

Using Laser Altimetry-based Segmentation to Refine Automated Tree Identification in Managed Forests of the Black Hills, South Dakota

Eric Rowell, Carl Seielstad, Lee Vierling, LLOYD Queen, and Wayne Shepperd

Abstract

The success of a local maximum (LM) tree detection algorithm for detecting individual trees from lidar data depends on stand conditions that are often highly variable. A laser height variance and percent canopy cover (PCC) classification is used to segment the landscape by stand condition prior to stem detection. We test the performance of the LM algorithm using canopy height model (CHM) smoothing decisions and crown width estimation for each stand condition ranging from open savannah to multi-strata stands. Results show that CHM smoothing improves stem predictions for low density stands and no CHM smoothing better detects stems in dense even-aged stands, specifically dominant and co-dominant trees ($R^2 = 0.61$, $RMSE = 20.91$ stems with smoothing; $R^2 = 0.85$, $RMSE = 46.02$ stems with no-smoothing; combined smoothed CHM for low density and unsmoothed CHM for high density stands $R^2 = 0.88$, $RMSE = 28.59$ stems). At a threshold of approximately 2,200 stems ha^{-1} , stem detection accuracy is no longer obtainable in any stand condition.

Introduction

Small-footprint airborne laser altimetry is currently the focus of considerable attention in forest inventory because of its demonstrated potential to provide both precise forest measurements (height) and accurate spatial renderings of them. One emphasis in forestry-related laser altimetry research is the automated extraction and measurement of individual trees across large tracts of forested land. In Scandinavia, several researchers have demonstrated the efficacy of individual tree stem detection from high spatial resolution (helicopter based) laser altimetry (Maltamo *et al.*, 2004; Holmgren and Persson, 2004), and in North America, delineation of individual trees has been demonstrated in the Piedmont region of Virginia using a conventional fixed

wing laser altimetry acquisition (Popescu *et al.*, 2003). The potential importance of tree stem detection is two-fold: First, by segmenting laser altimetry to single stem estimations, silviculturally important variables can be assessed at a tree scale, and second, individual tree estimates can be aggregated into coarse-scale data products for evaluation of stand and patch ecosystem dynamics.

Numerous studies have utilized passive satellite and/or airborne imagery (conventional aerial photography, multi-spectral, and hyperspectral imagery) to segment individual stems with laser altimetry data and to refine estimates of forest height and tree crown analysis (St-Onge *et al.*, 2004; Leckie *et al.*, 2003; Wulder and Seaman, 2003; Wulder *et al.*, 2000). Wulder *et al.* (2000) discussed the application of static and variable-window-size local maximum filters (LM) to high-resolution imagery for locating individual trees in plantation and naturally regenerating conifer stands on Vancouver Island, British Columbia. Popescu *et al.* (2002) transposed local maximum filters (specifically, the variable window-size), to identify individual trees and to estimate plot level tree heights in Virginia, U.S.A. using laser altimetry. The impetus for using a variable window-size is the assertion that successful detection of individual trees is predicated on proper selection of the LM filter window size (Popescu *et al.*, 2002). Using a filter that is too small or too large relative to expected crown width will propagate errors of commission or omission. Window size selection is derived from the relationship between tree height and expected crown width, in which window size is adjusted according to a local maximum (tree top) elevation. Significant progress has been made in the last three years in identifying individual stems and estimating stem densities using laser altimetry alone (Maltamo *et al.* 2004; Popescu and Wynne, 2004; Persson *et al.*, 2002) and using a combination of digital imagery and laser altimetry (Popescu *et al.* 2004).

To improve stem detection, smoothing filters have been incorporated to reduce the number of local maxima, thus mitigating the potential of detecting multiple stems within single tree crowns (Maltamo *et al.*, 2004). Focal filters have also been employed to reduce commission errors that originate from abrupt height changes in the canopy height model (CHM) (Parker and Mitchel, 2005). Popescu *et al.*

Eric Rowell, Carl Seielstad, and LLOYD Queen are with The National Center for Landscape Fire Analysis, 32 Campus Drive, University of Montana, Missoula, Montana 59812 (rowell@ntsg.umt.edu).

Lee Vierling is with the College of Natural Resources, Department of Rangeland Ecology and Management, University of Idaho, Moscow, ID 83844.

Wayne Shepperd is with the U.S. Forest Service, Rocky Mountain Research Station, 240 West Prospect, Ft. Collins, CO 80526.

Photogrammetric Engineering & Remote Sensing
Vol. 72, No. 12, December 2006, pp. 1379–1388.

0099-1112/06/7212-1379/\$3.00/0
© 2006 American Society for Photogrammetry
and Remote Sensing

(2002) used a median filter to smooth the CHM for improved estimation of individual tree crown width by fitting a polynomial line over the CHM for each tree location. Smoothing the CHM in open stands where individual trees are clearly identifiable or in stands where identification of dominant and co-dominant trees is desired is proving to be very promising. However, in heavily stocked stands and in even-aged stands, there are significant shortfalls to this method. Passing a smoothing array over stands with high stocking densities ($>1,000$ stems ha^{-1}) generalizes important subtleties that could be used to differentiate one tree from the next (Rowell, 2005).

