 5 |
| | C | 1 | 100 |
| | EPSILON | 0.0 | 0.2 |
| RF | NT | 1 | 100 |
| | NF | 1 | 5 |

*Note: k = number of neighbors; EPSILON = determines the risk of overfitting; GAMMA = controls the transformation produced by the kernel; EXP = kernel's exponent; C = penalty factor per instance of misclassification in training; NT = number of trees that form each ensemble; NF = number of attributes selected for constructing each tree

The general scheme described in Figure 2 was modified slightly according to the specific regression technique. Thus, we used a specific design for MLR (see García-Gutiérrez et al. 2014) and an adaptation of the genetic algorithm of Huang and Wang (2006) for the nonparametric techniques (kNN, SVM, and RF). In the kNN method, pure selection (coefficients associated with each feature as 1 or 0 depending on whether the predictor is selected or not) was substituted by weighting each attribute (real value between 0.0 and 1.0), which enables better adaptation of the algorithm to the characteristics of kNN (see Mateos et al. 2012). In SVMs, the type of kernel is another parameter to be optimized and had 2 possible values (radial basis function and polynomial). The parameters optimized for each machine learning technique are included in Table 3.

For comparison of the different techniques, validation was based on the leave-one-out CV technique. This is a special case of $k$-fold CV in which $k$ is equal to the number of observations and a prediction is obtained as many times as there are observations in the dataset (Packalén et al. 2012). In other words, an observation is excluded (target observation), and a prediction is computed with the other observations (reference observations). The prediction can be evaluated by the target observation. This procedure is repeated for every single observation. The final quality of a technique evaluated with CV is based on the averaged error obtained. A general description of the procedure is provided in Figure 3.

Parameterization of each submodel at the different stages of the CV was repeated 5 times for each technique to prevent skew (due to the random nature of the evolutionary algorithms applied to predictor selection and parameterization). The best submodel and the average submodel for the 5

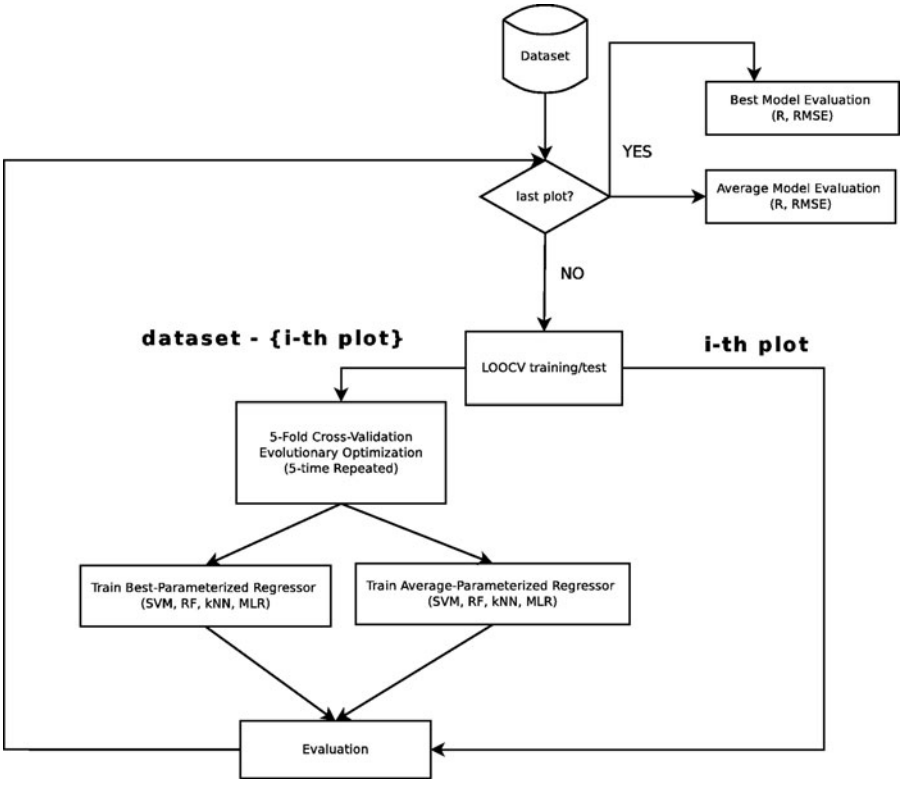


FIG. 3. Description of the leave-one-out CV evaluation of the techniques compared in the text.

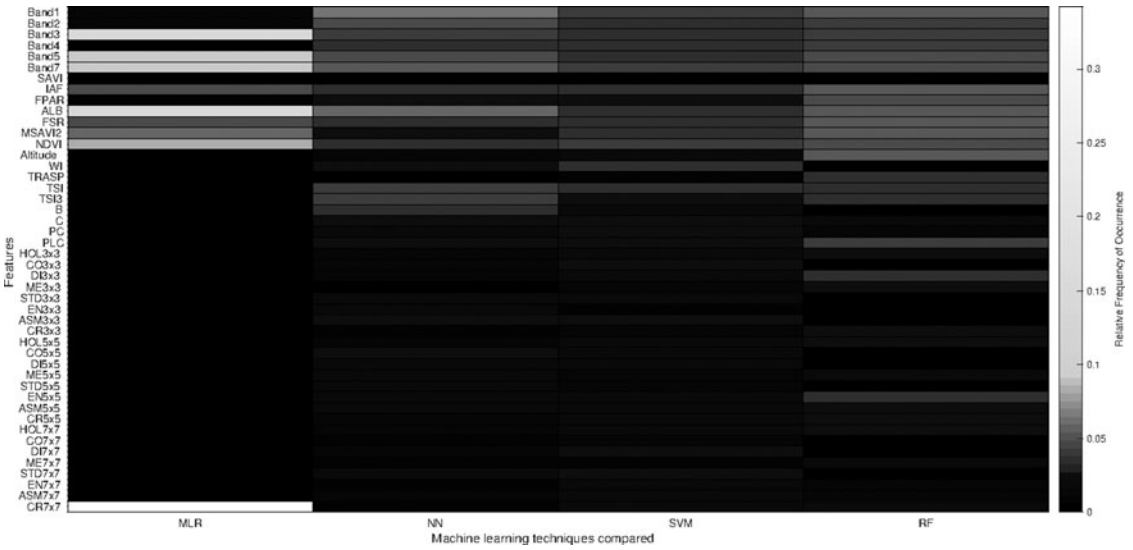


FIG. 4. Relative frequency of ocurrence (importance) of each attribute in the best models obtained by each technique (in terms of the sum of residuals).

executions, ranked in terms of the RMSE reached in the evolutionary procedure, were used to calculate the goodness-of-fit statistics.

