ISPRS Journal of Photogrammetry and Remote Sensing 64 (2009) 683–693

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

Tree species identification in mixed coniferous forest using airborne laser scanning

Agus Suratno *, Carl Seielstad [1], Lloyd Queen [2]

The University of Montana, College of Forestry and Conservation, 32 Campus Drive, Missoula, MT 59801, United States

ARTICLE INFO

Article history:
Received 26 March 2009
Received in revised form
29 July 2009
Accepted 30 July 2009
Available online 11 August 2009

Keywords:
North America
Conifer
Laser scanning
Intensity
Tree species

ABSTRACT

This study tests the capacity of relatively low density (<1 return/m^2) airborne laser scanner data for discriminating between Douglas-fir, western larch, ponderosa pine, and lodgepole pine in a western North American montane forest and it evaluates the relative importance of intensity, height, and return type metrics for classifying tree species. Collectively, Exploratory Data Analysis, Pearson Correlation, ANOVA, and Linear Discriminant Analysis show that structural and intensity characteristics generated from LIDAR data are useful for classifying species at dominant and individual tree levels in multi-aged, mixed conifer forests. Proportions of return types and mean intensities are significantly different between species (p-value < 0.001) for plot-level dominant species and individual trees. Classification accuracies based on single variables range from 49%–61% at the dominant species level and 37%–52% for individual trees. The accuracy can be improved to 95% and 68% respectively by using multiple variables. The inclusion of proportion of return type greatly improves the classification accuracy at the dominant species level, but not for individual trees, while canopy height improves the accuracy at both levels. Overall differences in intensity and return type between species largely reflect variations in the physical structure of trees and stands. These results are consistent with the findings of others and point to airborne laser scanning as a useful source of data for species classification. However, there are still many knowledge gaps that prevent accurate mapping of species using ALS data alone, particularly with relatively sparse datasets like the one used in this study. Further investigations using other datasets in different forest types will likely result in improvements to species identification and mapping for some time to come.

1. Introduction

Airborne laser scanning (ALS) has recently become a key technology for generating measurements to support forest inventories (Means et al., 2000; Gobakken and Naesset, 2004; Maas et al., 2008). The ability of laser scanning to provide direct estimates of canopy structural features such as height, canopy closure, and stem density is highly advantageous in the inventory data collection. However, one important limitation of ALS-derived inventories is the difficulty in identifying tree species. Consequently, foresters continue to rely on passive remote sensing classifications, field surveys, and *a priori* knowledge of vegetation distributions to generate species data (Cochrane, 2000; Asner et al., 2002). Because landscape-level acquisitions of traditional remote sensing (imagery) and ALS data combined with field measurement are considered expensive, an alternative scheme is to optimize the usefulness of laser scanning data, by expanding its ability to provide additional information content such as tree species. Several researchers have speculated that species discrimination is possible using laser data directly (Holmgren and Persson, 2004; Moffiet et al., 2005; Brandtberg, 2007; Ørka et al., 2007), in part because most modern laser scanning systems record the intensity of individual returns.

The most recent studies conducted in Scandinavian, Australian and Scottish forests respectively, have shown that intensity is useful to distinguish between different tree species, particularly when used in conjunction with structural variables such as return type and proportions of heights and canopy hits (Holmgren and Persson, 2004; Moffiet et al., 2005; Donoghue et al., 2007; Ørka et al., 2007). These analyses were mostly exploratory and conducted to discriminate between conifer and deciduous tree species using high density data (3–10 returns/m^2). For example, Holmgren and Persson (2004) applied linear and quadratic discriminant analyses to differentiate between individual Scots pine, Norway spruce and deciduous trees in a Scandinavian boreal

* Corresponding author. Tel.: +1 406 243 6288; fax: +1 406 243 2011.
E-mail addresses: agus.suratno@umontana.edu (A. Suratno), carl@ntsg.umt.edu (C. Seielstad), lpqueen@ntsg.umt.edu (L. Queen).
[1] Tel.: +1 406 243 2011; fax: +1 406 243 2011.
[2] Tel.: +1 406 243 2709; fax: +1 406 243 2011.

0924-2716/$ – see front matter © 2009 International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS). Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.isprsjprs.2009.07.001

forest using both lidar intensity data and tree crown shape. The proportion of returns located above crown base height, the standard deviation of intensity and height, and the proportion of first returns were helpful for species identification. Moffiet et al. (2005) used vegetation permeability (vegetation points/all points) and singular vegetation returns to distinguish between White cypress pine and Poplar box at the dominant species level (the species having the greatest proportion of canopy area for each subplot by stand) and found that the proportion of singular returns contributed most of the discriminatory power (overall classification accuracy ~77%). A clear distinction between the two species was not apparent at the individual tree scale. Donoghue et al. (2007) compared normalized intensity for spruce and pine trees in Galloway forest of Scotland using one-way ANOVA to show that discrimination was possible using mean intensity (p-value $<$ 0.05). The most recent study, conducted by Ørka et al. (2007), identified spruce, birch and aspen in Norway with Principal Component Analysis (PCA) and linear discriminant analysis (LDA). These authors used a combination of five components (mean intensities of first and second returns, standard deviations of first, second and single returns) to generate a classification with overall accuracy of 74%.

