

Article

Object-Based Tree Species Classification in Urban Ecosystems Using LiDAR and Hyperspectral Data

Zhongya Zhang ^{1,2,*}, Alexandra Kazakova ³, Ludmila Monika Moskal ² and Diane M. Styers ⁴

¹ School of Land Sciences and Technology, China University of Geosciences (Beijing), 29 Xueyuan Road, Haidian District, Beijing 100083, China

² Remote Sensing and Geospatial Analysis Laboratory, Precision Forestry Cooperative, School of Environmental and Forest Sciences, College of the Environment, University of Washington, Box 352100, Seattle, WA 98195, USA; lmmoskal@uw.edu

³ Bing Maps, Microsoft, 555 110th Ave NE, Bellevue, WA 98004, USA; Alexkaz@microsoft.com

⁴ Department of Geosciences and Natural Resources, College of Arts and Sciences, Western Carolina University, Cullowhee, NC 28723, USA; dmstyers@wcu.edu

* Correspondence: zhongyazzy@126.com; Tel.: +86-10-8232-2205

Academic Editors: Chris Hopkinson, Laura Chasmer and Craig Mahoney

Received: 15 April 2016; Accepted: 4 June 2016; Published: 11 June 2016

Abstract: In precision forestry, tree species identification is key to evaluating the role of forest ecosystems in the provision of ecosystem services, such as carbon sequestration and assessing their effects on climate regulation and climate change. In this study, we investigated the effectiveness of tree species classification of urban forests using aerial-based HyMap hyperspectral imagery and light detection and ranging (LiDAR) data. First, we conducted an object-based image analysis (OBIA) to segment individual tree crowns present in LiDAR-derived Canopy Height Models (CHMs). Then, hyperspectral values for individual trees were extracted from HyMap data for band reduction through Minimum Noise Fraction (MNF) transformation which allowed us to reduce the data to 20 significant bands out of 118 bands acquired. Finally, we compared several different classifications using Random Forest (RF) and Multi Class Classifier (MCC) methods. Seven tree species were classified using all 118 bands which resulted in 46.3% overall classification accuracy for RF *versus* 79.6% for MCC. Using only the 20 optimal bands extracted through MNF, both RF and MCC achieved an increase in overall accuracy to 87.0% and 88.9%, respectively. Thus, the MNF band selection process is a preferable approach for tree species classification when using hyperspectral data. Further, our work also suggests that RF is heavily disadvantaged by the high-dimensionality and noise present in hyperspectral data, while MCC is more robust when handling high-dimensional datasets with small sample sizes. Our overall results indicated that individual tree species identification in urban forests can be accomplished with the fusion of object-based LiDAR segmentation of crowns and hyperspectral characterization.

Keywords: LiDAR; hyperspectral; tree species classification; urban forests

1. Introduction

Tree species identification is important for forest resource management and monitoring [1]. Individual tree level information of species distribution is used in a multitude of forestry applications, including wildlife habitat mapping and assessment [2] and biodiversity monitoring [3]. Timely and accurate species-specific tree mapping is an indispensable component for studying aboveground biomass and carbon stock in forests [4]. In urban environments, forests also provide social and economic benefits, and other vital ecosystem services upon which public health and welfare depend. For example, trees can reduce adverse effects on the environment in urban areas by improving air

and water quality, absorbing atmospheric carbon dioxide (CO₂), and moderating energy use [5]. Knowledge of tree species information plays a key role in assessing and valuing the quantity and quality of various services provided by forest ecosystems.

Tree species data have been traditionally collected as part of a time- and labor-intensive field-based forest inventory or manual aerial image interpretation [6,7]. With the advantages of broad spatial coverage and rapid revisiting ability, remote sensing has provided a useful tool for identifying and classifying tree species. The high spectral resolution of hyperspectral remotely-sensed data has been successfully used to capture differences in biochemical characteristics of various tree species from leaf to crown scales [8–11]. Additionally, hyperspectral sensors have proven useful for tree species differentiation in a variety of forest types located in tropical [12], temperate [13], and boreal [14] forest ecosystems. Traditionally, tree species differentiation was conducted using per-pixel classification methods based on the different spectral responses of individual pixels. However, topographic variation and different sun illumination conditions can change the spectral signals of pixels received by optical remote sensors, which can alter final classification results. Further, using per-pixel classification methods with hyper-spatial imagery often results in the noisy “salt-and-pepper effect,” meaning that individual pixels are differently classified from their neighboring pixels [15]. Object-based image analysis (OBIA) uses image objects (or segments) rather than individual pixels as basic classification units. This approach has been widely used in different applications of remotely-sensed image analysis, such as vegetation mapping [16–18], land cover classification [19], and landscape change analysis [20,21]. Although OBIA has been used successfully in these applications, shadow effects and the occlusion of tree crowns still makes delineating tree crowns inherently difficult with two-dimensional (2-D) aerial or satellite imagery. By utilizing three-dimensional (3-D) light detection and ranging (LiDAR) data it is possible to isolate individual tree crowns from a 3-D perspective [22–24].

As one of the most popular active remote sensing technologies, LiDAR has proven to be a powerful tool for extracting the 3-D biophysical parameters of forests [25,26]. In the Pacific Northwest (PNW) region of the United States, LiDAR-returned intensity information has been used successfully to differentiate between broadleaved and coniferous tree species types in urban forests [27]. In continuation of this work, Kim *et al.* [28] used LiDAR-based structural variables combined with intensity information to discriminate deciduous from coniferous tree genera. In the same study area, Vaughn *et al.* [29] demonstrated that overall detection accuracy for five species improved from 79.2% with discrete-return point cloud data to 85.4% with full-waveform information recorded for the forest canopy. We draw on the field validation and species list of Kim *et al.* [28] in this research for comparative purposes.

Many recent studies have attempted to integrate LiDAR and hyperspectral data for the purpose of tree species identification. Thematic raster maps, such as height maps generated from LiDAR data, can provide very detailed spatial information about tree crowns [30], especially in comparison to that of conventional true color aerial images. Hyperspectral and LiDAR data offer the possibility of combining vertical and horizontal profile information and, thus, the synergistic use of these data is a promising new approach to tree species identification [31]. La *et al.* [32] were successful at extracting individual tree crowns using hyperspectral imagery and LiDAR data. Alonzo *et al.* [33] experimented with 29 different tree species in Santa Barbara, California, USA, and were able to increase overall classification accuracy by 4.2% by adding LiDAR structural metrics to AVIRIS image analysis. Although the inclusion of LiDAR data did not demonstrate a substantial increase in overall accuracy, species with small crowns and those with unique morphological characteristics gained significant improvements in classification accuracy. Matsuki *et al.* [34] also improved the overall accuracy of their tree species classification by adding LiDAR-derived metrics to hyperspectral data. Jones *et al.* [35] reported an increase in both producer's (+5.1%–11.6%) and user's (+8.4%–18.8%) accuracy of species-level maps for the coastal PNW, using both airborne hyperspectral and LiDAR data. Dalponte *et al.* [36] reported that hyperspectral data allowed one to distinguish between similar species, and that the addition of LiDAR data increased classification accuracy. Further work indicated that tree species classification

at the individual tree crown level is driven by spectral values of individual pixels rather than a LiDAR-derived canopy height model and structural metrics [37]. Moreover, the fusion of active (LiDAR) and passive (hyperspectral) remote sensing data for forestry applications continues to be an active area of research [30–37].

Most of the existing research focuses on the identification of tree species in natural forests with dense canopies where spectral features are relatively homogeneous. Compared to the natural forest environment, forests in urban areas has high frequency detail at very fine spatial resolutions [38]. This difference makes it challenging to classify tree species in urban ecosystems, and to the author's knowledge few studies to date have applied these methods to both urban and semi-urban forests. Therefore, the specific aims of this paper are:

- (1) To develop an approach to identify tree species by combining structural, canopy height model crown segmentation information derived from LiDAR data with spectral information extracted from hyperspectral data for urban forests; and
- (2) to test and compare the robustness and accuracy of two different classifiers commonly used for tree species identification: Random Forest (RF) and Multi Class Classifier (MCC).