*Statistical Analysis*

The error of the predictions in the CV was compared for each technique in terms of $R^2$ and RMSE. In addition, for statistical analysis of differences between the methods, the absolute errors of the predictions made by each technique throughout the 99 iterations in the CV were compared (the number of iterations is equal to the number of instances in the database, which, in this case, refers to the 99 plots available). In theory, this should be carried out by Analysis of Variance (ANOVA), if the data comply with the underlying assumptions of independence, normality, and homoscedasticity required for parametric tests. These conditions can be tested by, respectively, the Shapiro-Wilk test, Lilliefor's test, and Levenes' test. If the data do not comply with these conditions, a nonparametric test such as the Friedman's (aligned) test (described by García et al. 2010) should be used.

Friedman's (aligned) test first obtains the mean ranking for each technique by taking into account the position obtained for

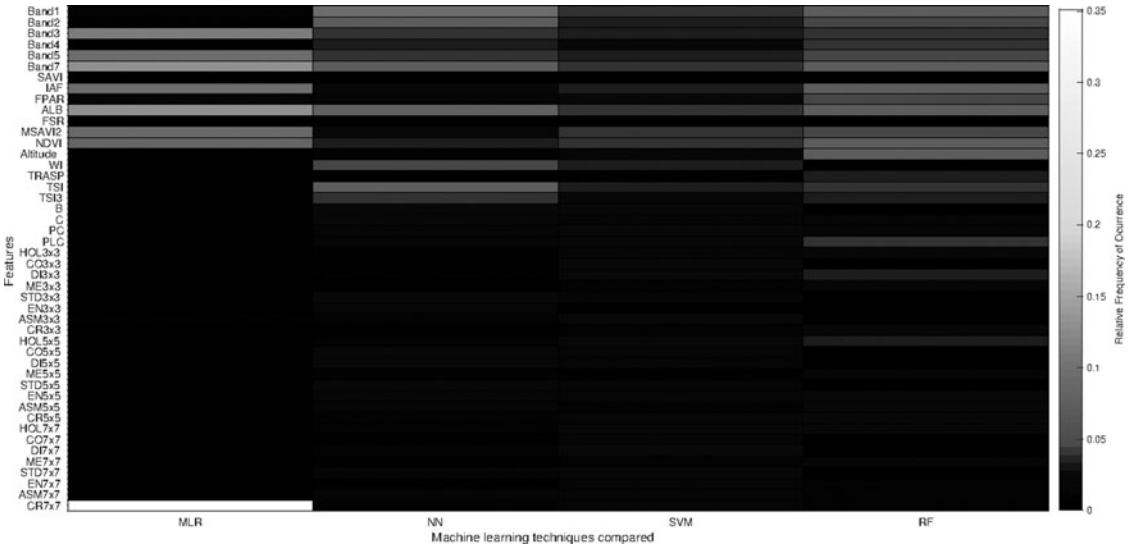


FIG. 5. Relative frequency of ocurrence (importance) in the averaged models obtained by each technique (in terms of the sum of residuals).

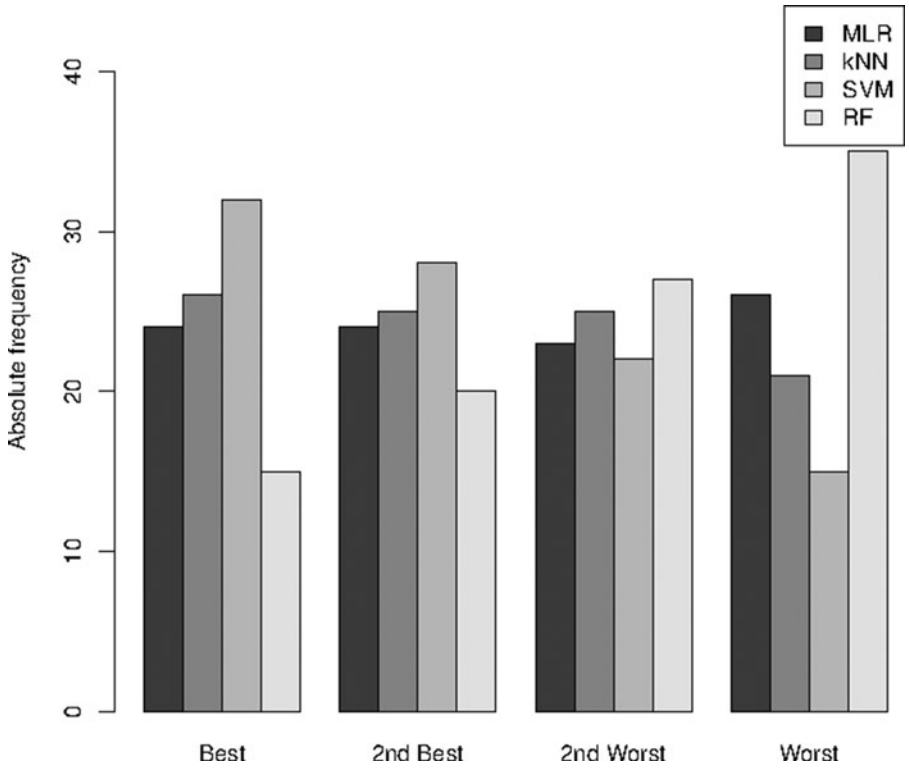


FIG. 6. Absolute frequency of relative position achieved by each technique (ranking) with the best parameterization of 5 executions.

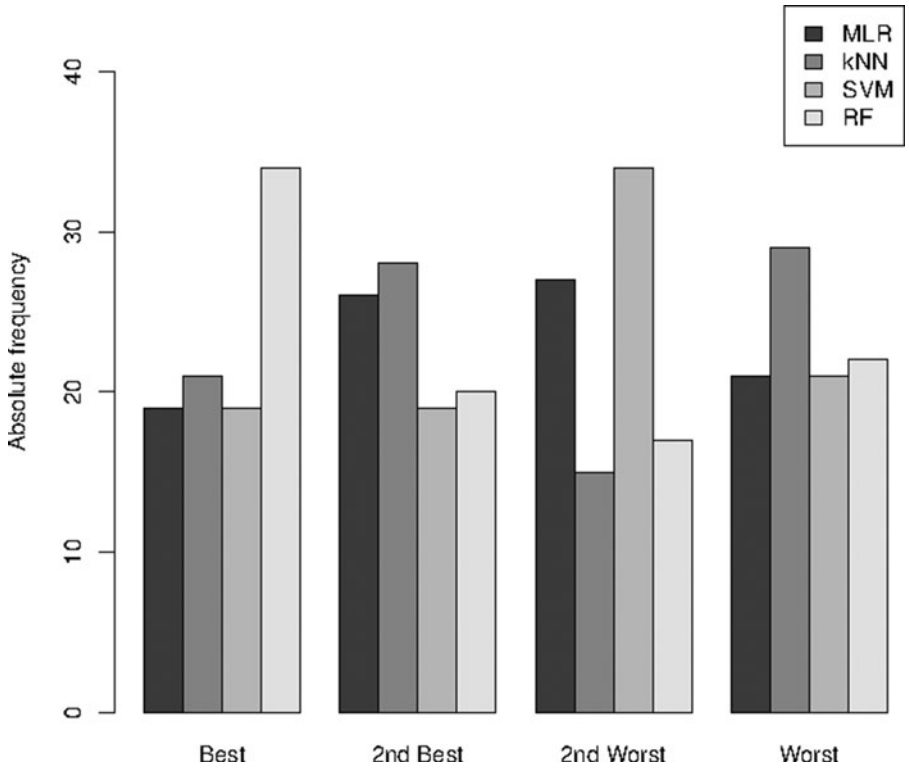


FIG. 7. Absolute frequency of the relative position achieved (ranking) by each technique with average parameterization.