Additional studies have also suggested that a minimum density of laser altimetry points is needed to accurately detect both individual tree crowns and to provide stand-level delineation of tree attributes. Gobakken and Næsset (2004) assert that at least five pulses per m^2 are required for tree stem identification, whereas stand-level delineation of forest biophysical variables can be accomplished with one point per m^2 (Næsset, 2002 and 2004). The primary concern for predicting individual tree heights from stem locations is that the probability that laser pulses are hitting the tops of coniferous trees is reduced as the laser post spacing increases, resulting in laser pulses that instead characterize multiple positions on the crown (Leckie *et al.*, 2003). However, both Popescu *et al.* (2002) and Rowell (2005) have demonstrated that individual tree stem detection is possible at post-spacings approximating two points per m^2 in stands with low stem densities (approximately 100 stems ha^{-1}) and in stands with moderate densities (approximately 800 stems ha^{-1}) that exhibit regular stem spacing. Conversely, extraction of individual stems from stands with irregular stocking densities (e.g., multiples structures within a single stand) remains problematic.

What is clear from these studies is that the performance of a tree detection algorithm is largely dependent on stand attributes. Consequently, an algorithm that can be parameterized by stand characteristics is likely to result in improved tree stem detection for a landscape. Segmentation of the landscape prior to tree stem detection would permit tailoring of the algorithm to suit conditions on the ground. We see three general approaches to segmentation: (a) segment by field measurement/observation (plots or stands), (b) utilize classified optical imagery, or (c) use the laser altimetry data itself. The latter option is perhaps the most desirable given the potential gains in efficiency and consistency that such an approach could afford, and this premise forms the basis for the research presented here.

This study investigates the refinement of individual stem detection in the Black Hills, South Dakota, using small footprint laser altimetry at a nominal post spacing of approximately 1.5 meters. The primary focus of research is to demonstrate that tree stem extraction can be optimized across a range of stand conditions by segmenting the laser altimetry data set according to variance of the laser altimetry height and by percent canopy cover. Segmentation promotes site-specific refinement of the crown width predictions that drive the variable window-size algorithm introduced by Popescu *et al.* (2002).

Because heavily stocked stands in the Black Hills are generally even-aged with low height variance, they are distinguishable from multi-aged/multi-strata stands that demonstrate much greater height variance (Rowell, 2005). Zimble *et al.* (2003) also distinguished between single-story and multistory vertical structure classes at fine scales using tree height variance in Idaho coniferous ecosystems. Although classification based on height variance is useful for gross characterization of vertical structure, the variance of height alone neglects a horizontal element (e.g., stem

distribution/density) that potentially influences the performance of a tree stem detection algorithm. We propose that utilizing a second metric (e.g., percent canopy cover (PCC)) in tandem with a vertical structure metric (height variance) will facilitate forest segmentation based on conditions that influence tree stem detection algorithms. The individual tree detection algorithm could then be customized by *stand* condition to optimize performance.

Following, we present the results of a study designed around three primary objectives:

1. Classify the Canopy Height Model (CHM) by height variance and percent canopy cover based on natural breaks that occur in heavily managed Black Hills stands.
2. Implement a local maximum (LM) tree stem location algorithm adapted from Popescu *et al.* (2002) on each stand condition class.
3. Test LM model performance as a function of both differential CHM smoothing and crown width estimation in each stand condition, and explore the ability of the LM approach to characterize dominant, co-dominant, and intermediate trees.

Materials and Methods

Study Area

The study area encompasses approximately 2,600 ha in the Black Hills Experimental Forest (BHEF) in west central South Dakota ($44^{\circ}94'N$, $103^{\circ}40'W$) (Figure 1). The area ranges in elevation from 1,500 to 1,800 m above sea level and is dominated by ponderosa pine (*Pinus ponderosa* Douglas ex. Lawson), with lesser concentrated populations of Black Hills white spruce (*Picea glauca*), quaking aspen (*Populus tremuloides*), and paper birch (*Betula papyrifera*). The research sites selected within the study area represent a heterogeneous mosaic of structural stand stages and age classes that result from numerous silvicultural management activities performed by the Rocky Mountain Research Station (RMRS), U.S. Forest Service.