2. Methods

The classification procedure consisted of the following steps: (1) object-based segmentation and visual validation; (2) spectral information extraction from delineated tree crowns and spectral band reduction; (3) tree species identification using the RF and MCC classifiers; and (4) accuracy assessment of classifications created using different algorithms and datasets. A more detailed workflow for this study is summarized in Figure 1.

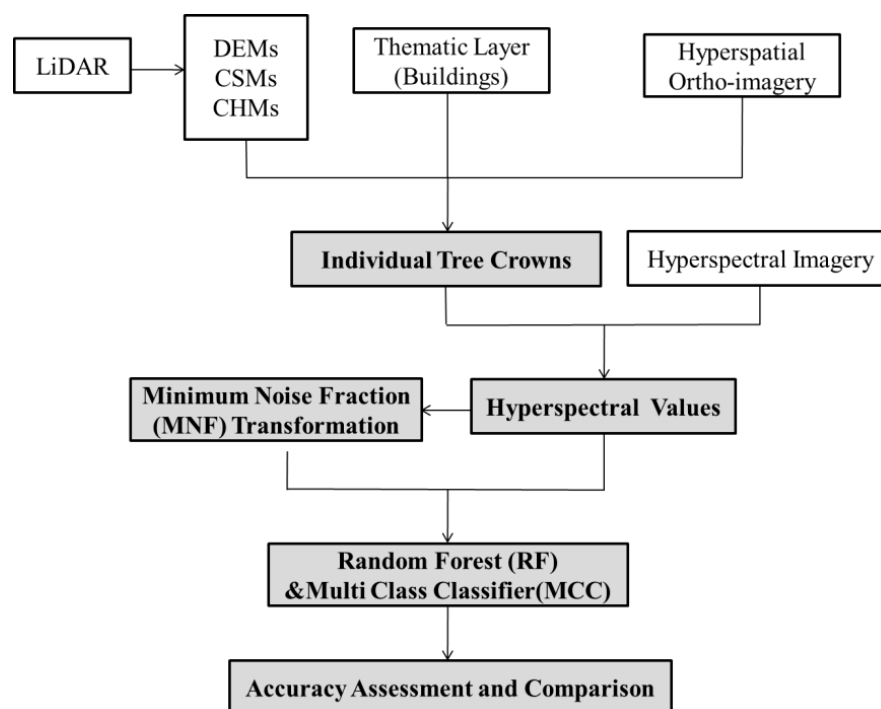


Figure 1. Flowchart for integrating LiDAR data with hyperspectral imagery for tree species identification at the individual tree level.

2.1. Study Area

Two study areas were chosen for the current study; an area of southeast Seattle, Washington located within the 98118 zip code (ZIP) and the Washington Park Arboretum (WPA) (Figure 2). The ZIP

study area spans approximately 1640 hectares (ha), and mostly consists of residential areas with some industrial land use. This area contains a wide variety of native and non-native tree species. There are about 60 different trees species among the 473 individual trees inventoried in this study area. Thus, a limited number of samples for individual species were available due to the high heterogeneity of species in this region. The ZIP study area contains street trees, dense canopies within parks, and trees on private properties. The WPA study area is managed by the Botanical Gardens at the University of Washington. The WPA represents a managed, semi-urban or park-like ecosystem, and is ideal for this study since it contains many of the species found in planted urban environments, as well as native species found in natural forests of the PNW. The two study sites allowed us to test our methods in two locations within the same city, one is a continuous urban forest, the other represents a pattern of urban street trees; each is a typical urban forest but with differing canopy structures, ground covers, and canopy overlapping degree.

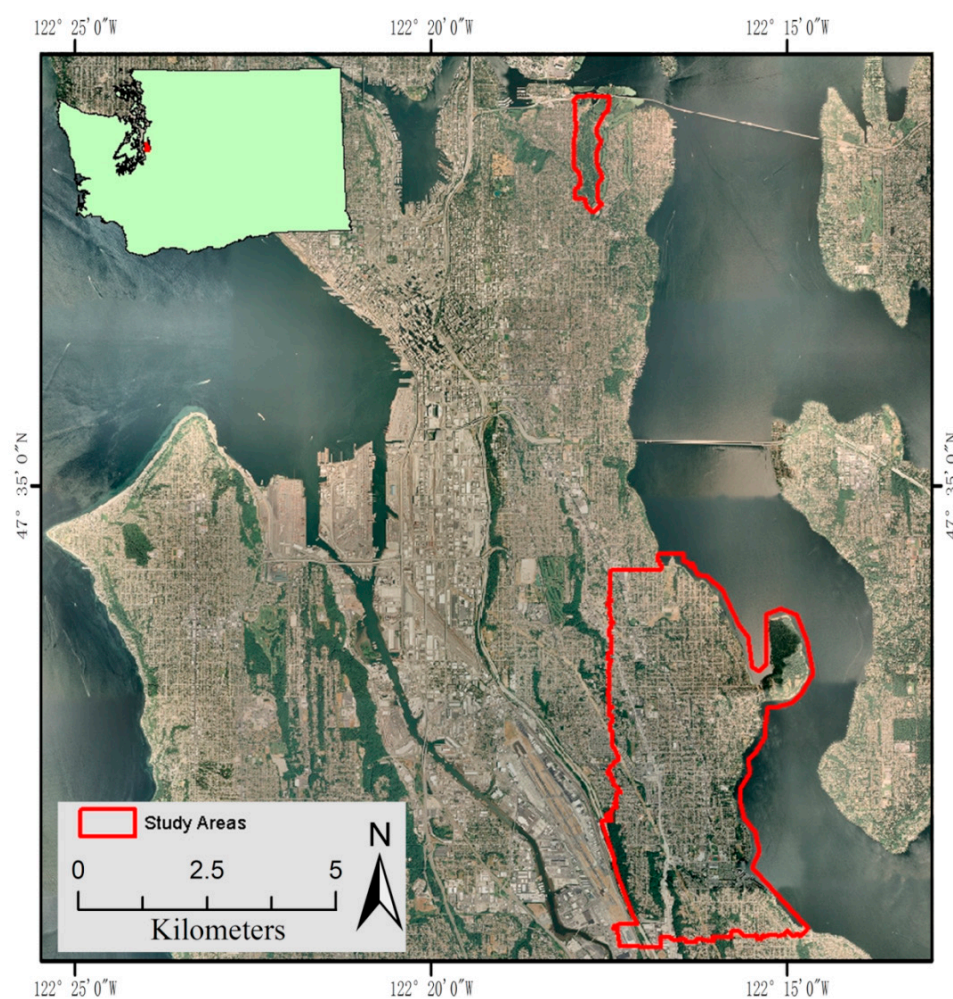


Figure 2. Location of the ZIP and WAP study areas (outlined in red).

We chose to identify the seven most common tree species found in these two sites. Four broadleaved species consisted of American Sweetgum (*Liquidambar styraciflua* L.), Red Maple (*Acer rubrum* L.), Japanese Zelkova (*Zelkova serrata* (Thunb.) Makino), and Sycamore (*Platanus* L.). Three coniferous species were Douglas Fir (*Pseudotsuga menziesii* (Mirb.) Franco), Western Red Cedar (*Thuja plicata* Donn ex D. Don), and European Larch (*Larix decidua* Mill.). Originally, we collected 41 sweetgums, 18 sycamores, 12 Japanese Zelkovas, 18 red maples, 16 Douglas-firs, 23 European larches, and 28 western red cedars. Since uneven sample sizes can bias the species with small sample

sizes, we randomly selected samples from each species group to create an evenly distributed sample dataset. The minimum number of samples available for each species in this study was 16, except for Japanese Zelkova, which only had 12 samples. Therefore, all of the samples of Japanese Zelkova were used for classification, whereas for the other six species, 16 samples were selected at random. This reduced the total number of trees used in the study from the much larger available pool, but assured unbiased results in our statistical analysis.