TABLE 4
Mean rankings for models obtained by each technique; best mean ranks indicated in bold

| Technique | Ranking | |
|---|---|---|
| | Best Models Only | Averaged Models |
| MLR | 197.98 | 188.11 |
| kNN | 186.69 | 184.37 |
| RF | 232.16 | **176.04** |
| SVM | **161.18** | 197.48 |

each of the results relative to the other. Thus, a ranking of 1 for one of the techniques signifies that the result is the best of all results obtained in the procedure, whereas a rank of $m * n$, where $m$ is the number of techniques being compared and $n$ is the number of tests, indicates the poorest result obtained. After establishing the mean rankings for each technique, the Friedman's (aligned) test and Holm's post hoc procedure (see Luengo et al. 2009; and García et al. 2010, for a complete description) are used for statistical validation of the differences between the methods compared.

## RESULTS AND DISCUSSION

The results indicated that the features for estimating AGB by the different machine learning algorithms evaluated can be classified into 3 different groups. In order of decreasing importance, the 1st group comprises the spectral bands and the spectral indices, the 2nd group comprises the first- and second-order terrain topographical variables derived from DEM, and the 3rd group comprises the texture features derived from the NDVI. The correlation derived from the texture image with a moving window of 7 pixels x 7 pixels (CR7 $\times$ 7, see Table 2 for acronyms and abbreviations) was also a key feature in the MLR technique (as was also reported by Kelsey and Neff 2014), although it was not important in the other techniques (see Figures 4 and 5).

The results obtained in terms of the RMSE were used for statistical comparison of the techniques. The comparison is summarized in histograms showing the relative positions reached (rankings) for each technique (Figures 6 and 7). Qualitatively, the SVM technique yielded the best results when the parameterization and selection of predictors were relatively optimal, whereas, on average, the RF technique produced the best results.

The rankings associated with the Friedman's (aligned) test and a post hoc Holm's test for paired comparison of the best algorithms (Table 4) confirm the idea (previously outlined in Figures 6 and 7) that SVM and RF techniques yielded the best results, considering, respectively, the best model and the averaged models for the 5 executions per plot. Friedman's (aligned) test yielded a $p$-value $< 0.0001$, thus confirming rejection of the null hypothesis (i.e., that the overall performance of the methods was not significantly different).

Application of Holm's procedure revealed that the results yielded by the SVM technique were significantly different from those produced by all other techniques except kNN ($p = 0.1131$, higher than the significance levels of the test; $\alpha = 0.05$). Comparison of RF and the other techniques for the averaged models showed that none of the comparisons was statistically significant, and it was, therefore, not possible to infer that RF performed better than the other models. The results of both procedures are summarized in Table 5.

Finally, the results for all plots were used to calculate the goodness-of-fit statistics: $R^2$ and RMSE. Table 6 summarizes

TABLE 5
Results of post hoc Holm's test of paired comparisons for SVM (best models only) and RF (averaged models); not significantly different comparisons indicated in bold

| Best Models Only | | | | Averaged Models | | | |
|---|---|---|---|---|---|---|---|
| Technique | p | z | Holm | Technique | p | z | Holm |
| RF | 0.000 | 4.408 | 0.0167 | SVM | **0.174** | 1.36 | 0.0167 |
| MLR | 0.022 | 2.285 | 0.025 | MLR | **0.444** | 0.77 | 0.025 |
| kNN | **0.113** | 1.584 | 0.05 | kNN | **0.598** | 0.53 | 0.05 |

TABLE 6
Summary of goodness-of-fit statistics, taking into account overall results for 99 plots; best models indicated in bold

| | | MLR | kNN | SVM | RF |
|---|---|---|---|---|---|
| Best models only | $R^2$ | 0.54 | **0.66** | 0.62 | 0.48 |
| | RMSE (Mg ha$^{-1}$) | 29.61 | **26.64** | 27.28 | 31.61 |
| Averaged models | $R^2$ | 0.36 | **0.41** | 0.30 | 0.29 |
| | RMSE (Mg ha$^{-1}$) | 34.67 | **33.53** | 36.15 | 39.20 |

