OPEN ACCESS

Remote Sensing

Article

Semi-Supervised Methods to Identify Individual Crowns of Lowland Tropical Canopy Species Using Imaging Spectroscopy and LiDAR

Jean-Baptiste Féret * and Gregory P. Asner

Department of Global Ecology, Carnegie Institution for Science, Stanford, CA 94305, USA;
E-Mail: gpa@stanford.edu

* Author to whom correspondence should be addressed; E-Mail: jbferet@stanford.edu;
Tel.: +1-650-462-1047; Fax: +1-650-462-5968.

Abstract: Our objective is to identify and map individuals of nine tree species in a Hawaiian lowland tropical forest by comparing the performance of a variety of semi-supervised classifiers. A method was adapted to process hyperspectral imagery, LiDAR intensity variables, and LiDAR-derived canopy height and use them to assess the identification accuracy. We found that semi-supervised Support Vector Machine classification using tensor summation kernel was superior to supervised classification, with demonstrable accuracy for at least eight out of nine species, and for all combinations of data types tested. We also found that the combination of hyperspectral imagery and LiDAR data usually improved species classification. Both LiDAR intensity and LiDAR canopy height proved useful for classification of certain species, but the improvements varied depending upon the species in question. Our results pave the way for target-species identification in tropical forests and other ecosystems.

Keywords: biodiversity; Carnegie Airborne Observatory; species mapping; hyperspectral imagery; canopy height; LiDAR intensity; tropical forests; semi-supervised classification

1. Introduction

Large-scale mapping of tree species composition is of growing interest in ecology, conservation, and ecosystem management. Imaging spectroscopy, also known as hyperspectral imaging, has proven

to be among the most useful technologies for species mapping [1–10]. By comparison, active remote sensing instruments, such as Light Detection and Ranging (LiDAR) scanners, have played a relatively small role in efforts to detect and map tree species [11,12]. Studies using LiDAR in conjunction with spectral data [13–16] have yielded promising results, indicating the potential of combining remote sensing technologies for species mapping.

Despite the demonstrable value of combined spectral and LiDAR data for species mapping, the application of these approaches to high-diversity tropical forests presents a unique challenge. Recent work has shown that multiple species can be detected in tropical forests [2,5], yet accuracies and the potential for automation remain highly uncertain. Most studies have been conducted at the leaf level, and have found that species often have spectral signatures that might render them identifiable in high-resolution airborne or space-based imagery [17–19]. Other studies, however, have found substantial spectral overlap between species, either based on their chemical traits or phenology or both [20]. All studies agree that the risk of spectral confusion will rise with increased species richness. At the canopy level, a combination of chemical (leaf level) and structural (canopy level) characteristics can render species relatively distinct [11,21], but much more research is needed to assess the potential of this approach, and to develop methods for mapping species in tropical environments.

Because of these challenges, mapping species in high-diversity tropical forests is probably best approached on a species-by-species basis. First, in lieu of an exhaustive mapping of all species, which is technically unlikely, focusing on a limited number of ecologically important or indicator species may be more tractable. In addition, target species can be chosen that are important to ecosystem functions such as carbon and nutrient cycling [22], or to food-web dynamics [23].

Focusing on a limited number of target species can only improve classification accuracy, however, if species are clearly and repeatably separable. A wide range of classification methods are available to identify tropical species using hyperspectral data [2,3,5], but to our knowledge only supervised classification methods have been used for this purpose. Despite relatively good performance with a moderate number of classes, supervised methods suffer from one important weakness in regards to species mapping: the classifier only learns from input training data. This is a problem because, due to the time and cost of data collection in tropical forests, training datasets may not capture the full range of spectral variability in each class (*i.e.*, each species), resulting in an inaccurate description of each class and a poor generalization of the classification model. An alternative to supervised classification is semi-supervised classification, which uses unlabeled (*i.e.*, unidentified) individuals to supplement the training data during the training stage. This is done by assigning randomly-selected pixels to existing classes on the basis of their spectral similarity, thus improving the representativeness of the training data. Because these unlabeled samples exist in large quantities and can improve discrimination of tropical tree species, the ability of semi-supervised classification to improve image classification needs to be explored.

Here we compare one supervised and two semi-supervised methods for target-species detection in a humid tropical forest. These classification methods belong to the Support Vector Machine (SVM) framework and the very first task is a parameter optimization. We also combine airborne imaging spectrometer and LiDAR measurements to determine which data type provides the highest accuracy for forest species mapping. Finally, we determine which combination of classifier and data type performs the best for the classification of nine tree species, with a constraint on the training data: we

systematically assume that only one individual tree crown from the target species and one individual tree crown from half of the eight non-target species are available for training.

2. Materials

2.1. Hyperspectral Imagery and LiDAR Data

To compare the different methods for classification, we used field and remotely-sensed data from the Nanawale Forest Reserve, Hawaii. We acquired imaging spectrometer data with the Carnegie Airborne Observatory (CAO)-Alpha sensor package [24] in September 2007, flying at an altitude of 1,000 m above ground surface. For this study we used a 1,980-by-1,420 pixel image, with a spatial resolution of 0.56 m and covering an area of about 70 ha (Figure 1). This image has 24 spectral bands of 28 nm in width, evenly spaced between 390 nm and 1,044 nm. The image thus covers both visible (VIS) wavelengths and part of the near-infrared (NIR) domain, and spectra were radiometrically calibrated in the laboratory following the flight. Atmospheric corrections were applied using the ACORN 5LiBatch (Imspec LLC) model, and a MODTRAN look-up table was used to correct for Rayleigh scattering and aerosols [25].

Figure 1. Carnegie Airborne Observatory (CAO) image of the Nanawale Forest Reserve (HI). The three channels used to display the image are (R = 646.0 nm; G = 560.7 nm; B = 447.0 nm). The red contours correspond to the individual tree crowns delineated after the field survey.
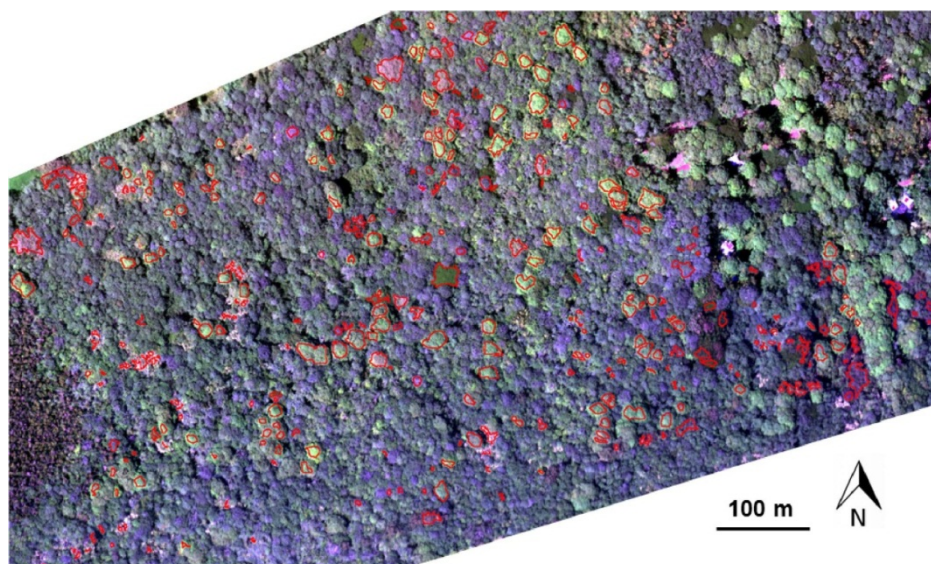
LiDAR data were acquired during the same flight as the hyperspectral imagery, and the LiDAR was operated in discrete-return mode, providing up to four laser returns per shot. The LiDAR beam divergence was matched to that of the spectrometer [24], and thus the laser spot spacing was 0.56 m both across and down-track. We additionally collected 50% overlap between adjacent flightlines, resulting in two laser shots per 0.56 m. From the LiDAR point cloud data, a physically-based model was used to estimate top-of-canopy and ground surfaces using REALM (Optech Inc., Toronto, ON, Canada) and Terrascan/Terramatch (Terrasolid Ltd., Jyväskylä, Finland) software packages. We

computed vegetation height by differencing the top-of-canopy and ground surface DEMs [26], and three intensity variables, corresponding to the first-and-only return, first-of-many returns, and first return.