2.2. LiDAR Data

The LiDAR data for the ZIP area were downloaded from the Puget Sound LiDAR Consortium [39]. LiDAR data were collected on 1 April 2000 at an elevation of 1000 m above ground surface. The laser was pulsed at 30+ KHz, and the rangefinder recorded up to four returns per pulse. Most areas were covered by two swaths, resulting in a nominal pulse density of about one pulse per square meter.

The LiDAR data for the WPA area were acquired on 31 August 2004 using an Optech ALTM 30/70 laser scanner (Optech Inc., Vaughan, ON, Canada). The coverage was collected at an elevation of 1200 m above ground surface with a maximum scan angle of $\pm 10^\circ$ from nadir. This resulted in a discrete point dataset containing four returns per pulse and at least eight pulses per square meter.

Processing of the LiDAR data involved converting raw point cloud data into various raster models. Three types of models were created and included digital elevation models (DEMs), canopy surface models (CSMs), and ground-normalized canopy height models (CHMs) that were generated using Fusion software (Pacific Northwest Research Station, Seattle, WA, USA). The approach has been successfully applied in previous research by Kim *et al.* [27,28] for creating the LIDAR-based CHMs. It was equivalent for both datasets to make the resulting CHMs comparable. LiDAR data were delivered with ground points identified and validated by the vendor. The elevation of each grid cell was computed using the average elevation of all points within the 5×5 cell to produce DEMs. This method seems to work well with LIDAR data that has been filtered to identify bare-earth points. For the CSMs extraction, an initial surface is computed using the highest return elevation for each 5×5 cell. Then the initial surface is used to produce the final, smoothed surface by a 5×5 cell median filter. CHMs were computed by subtracting the DEMs from the CSMs. We did not test the preciseness of the tree peaks as this question is beyond the scope of this research; however, Popescue *et al.* [40] provide a good exploration of variable window size and its impacts on tree height and, thus, relates to the general precision of canopy height models. The CHMs were then used for segmenting and classifying forest canopies into individual crowns. We chose to use CHMs for segmentation because this approach has been previously established by Alonzo *et al.* [33] and Dian *et al.* [41] in projects using aerial LiDAR and hyperspectral data.

2.3. Airborne Hyperspectral Data

Hyperspectral data for both study sites were collected with an Integrated Spectronics HyMap sensor owned and operated by HyVista Corporation and flown by Watershed Sciences, Inc. (now Quantum Spatial Inc., Corvallis, OR, USA). The sensor consisted of four spectrometers (Table 1) including visible (VIS), near-infrared (NIR), shortwave infrared sensor 1 (SWIR1), and shortwave infrared sensor 2 (SWIR2). Each spectrometer produced 32 bands of imagery (128 bands in total); however, data provided by the vendor consisted of only 125 bands. All HyMap data were collected with a 3 m spatial resolution. The acquisition of data for both sites took place on 15th August 2010. The HyMap imagery were georeferenced based on the geographic look-up tables and input geometry files included in the data package by the provider. The data were then geo-registered using the aforementioned inputs in ENVI 4.7 software (ITT, Boulder, CO, USA).

Table 1. The spectral configuration of the HyMap sensor is shown below for each of the four spectral modules containing 32 bands each, totaling 128 spectral bands.

Module	Spectral Range	Bandwidth across Module	Average Spectral Sampling Interval
VIS	0.45–0.89 μm	15–16 μm	15 μm
NIR	0.89–1.35 μm	15–16 μm	15 μm
SWIR1	1.40–1.80 μm	15–16 μm	13 μm
SWIR2	1.95–2.48 μm	18–20 μm	17 μm

The radiance images were converted to apparent surface reflectance using HyCorr software ((HyVista Corporation, Sydney, Australia) and EFFORT processing techniques. Empirical Line Calibration (ELC) using field reflectance spectra was conducted for atmospheric correction. Field-collected spectra were acquired at the time of hyperspectral imagery capture for the purpose of collecting reflectance signatures with less atmospheric distortions than those obtained from satellite or aircraft-mounted sensors because of the close proximity of the ground sensor to the surface [42]. Ground targets were placed to enable spectral calibrations for hyperspectral imagery, and a ground-based spectral radiometer was used to collect spectral data 1 m above the ground targets during the same time frames of the flights. Target spectral information was applied to ELC using the dark object subtraction method in ENVI 4.7 software (ITT Visual Information Solutions, Boulder, CO, USA). It has been shown that for HyMap imagery, a combination of HyCorr, EFFORT, and ELC is the best atmospheric correction method for creating a less noisy spectrum that corresponds best to reference spectra [43].

2.4. Validation Data

Additional data were collected for model calibration and validation. Hyperspatial, sub-meter per-pixel spatial resolution, ortho-imagery was obtained from the King County digital data archive. This RGB imagery was flown during the summer of 2009 and has a resolution of 0.09 m. This highly detailed imagery was utilized to create a tree-top map for visual validation of OBIA results. Field data were also available for both sites. After each sample tree was identified, its exact location was recorded with a survey grade JAVAD GNSS GPS. The GPS data were differentially corrected to an average accuracy of 10 cm. These sample locations were used for identifying crowns used for individual species classification.

2.5. Image Segmentation

The CHMs were used to delineate individual tree crowns. The major obstacle in extracting canopy areas involved excluding objects that share similar height information. An ancillary GIS thematic layer (buildings) was added to mask them out of the segmentation process. In the WPA area, the remaining pixels with heights above 2 m were classified as canopy. Some pixels with small areas around the border of the canopy were wrongly classified as ground and these pixels were reclassified as canopy. In the ZIP area, the threshold of CHMs was chosen as 6 m. After a preliminary classification, misclassified pixels were removed based on NDVI values ($\text{NDVI} = (\text{Mean NIR} - \text{Mean Red}) / (\text{Mean NIR} + \text{Mean Red})$), relative border to building, and relative border to canopy.

Once the canopy was extracted, the tops of trees were located by using a local maxima algorithm, which created a 1 m by 1 m grid, and applied a median filter with 5×5 moving windows. Crown segments were then generated using a region growing algorithm based on height increments. OBIA was performed automatically in eCognition Developer software (Trimble Navigation Ltd., Sunnyvale, CA, USA). An example of the segmentation process is shown in Figure 3.

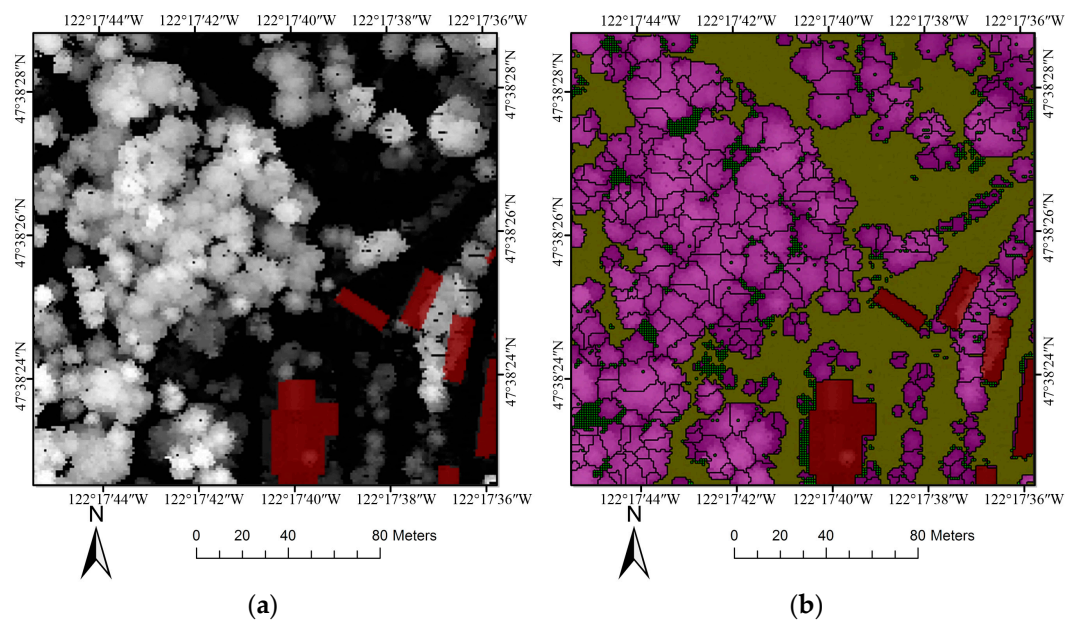


Figure 3. Segmentation process: (a) buildings were excluded by a thematic layer; and then (b) canopies were segmented into individual crowns.