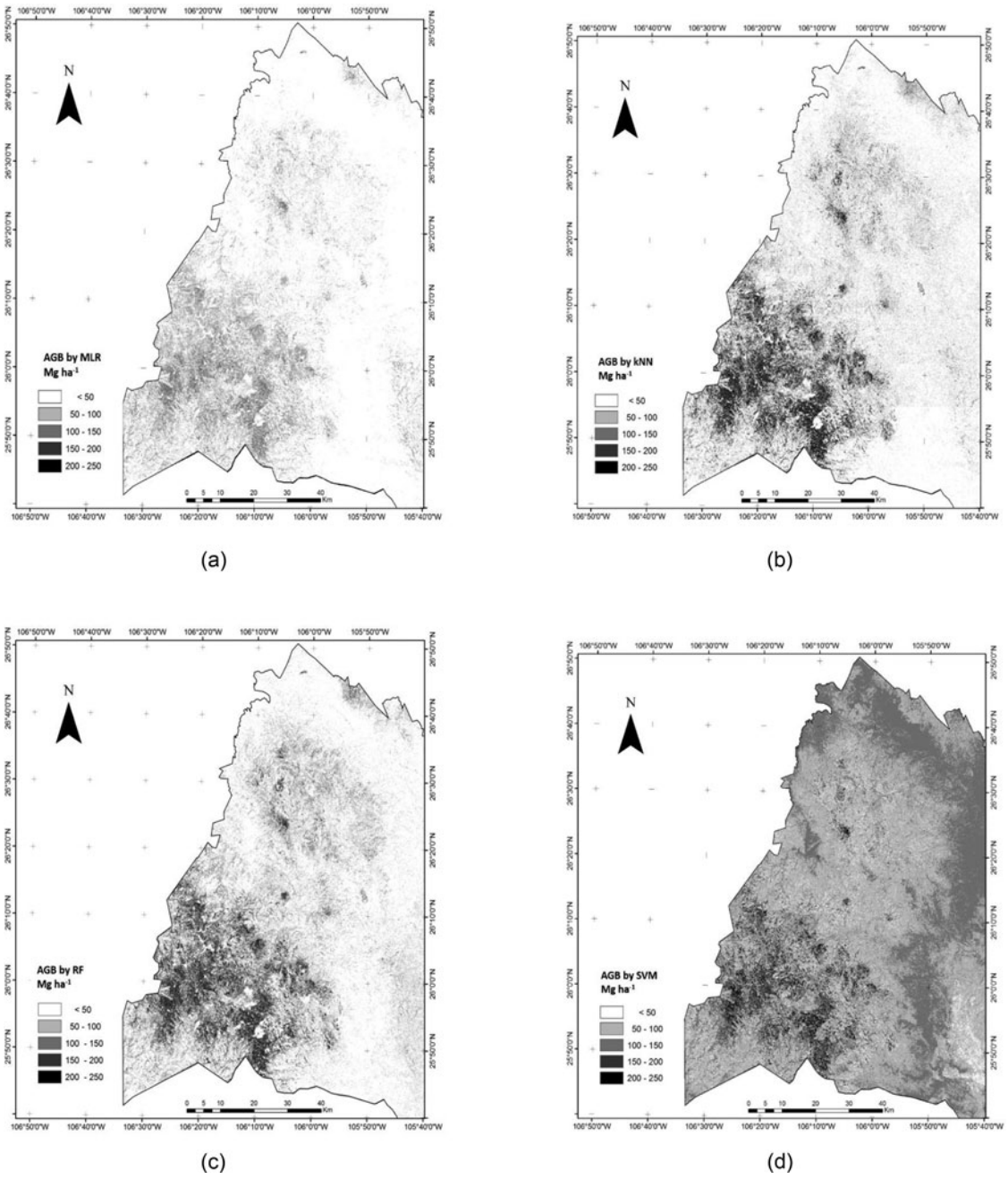


FIG. 8. Biomass maps derived by each technique: (a) MLR, (b) kNN , (c) RF, and (d) SVM.

the application of these to the best models and the averaged models, in which kNN was the best technique in both cases. Maps of the AGB estimations obtained for the study area by each technique are shown in Figure 8.

The results showed that the features that were most important for estimating AGB by the different machine learning techniques evaluated (kNN, RF, and SVM) correspond to the bands and spectral indices derived from the Landsat-5 TM sensor (Band 1, Band 5, and Band 7, IAF, ALB, MSAVI2, and NDVI, see Table 2 for acronyms and abbreviations); which are correlated with many ecosystem attributes, such as photosynthetic activity, total plant cover, plant and soil moisture, plant stress, and biomass (Lu et al. 2004; Günlü et al. 2014). Several studies have demonstrated that spectral bands and vegetation indices are usually good predictors for estimating AGB (Lu et al. 2012; Castillo-Santiago et al., 2013; Lu et al., 2016; López-Serrano, Corral-Rivas, et al. 2016; López-Serrano, López-Sánchez, Díaz-Varela, et al. 2016; López-Serrano, López-Sánchez, Solís-Moreno, et al. 2016). Terrain features are potentially related to key features for forest stand development, such as overall climate characteristics,

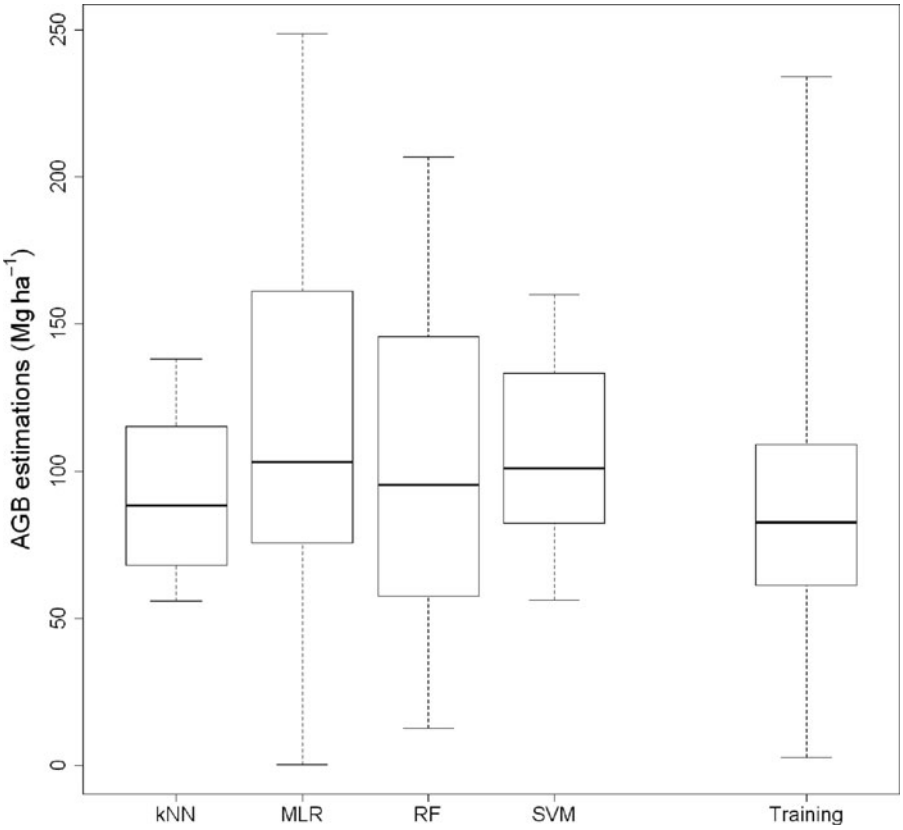


FIG. 9. Box-plot of AGB estimations in the study area for the 4 techniques used and AGB observed values in the sample plots (training data). Boxes represent the interquartile range, and maximum and minimum of AGB estimations are represented by upper and lower whiskers, respectively.

insolation, evapotranspiration, run-off, infiltration, wind exposure, and site productivity (McNab 1989; Roberts and Cooper 1989; Wilson and Gallant 2000). Finally, the texture features could address some of the existing problems with vegetation index saturation and the data acquisition constraints related to mapping forest biomass at regional scales (Kelsey and Neff 2014).

The results obtained in the statistical study of the 99 plots showed that the SVM technique yielded the best fits once the parameterization had been optimized (averaged ranking of 161.18, which is about 15% better than kNN, the 2nd best technique), 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 (e.g., Zhao et al. 2011; García-Gutiérrez et al. 2015). However, the results show that SVMs are very sensitive to parameterization, which hampers their use by nonexperts. For nonexperts, an autoparameterization procedure such as Grid Search, which is a classic technique used to fit machine learning models (Gleason and Im 2012), could be applied. Unfortunately, this type of procedure has an important drawback in that it separates optimization of parameters (specific to each technique) from feature selection. Both concepts (parameterization and feature selection) are closely related and should occur simultaneously (Huang and Wang 2011). Nonetheless, Grid Search represents a simpler alternative to more complex procedures such as metaheuristics.