2.2. Study Site and Field Data Collection

The study area contains lowland humid tropical forest, at an average elevation of 150 m above sea level, and with a mean annual precipitation and temperature of 3,200 mm·yr^{-1} and 23 °C, respectively. The forest emergent canopy at Nanawale is comprised of approximately 17 species, most of which are invasive, non-native trees, though some native species do remain in this forest [27].

We did field data collection and identified a total of 791 individual tree crowns (ITCs) from 17 species in the field in November 2010 using a tablet PC with integrated, differentially-corrected Global Position System (PC-GPS). Using this system, we delineated tree crowns on the PC-GPS using the hyperspectral imagery as a guide, and tree crowns varied in size from a few pixels to several thousand pixels. For classification purposes (see Section 3 for more details), we reduced the dataset by discarding ITCs smaller than 50 pixels, and species with less than 12 ITCs. The final dataset used for this study encompassed 333 ITCs from nine different species (Table 1, Figure 1).

Table 1. Listing of the species studied and mapped, along with the number of crowns and pixels per crown delineated in the field.

Species	Code	Crowns	$\bar{X}_{\text{Pixels/crown}}$ (±SD)
Aleurites moluccana	ALEMOL	40	105.0 ± 108.5
Cecropia peltata	CECPEL	37	131.6 ± 146.2
Eucalyptus robusta	EUCROB	12	303.4 ± 263.9
Mangifera indica	MANIND	135	374.4 ± 327.7
Pandanus tectorius	PANTEC	18	113.3 ± 78.5
Pithecellobium saman	PITSAM	17	367.7 ± 424.7
Psidium cattleianum	PSICAT	19	113.3 ± 90.6
Syzygium jambos	SYZJAM	16	305.3 ± 412.2
Trema orientalis	TREORI	39	367.7 ± 441.6

Despite the 3-year time lag between the acquisition of the image (September 2007) and the collection of the ground truth (November 2010), the relationships between the pixel label and the spectral signature are still reliable because there is a slow evolution of these tropical tree species over time. Moreover, we paid particular attention to delineate sunlit parts and avoid shaded parts and edges where the risk of overlap between ITCs increases.

3. Methods for Classification

3.1. Generation of the Datasets Used during Classification

All of the classification methods tested required two datasets: a training dataset used to build the classification model; and a test dataset, independent from the training dataset. For each of these datasets, we defined two classes corresponding to the target species and the non-target species. Under optimal conditions, the non-target class would include samples from all species except the target

species. However, often in tropical forests, only a fraction of the species can be identified during a field campaign. As such, our goal was to compare the performance of our species classification techniques using only limited information about the non-target species present in the study area, so we deliberately included only half of the non-target species in the training dataset. The typical training dataset was thus assembled by randomly selecting one ITC from the target species and one ITC from four non-target species (randomly selected among the eight available species). We then randomly selected fifty samples (*i.e.*, pixels) from the target ITC and 30 samples from each non-target ITC. The whole training dataset for each target species thus included 170 samples. The test dataset included 99 ITCs (11 for each of the nine species). We performed the classification on a pixel-by-pixel basis, and identified each ITC using a majority vote rule. The performance of the classifier was assessed based on the successful classification of the target species and the non-target species.

3.2. Estimation of the Classification Performance

Because non-target pixels were more abundant than target pixels, a simple measure of overall accuracy would have been strongly biased to favor the non-target class and would thus have substantially over-reported the accuracy of the method. Therefore we calculated the balanced accuracy (BAC) defined as:

$$BAC = \frac{Sensitivity + Specificity}{2} \tag{1}$$

where sensitivity is the proportion of pixels of target species that were correctly identified as target species (*i.e.*, the number of correctly identified target species pixels divided by the total number of target species pixels), and specificity is the proportion of pixels of non-target species that were correctly identified as non-target species (*i.e.*, the number of correctly identified non-target species pixels divided by the total number of non-target species pixels). We maximized BAC during the optimization of the classifiers and the training stage, and used sensitivity, specificity, and BAC to compare the performance of the classifiers tested here.

3.3. General Framework for Supervised and Semi-Supervised Classification

Both the supervised and semi-supervised classifications performed in this study were based on the SVM, a popular non-parametric classifier widely used in the machine learning and remote sensing communities [28]. SVM uses decision boundaries (hyperplanes) derived from a training dataset to separate classes in the feature space. These linear boundaries between classes are generated by maximizing the margins between the hyperplane and the closest training samples (*i.e.*, the support vectors) and minimizing the error of the training samples that cannot be differentiated. As the classes are rarely linearly separable in the original feature space, SVM projects the training dataset into a kernel feature space of higher dimensionality. This is performed nonlinearly, using the following kernel function:

$$\boldsymbol{K}_{ij} = K\left(x_i, x_j\right) = \langle \Phi(x_i), \Phi(x_j) \rangle \tag{2}$$

where $\boldsymbol{K}_{ij}$ is the value of the kernel function for the pair (i, j) of training samples, and Φ is the mapping function used to project data from the original feature space to the higher-dimensionality feature space [29–31].

SVMs are highly efficient and usually produce comparable or better results than other methods (see references from [32–35]). They are known to show good generalized performance and require no prior knowledge of the problem [31]. They are also particularly useful for hyperspectral imagery classification because they can handle a large input space efficiently, and are less subject to the Hughes phenomenon than other classifiers [36]. Féret and Asner [5] have investigated the ability of these classifiers to identify tree species in tropical forests, and showed that linear (L-) SVM and radial basis function (RBF-) SVM both outperform other non-parametric classifiers such as the *k*-nearest neighbor or artificial neural network approaches, and have comparable or better performance than discriminant analysis. Our study focused on RBF-SVM, and all classification tasks were performed using the MATLAB interface of the *LIBSVM* package. The RBF function is written as follows:

$$K\left(x_i, x_j\right) = exp\left(-\frac{\left\|x_i - x_j\right\|^2}{2\sigma^2}\right) \tag{3}$$

where $\sigma \in \mathbb{R}^+$ is the kernel width which controls the tradeoff between over-fitting (small values) and under-fitting (large values). A second parameter, the penalty parameter C (also called the error term), must also be optimized. This term controls the trade-off between complexity of decision rule and frequency of training error [29]. In this study, we first rescaled hyperspectral and LiDAR data between 0 and 1, and then calculated the optimal parameters using an exhaustive grid search. The optimal σ and C were tuned in the range from 10^{-3} to 10^3 using a 3-fold cross-validation applied to a training dataset.