2.6. Extraction of Hyperspectral Values

The reflectance values of each band were extracted for each tree using crown objects as regions of interest (ROIs) in ENVI 4.7 (ITT, Boulder, CO, USA). Manually-recorded GPS locations of 30 trees with vertically-grown stems were used to obtain a stem-map layer. The stem-map layer was compared with the tree-top map produced from the 0.09 m spatial resolution aerial ortho-imagery. The RMSE of this registration was less than 0.5 m. The delineated tree crown map was then overlaid with the hyperspectral images to extract the reflectance values from each crown object to produce spectral curves for each individual tree. Bands centered at 1.3897 μm , 1.4046 μm , 1.4192 μm , 1.4335 μm , 1.9484 μm , 2.4714 μm , and 2.4867 μm were deleted because these were either noisy bands in the shortwave infrared, or their values were very low or negative. The removal of these seven bands resulted in a total of 118 bands that were useful for creating spectral curves for tree species identification.

2.7. Hyperspectral Band Reduction

The significant correlation between bands and the redundant and noisy signals within hyperspectral datasets require the preprocessing of hyperspectral data. When all original bands are used as predictive variables for a classification, it may cause the following consequences: (1) the extraction of important values for classification and model training becomes time-consuming; and (2) the high dimensionality of the data complicates the classification process and hampers transferability of the algorithm [36,44]. Supervised classification algorithms are often hindered by the large dimensionality of hyperspectral imagery. The curse of dimensionality (or Hughes phenomenon) [45] may occur when there is a small ratio between the number of training samples and the number of features. Non-redundant data become sparse with the increase of dimensionality. This sparsity makes it difficult to estimate the parameters of the classifiers and causes problems for any method requiring statistical significance.

A good solution is to preprocess the data into a lower-dimensional form to reduce redundancy and irrelevancy, while retaining the most vital and useful information in the original dataset. Minimum Noise Fraction (MNF) transformation is a spectral tool that has been widely used to segregate spectral noise and reduce data dimensionality in hyperspectral data [46]. It is a linear transformation of the

original bands applying two cascaded principal components analysis that maximizes the ratio of signal to noise. The first transformation decorrelates and rescales noise in the data. The noise in the data after the first transformation has unit variance and no band-to-band correlations. The second transformation is a principal components analysis of the noise-whitened data which produces coherent MNF eigenimages that contain useful information and noise-dominated MNF eigenimages. The MNF transformation produces a set of principal component images ordered in terms of decreasing signal quality with the decreasing eigenvalues and lower ratio of signal to noise. Usually, the first few MNF bands separate out most of the noise and explain most of the surface reflectance variation in the image while the rest of MNF bands are primarily dominated by noise. The MNF procedure was conducted in ENVI 4.7 (ITT, Boulder, CO, USA). The eigenvalues plot indicated that more than 80% of the variance was found within the first 20 bands of the MNF image; therefore, the first 20 bands were chosen as the MNF bands dataset [47].

2.8. Classification

Several classification algorithms were initially tested as part of this project, including Support Vector machines (SVM) and RF, but were ineffective due to the small sample sizes and high dimensionality of the data. The classification accuracy of the tested algorithms was below 60%, which does not meet the expectations of classification accuracy needed for forest planning and management in urban areas. Thus, we introduced a new simple classification method, MCC, which is capable of handling classifications of high-dimensional data with small sample sizes. Although many of the tested algorithms were deemed unsuccessful, we will present the results generated from RF for the purpose of comparison between this widely-used classification method and MCC.

2.8.1. Random Forests (RF) Approach

RF is an ensemble learning method for classification that constructs a multitude of decision trees at training time [48]. To classify a new object, RF grows many trees and then puts the input vector down each of these trees. Each tree votes for that class and the forest chooses the classification having the most votes. RF uses out-of-bag (OOB) error as an estimate of the generalization error. During the OOB procedure, a different bootstrap sample from the original data is used for the construction of each tree. About one-third of the cases are left out of the bootstrap sample and not used. In addition to classification, RF can be used to rank the importance of variables. RF has been successfully implemented in the forestry community to handle the high dimensionality of hyperspectral datasets for classification and band reduction [49,50]. In this research the band selection function of RF was not considered and only MNF was performed to make comparisons between different classifiers. The RF algorithm was constructed using Weka software (University of Waikato, Hamilton, New Zealand).

2.8.2. Multi Class Classifier (MCC) Approach

When designing classification algorithms it is often easier to develop algorithms that deal with two classes, and in the literature there exists a variety of binary classification algorithms. Yet, publications on tree species classifications usually address the problem of classifying multiple instances into multiple classes [51,52]. Therefore, a classification algorithm suitable for handling a multi-class problem should be developed to effectively process the hyperspectral data for tree species identification. Typically in such cases the multi-class problem is divided into multiple binary classification problems that can be solved separately [52]. This is a very general problem that is associated with many classification tasks. Traditional strategies such as one-*vs.*-all and all-*vs.*-all approaches can divide multi-class problems into several two-class ones, and these have worked well with low-dimensional datasets. However, hyperspectral data have high dimensionality and the spectral reflectance is influenced by many factors. For that reason hyperspectral data requires decorrelations and error-correcting to improve the final classification accuracy.

Error Correcting Output Code (ECOC) [53,54] was initially proposed to improve the reliability of binary signals by detecting and correcting the errors in data transmission through noisy channels. This method is able to improve the generalization ability of classification algorithms [53]. ECOC assigns a unique code to a class instead of assigning each class a label. This code is a unique binary string of length n . The n -binary functions are learned with one for each bit position in these binary strings. The code for class i is used to specify the desired outputs of these n binary functions during training, for example, from class k . Due to its ability to reduce the correlation and correct errors of the datasets, ECOC can solve classification problems in high-dimensional hyperspectral datasets.

Dietterich *et al.* [54] proposed two main criteria to take advantage of an ECOC-based classification: (1) row separation: codes should be as far apart from one another as possible; and (2) column separation: each binary classifier function defined by each output bit should not be correlated to the functions of another output's node to be learned. Highly-correlated bits will enforce similar decision boundaries on the data and are useless for classification. Four forms of ECOC were proposed according to the number of class k : (1) exhaustive codes when $3 \leq k \leq 7$; (2) column selection from exhaustive code when $8 \leq k \leq 11$; (3) randomized hill climbing when $k > 11$; and (4) Bose-Chaudhuri-Hocquenghem (BCH) codes when $k > 11$ [54]. Since we were classifying seven species exhaustive codes were chosen. For dataset having k classes when $3 \leq k \leq 7$, a code of length $2^{k-1} - 1$ can be constructed where:

- Class 1 All strings are one;
- Class 2 There are 2^{k-2} zeroes followed by $2^{k-2} - 1$ ones;
- Class 3 There are 2^{k-3} zeroes, followed by 2^{k-3} ones, followed by $2^{k-3} - 1$ zeroes, followed by $2^{k-3} - 1$ ones;
- Class i There are alternating runs of 2^{k-i} zeroes and ones.