A boxplot with the AGB estimations obtained for the sample plots with the different approaches used after just one evolutionary parameterization and feature selection is shown in Figure 9. Both MLR and RF present a range of AGB estimations in the study area similar to the values observed in the sample plots used as training data (2 Mg ha$^{-1}$ to 234 Mg ha$^{-1}$), especially MLR, although MLR tended to overestimate the values (Figure 9). However, the kNN and SVM techniques estimated a limited range of values of AGB (from 56 Mg ha$^{-1}$ to 138 Mg ha$^{-1}$ for kNN and from 56 Mg ha$^{-1}$ to 160 Mg ha$^{-1}$ for SVM). This was mainly due to inaccurate parameterization by the evolutionary procedure (see averaged models vs. optimal models in Figures 6 and 7). In the case of kNN, inadequate feature selection might lead to a decrease in accuracy due to the Hughes effect. For the SVM, the number of parameters was higher and the evolutionary procedure was, therefore, more complex. Note that if the penalty factor (parameter C) is not well fitted and the hyperplane is thus not optimized, problems related to over- or underfitting may occur (Xie et al. 2008). In addition to the

special random nature of evolutionary computation, this risk makes the automatic configuration for SVM difficult in a single optimization procedure (due to the random nature of evolutionary computation). Regardless of whether automatic or manual parameterization is selected, determination of the best configuration for remote sensing nonparametric techniques (especially for the most complex such as SVMs and not so much for others such as RF) is time consuming, scenario dependent and sometimes requires a priori knowledge (Camps-Valls and Bruzzone 2005).

The results obtained for the averaged models show that RF is a more robust technique, confirming the results reported by other authors (Latifi et al. 2010); however, the differences relative to the other techniques were not statistically significant. The lack of significance could be related to the fact that RF is a combination of techniques that depend on other regression techniques; this factor is often not taken into account, although it is essential to guarantee the quality of the regressions. In the present study, we used regression trees, as in the original study in which RFs were proposed (Breiman 2001), although the best results for biomass estimation involve the use of RF with kNN as an internal algorithm (Latifi et al. 2010).

Analysis of the overall results obtained using data from 99 plots showed that the classical kNN algorithm, which is well known and frequently used in the field of remote sensing, especially in forestry applications (Latifi et al. 2010; McRoberts 2012; Packalén 2012), yielded the best results in relation to the goodness-of-fit statistics (best parameterization: $R^2 = 0.66$, RMSE = 26.64 Mg ha$^{-1}$; average parameterization: $R^2 = 0.41$, RMSE = 33.35 Mg ha$^{-1}$). This difference relative to the statistical comparison is due to the fact that the rankings do not specifically take into account the size of the error, whereas the goodness-of-fit statistics do take this into account. In other words, although RF and SVM generally produce good results, in some cases they yield much larger errors than the mean, so that the kNN provided the best results. Although a priori this appears to contradict the statistical comparison, it does not invalidate it, because the results of the SVM or RF were not significantly different from those obtained by kNN in any of the cases. In addition, the goodness-of-fit statistics for the kNN technique obtained in the present study were better than those reported by Guo et al. (2014) in a study in which the AGB in a *Picea crassifolia* forest in NW China was estimated using Landsat TM data and 2 nonparametric methods (kNN: $R^2 = 0.54$, RMSE = 26.62 Mg ha$^{-1}$ and SVM: $R^2 = 0.51$, RMSE = 27.45 Mg ha$^{-1}$). Likewise, Tian et al. (2014), who estimated AGB with Landsat TM data, reported an optimized kNN method ($R^2 = 0.59$, RMSE = 24.92 Mg ha$^{-1}$) in comparison with the classic MLR ($R^2 = 0.42$, RMSE = 29.74 Mg ha$^{-1}$).

The findings of the present study with regard to the use of machine learning algorithms to estimate AGB from remote sensing data are encouraging, specifically because the spectral data used in this study are available to the public free of charge. The use of nonparametric machine learning methods in combination with multisensor data has been shown to be a valuable option for increasing the accuracy of aboveground forest biomass estimation (Li et al. 2012; Ali et al. 2016; Lu et al. 2016). For example, in a study of mangrove in southwest Thailand, Jachowsky et al. (2013) found that the SVM model performed better than another 18 parametric and nonparametric approaches for estimating forest biomass by using GeoEye-1 and ASTER data. Similar results were obtained by Zhang et al. (2014), who used the Geoscience Laser Altimeter System (GLAS) and MODIS data in a study in which SVM performed better than stepwise regression and partial least-squares regression for forest biomass mapping in northeastern China. Moreover, the new technologies associated with active sensors such as LiDAR might provide more accurate results ($R^2$ in [0.60,0.90]), depending on the vegetation conditions (Lu et al. 2012), the number of field observations, and the specific focus of the statistical modeling (Laurin et al. 2014). However, these multisensory alternatives are costly and, in the case of LiDAR, of limited accessibility. The present findings are, therefore, important because they demonstrate that medium resolution sensors, such as the Landsat-5 TM, together with machine learning algorithms, can represent a balanced solution between the costs of training data and the utility of the mapped outcomes for monitoring AGB in large areas.

## CONCLUSIONS

The kNN, RF and SVM machine learning algorithms are powerful tools for estimating aboveground forest biomass with remote sensing datasets, and they are all viable and accurate alternatives to the classic parametric MLR method. In addition to the usual sources of uncertainty associated with the accuracy of the AGB estimations from remote sensing data (field measurement errors, plot locations errors, errors of the individual tree biomass equations, or error caused by geometrical and radiometric correction of remotely sensed data), parameterization of machine learning algorithms also has an important influence on the final performance of the models. The choice of method used will largely depend on the user's capacity to carry out that parameterization, because the techniques (especially SVMs) are not easy to apply and require a certain degree of expertise. Our findings indicate that SVM is the best alternative for experts, whereas RF represents a balance between model accuracy and ease of use for nonexperts although differences with kNN could not be statistically demonstrated.

## FUNDING

This research was supported by SEP-PROMEP (project: Seguimiento y Evaluación de Sitios Permanentes de Investigación Forestal y el Impacto Socio-económico del Manejo Forestal en el Norte de México), the Spanish Ministry of Science and Technology (project TIN2011-28956-C02) and the Regional Ministry of Andalucía (project P11-TIC-7528).

## REFERENCES

Abedi, M., Norouzi, G.H., and Bahroudi, A. 2012. "Support vector machine for multi-classification of mineral prospecting areas." *Computers & Geosciences*, Vol. 46: pp. 272–283.