3.4. Characteristics Specific to the Semi-Supervised Classification

We compared two distinct classification approaches: supervised and semi-supervised classification. The main difference between these methods was the use of unlabeled samples, in addition to the training data, by semi-supervised methods. The semi-supervised classification takes advantage of complementary data corresponding to unlabeled samples in order to improve the estimation of the marginal data distribution during the training stage. To start the semi-supervised approach, we first randomly selected 500 unlabeled pixels from the total dataset. We then implemented the semi-supervised approach proposed by Tuia and Camps-Valls [37], which is based on the local regularization of the training kernel. The information contained in these unlabeled samples is used to create a bagged kernel combined with the training kernel in order to deform its base structure through a cluster-based method. This bagged kernel is obtained after successive *k-means* clustering (with different initialization but the same number of clusters) are performed on the combined training/unlabeled samples. The bagged kernel accounts for the number of time two samples *i* and *j* have been assigned to the same cluster. Here we compared two different kernels: the tensor product kernel which deforms the training kernel by multiplying it with the bagged kernel, and the tensor summation kernel which deforms the training kernel by adding it with the bagged kernel. A package including MATLAB source code for this method is publicly available [37] (http://www.uv.es/gcamps/code/bagsvm.htm).

Two parameters specific to the bagged kernels were defined prior to classification: t, the number of times running the *k-means* clustering and k, the number of clusters to create. We used $t = 20$. To determine the optimal number of clusters (k) used to create the bagged kernel we evaluated k values ranging from 5 to 40 (using an increment of 5) on the classification of target species, and selected the optimal k value for each tensor product and summation kernel. To identify this k value, we performed 100 iterations for each target species, each iteration starting with the generation of one training dataset, one test dataset, and one unlabeled dataset. Once these datasets were created, we trained a corresponding classification model for each possible combination of "data type/semi-supervised classifier/k value", resulting in 64 (4 × 2 × 8) semi-supervised classification models for each iteration and each species. We then averaged the BAC for all iterations and all possible target species in order to obtain one BAC for each value of k and each classifier. The value of k that maximized this average BAC was selected as the optimal k and used for the four different combinations hyperspectral and LiDAR data.

3.5. Comparison of the Classifiers

Once the optimal value of k was obtained, we compared the two semi-supervised approaches and one supervised classification. This experiment was similar to the one explained for the optimization of the k value. Here our aim was to assess the influence of two independent factors on the performance of the classification: (i) the type of data; and (ii) the method for classification (supervised SVM, semi-supervised SVM with tensor summation kernel, semi-supervised SVM with tensor product kernel). This results in 12 possible configurations of data type/classification method. As the results of the classification are highly dependent on the training data, we performed repetitions by generating 100 datasets (including training, test and unlabeled data) for each target species, and we created one classification model for each of the 12 configurations. We finally compared the global performance obtained for each configuration and the results specific to each species in each of the configurations. One-tailed t-tests were used to determine which configuration obtained significantly better results.

4. Results

4.1. Optimization of the k-Means Parameterization

We compared the influence of k on the sensitivity, specificity and BAC of the two semi-supervised classifiers, and observed several common responses. First, lower values of k resulted in decreased sensitivity, regardless of the classifier, while higher values of k resulted in higher specificity, (Table 2). A value of $k = 10$ maximized sensitivity for the tensor summation kernel, and a value of $k = 5$ for the tensor product kernel (Table 2). Furthermore, for both semi-supervised methods specificity increased with k, reaching its maximum value for $k = 40$. Despite these similarities, the optimal BAC differed substantially for the two methods: the maximum BAC for the tensor summation and tensor product kernels were obtained for $k = 35$ and $k = 10$, respectively. This is in agreement with the results obtained by Tuia *et al.* [37], who conclude that the optimal results for the tensor summation kernel are obtained for larger values of k, while the tensor product kernel performs better at small values of k.

Table 2. Sensitivity, specificity and balanced accuracy (BAC) as function of the number of clusters *k* for tensor summation kernel, tensor product kernel and supervised classification.

k	Tensor Summation Kernel			Tensor Product Kernel		
	Sensitivity	**Specificity**	**BAC**	**Sensitivity**	**Specificity**	**BAC**
5	50.11	88.23	69.17	**40.94**	91.23	66.09
10	**53.66**	88.32	70.99	40.91	91.61	**66.26**
15	53.50	88.94	71.22	40.04	92.34	66.19
20	53.61	89.24	71.42	39.52	92.80	66.16
25	53.22	89.46	71.34	38.33	93.20	65.76
30	52.78	89.94	71.36	37.23	93.68	65.45
35	52.71	90.25	**71.48**	36.57	94.38	65.48
40	51.74	**90.49**	71.12	35.47	**94.78**	65.13
Supervised SVM	36.47	93.06	64.77			

The optimal *k* value for the tensor product kernel was most influenced by sensitivity, which decreased faster than the increase in specificity for higher values of *k*; while the optimal *k* value for the tensor summation kernel was most influenced by specificity, which increased faster than the reduction in sensitivity for higher values of *k*. Strong discrepancies were also observed between the two semi-supervised methods: the most important difference was that sensitivity was 9–16 percentage points higher with the tensor summation kernel. This difference was balanced by the lower specificity of this method (3–4 percentage points). Despite this reduction in specificity, the tensor summation kernel outperformed the tensor product kernel by 3–6 percentage points based on BAC values (Table 2). The average specificity exceeded 90% for the value of *k* selected in our experiment. This result is encouraging because it means that the number of species incorrectly identified as a target species was usually very low. Based on these results, we selected *k* = 35 clusters for the tensor summation kernel, and *k* = 10 clusters for the tensor product kernel, for all following classifications.

4.2. Classification Method and Data Type

We found that the summation-bagged kernel outperformed the product-bagged kernel of 4.2 percentage points, and the supervised SVM of 6.5 percentage points, based on BAC averaged over 100 iterations performed on the nine species. Both semi-supervised methods improved the sensitivity to target species: they outperformed the supervised method by 7–15 percentage points in terms of sensitivity (Table 3). Among the two semi-supervised classifiers, the tensor summation kernel obtained 5–9 percentage points greater sensitivity than the tensor product kernel, depending on the data type used (Table 3).

From these results we conclude that the tensor summation kernel offers the best performance for target species discrimination in the forests studied here. Furthermore, fusing hyperspectral measurements with LiDAR canopy height and intensity data increased classification accuracy (BAC, Table 3), as averaged over all species. Combining hyperspectral data with either of the LiDAR data types alone, however, resulted in similar or slightly reduced accuracy in most of the cases, mainly due to lower sensitivity.

Table 3. Sensitivity, specificity and balanced accuracy (BAC) for the different classification methods and data sources (averaged after 100 repetitions). HS: hyperspectral; CH: canopy height.

Classification	(1)	(2)	(3)
Data Type	**Sensitivity**		
Hyperspectral	39.1	**46.4**	53.1
HS + Intensity	35.3	45.3	51.7
HS + CH	39.6	45.5	50.6
HS + Intensity+ CH	**39.9**	46.3	**55.1**
	Specificity		
Hyperspectral	92.5	90.6	90.8
HS + Intensity	**93.6**	91.6	91.4
HS + CH	93.3	**92.0**	91.5
HS + Intensity+ CH	**93.6**	91.9	**91.6**
	BAC		
Hyperspectral	65.8	68.5	71.9
HS + Intensity	64.4	68.5	71.5
HS + CH	66.4	68.7	71.1
HS + Intensity+ CH	**66.8**	**69.1**	**73.3**

(1): supervised SVM; (2): tensor product kernel; (3): tensor summation kernel.