Kong *et al.* [53] discussed why ECOC tends to generate higher accuracy with multi-class classification problems. Their research demonstrated that the ECOC method can reduce the variance of the learning algorithm. In addition, ECOC corrects errors induced by the bias of the learning algorithm. MCC is a meta-classifier specifically created for handling multi-class datasets with two-class classifiers. In this study, ECOC was set as the method to use in MCC for transforming the multi-class problem into several two-class ones. Logistic regression classifiers were used as base classifiers. The implementation of this MCC approach was conducted in Weka software.

2.8.3. Cross-Validation

When there is not enough data available, or when the distribution is not normal and the spread is too narrow to split the data into separate training and test sets for conventional validation methods (e.g., 80% of data is used for training and 20% for testing), the error rate in the training set is not a useful estimator of model performance. Thus, the error in the test data set does not properly assess the model performance. In such cases, cross-validation is a more powerful and proper technique to estimate the error rate [55]. In k -fold cross-validation, the original sample is randomly partitioned into k equally-sized subsamples. Then $k-1$ subsamples of the k subsamples are used as training data, and the remaining single subsample is retained as validation data for testing the model. Each of the k subsamples is used once as the validation data, and the cross-validation process is repeated k times to reduce variability among the results. The k results from the folds can then be combined to produce a single estimation of error. The advantage of k -fold cross-validation is that all the samples in the dataset are eventually used for both training and testing. The 10-fold cross-validation is commonly used in this process.

With the limited number of samples, classification training and testing for both RF and MCC were performed with 10-fold cross-validation [56,57] to assess the predictive performance of both classification algorithms. Theoretically, there is no need for cross-validation or a separate test set to get an estimate of the test set error when using RF because the OOB error estimate is generated internally. However, the internal validation of RF is somewhat overly positive. Accordingly, in this research the

validation of RF was also conducted according to 10-folds cross-validation procedure. This procedure allowed us to objectively compare the performance between RF and MCC from a rigorous perspective.

The 10-fold cross-validation of both classifiers was conducted using Weka software. RF and MCC approaches were used on both the original bands and MNF bands datasets. Four confusion matrices were generated: the original bands with RF and MCC, and the MNF bands with RF and MCC.

3. Results

3.1. Crown Segmentation and Validation

Visual inspection (Figure 4) revealed the majority of segment boundaries corresponded well to the crowns as seen on the ortho-imagery. Nevertheless, there are three main issues with the segmentation of tree crowns in this forest study: (1) omission of trees under subdominant canopies; (2) under-segmentation or overestimation of the size of segments; and (3) over-segmentation of large deciduous crowns. Since our focus is at the individual tree level, it is more crucial to eliminate as many under-segments as possible, and capture the majority of individual tree crowns. Over-segmentation is more tolerable due to the resolution of the hyperspectral imagery (3 m).

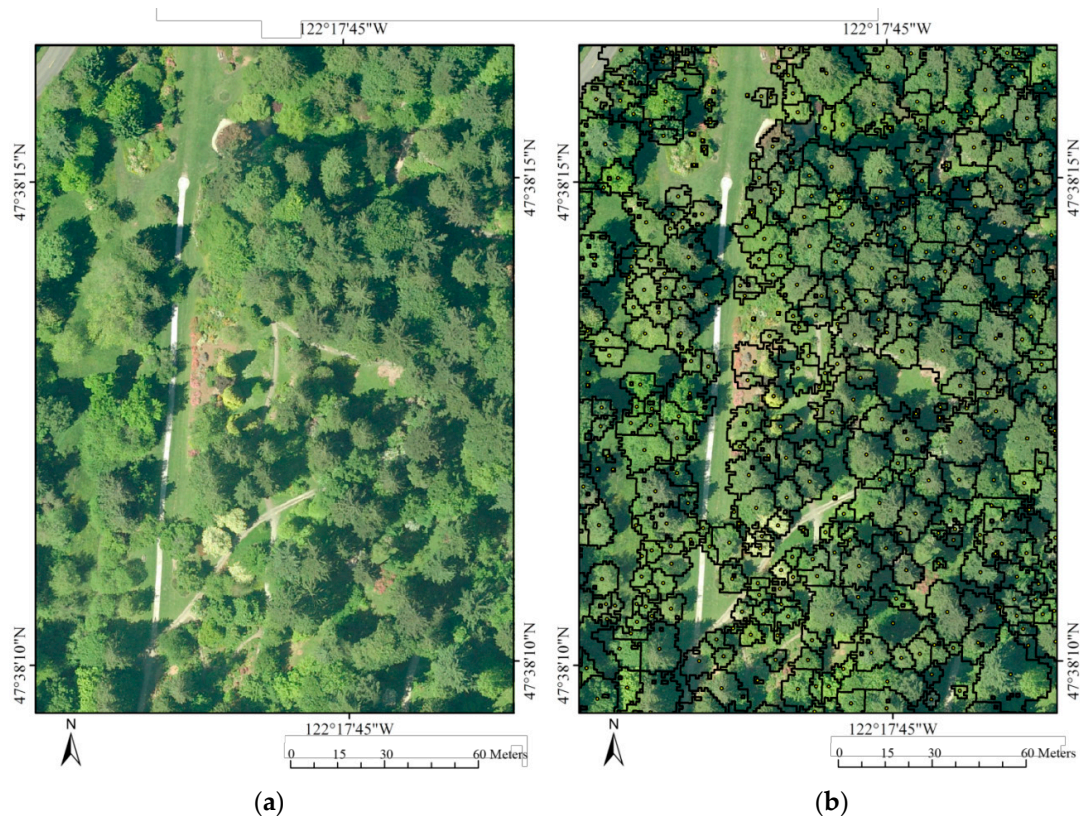


Figure 4. Visualization of segmentation results: (a) ortho-imagery with spatial resolution of 0.09 m; and (b) overlaying CHM-generated crown segments over the ortho-imagery.

3.2. Classification Results

When using the RF approach on the original bands dataset, the overall accuracy for the seven species was 46.3% (Kappa = 0.372). For the MNF bands dataset, the overall accuracy using the RF algorithm increased dramatically to 87.0% (Kappa = 0.849). Although the number of variables was sharply reduced from 118 to 20, the overall accuracy by contrast increased by 40.7%. Both producer's and user's accuracy of Sweetgum, Japanese Zelkova, and Western Red cedar improved to over 90%. Compared with RF, MCC yielded a better overall classification accuracy of 79.6% (Kappa = 0.762).

when using the original bands. Sweetgum, Japanese Zelkova, and European Larch achieved the highest producer's and user's accuracy. The overall accuracy increased to 88.9% (Kappa = 0.870) when applying MCC to the MNF bands dataset. Producer's and user's accuracy of all seven species was greater than 75.0%. All results from the classification accuracy assessment are shown in Table 2.

Table 2. Classification results in terms of PA (producer's accuracy), UA (user's accuracy) and OA (overall accuracy) for seven tree species. Accuracies were obtained using different combinations of datasets and classifiers of the 10-fold cross-validation data.

Datasets	Original				MNF			
Classifier	RF		MCC		RF		MCC	
Species	PA	UA	PA	UA	PA	UA	PA	UA
Sweetgum	56.3%	40.9%	100.0%	100.0%	100.0%	94.1%	100.0%	84.2%
Sycamore	43.8%	58.3%	81.3%	76.5%	93.8%	78.9%	93.8%	88.2%
Japanese Zelkova	66.7%	72.7%	83.3%	100.0%	100.0%	100.0%	91.7%	100.0%
Red Maple	50.0%	42.1%	87.5%	77.8%	75.0%	92.3%	75.0%	92.3%
Douglas-fir	12.5%	22.2%	31.3%	45.5%	68.8%	68.8%	75.0%	80.0%
European Larch	56.3%	50.0%	87.5%	87.5%	81.3%	86.7%	100.0%	94.1%
Western Red Cedar	43.8%	41.2%	87.5%	70.0%	93.8%	93.8%	87.5%	87.5%
OA	46.3%		79.6%		87.0%		88.9%	

4. Discussion

4.1. Factors Influencing Segmentation

The temporal discrepancies between the remotely sensed datasets biased the segmentation results. The LiDAR data were acquired for the ZIP and WPA areas in spring of 2000 and summer of 2004, respectively, while the HyMap datasets were acquired in summer of 2010. Specifically, the final accuracies of the segmentation and classification results were affected by changes in tree height and crown size, as well as the appearance or disappearance of trees due to planting or removal. These errors could be minimized, or possibly even eliminated, by acquiring remotely-sensed datasets at the same time.