Agarwal, R., Ranjan, P., and Chipman, H. 2014. "A new Bayesian ensemble of trees approach for land cover classification of satellite imagery." *Canadian Journal of Remote Sensing*, Vol. 39(No. 6): pp. 507–520.

Ali, I., Greifeneder, F., Stamenkovic, J., Neumann, M., and Notarnicola, C. 2015. "Review of machine learning approaches for biomass and soil moisture retrievals from remote sensing data." *Remote Sensing*, Vol. 7: pp. 16398–16421.

Andersen, H.-E., Stunk, J., Temesgen, H., Atwood, D., and Winterberger, K. 2011. "Using multi-level remote sensing and ground data to estimate forest biomass resources in remote regions: a case study in the boreal forests of interior Alaska." *Canadian Journal of Remote Sensing*, Vol. 37(No. 6): pp. 596–611.

Asrar, G. 1989. "*Theory and applications of optical remote sensing*." New York, NY: J. Wiley & Sons.

Asrar, G., Fuchs, M., Kanemasu, E.T., and Hatfield, J.L. 1984. "Estimating absorbed photosynthetically active radiation and leaf area index from spectral reflectance in wheat." *Agronomy Journal*, Vol. 76(No. 2): pp. 300–306.

Baret, F., and Guyot, G. 1991. "Potentials and limits of vegetation indices for LAI and APAR assessment." *Remote Sensing of Environment*, Vol. 35(No. 23): pp. 161–173.

Bayoudh, M., Roux, E., Richard, G., and Nock, R. 2015. Structural knowledge learning from maps for supervised land cover/use classification: application to the monitoring of land cover/use maps in French Guiana. *Computers & Geosciences*, Vol. 76: pp. 31–40.

Breidenbach, J., Næsset, E., and Gobakken, T. 2012. "Improving k-nearest neighbor predictions in forest inventories by combining high and low density airborne laser scanning data. *Remote Sensing of Environment*, Vol. 117: pp. 358–365.

Breiman, L. 2001. "Random forests." *Machine Learning*, Vol. 45(No. 1): pp. 5–32.

Brutsaert, W. 1975. "On a derivable formula for long-wave radiation from clear skies." *Water Resources Research*, Vol. 11(No. 5): pp. 742–744.

Bunker, D.E., DeClerck, F., Bradford, J.C., Colwell, R.K., and Perfecto, I. 2005. "Species loss and aboveground carbon storage in a tropical forest." *Science*, Vol. 310: pp. 1029–1031.

Camps-Valls, G., and Bruzzone, L. 2005. "Kernel-based methods for hyperspectral image classification." *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 43: pp. 1351–1362.

Castillo-Santiago, M.A, Ghilardi, A., Oyama, K., Hermadez-Stefanoni, J.L., Torres, I., Flamenco-Sandoval, A., Fernández, A., and Jean-Francois, M. 2013. "Estimating the spatial distribution of woody biomass suitable for charcoal making from remote sensing and geostatistics in central Mexico." *Energy for Sustainable Development*, Vol. 17(No. 2): pp. 177–188.

Corral-Rivas, J., Vargas, B., Wehenkel, C., Aguirre, O., Álvarez, J.G., and Rojo, A. 2009. *Guía para el Establecimiento de Sitios de Inventario Periódico Forestal y de Suelos del Estado de Durango*. Durango, Mexico: Universidad Juárez del Estado de Durango.

Cortes, C., and Vapnik, V. 1995. "Support-vector networks." *Machine Learning*, Vol. 20(No. 3): pp. 273–297.

Cover, T., and Hart, P. 1967. "Nearest neighbor pattern classification." *IEEE Transactions on Information Theory*, Vol. 13(No. 1): pp. 21–27.

Cracknell, J.M., and Reading, M.A. 2014. "Geological mapping using remote sensing data: a comparison of five machine learning algorithms, their response to variations in the spatial distribution of training data and the use of explicit spatial information." *Computers & Geosciences*, Vol. 63: pp. 22–33.

Deschamps, B., McNairn, H., Shang, J., and Jiao, X. 2012. "Towards operational radar-only crop type classification: comparison of a traditional decision tree with a random forest classifier." *Canadian Journal of Remote Sensing*, Vol. 38(No. 1): pp. 60–68.

Dube, T., and Mutanga, O. 2015. "Evaluating the utility of the medium-spatial resolution Landsat 8 multispectral sensor in quantifying aboveground biomass in Mgeni catchment, South Africa." *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 101: pp. 36–46.

Eckert, S. 2012. "Improved forest biomass and carbon estimations using texture measures from worldview-2 satellite data." *Remote Sensing*, Vol. 4(No. 4): pp. 810–829.

ERDAS Inc. 2014. Erdas Imagine, accessed October 6, 2014, http://www.hexagongeospatial.com/products/ERDAS-IMAGINE/details.aspx.

FAO. 2001. *Global Forest Resources Assessment 2000*, Main Report. FAO Forestry Paper 140. Rome: FAO.

FAO. 2006. *Global Forest Resources Assessment 2005*. FAO Forestry Paper 147. Rome: FAO.

Fassnacht, F.E., Hartig, F., Latifi, H., Berger, C., Hernández, J., Corvalán, V., and Koch, B. 2014. "Importance of sample size, data type and prediction method for remote sensing-based estimations of aboveground forest biomass." *Remote Sensing of Environment*, Vol. 154: pp. 102–114.

Frazier, R. J., Coops, C.N., Wulder, A.M., and Kennedy, R. 2014. "Characterization of aboveground biomass in an unmanaged boreal forest using Landsat temporal segmentation metrics." *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 92: pp. 137–146.

García, S., Fernández, A., Luengo, J., and Herrera, F. 2010. "Advanced non-parametric tests for multiple comparisons in the design of experiments in computational intelligence and data mining: experimental analysis of power." *Information Sciences*, Vol. 180(No. 10): pp. 2044–2064.

García-Gutiérrez, J., González-Ferreiro, E., Riquelme-Santos, J.C., Miranda, D., Dieguez-Aranda, U., and Navarro-Cerrillo, R.M. 2014. "Evolutionary feature selection to estimate forest stand variables using LiDAR." *International Journal of Applied Earth Observation and Geoinformation*, Vol. 26: pp. 119–131.

García-Gutiérrez, J., Martinez-Alvarez, F., Troncoso, A., and Riquelme, J. C. 2015. "A comparison of machine learning regression techniques for LiDAR-derived estimation of forest variables." *Neurocomputing*, Vol. 167: pp. 24–31.

GCOS. 2010. Global Climate Observing System. *Essential Climate Variables*, accessed April 13, 2016, http://www.wmo.int/pages/prog/gcos/index.php?name=Essential ClimateVariables.

Geosystems. 2013. *Atcor for Erdas Imagine*. Switzerland: Geosystems GmbH.