We reviewed species-specific results throughout our analysis. First we compared the different classifiers for a given data type. As observed overall species, the individual difference in BAC between classifiers (Figure 2) was mainly driven by the difference in sensitivity of the classifier for all data types (Figure 3). The increased overall sensitivity obtained with the tensor summation kernel (Table 3) was also observed when studying species separately, with the exception of *Syzygium jambos* which was usually classified more accurately when using a supervised method (Figure 3, Table 2). This result was observed for all four data types, but the difference in sensitivity between the three classification methods was species-dependent, with *Aleurites moluccana* showing the highest increase in sensitivity of all species when using semi-supervised classification.

Figure 2. BAC measured for each species and three classification methods (circles: supervised RBF-SVM; squares: tensor product kernel; triangles: tensor summation kernel) when using (**a**) hyperspectral data only; (**b**) hyperspectral data+LiDAR intensity; (**c**) hyperspectral data+LiDAR CH; (**d**) hyperspectral data+all LiDAR data.

Figure 2. *Cont.*

Figure 2 legend:
- ● BAC with RBF-SVM
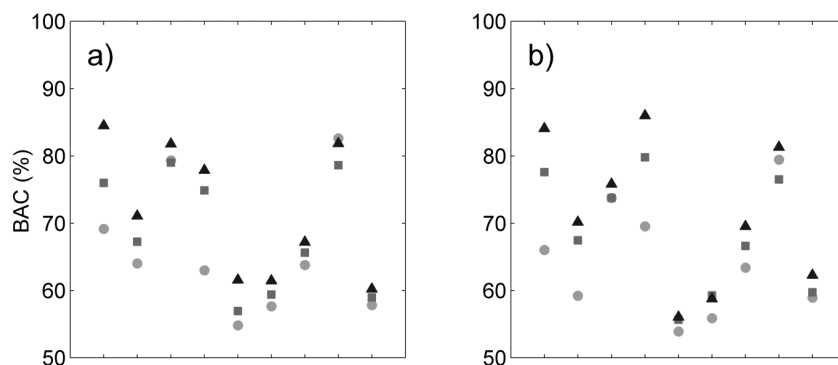- ■ BAC with SS-SVM (tensor product kernel)
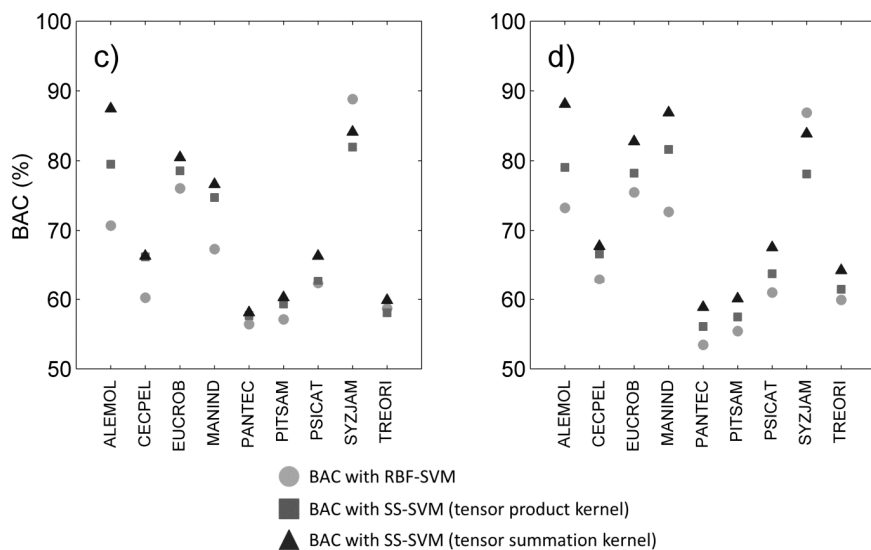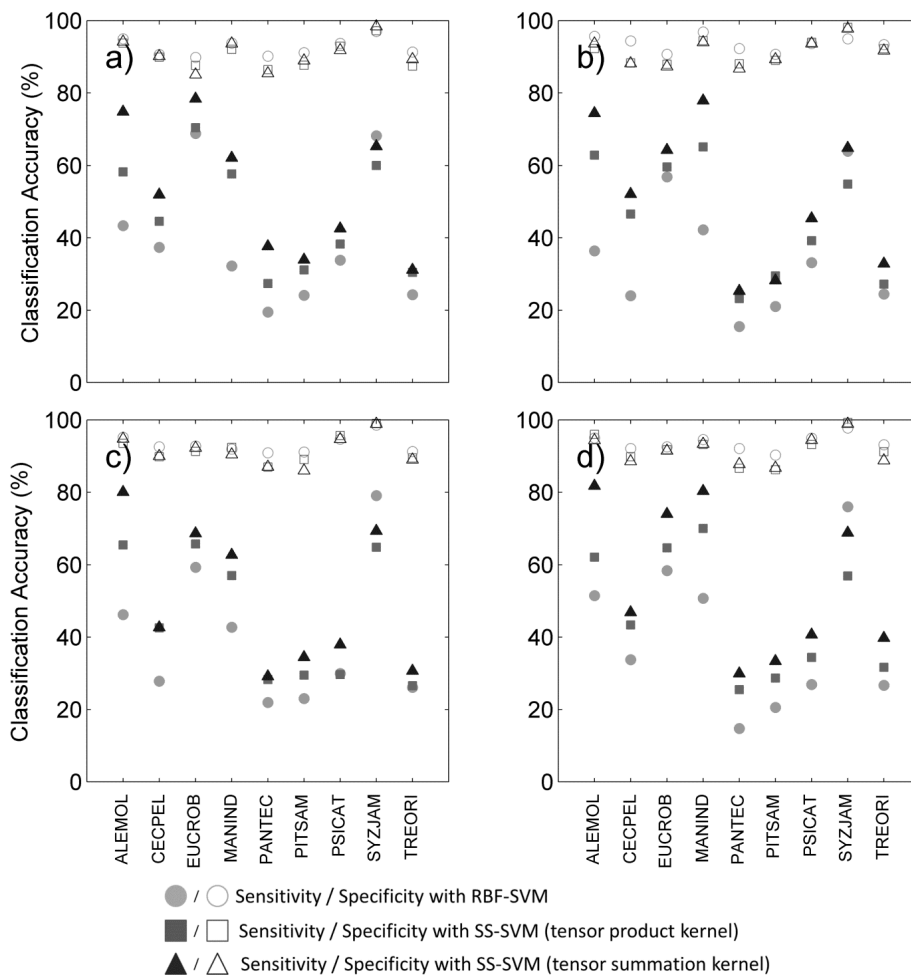- ▲ BAC with SS-SVM (tensor summation kernel)

Figure 3. Sensitivity (filled symbols) and specificity (unfilled symbols) for each species and three classification methods (circles: supervised RBF-SVM; squares: tensor product kernel; triangles: tensor summation kernel) when using (**a**) hyperspectral data only; (**b**) hyperspectral data + LiDAR intensity; (**c**) hyperspectral data + LiDAR CH model; (**d**) hyperspectral data + all LiDAR data.