Only the height information extracted from LiDAR data was used in the current OBIA segmentation criteria system. Some additional characteristics of each segment such as texture or ratio of perimeter to area was not utilized in this research. However, Vaughn *et al.*'s [29] study at WPA suggests that LIDAR structure is important in species classification, thus, this is a direction that deserves further exploration in future studies. Additionally, quantitative validation in addition to visual-based validation procedures used in the current study would benefit comparisons between different segmentation schemes.

Due to the 3 m spatial resolution of the HyMap dataset, only tree centers were extracted for classification. The georeferenced errors and/or the small crown sizes introduced errors into the hyperspectral information extracting process. Hyperspectral data with higher spatial resolution would be preferable for extraction of more indices and improvement in classification accuracy. With finer spatial resolution data the sun-lit peaks of trees should be compared to the crown centers for classification purposes.

4.2. Classification Accuracy

The original bands dataset produced lower accuracy than the MNF bands dataset. The overall classification accuracy for RF increased sharply from 46.3% to 87.0%, and also increased for MCC from 79.6% to 88.9%. This accuracy is similar to the results of work by Kim *et al.* [27] that utilized LiDAR intensity information, which represents the reflective characteristics of the canopies. However, the scale of our research is different. Moreover, our hyperspectral analysis results outperformed Vaughn *et al.* [29], who combined discrete point LiDAR data and Fourier transformation variables

of full-waveform LiDAR data to classify five individual species. These results suggest that reflective characteristics of the canopy are of importance for urban tree species classification. Additional work comparing the performance of structural features and reflective characteristics in classifying tree species should be conducted. Alonzo *et al.* [33] conducted a watershed segmentation algorithm to segment individual crowns. Spectra exceeding a certain NDVI threshold were extracted and structural metrics were computed directly from LiDAR data. Different combinations of all spectral bands and seven FFS (Forward Feature Selection) selected structural metrics were reduced to canonical variates and classified. Overall accuracy for 29 urban tree species classified by fusion of seven FFS-selected structural metrics and all hyperspectral variables was 83.4%, while classification accuracy using spectral data alone was 79.2%. The accuracy of the MNF bands datasets of this research is higher than Alonzo *et al.* [33]; however, the results should be compared carefully, as different research sites and tree species were tested. Moreover, Alonzo *et al.* [33] utilized high point density LiDAR and 224 channel AVIRIS hyperspectral data while our research used low point density LiDAR and 128 channel HyMap data. Finally, this study has a dataset of small sample sizes, thus, an algorithm favorable in small sample size was used here. Alonzo *et al.* [33] used at least 50 samples for each species in their study. In addition, our research has two study sites including continuous urban forest and street-trees, while Alonzo *et al.* [33] acquired field data in downtown area dominated by street trees, which have crowns easier to delineate and allow for extraction of tree spectra without contamination by neighboring trees as compared to our sometimes heavily-overlapping park trees.

There are some potential factors influencing the classification accuracy. The spatial resolution of HyMap data is 3 m and, thus, it is unavoidable that pixels of some small crowns are influenced by other land use cover types; a common spectral mixture problem in data of this resolution. Possible solutions might include resampling the hyperspectral data or acquiring higher spatial resolution hyperspectral data to examine how spatial resolution influences classification accuracy. In addition, the spectral reflectance of the background can penetrate tree crowns and be acquired by imaging sensors. For example, in the ZIP area reflectance spectra are impacted by roads, buildings, impervious surfaces, and other human artifacts. A spectral mixture analysis could be conducted to evaluate how background context in these two areas influenced classification accuracy.

4.3. Classification Algorithms

In this research the RF algorithm was unsuccessful at handling a dataset with high dimensionality without band reduction. In contrast, MCC was able to achieve an acceptable overall accuracy using all bands in the original dataset. After MNF transformation, RF and MCC yielded similar overall accuracies, and compared to the original bands dataset, MNF contributed to a substantial increase in classification accuracy and is thus preferable for our image classification. These results are consistent with previous studies [58,59] that have reported the MNF transformation is successful at reducing noise and dimensionality present in hyperspectral datasets. MNF is capable of significantly decreasing classifier complexity while retaining useful information with lower dimensionalities.

The RF approach did not work well with the original bands' dataset, and can probably be attributed to inherent noise in this dataset. RF is prone to overfitting for some datasets which is even more pronounced in noisy classification tasks. This hypothesis is validated by the results obtained from applying the RF algorithm to the MNF bands dataset. When used on datasets in which spectral noise has been segregated, RF yielded similar accuracy to MCC. The MCC approach, by contrast, resulted in higher overall accuracy for both the original and MNF bands datasets. Kong *et al.* [53] discussed why ECOC produces high accuracy and how it corrects for errors and bias. Their research reported that ECOC reduces the variance of the learning algorithm. Additionally, ECOC corrects for errors caused by bias of the learning algorithm instead of simply combining multiple runs of the learning algorithm. Dietterich *et al.* [54] proposed that ECOC produces higher accuracy with small samples. This is because ECOC works by reducing variance of the learning algorithm while smaller samples tend to have higher variance. ECOC can, thus, provide more benefits for datasets with small

samples. As for our research, we had a small-sample problem, and thus, the use of ECOC was more appropriate in our situation. This result demonstrates that MCC with ECOC is a robust classifier capable of handling high dimensional data with small sample sizes.

The robustness of the algorithm should be tested further for different forest stands in urban areas since it has mainly been applied to image analysis in the computer science field.

5. Conclusions

In this study, we proposed an approach to classify forest tree species by combining the height information from LiDAR data and the spectral information from hyperspectral imagery in urban forests. The main conclusions derived from our analysis were:

- (1) LiDAR-based tree height information can effectively delineate tree crowns, and a visual examination showed good alignment between crown segments and individual tree crowns on ortho-imagery. The MNF transformation can effectively save most of the useful spectral information within a few transformed bands, which removes noise and highly-correlated information between different bands.
- (2) Classifications created using the original bands dataset revealed that when handling a dataset with large amount of noise, RF tends to over fit the data. Compared with RF, MCC is better at handling a dataset with a lot of noise. These conclusions are confirmed by the overall accuracies of the classifications created using the MNF bands dataset, where the RF approach produced similar results to MCC after the spectral noise had been segregated.
- (3) The sample size and data dimensionality of this research made it challenging to utilize widely-accepted supervised classification methods. The MCC method with ECOC proposed in this paper is capable of handling high-dimensional datasets with small sample sizes. The inherent working mechanism of ECOC is reducing the variance of the learning algorithm. Smaller sample sizes tend to have greater variance and, thus, benefit from the use of ECOC.

Compared with conventional methods, the combination of the structural information, such as crown delineation, derived from LiDAR height data and the spectral information extracted from hyperspectral data will effectively improve forest classification results. In urban areas, however, the background context may include considerable spectral diversity at very fine spatial resolutions which induces spectral mixture and can lead to misclassification. The research discussed in this paper established a framework of classifying tree species with high accuracy and can be applied to study areas including both urban and semi-urban areas, which have different canopy structures representative of these environments. The approaches in this study holds the potential to provide fast and efficient tree species information for municipal and utility foresters, sustainability officers, city planners and other stakeholders. Tree species classification at the individual tree level will greatly facilitate the accurate simulation and prediction of the spatiotemporal distribution of aboveground forest biomass and its capacity for carbon sequestration to guarantee the quantity of organic carbon pools in urban areas. In conclusion, this study demonstrates practicability of detecting and classifying individual tree species, which is a key consideration in precision forestry and when making urban forest management and conservation decisions.