Each of the aforementioned studies utilized relatively high density datasets that are generally obtained using slow, low-flying, rotor-wing aircraft. In western North American forests, ALS systems are typically flown aboard fixed-wing aircraft at higher altitudes, usually resulting in data densities more than an order of magnitude smaller in size than those obtained from typical helicopter acquisitions. Additionally, complex terrain, large environmental gradients, and mixed species composition/ forest structure complicate classification of vegetation properties. In western Montana, USA, the location of our study area, forests occupy 7.5 million hectares, and are mostly dominated by multi-aged ponderosa pine (*Pinus ponderosa*), Douglas-fir (*Pseudotsuga menziesii*), western larch (*Larix occidentalis*) and even-aged lodgepole pine (*Pinus contorta*) (Arno et al., 1985). Correctly classifying these four conifer species is a pre-requisite of forest inventory in the region, and if ALS is to be used, the size of the landscape dictates the use of relatively sparse laser scanner data.

This study tests the capacity of relatively low density (<1 return/m^2) ALS structure and intensity data for discriminating between the four previously cited tree species in a montane mixed conifer forest at dominant species and individual tree scales. Exploratory Data Analysis (EDA), Pearson correlation, one-way ANOVA, and Linear Discriminant Analysis (LDA) are used at the levels of plot-level dominant species and individual tree to characterize (1) differences/similarities in the proportion of vegetation return types as a function of tree species, (2) the relationships between return types and intensities, and (3) many of the intensity variables that are most useful for classifying species. This work corroborates the findings of others studying tree species discrimination in Europe and Australia, but differs from previous research in that all combinations of discriminating variables are examined.

2. Materials

The study site is The University of Montana's Lubrecht Experimental Forest (LEF) in the Blackfoot River drainage, 40 km northeast of Missoula, Montana (approximately N 46° 53' W 113° 27', with elevations from 1160 to 1930 m) as shown in Fig. 1. Approximately 45% of the area has a slope gradient over 30%, with the steepest gradients exceeding 90% (Nimlos, 1986). LEF covers 11,300 hectares and is dominated by western larch and Douglas-fir on the north facing slopes and ponderosa pine on south facing slopes, with a substantial intermixing of species. The eastern portion of the forest is represented by even-aged stands of lodgepole pine, with subalpine fir (*Abies lasiocarpa*) dominant in higher elevations.

In LEF, most ponderosa pine trees are represented by high open crowns with needle-like leaves in bundles of two or three that are green in color. The crown is irregular and flat although some trees are slightly conical near their tops. Douglas-fir typically has a pyramidal, symmetrical crown with needle-like leaves that are bluish-green in color. Mostly, the canopy is dense, depending on site conditions. Meanwhile, lodgepole pine is characterized by a long, slender trunk. The crown is thin on the top with needle-like leaves in bundles of two that are shorter than those of the ponderosa pine. In dense, even-aged stands, the canopy top is green while the middle and lower parts are dominated by dead branchwood. Western larch is the only deciduous conifer tree in the study area, characterized by an open, narrow conic crown with leaves clustered on branches that turn yellow in autumn and fall to the ground.

2.1. Field data

A total of 61 rectangular plots (20×20 m^2) were established using a stratified random sample in two steps. First, lidar canopy height model (CHM) characteristics (canopy density and height variance) were used to generate five different canopy structures: (1) dense single strata stands (DSS), (2) dense multi strata (DMS), (3) moderate multi strata (MMS), (4) Moderate single strata (MSS), and (5) open (OPEN) (after Rowell et al., 2006). Second, a point generator placed on the study area was used to generate plots, which were subsequently stratified across the structure classes. It is worth acknowledging that the plots, on which 1555 trees were measured, were originally located to validate an automated tree stem identification algorithm (not directly related to this research). For the purposes of this study, 43 of the 61 plots meet a dominant species criteria, where dominant species is >70% on a tree count basis. The 70% threshold was selected to balance the need for within-plot species homogeneity against number of plots available for analysis. Four plots were located in DSS with an average of 60 trees/plot and mean tree heights of ~10 m, five plots were in DMS consisting ~50 trees/plot and mean tree heights of ~13 m, and the remaining plots were distributed in MSS, MMS, and OPEN as shown in Table 1. MMS has the highest average crownbase height (CBH) while DSS has a smaller mean height than other classes.

Trees having diameter at breast height (DBH) > 7 cm were measured and canopy widths in two perpendicular directions (intersected approximately in the middle of a stem) were recorded. Trees with DBH < 7 cm were tallied by species. Tree heights and CBH were measured using a Forest Pro Laser with an effective accuracy of 0.3 m (Laser Technology, 2007) and tree positions were recorded by measuring range and distance from monumented plot corners, which were located using differentially corrected GPS measurements (average accuracy of ±1.5 m). Additionally, each tree was classified using Kraft-class as dominant, codominant, intermediate, and suppressed (Oliver and Larson, 1996). At the dominant species level, 19 plots were represented by Douglas-fir with an average of 33 trees/plot followed by ponderosa pine (11) consisting of 21 trees/plot, western larch (7) and lodgepole pine (6). The range of tree heights within plots dominated by Douglas-fir was 8–26 m with a mean of ~14 m. Lodgepole pine plots contained mostly even-aged trees with a mean height of 13 m and CBH of 7 m as shown in Table 2. At the level of individual tree, 225 trees measured in the field were co-identified in the laser returns (process described below). Of these, Douglas-fir was the most abundant (84 trees), distributed within five different canopy structure classes with an average tree height of ~18 m followed by western larch (63), lodgepole pine (44), and ponderosa pine (37). In general, western larch exhibited larger CBH than the other species.