Figure 3 legend:
- ● / ○ Sensitivity / Specificity with RBF-SVM
- ■ / □ Sensitivity / Specificity with SS-SVM (tensor product kernel)
- ▲ / △ Sensitivity / Specificity with SS-SVM (tensor summation kernel)

The influence of data type on the classification varied with the species targeted, and no common trend was observed. For example, *Cecropia pelata* was classified more accurately with hyperspectral data alone when using supervised classification. However, combining hyperspectral data with LiDAR intensity slightly improved the sensitivity obtained for this species for semi-supervised methods, but combining hyperspectral data with LiDAR canopy height reduced sensitivity. Another example is *Mangifera indica*, for which combining hyperspectral data and LiDAR intensity resulted in a greater than 15 percentage points improvement in sensitivity (Figure 3(a,b)) and 8 percentage points improvement in BAC (Figure 2(a,b)) when using a tensor summation kernel. In contrast, combining LiDAR intensity with hyperspectral data reduced the BAC for *Aleurites moluccana* and *Syzygium jambos*, but combining canopy height with hyperspectral data improved their classification accuracy (Figure 2(a,c)).

These species-specific differences were partly explained by further examining the LiDAR canopy height and intensity values. Among the nine species studied here, some show very specific behavior in terms of intensity or height, which may help with the discrimination of these species despite their high within-species variability (Table 4). For example, in the case of *Mangifera indica*, the large difference measured between the mean intensity value of the first and only returns (the highest of all species) and the mean intensity value of the first of many returns (the lowest of all species) explains the significant improvement observed for its discrimination when including LiDAR intensity to the classification (see Table 4). Almost no first of many returns points were recorded for *Mangifera indica*, leading to very low intensity values. The low mean height measured for *Aleurites moluccana* and *Syzygium jambos* also explains their improved discrimination when combining canopy height with hyperspectral data.

Table 4. Mean value (±SD) of LiDAR-based measurements on a species basis.

Species	Spectral Variables			Spatial Variable
	First and Only	First of Many	First	Height
Aleurites moluccana	27.8 ± 13.9	9.5 ± 19.2	33.8 ± 8.8	13.5 ± 1.7
Cecropia peltata	9.7 ± 13.5	41.7 ± 26.7	29.4 ± 8.8	20.0 ± 3.3
Eucalyptus robusta	7.1 ± 10.9	19.9 ± 14.0	21.3 ± 7.2	22.1 ± 2.8
Mangifera indica	40.8 ± 12.1	1.6 ± 8.6	42.1 ± 10.3	19.8 ± 2.4
Pandanus tectorius	9.5 ± 13.6	29.6 ± 21.9	25.7 ± 9.2	16.1 ± 3.1
Pithecellobium saman	18.7 ± 18.6	18.6 ± 23.3	29.7 ± 12.8	18.0 ± 3.3
Psidium cattleianum	29.7 ± 14.8	12.0 ± 21.2	35.4 ± 9.2	17.1 ± 5.3
Syzygium jambos	15.7 ± 13.6	16.1 ± 16.3	25.3 ± 7.0	10.8 ± 2.3
Trema orientalis	2.5 ± 6.9	25.0 ± 17.7	19.7 ± 8.35	18.9 ± 2.6

Finally we studied which combination of data significantly improved BAC compared to any given data type (Table 5). We tested all classifiers and data types, but the results showed here are obtained with semi-supervised classification and tensor summation kernel. The classification accuracy for *Pandanus tectorius* was significantly higher when LiDAR information was not used, but this species exhibits the lowest BAC of all nine species (Figure 2), which suggests that this is an inappropriate candidate for species targeting. The classification accuracy for *Cecropia peltata* was also significantly improved when the LiDAR height was not used, either by itself or in combination with intensity, but LiDAR intensity did not impact the classification. This LiDAR intensity data significantly improved

the classification of *Mangifera indica*, *Trema orientalis* and *Psidium cattleianum*; for the latter, however, adding LiDAR height did not lead to any improvement. The addition of either one or the other LiDAR variable did not improve the classification accuracy of the same species, highlighting the different discriminating capabilities related to each of these data types. Finally these results show that the accuracy obtained when using all data types was comparable to or significantly better than when the LiDAR data were not used or only partly used for six species whereas only three species were significantly better classified when not using all data types together: *Cecropia peltata*, *Pandanus tectorius*, *Psidium cattleianum*. Notice that these three species show a relatively low BAC, suggesting that they would not be good candidates for target species classification. These results obtained on these three species are not surprising as Féret and Asner [5] obtained particularly low producer's and user's accuracy with these species (40 to 60%) when performing multiclass discrimination including these nine species and eight other ones. These results confirm the adding of multi-sensor information combining LiDAR and hyperspectral imagery for species classification.

Table 5. Significant increase in BAC induced by the data type for each species after semi-supervised classification with tensor summation kernel. A right-tailed paried *t*-test is used for statistical significance ($p = 0.05$). The BAC obtained with the data type indicated in the column is significantly superior to the BAC obtained with the data type indicated in the line for the species reported in the corresponding box. HSI = hyperspectral imagery; Int. = LiDAR intensity; CHM = canopy height model.

	HSI	HSI + Int.	HSI + CHM	HSI + Int + CHM
HSI		4; 7; 9	1; 8	1; 4; 9
HSI + Int.	3; 5; 6		1; 3; 5; 8	1; 3; 5; 8; 9
HSI + CHM	2; 5	2; 4; 7; 9		4; 9
HSI + Int + CHM	2; 5	7		

1: *Aleurites moluccana*; 2: *Cecropia peltata*; 3: *Eucalyptus robusta*; 4: *Mangifera indica*; 5: *Pandanus tectorius*; 6: *Pithecellobium saman*; 7: *Psidium cattleianum*; 8: *Syzygium jambos*; 9: *Trema orientalis*.

4.3. Application to a Practical Situation

We performed target species classification on the whole site for each species, using semi-supervised classification with tensor summation kernels on the 333 ITCs available, with the full data combination (Figure 4). For each species, we used the classification model that produced the highest BAC among the 100 iterations performed in Section 4.2. All species except *Pandanus tectorius* showed sensitivity values greater than 80% ITCs (Table 6). These values were surprisingly high for some species such as *Pithecellobium saman*, *Psidium cattleianum* and *Trema orientalis*, as they were dramatically higher than the mean sensitivity measured after 100 iterations, which is about 40% for these three species (Figure 3(d)). These findings confirm the importance of the training dataset, and the very high variability of the discriminating capabilities of individual tree crowns from the same species. The specificity was also higher than 80% for all species, and superior to 95% for five of them, which was expected given the results obtained previously (see Figure 3(d)). The high specificity observed here is promising for future applications targeting individual species, but these findings must be placed in the perspective of operational applications.