Acknowledgments: This research was funded by the USDA McIntire-Stennis Cooperative Forestry Research Program and the Precision Forestry Cooperative at the University of Washington. Support has also been provided by the Green Cities Research Alliance and the Urban Natural Resources Stewardship program at the U.S. Forest Service PNW Research Station. The hyperspectral data were kindly provided by Watershed Sciences, Inc. (now Quantum Spatial). This work was conducted at the University of Washington Remote Sensing and Geospatial Analysis Laboratory. The authors would like to thank the China Scholarship Council for their support. Finally, we kindly thank the anonymous reviewers who helped us improve this paper.

Author Contributions: Zhongya Zhang was involved in the whole study (image processing, algorithm construction, data analysis and writing the manuscript); Alexandra Kazakova performed image preprocessing; Ludmila Monika Moskal supervised the entire research project and revised drafts of the paper; Diane M. Styers collected the field data, revised drafts of the paper, and polished the English.

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

Abbreviations

The following abbreviations are used in this manuscript:

LiDAR	Light detection and ranging
OBIA	Object-based image analysis
PNW	Pacific Northwest
DEMs	Digital elevation models
CSMs	Canopy surface models
CHMs	Canopy height models
MNF	Minimum Noise Fraction
RF	Random Forest
MCC	Multi Class Classifier
ECOC	Error Correcting Output Code

References

1. Plourde, L.C.; Ollinger, S.V.; Smith, M.L.; Martin, M.E. Estimating species abundance in a northern temperate forest using spectral mixture analysis. *Photogramm. Eng. Remote Sens.* **2007**, *73*, 829–840. [[CrossRef](#)]
2. Hill, R.A.; Wilson, A.K.; George, M.; Hinsley, S.A. Mapping tree species in temperate deciduous woodland using time-series multi-spectral data. *Appl. Veg. Sci.* **2010**, *13*, 86–99. [[CrossRef](#)]
3. Mathieu, M.; Asner, R.; Naidoo, G.; van Aardt, L.; Ramoelo, J.; Debba, A.; Wessels, P.; Main, K.; Smit, R.; Erasmus, I. Mapping tree species composition in South African savannas using an integrated airborne spectral and LiDAR system. *Remote Sens. Environ.* **2012**, *125*, 214–226.
4. Palosuo, T.; Peltoniemi, M.; Mikhailov, A.; Komarov, A.; Faubert, P.; Thürig, E.; Lindner, M. Projecting effects of intensified biomass extraction with alternative modelling approaches. *For. Ecol. Manag.* **2008**, *255*, 1423–1433. [[CrossRef](#)]
5. Dwyer, J.; McPherson, E.; Schroeder, H.; Rowntree, R. Assessing the benefits and costs of the urban forest. *J. Arboric.* **1992**, *18*, 227–234.
6. Gong, P.; Pu, R.L.; Yu, B. Conifer species recognition: An exploratory analysis of *in situ* hyperspectral data. *Remote Sens. Environ.* **1997**, *62*, 189–200. [[CrossRef](#)]
7. Lucas, R.M.; Lee, A.C.; Bunting, P.J. Retrieving forest biomass through integration of CASI and LiDAR data. *Int. J. Remote Sens.* **2008**, *29*, 1553–1577. [[CrossRef](#)]
8. Pu, R.; Liu, D. Segmented canonical discriminant analysis of *in situ* hyperspectral data for identifying 13 urban tree species. *Int. J. Remote Sens.* **2011**, *32*, 2207–2226. [[CrossRef](#)]
9. 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. [[CrossRef](#)]
10. Ghiyamat, A.; Shafri, H.; Mahdiraji, G.; Shariff, A.; Mansor, S. Hyperspectral discrimination of tree species with different classifications using single- and multiple-endmember. *Int. J. Appl. Earth Obs. Geoinform.* **2013**, *23*, 177–191. [[CrossRef](#)]
11. 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. [[CrossRef](#)]
12. Ferreira, M.P.; Grondona, A.E.; Rolim, S.B.A.; Shimabukuro, Y.E. Analyzing the spectral variability of tropical tree species using hyperspectral feature selection and leaf optical modeling. *J. Appl. Remote Sens.* **2013**, *7*. [[CrossRef](#)]
13. Boschetti, M.; Boschetti, L.; Oliveri, S.; Casati, L.; Canova, I. Tree species mapping with Airborne hyper-spectral MIVIS data: The Ticino Park study case. *Int. J. Remote Sens.* **2007**, *28*, 1251–1261. [[CrossRef](#)]
14. Dalponte, M.; Orka, H.O.; Gobakken, T.; Gianelle, D.; Naesset, E. Tree species classification in boreal forests with hyperspectral data. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 2632–2645. [[CrossRef](#)]
15. Yu, Q.; Gong, P.; Clinton, N.; Biging, G.; Kelly, M.; Schirokauer, D. Object-based detailed vegetation classification with airborne high spatial resolution remote sensing imagery. *Photogramm. Eng. Remote Sens.* **2006**, *72*, 799–811. [[CrossRef](#)]
16. Zhang, C.; Xie, Z. Object-based vegetation mapping in the Kissimmee river watershed using HyMap data and machine learning techniques. *Wetlands* **2013**, *33*, 233–244. [[CrossRef](#)]