Ghosh, A., Fassnacht, F.E., Joshi, P.K., and Koch, B. 2014. "A framework for mapping tree species combining hyperspectral and LiDAR data: role of selected classifiers and sensor across three spatial scales." *International Journal of Applied Earth Observation and Geoinformation*, Vol. 26: pp. 49–63.

Gleason, C.J., and Im, J. 2012. Forest biomass estimation from airborne LIDAR data using machine learning approaches. *Remote Sensing of Environment*, Vol. 125: pp. 80–91.

Günlü, A., Ercanli, A., Baskent, I., and Cakir, G. 2014. "Estimating aboveground biomass using Landsat TM imagery: a case study of Anatolian Crimean pine forests in Turkey." *Annals of Forest Research*, Vol. 57: pp. 289–298.

Guo, Y., Tian, X., Li, Z., Ling, F., Chen, E., Yan, M., Li, C., 2014. Comparison of estimating forest above-ground biomass over montane area by two non-parametric nonparametric methods. In *Proceedings of 2014 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 741–744. IEEE.

Hall, M. 1999. *Correlation-Based Feature Selection for Machine Learning*. PhD thesis, University of Waikato, Hamilton, New Zealand.

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., and Witten, I.H. 2009. The WEKA data mining software: An update. *SIGKDD Explorations*, Vol. 11(No. 1): pp. 10–18.

Haralick, R.M., Shanmugam, K., and Dinstein, I. 1973. "Textural features for image classification." *IEEE Transactions on Systems, Man and Cybernetics*, Vol. 3(No. 6): pp. 610–621.

Hartig, F., Dyke, J., Hickler, T., Higgins, S.I., O'Hara, R.B., Scheiter, S., and Huth, A. 2012. "Connecting dynamic vegetation models to data - an inverse perspective." *Journal of Biogeography*, Vol. 39(No. 12): pp. 2240–2252.

Huang, C.L., and Wang, C.J. 2006. "A GA-based feature selection and parameters optimization for support vector machines." *Expert Systems with Applications*, Vol. 31(No. 2): pp. 231–240.

Huete, A.R. 1988. "A soil-adjusted vegetation index (SAVI)." *Remote Sensing of Environment*, Vol. 25(No. 3): pp. 295–309.

Hughes, G.F. 1968. "On the mean accuracy of statistical pattern recognizers." *IEEE Transactions on Information Theory*, Vol. 14(No. 1): pp. 55–63.

INEGI - Instituto Nacional de Estadística Geográfica e Informática. 2012. Uso del suelo y vegetación escala 1: 250 000 serie v, información vectorial. México.

INEGI - Instituto Nacional de Estadística Geográfica e Informática. 2014. Continuo de elevaciones mexicano 3.0 (CEM 3.0). México.

Jachowski, N.R.A., Quak, M.S.Y., Friess, D.A., Duangnamon, D., Webb, E.L., and Ziegler, A.D. 2013. "Mangrove biomass estimation in Southwest Thailand using machine learning." *Applied Geography*, Vol. 45: pp. 311–321.

Jung, J., Kim, S., Hong, S., Kim, K., Kim, E., Im, J., and Heo. J. 2013. "Effects of national forest inventory plot location error on forest carbon stock estimation using k-nearest neighbor algorithm." *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 81: pp. 82–92.

Kelsey, K.C., and Neff, J.C. 2014. "Estimates of aboveground biomass from texture analysis of Landsat imagery." *Remote Sensing*, Vol. 6(No. 7): pp. 6407–6422.

Latifi, H., Nothdurft, A., and Koch, B. 2010. "Non-parametric prediction and mapping of standing timber volume and biomass in a temperate forest: application of multiple optical/lidar-derived predictors." *Forestry*, Vol. 83(No. 4): pp. 395–407.

Laurin, V.G., Chen, Q., Lindsell, A.J., Coomes, A.D., Del Frate, F., Guerriero, L., Pirotti, F., and Valentini, R. 2014. "Above ground biomass estimation in an African tropical forest with LiDAR and hyperspectral data." *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 89: pp. 49–58.

Le Toan, T., and Quegan, S. 2015. *BIOMASS Biomass Monitoring Mission for Carbon Assessment*. Toulouse, France: CESBIO.

López-Serrano, P.M., Corral-Rivas, J.J., Díaz-Varela, R.A., Álvarez-González, J.G., and López-Sánchez, C.A. 2016. "Evaluation of radiometric and atmospheric correction algorithms for aboveground forest biomass estimation using Landsat 5 TM data." *Remote Sensing*, Vol. 8: p. 369.

López-Serrano, P.M., López-Sánchez, C.A., Díaz-Varela, R.A., Corral-Rivas, J.J., Solís-Moreno, R., Vargas-Larreta, B., Álvarez-González, J.G. 2016. "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*, Vol. 9: pp. 226–234.

López-Serrano, P.M., López-Sánchez, C.A., Solís-Moreno, R., Corral-Rivas, J.J. 2016. "Geospatial estimation of above ground forest biomass in the Sierra Madre Occidental in the state of Durango, Mexico." *Forests*, Vol. 7: 70.

Lu, D. 2006. "The potential and challenge of remote sensing-based biomass estimation." *International Journal of Remote Sensing*, Vol. 27: pp. 1297–1328.

Lu, D., Chen, Q., Wang, G., Liu, L., Li, G., and Moran, E. 2016. A survey of remote sensing-based aboveground biomass estimation methods in forest ecosystems. *International Journal of Digital Earth*, Vol. 9(No. 1): pp 63–105.

Lu, D., Mausel, P., Brondízio, E., and Moran, E. 2004. "Relationships between forest stand parameters and Landsat TM spectral responses in the Brazilian Amazon Basin." *Forest Ecology and Management*, Vol. 198: pp. 149–167.

Lu, D., Qi, C., Guangxing, W., Moran, E., Batistella, M., Maozhen Z., Vaglio, L.G., and Saah, D. 2012. "Aboveground forest biomass estimation with Landsat and LiDAR data and uncertainty analysis of the estimates." *International Journal of Forestry Research*, Vol. 2012: pp. 1–16.

Luengo, J., García, S., and Herrera, F. 2009. "A study on the use of statistical tests for experimentation with neural networks: analysis of parametric test conditions and non-parametric tests." *Expert Systems with Applications*, Vol. 36(No. 4): pp. 7798–7808.

Malhi, Y., Meir, P., and Brown, S. 2002. "Forests, carbon and global climate." *Philosophical Transactions of the Royal Society A*, Vol. 360: pp. 1567–1591.

Malhi, Y., and Phillips, O.L. 2004. "Tropical forests and global atmospheric change: a synthesis." *Philosophical Transactions of the Royal Society*, Vol. 359: pp. 549–555.

Mateos, G.J., García, G., and Riquelme, J.C.S. 2012. "On the evolutionary optimization of k-NN by label-dependent feature weighting." *Pattern Recognition Letters*, Vol. 33(No. 16): pp. 2232–2238.