Figure 4. Detection of the nine species over the whole site. (**Green**): target species correctly identified (true positive); (**Blue**): non-target species correctly identified (true negative); (**White**): target species misidentified (false negative); (**Red**): non-target species misidentified (false positive).
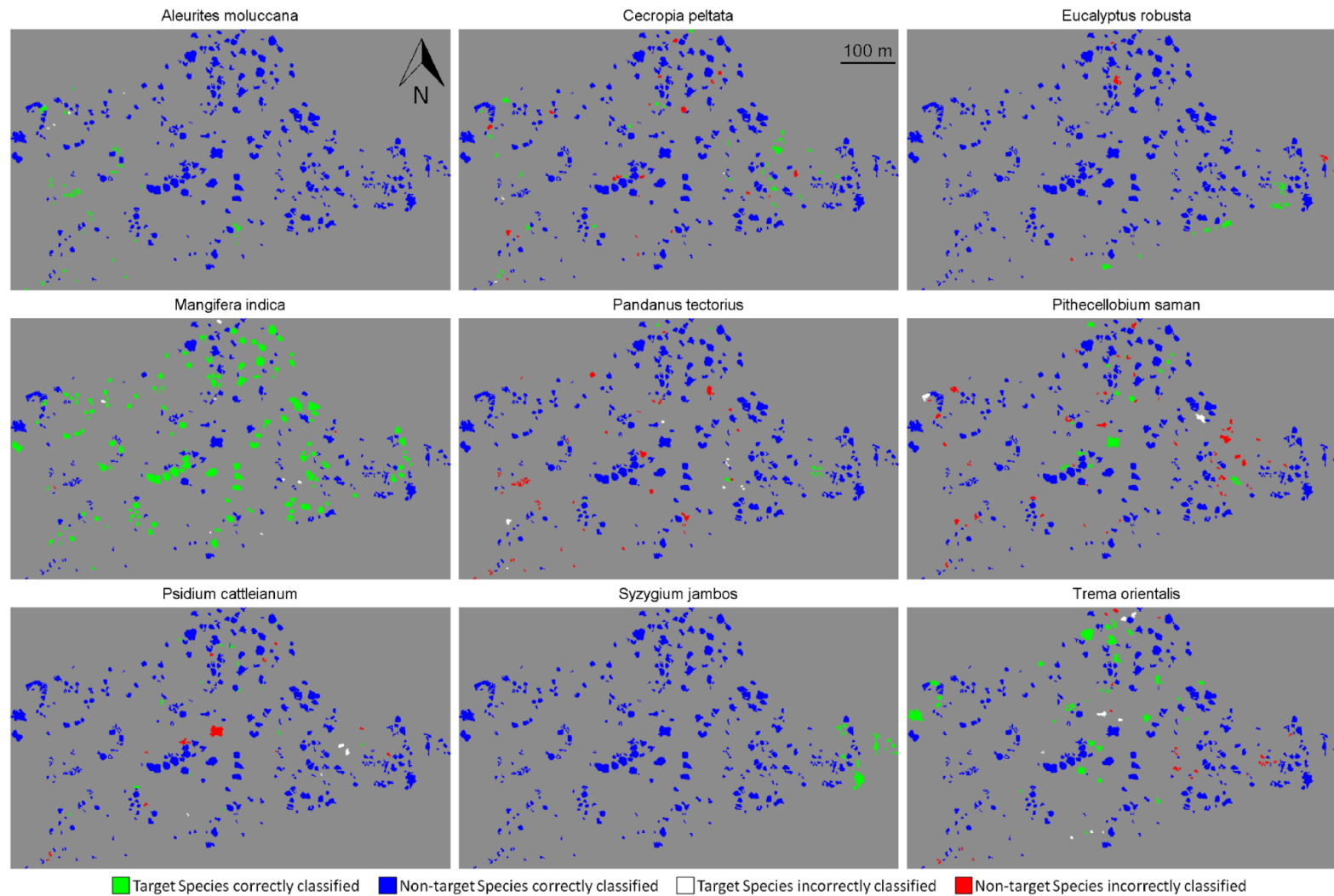
Table 6. Mapping sensitivity, specificity and balanced accuracy (BAC) (%) based on the correct classification among 333 ITCs using combined hyperspectral imagery, LiDAR intensity and LiDAR canopy height with tensor summation bagged semi-supervised SVM classification, for each species on the complete study site.

Species	Sensitivity	Specificity	BAC
Aleurites moluccana	84.4	100.0	92.2
Cecropia peltata	89.2	93.2	90.4
Eucalyptus robusta	100.0	98.0	97.8
Mangifera indica	94.4	97.7	96.0
Pandanus tectorius	57.9	88.7	73.3
Pithecellobium saman	85.7	85.2	85.4
Psidium cattleianum	81.0	95.5	88.2
Syzygium jambos	100.0	100.0	100.0
Trema orientalis	81.0	92.1	86.5

5. Discussion

We compared different methods and data for mapping target species in a Hawaiian tropical forest, obtaining optimal results with a semi-supervised SVM using the combination of LiDAR intensity, LiDAR-derived canopy height and hyperspectral data. In this case, the balanced accuracy was superior at 80% for eight of nine species. These results are very promising from the perspective of targeting species of interest in tropical rainforest. Nonetheless, the overall approach requires additional review and work in order to bring it to an operational level. Our priority now is to validate the method on tropical forest sites with higher species richness. As the probability of spectral similarity between species increases with species richness [5,18,19], the specificity of the methods proposed here may not be as high in a different context with high biodiversity. However, it may be possible to reduce confusion between species using different approaches. These approaches include the use of prior knowledge about the sites under study to choose relevant species to target, the adaptation of the classification method to finely tune the criterion to optimize the automatic tree crowns delineation with a segmentation method to determine, and also the use of a broader spectral domain including short wave infrared (SWIR).

5.1. Prior Information Used to Select the Right Species to Target

The first way to minimize confusion between species involves prior knowledge about the spectral specificity and/or physiological characteristics of a target species. A preliminary study may show very specific spectral properties associated with a species or a group of species of interest. Sanchez-Azofeifa *et al.* [38] found for example, promising results for the discrimination of *Tabebuia guayacan*. This discrimination was performed using very high spatial resolution images acquired by Quickbird, and was enabled by the inflorescence synchronization and the yellow flowers characteristic of this species. Hyperspectral imagery coupled with LiDAR may improve the species discrimination by detection of seasonal cycle or synchronized phenological processes, as well as the capacity to measure more subtle species-specific spectral and structural properties related to leaf chemistry, foliage density, and crown architecture. Several large-scale projects are currently accumulating ecological data from tropical forests using a variety of methods and approaches to better describe tree

species (spatial distribution, chemical composition, spectroscopy), their interaction with the whole ecosystem, and help in deciding of their potential interest as target species. Such projects are needed for the validation of our study and its evolution to an operational level of application. These projects include Spectranomics (http://www.spectranomics.ciw.edu/) [21,39] and the Carnegie Airborne Observatory (http://www.cao.stanford.edu/), both developed and managed by the research team from the Carnegie Institution for Science, as well as the Center for Tropical Forest Science global network (http://www.ctfs.si.edu/) led by the Smithsonian Tropical Research institute.

5.2. Metric Used for the Evaluation of Classifier's Performance

We showed that semi-supervised classifications significantly improve detection sensitivity as compared to the standard supervised SVM, but this was achieved at the cost of reduced specificity. However, the interpretation of the high BAC accuracy should be considered relative to the abundance of the target species compared to the non-target species: a low percentage of misclassified non-target individuals may represent a larger number of ITCs than the total number of target individuals. In such situations, high sensitivity and specificity may be misleading (see for example the high sensitivity and specificity of *Cecropia peltata* and *Pithecellobium saman* in Tables 6 and 7, to compare to their spatial distribution after classification in Figure 4). The performance evaluation metric of the classification may need to be changed in order to favor sensitivity, specificity, precision (the proportion of true positives among trees classified as target species), or any other measures of relevance such as F-measure, receiver operating characteristic (ROC) or Cohen's Kappa statistic [40,41]. The aim of species targeting may vary depending on the species of interest, its expected distribution, similarity with other species or any prior knowledge available. This information may help in deciding on the most suitable performance evaluation metric. In our study for example, we implicitly allowed false positives by balancing sensitivity and specificity. Including precision as an evaluation metric may help decrease the high number of false positive cases with respect to the number of true positives, as observed with *Cecropia peltata* and *Pithecellobium saman*, but may also risk decrease the number of correctly classified target individuals.

Table 7. Mapping sensitivity, specificity and balanced accuracy (BAC) (%) based on the correct pixelwise classification using combined hyperspectral imagery, LiDAR intensity and LiDAR canopy height with tensor summation bagged semi-supervised SVM classification, for each species on the complete study site.