17. Moskal, L.M.; Styers, D.M.; Halabisky, M. Monitoring urban tree Cover using object-based image analysis and public domain remotely sensed data. *Remote Sens.* **2011**, *3*, 2243–2262. [[CrossRef](#)]
18. Richardson, J.J.; Moskal, L.M. Uncertainty in urban forest canopy assessment: Lessons from Seattle, WA, USA. *Urban For. Urban Green.* **2014**, *13*, 152–157. [[CrossRef](#)]
19. Styers, D.M.; Moskal, L.M.; Richardson, J.J.; Halabisky, M.A. Evaluation of the contribution of LiDAR data and postclassification procedures to object-based classification accuracy. *J. Appl. Remote Sens.* **2014**, *8*. [[CrossRef](#)]
20. Yang, G.; Pu, R.; Zhang, J.; Zhao, C.; Feng, H.; Wang, J. Remote sensing of seasonal variability of fractional vegetation cover and its object-based spatial pattern analysis over mountain areas. *ISPRS J. Photogramm. Remote Sens.* **2013**, *77*, 79–93. [[CrossRef](#)]
21. Huth, J.; Kuenzer, C.; Wehrmann, T.; Gebhardt, S.; vo Quoc, T.; Dech, S. Land cover and land use classification with TWOPAC: Towards automated processing for pixel- and object-based image classification. *Remote Sens.* **2012**, *4*, 2530–2553. [[CrossRef](#)]
22. Zhang, C.; Zhou, Y.; Qiu, F. Individual tree segmentation from LiDAR point clouds for urban forest inventory. *Remote Sens.* **2015**, *7*, 7892–7913. [[CrossRef](#)]
23. Moskal, L.M.; Zheng, G. Retrieving forest inventory variables with terrestrial laser scanning (TLS) in urban heterogeneous forest. *Remote Sens.* **2012**, *4*, 1–20. [[CrossRef](#)]
24. Jing, L.; Hu, B.; Li, J.; Noland, T. Automated delineation of individual tree crowns from lidar data by multi-scale analysis and segmentation. *Photogramm. Eng. Remote Sens.* **2012**, *78*, 1275–1284. [[CrossRef](#)]
25. Zheng, G.; Moskal, L.M.; Kim, S.H. Retrieval of effective leaf area index in heterogeneous forests with terrestrial laser scanning. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 777–786. [[CrossRef](#)]
26. Koetz, B.; Morsdorf, F.; Sun, G.; Ranson, K.J.; Itten, K.; Allgower, B. Inversion of a lidar waveform model for forest biophysical parameter estimation. *IEEE Geosci. Remote Sens. Lett.* **2006**, *3*, 49–53. [[CrossRef](#)]
27. Kim, S.; McGaughey, R.J.; Andersen, H.E.; Schreuder, G. Tree species differentiation using intensity data derived from leaf-on and leaf-off airborne laser scanner data. *Remote Sens. Environ.* **2009**, *113*, 1575–1586. [[CrossRef](#)]
28. Kim, S.; Hinckley, T.; Briggs, D. Classifying individual tree genera using stepwise cluster analysis based on height and intensity metrics derived from airborne laser scanner data. *Remote Sens. Environ.* **2011**, *115*, 3329–3342. [[CrossRef](#)]
29. Vaughn, N.R.; Moskal, L.M.; Turnblom, E.C. Tree species detection accuracies using discrete point lidar and airborne waveform lidar. *Remote Sens.* **2012**, *4*, 377–403. [[CrossRef](#)]
30. Anderson, J.; Plourde, L.; Martin, M.; Braswell, B.; Smith, M.; Dubayah, R.; Hofton, M.; Blair, J. Integrating waveform lidar with hyperspectral imagery for inventory of a northern temperate forest. *Remote Sens. Environ.* **2008**, *112*, 1856–1870. [[CrossRef](#)]
31. Ghosh, A.; Fassnacht, F.; Joshi, P.; Koch, B. A framework for mapping tree species combining hyperspectral and LiDAR data: Role of selected classifiers and sensor across three spatial scales. *Int. J. Appl. Earth Obs.* **2014**, *26*, 49–63. [[CrossRef](#)]
32. La, H.P.; Eo, Y.D.; Chang, A.; Kim, C. Extraction of individual tree crown using hyperspectral image and LiDAR data. *KSCE J. Civ. Eng.* **2015**, *19*, 1078–1087. [[CrossRef](#)]
33. Alonzo, M.; Bookhagen, B.; Roberts, D.A. Urban tree species mapping using hyperspectral and lidar data fusion. *Remote Sens. Environ.* **2014**, *148*, 70–83. [[CrossRef](#)]
34. Matsuki, T.; Yokoya, N.; Iwasaki, A. Hyperspectral tree species classification of Japanese complex mixed forest with the aid of lidar data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 2177–2187. [[CrossRef](#)]
35. Jones, T.G.; Coops, N.C.; Sharma, T. Assessing the utility of airborne hyperspectral and LiDAR data for species distribution mapping in the coastal Pacific Northwest, Canada. *Remote Sens. Environ.* **2010**, *114*, 2841–2852. [[CrossRef](#)]
36. Dalponte, M.; Bruzzone, L.; Gianelle, D. Tree species classification in the Southern Alps based on the fusion of very high geometrical resolution multispectral/hyperspectral images and LiDAR data. *Remote Sens. Environ.* **2012**, *123*, 258–270. [[CrossRef](#)]
37. Dalponte, M.; Orka, H.O.; Ene, L.T.; Gobakken, T.; Naesset, E. Tree crown delineation and tree species classification in boreal forests using hyperspectral and ALS data. *Remote Sens. Environ.* **2014**, *140*, 306–317. [[CrossRef](#)]
38. Welch, R. Spatial resolution requirements for urban studies. *Int. J. Remote Sens.* **1982**, *3*, 139–146. [[CrossRef](#)]

39. Puget Sound Lidar Consortium. Available online: <http://pugetsoundlidar.ess.washington.edu/> (accessed on 8 February 2012).
40. Popescu, S.; Wynne, R.; Nelson, R. Estimating plot-level tree heights with lidar: Local filtering with a canopy-height based variable window size. *Comput. Electron. Agric.* **2002**, *37*, 71–95. [[CrossRef](#)]
41. Dian, Y.; Pang, Y.; Dong, Y.; Li, Z. Urban tree species mapping using airborne lidar and hyperspectral data. *J. Indian Soc. Remote Sens.* **2016**, *44*, 1–9. [[CrossRef](#)]
42. Harris, T. *Ground Truth for Calibrating Optical Imagery to Reflectance: The Empirical Line Calibration in ENVI*; ITT Visual Information Solutions: Boulder, CO, USA, 2009.
43. Lau, I.C. Application of atmospheric correction to hyperspectral data: Comparisons of different techniques on Hymap data. In Proceedings of the 12th Australasian Remote Sensing and Photogrammetry Conference (ARSPC), Fremantle, Australia, 18–22 October 2004.
44. 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. [[CrossRef](#)]
45. Hughes, G.F. On the mean accuracy of statistical pattern recognizers. *IEEE Trans. Inf. Theory* **1968**, *14*, 55–63. [[CrossRef](#)]
46. Guan, L.X.; Xie, W.X.; Pei, J.H. Segmented minimum noise fraction transformation for efficient feature extraction of hyperspectral images. *Pattern Recognit.* **2015**, *48*, 3216–3226.
47. Sugumaran, R.; Voss, M. Object-oriented classification of LIDAR-fused hyperspectral imagery for tree species identification in an urban environment. In Proceedings of the IEEE Urban Remote Sensing Joint Event, Paris, France, 11–13 April 2007; pp. 1–6.
48. Liaw, A.; Wiener, M. Classification and regression by randomForest. *R News* **2002**, *2*, 18–22.
49. Naidoo, L.; Cho, M.A.; Mathieu, R.; Asner, G. Classification of savanna tree species, in the Greater Kruger National Park region, by integrating hyperspectral and LiDAR data in a Random Forest data mining environment. *ISPRS J. Photogramm. Remote Sens.* **2012**, *69*, 167–179. [[CrossRef](#)]
50. Lopatin, J.; Dolos, K.; Hernández, H.J.; Galleguillos, M.; Fassnacht, F.E. Comparing generalized linear models and random forest to model vascular plant species richness using LiDAR data in a natural forest in central Chile. *Remote Sens. Environ.* **2016**, *173*, 200–210. [[CrossRef](#)]
51. Mehra, N.; Gupta, S. Survey on multi-class classification methods. *Int. J. Comput. Sci. Inf. Technol.* **2013**, *4*, 572–576.
52. Allwein, E.; Schapire, R.; Singer, Y. Reducing multiclass to binary: A unifying approach for margin classifiers. *J. Mach. Learn. Res.* **2001**, *1*, 113–141.
53. Kong, E.B.; Dietterich, T.G. Error-correcting output coding corrects bias and variance. In Proceedings of the 12th International Conference on Machine Learning, Tahoe City, CA, USA, 9–12 July 1995; pp. 313–321.
54. Dietterich, T.G.; Bakiri, G. Solving multiclass learning problems via error-correcting output codes. *J. Artif. Intell. Res.* **1995**, *2*, 263–286.
55. Seni, G.; Elder, J. Ensemble methods in data mining: Improving accuracy through combining predictions. *Synth. Lect. Data Mining Knowl. Discov.* **2010**, *1*, 1–126. [[CrossRef](#)]
56. You, W.; Zang, Z.; Zhang, L.; Li, Z.; Chen, D.; Zhang, G. Estimating ground-level PM10 concentration in northwestern China using geographically weighted regression based on satellite AOD combined with CALIPSO and MODIS fire count. *Remote Sens. Environ.* **2015**, *168*, 276–285. [[CrossRef](#)]
57. Rodriguez, J.D.; Perez, A.; Lozano, J.A. Sensitivity analysis of *k*-fold cross validation in prediction error estimation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2010**, *32*, 569–575. [[CrossRef](#)] [[PubMed](#)]
58. Fassnacht, F.E.; Neumann, C.; Forster, M.; Buddenbaum, H.; Ghosh, A.; Clasen, A.; Joshi, P.K.; Koch, B. Comparison of feature reduction algorithms for classifying tree species with hyperspectral data on three central European test sites. *IEEE J. Sel. Top. Appl.* **2014**, *7*, 2547–2561. [[CrossRef](#)]
59. Zhang, C.; Xie, Z. Combining object-based texture measures with a neural network for vegetation mapping in the Everglades from hyperspectral imagery. *Remote Sens. Environ.* **2012**, *124*, 310–320. [[CrossRef](#)]