McNab, W.H. 1989. "Terrain shape index: quantifying effect of minor landforms on tree height." *Forest Science*, Vol. 35: pp. 91–104.

McRoberts, R.E. 2012. "Estimating forest attribute parameters for small areas using nearest neighbors techniques." *Forest Ecology and Management*, Vol. 272: pp. 3–12.

Molinaro, A.M., Simon, R., and Pfeer, R.M., 2005. "Prediction error estimation: a comparison of resampling methods." *Bioinformatics*, Vol. 21(No. 15): pp. 3301–3307.

Moore, I.D., and Nieber, J.L. 1989. "Landscape assessment of soil erosion and non-point source pollution." *Journal of the Minnesota Academy of Science*, Vol. 55: pp. 18–24.

Mutanga, O., Adam, E., and Cho, M.A., 2012. "High density biomass estimation for wetland vegetation using worldview-2 imagery and random forest regression algorithm." *International Journal of Applied Earth Observation and Geoinformation*, Vol. 18: pp. 399–406.

Mutanga, O., and Skidmore, A.K. 2004. "Narrow band vegetation indices overcome the saturation problem in biomass esti-

mation." *International Journal of Remote Sensing*, Vol. 25: pp. 3999–4014.

Næsset, E., Gobakken, T., Bollandsås, O.M., Gregoire, T.G., Nelson, R., and Ståhl. G. 2013. "Comparison of precision of biomass estimates in regional field sample surveys and airborne LiDAR-assisted surveys in Hedmark County, Norway." *Remote Sensing of Environment*, Vol. 130: pp. 108–120.

NASA 2011. *Landsat 7 Science Data Users Handbook*, accessed August 16, 2016, http://landsathandbook.gsfc.nasa.gov/pdfs/Landsat7_Handbook.pdf.

Packalén, P., Temesgen, H., and Maltamo, M. 2012. "Variable selection strategies for nearest neighbor imputation methods used in remote sensing based forest inventory." *Canadian Journal of Remote Sensing*, Vol. 38(No. 6): pp. 557–569.

Pflugmacher, D., Cohen, W.B., Kennedy, R.E., and Yang, Z. 2014. "Using Landsat-derived disturbance and recovery history and LiDAR to map forest biomass dynamics." *Remote Sensing of Environment*, Vol. 151: pp. 124–137.

Pretzsch, H., Biber, P., Ïurský, J., von Gadow, K., Hasenauer, H., Kändler G., Kenk, G., et al. 2002. "Recommendations for standardized documentation and further development of forest growth simulators." *Forstwissenschaftliches Centralblatt*, Vol. 121(No. 3): pp. 138–151.

Qi, J., Chehbouni, A., Huete, A.R., and Kerr, Y.H. 1994. "A modified soil adjusted vegetation index." *Remote Sensing of Environment*, Vol. 48(No. 2): pp. 119–126.

R Core Team. 2014. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Roberts, D.W., and Cooper, S.V. 1989. "Concepts and techniques of vegetation mapping." In *Land Classifications Based on Vegetation: Applications for Resource Management, USDA Forest Service GTR INT-257*, edited by D. Ferguson, P. Morgan, and F.D. Johnson, pp. 90–96. Ogden, UT: USDA Forest Service.

Rouse, J.W., Haas, R.H., Schiell, J.A., Deferino, D.W., and Harlan, J.C. 1974. *Monitoring the vernal advancement of retrogradation of natural vegetation. NASA/GSFC Type III*, Greenbelt, MD: NASA.

Samadzadegan, F., Hasani, H., and Schenk, T. 2012. "Simultaneous feature selection and SVM parameter determination in classification of hyperspectral imagery using ant colony optimization." *Canadian Journal of Remote Sensing*, Vol. 38(No. 2): pp. 139–156.

Shataee, S. 2013. "Forest attributes estimation using aerial laser scanner and TM Data." *Forest Systems*, Vol. 22(No. 3): 484–496.

Tian, X., Li, Z., Su, Z., Chen, E., Van der Tol, C., Li, X., Guo, Y., Li L., and Ling, F. 2014. "Estimating montane forest above-ground biomass in the upper reaches of the Heihe River Basin using Landsat-TM data." *International Journal of Remote Sensing*, Vol. 35(No. 21): pp. 7339–7362.

Tian, X., Su, Z., Chen, E., Li, Z., van der Tol, C., Guo, J., and He, Q. 2012. "Estimation of forest above-ground biomass using multi-parameter remote sensing data over a cold and arid area." *International Journal of Applied Earth Observation and Geoinformation*, Vol. 14(No. 1): pp. 160–168.

Vargas-Larreta, B. 2013. Estimación del potencial de los bosques de Durango para la mitigación del cambio climático. Modelización de la biomasa forestal. Proyecto: FOMIX-DGO-2011-C01-165681. ITES, EL Salto (México).

Wang, L., Zhou, X., Zhu, X., Dong, X., and Guo, W. 2016. "Estimation of biomass in wheat using random forest regression algorithm and remote sensing data." *The Crop Journal* (in press).

Wehenkel, C., Corral-Rivas, J.J., Hernández-Díaz, J.C., von Gadow, K. 2011. "Estimating balanced structure areas in multi-species forests on the Sierra Madre Occidental, Mexico." *Annals of Forest Science*, Vol. 68: pp. 385–394.

Wilson, J.P., and Gallant, J.C. 2000. "Digital terrain analysis." In *Terrain Analysis: Principles and Applications*. New York, NY: John Wiley & Sons.

Wu, C., Shen, H., Wang, K., Shen, A., Deng, J., and Gan, M. 2016. "Landsat imagery-based above ground biomass estimation and change investigation related to human activities." *Sustainability*, Vol. 8(No. 2): 159.

Xie, X., Liu, W.T., and Tang, B. 2008. "Space-based estimation of moisture transport in marine atmosphere using support vector regression." *Remote Sensing of Environment*, Vol. 112(No. 4): pp. 1846–1855.

Yang, Y., Monserud, R.A., and Huang, S. 2004. "An evaluation of diagnostic tests and their roles in validating forest biometrics models." *Canadian Journal of Forest Research*, Vol. 34: pp. 619–629.

Zhang, Y., Liang, S., and Sun, G. 2014. "Forest biomass mapping of northeastern China using GLAS and MODIS data." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, Vol. 7: pp. 140–152.

Zhao, K., Popescu, S., Meng, X., Pang, Y, and Agca, M. 2011. "Characterizing forest canopy structure with LiDAR composite metrics and machine learning." *Remote Sensing of Environment*, Vol. 115(No. 8): pp. 1978–1996.

Zhu, X., and Liu, D. 2015. "Improving forest aboveground biomass estimation using seasonal Landsat NDVI time-series." *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 102: pp. 222–231.