Species	Sensitivity	Specificity	BAC
Aleurites moluccana	93.9	100.0	96.9
Cecropia peltata	96.8	96.0	96.4
Eucalyptus robusta	100.0	98.5	98.3
Mangifera indica	98.4	99.7	99.1
Pandanus tectorius	57.6	94.2	75.9
Pithecellobium saman	80.9	91.1	86.0
Psidium cattleianum	67.1	96.3	81.7
Syzygium jambos	100.0	100.0	100.0
Trema orientalis	89.1	97.1	93.1

5.3. Individual Tree Crown Delineation

Our results are based on individual tree crowns manually delineated through extensive field work. An operational, automated version of this species targeting method thus requires an efficient tree crown segmentation method, which is extremely challenging for many forest types. To our knowledge no method has been proposed to accurately delineate individual tree crowns in very dense tropical forests. The combination of high spatial resolution hyperspectral imagery and LiDAR data for image segmentation has not been fully investigated until now, and it may help to improve tree crown delineation as compared to existing methods based on canopy height or on a limited amount of spectral bands. We are currently working on the development of such a method. The accurate tree crown delineation also has the advantage of allowing the creation of variables obtained at the tree crown scale, which may improve tree species classification. These crown-scale variables can be derived from hyperspectral or LiDAR data. LiDAR-derived variables integrated over spatial units such as land cover type or individual tree crown improve classification accuracy [11,15].

5.4. Improving Species Classification with Enhanced Spectral Range and Resolution

Studies show the important contribution of the shortwave-infrared domain for the characterization of vegetation, and for species differentiation at both leaf and canopy scales [2,19,42,43]. The results proposed here were obtained with hyperspectral imagery covering only visible and near-infrared domains. Hence, improvements in both classification and tree crown delineation can be expected with a new generation of sensor systems including hyperspectral imagery with higher spectral range and resolution and LiDAR with better performance such as the CAO AToMS system (http://www.cao.ciw.edu), which is operational since June 2011.

6. Conclusion

Our study provides a comparison of several methods and data types for detecting and mapping target species in tropical forests. The results indicate that imaging spectroscopy, combined with the full suite of LiDAR variables, can provide highly accurate species identification. The newer generation of semi-supervised classification approaches appears promising for enhanced species mapping as seen in the high classification accuracy for canopy species identification in a Hawaiian tropical forest, particularly when only a limited number of tree crowns is available for the training stage. We obtained very good prediction of the spatial distribution for most of them, based on a training dataset including only one individual tree crown from the target species, and incomplete information about the non-target species included in the test dataset, limited to one individual tree crown from four species randomly selected among the eight non-target species available. We compared supervised and semi-supervised classification approaches, and found that including information from unlabeled samples during the training stage significantly improves classification sensitivity for eight tropical species out of nine. We found that semi-supervised Support Vector Machine classification using tensor summation kernel provided the highest balanced accuracy averaged over nine species, with demonstrable accuracy for at least eight out of nine species, and for all combinations of data types tested. The balanced accuracy obtained with this classifier ranged between 71.1% and 73.3%

depending on the type of data used for the analysis, whereas it ranged between 64.4% and 66.8% for supervised classification, and between 68.5% and 69.1% for the semi-supervised Support Vector Machine classification using tensor product kernel. This global improvement averaged over nine species corresponded to an improvement in balanced accuracy higher than 10 percentage points compared to supervised classification for species such as *Aleurites moluccana* and *Mangifera indica*.

We also showed that the different variables derived from LiDAR with multiple returns can help to improve species discrimination. We studied the separate and combined influence of LiDAR intensity and canopy height coupled with hyperspectral data, and we concluded that each data type boosts accuracy for some species. These two types of LiDAR variables complement one another, and we recommend combining them with hyperspectral data whenever possible as the full combination of hyperspectral imagery, LiDAR intensity and canopy height outperformed any other combination tested here when averaged on the nine species studied, and showed significant improvements compared to hyperspectral data only or combined with one of the two LiDAR data types studied for six of these species.

Finally, the application of the best classifier with the optimal combination of data produced a balanced accuracy between 80% and 100% for eight of nine species. *Eucalyptus robusta*, *Mangifera indica* and *Syzygium jambos* were particularly well identified, with sensitivity, specificity and balanced accuracy between 98% and 100%, whereas the sensitivity obtained for *Pandanus tectorius* was lower than 60%, which implies a high number of false positives. These results are very promising from the perspective of applications aimed at targeting species of interest in tropical rainforest. Our next goal is to improve this method for operational applications at a larger scale.

Acknowledgements

References

1. Goodwin, N.; Turner, R.; Merton, R. Classifying eucalyptus forests with high spatial and spectral resolution imagery: An investigation of individual species and vegetation communities. *Aust. J. Bot.* **2005**, *53*, 337–345.
2. Clark, M.L.; Roberts, D.A.; Clark, D.B. Hyperspectral discrimination of tropical rain forest tree species at leaf to crown scales. *Remote Sens. Environ.* **2005**, *96*, 375–398.
3. Zhang, J.; Rivard, B.; Sánchez-Azofeifa, A.; Castro-Esau, K. Intra- and inter-class spectral variability of tropical tree species at La Selva, Costa Rica: Implications for species identification using HYDICE imagery. *Remote Sens. Environ.* **2006**, *105*, 129–141.

4. Thomas, V. Hyperspectral Remote Sensing for Forest Management. In *Hyperspectral Remote Sensing of Vegetation*; Thenkabail, A., Lyon, P.S., Huete, J.G., Eds.; CRC Press: Boca Raton, FL, USA, 2011; pp. 469–486.

5. Féret, J.-B.; Asner, G.P. Tree species discrimination in tropical forests using airborne imaging spectroscopy. *IEEE Trans. Geosci. Remote Sens.* **2012**, in press.

6. Underwood, E.; Ustin, S.; DiPietro, D. Mapping nonnative plants using hyperspectral imagery. *Remote Sens. Environ.* **2003**, *86*, 150–161.

7. Hamada, Y.; Stow, D.A.; Coulter, L.L.; Jafolla, J.C.; Hendricks, L.W. Detecting tamarisk species (Tamarix spp.) in riparian habitats of Southern California using high spatial resolution hyperspectral imagery. *Remote Sens. Environ.* **2007**, *109*, 237–248.

8. Im, J.; Jensen, J.R.; Jensen, R.R.; Gladden, J.; Waugh, J.; Serrato, M. Vegetation cover analysis of hazardous waste sites in Utah and Arizona using hyperspectral remote sensing. *Remote Sens.* **2012**, *4*, 327–353.

9. Olsson, A.D.; van Leeuwen, W.J.D.; Marsh, S.E. Feasibility of invasive grass detection in a desertscrub community using hyperspectral field measurements and Landsat TM imagery. *Remote Sens.* **2011**, *3*, 2283–2304.

10. Clark, M.L.; Roberts, D.A. Species-level differences in hyperspectral metrics among tropical rainforest trees as determined by a tree-based classifier. *Remote Sens.* **2012**, *4*, 1820–1855.

11. Brandtberg, T. Classifying individual tree species under leaf-off and leaf-on conditions using airborne lidar. *ISPRS J. Photogramm.* **2007**, *61*, 325–340.

12. Korpela, I.; Ole Ørka, H.; Maltamo, M.; Tokola, T.; Hyyppä, J. Tree species classification using airborne LiDAR—Effects of stand and tree parameters, downsizing of training set, intensity normalization, and sensor type. *Silva Fenn.* **2010**, *44*, 319–339.

13. Liu, L.; Pang, Y.; Fan, W.; Li, Z.; Li, M. Fusion of Airborne Hyperspectral and LiDAR Data for Tree Species Classification in the Temperate Forest of Northeast China. In *Proceedings of the 2011 19th International Conference on Geoinformatics*, Shanghai, China, 24–26 June 2011; pp. 1–5.

14. Dalponte, M.; Bruzzone, L.; Gianelle, D. Fusion of hyperspectral and LiDAR remote sensing data for classification of complex forest areas. *IEEE Trans. Geosci. Remote Sens.* **2008**, *46*, 1416–1427.

15. Holmgren, J.; Persson, Å.; Söderman, U. Species identification of individual trees by combining high resolution LiDAR data with multi-spectral images. *Int. J. Remote Sens.* **2008**, *29*, 1537–1552.

16. Dinuls, R.; Erins, G.; Lorencs, A.; Mednieks, I.; Sinica-Sinavskis, J. Tree species identification in mixed baltic forest using LiDAR and multispectral data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2012**, *5*, 594–603.

17. Fung, T.; Ma, F.Y.; Siu, W.L. Hyperspectral Data Analysis for Subtropical Tree Species Recognition. In *Proceedings of the 1998 IEEE International Geoscience and Remote Sensing Symposium Proceedings*, Seattle, WA, USA, 6–10 July 1998; Volume 3, pp. 1298–1300.

18. Castro-Esau, K.L.; Sánchez-Azofeifa, G.A.; Rivard, B.; Wright, S.J.; Quesada, M. Variability in leaf optical properties of Mesoamerican trees and the potential for species classification. *Am. J. Bot.* **2006**, *93*, 517–530.

19. Féret, J.-B.; Asner, G.P. Spectroscopic classification of tropical forest species using radiative transfer modeling. *Remote Sens. Environ.* **2011**, *115*, 2415–2422.

20. Cochrane, M.A. Using vegetation reflectance variability for species level classification of hyperspectral data. *Int. J. Remote Sens.* **2000**, *21*, 2075–2087.

21. Asner, G.P.; Martin, R.E. Canopy phylogenetic, chemical and spectral assembly in a lowland Amazonian forest. *New Phytol.* **2011**, *189*, 999–1012.

22. Asner, G.P.; Vitousek, P.M. Remote analysis of biological invasion and biogeochemical change. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 4383–4386.

23. Bergen, K.M.; Goetz, S.J.; Dubayah, R.O.; Henebry, G.M.; Hunsaker, C.T.; Imhoff, M.L.; Nelson, R.F.; Parker, G.G.; Radeloff, V.C. Remote sensing of vegetation 3-D structure for biodiversity and habitat: Review and implications for LiDAR and RADAR spaceborne missions. *J. Geophys. Res.* **2009**, *114*, doi:10.1029/2008JG000883.

24. Asner, G.P.; Knapp, D.E.; Kennedy-Bowdoin, T.; Jones, M.O.; Martin, R.E.; Boardman, J.; Field, C.B. Carnegie Airborne Observatory: In-flight fusion of hyperspectral imaging and waveform light detection and ranging for three-dimensional studies of ecosystems. *J. Appl. Remote Sens.* **2007**, *1*, 013536.

25. Green, R.O.; Pavri, B.E.; Chrien, T.G. On-Orbit radiometric and spectral calibration characteristics of EO-1 Hyperion derived with an underflight of AVIRIS and *in situ* measurements at Salar de Arizaro, Argentina. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 1194–1203.

26. Clark, M.L.; Clark, D.B.; Roberts, D.A. Small-Footprint LiDAR estimation of sub-canopy elevation and tree height in a tropical rain forest landscape. *Remote Sens. Environ.* **2004**, *91*, 68–89.

27. Carlson, K.; Asner, G.; Hughes, R.; Ostertag, R.; Martin, R. Hyperspectral remote sensing of canopy biodiversity in Hawaiian lowland rainforests. *Ecosystems* **2007**, *10*, 536–549.

28. Licciardi, G.; Pacifici, F.; Tuia, D.; Prasad, S.; West, T.; Giacco, F.; Thiel, C.; Inglada, J.; Christophe, E.; Chanussot, J.; *et al*. Decision fusion for the classification of hyperspectral data: Outcome of the 2008 GRS-S data fusion contest. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 3857–3865.

29. Cortes, C.; Vapnik, V. Support-Vector networks. *Mach. Learn.* **1995**, *20*, 273–297.

30. Vapnik, V.N. An overview of statistical learning theory. *IEEE Trans. Neural Netw.* **1999**, *10*, 988–999.

31. Burges, C.J.C. A tutorial on support vector machines for pattern recognition. *Data Min. Knowl. Discov.* **1998**, *2*, 121–167.

32. Knorn, J.; Rabe, A.; Radeloff, V.C.; Kuemmerle, T.; Kozak, J.; Hostert, P. Land cover mapping of large areas using chain classification of neighboring Landsat satellite images. *Remote Sens. Environ.* **2009**, *113*, 957–964.

33. Pal, M.; Mather, P.M. Support vector machines for classification in remote sensing. *Int. J. Remote Sens.* **2005**, *26*, 1007–1011.

34. Dixon, B.; Candade, N. Multispectral landuse classification using neural networks and support vector machines: One or the other, or both? *Int. J. Remote Sens.* **2007**, *29*, 1185–1206.

35. Huang, C.; Davis, L.S.; Townshend, J.R.G. An assessment of support vector machines for land cover classification. *Int. J. Remote Sens.* **2002**, *23*, 725–749.

36. Melgani, F.; Bruzzone, L. Classification of hyperspectral remote sensing images with support vector machines. *IEEE Trans. Geosci. Remote Sens.* **2004**, *42*, 1778–1790.

37. Tuia, D.; Camps-Valls, G. Semi-supervised remote sensing image classification with cluster kernels. *IEEE Geosci. Remote Sens. Lett.* **2009**, *6*, 224–228.

38. Sánchez-Azofeifa, A.; Rivard, B.; Wright, J.; Feng, J.-L.; Li, P.; Chong, M.M.; Bohlman, S.A. Estimation of the distribution of Tabebuia guayacan (Bignoniaceae) using high-resolution remote sensing imagery. *Sensors* **2011**, *11*, 3831–3851.

39. Asner, G.; Martin, R.; Suhaili, A. sources of canopy chemical and spectral diversity in lowland Bornean forest. *Ecosystems* **2012**, *15*, 504–517.

40. Allouche, O.; Tsoar, A.; Kadmon, R. Assessing the accuracy of species distribution models: Prevalence, kappa and the true skill statistic (TSS). *J. Appl. Ecol.* **2006**, *43*, 1223–1232.

41. Congalton, R.G. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sens. Environ.* **1991**, *37*, 35–46.

42. Rivard, B.; Sanchez-Azofeifa, G.A.; Foley, S.; Calvo-Alvarado, J.C. Species Classification of Tropical Tree Leaf Reflectance and Dependence on Selection of Spectral Bands. In *Hyperspectral Remote Sensing of Tropical and Sub-Tropical Forests*; Kalacska, M., Sanchez-Azofeifa, G.A., Eds.; CRC Press: Boca Raton, FL, USA, 2008; pp. 141–159.

43. Thenkabail, P.S.; Enclona, E.A.; Ashton, M.S.; van der Meer, B. Accuracy assessments of hyperspectral waveband performance for vegetation analysis applications. *Remote Sens. Environ.* **2004**, *91*, 354–376.