Numerous management and thinning strategies have been implemented in the BHEF over the last thirty years. As

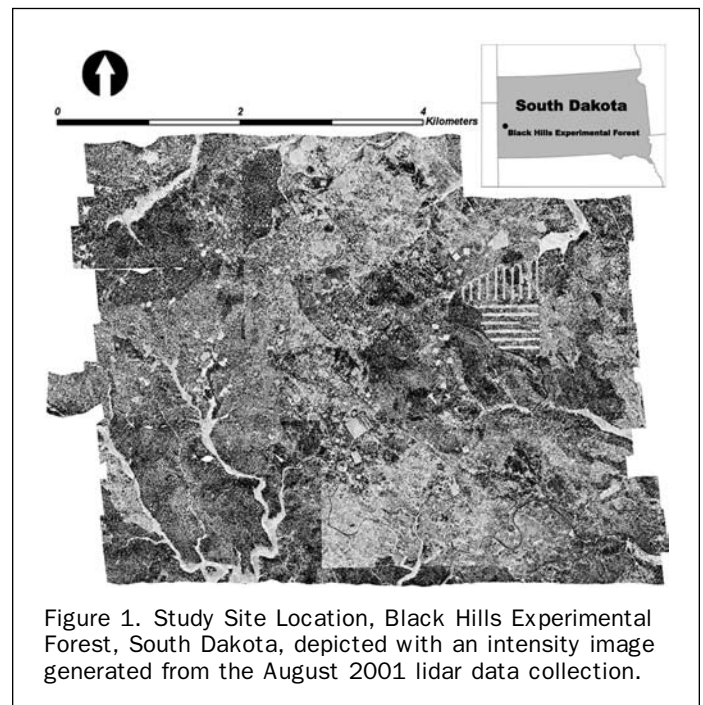


Figure 1. Study Site Location, Black Hills Experimental Forest, South Dakota, depicted with an intensity image generated from the August 2001 lidar data collection.

a result of these management activities, stands that were previously dense *dog hair* stands of ponderosa pine ($>1,000$ stems hectare⁻¹) have been thinned to stem densities <500 stems hectare⁻¹. Additionally, species composition and age class affect what type of expected biophysical variables will occur in Black Hills stands.

Ground Inventory Data Collection

Ground data were collected at seventy permanent plots in 2002 and 2003 using procedures outlined by the Fire Effects Monitoring and Inventory Protocol (FIREMON) tree data sampling methods (<http://fire.org/firemon/td.htm>). Plots were selected by generating random points using a GIS-based random point generator and then stratified according to structure type based on a statistical cluster classification map generated from the laser altimetry data. Data collected at each square 900 m² plot consist of detailed tree measurements including tree count, tree height, diameter at breast height (DBH), height to crown base (HTCB), species type, status of tree (dead or alive), and crown diameter (CD) measurements for four 100 m² nested subplots in each larger plot. A laser range finder (LaserTech Impulse 200 LR) attached to a digital compass and range staff (Mapstar Compass Module II) was used to measure offset and bearing from the point of collection to the tree with the associated biophysical variables included as an attached database. This technique produced a GPS point for each stem. A total of 4,996 individual trees were measured over the summers of 2002 and 2003. No thinning or fire occurred within the study area between the laser altimetry acquisition and the ensuing field campaign.

Airborne Laser Altimetry Data Collection and Data Processing

Laser altimetry data were collected over the study area on 16 August 2002 using a small footprint, discrete return instrument. The laser instrument employs a laser transmitter and receiver that are combined with inertial measurements and data from a dual frequency onboard GPS antenna. These data were collected using a 1,064 nm Leica Geosystems ALS-40 instrument on a Cessna 310 plane flown at 1,300 meters above mean terrain (AMT). The August 2002 acquisition collected laser returns at a nominal post spacing of 1.5 meters. Flown with a forward speed of ± 56.9 ms⁻¹ and a scan swath of 20°, laser x, y, and z reflectances were recorded at a rate of 20 kHz with an illuminated footprint of approximately 45 cm in diameter that registers up to three returns per laser pulse. Flight lines were corrected for roll, pitch, and yaw to yield data with a vertical accuracy of ± 15 cm and a horizontal accuracy ± 25 cm. Vertical accuracy of the lidar data was confirmed by comparing airport lines with a static survey GPS point, and project flight lines against four widely spaced GPS surveys throughout the study area.

Aboveground laser return points were separated from ground *bare earth* points using the virtual deforestation algorithm (VDF) (Haugerud and Harding, 2001) followed by a semi-automatic ground/above-ground separation algorithm in the TerraScan software suite. The virtual deforestation process uses a despiking algorithm to smooth the bare earth digital elevation model (DEM). A pitfall in the process is that some points are classified as bare earth because there are no points below them; these points might actually represent thickets of understory shrubs, seedlings, or tops of tall grass. (Haugerud and Harding, 2001). To further refine the DEM, we employed ground separation procedures from TerraScan to calculate an assumed bare earth surface from the lowest laser return point elevation. Remaining points are then classified as bare earth or canopy based on manually entered tolerances and the height from the assumed bare earth

surface. The combination of the VDF and TerraScan procedures may not accurately clean the bare earth DEM of all vegetation points. Therefore, the DEM is further refined by generating a temporary raster from the final bare earth surface that is adjusted higher or lower, with points above the bare earth surface classified as canopy and points below retaining a bare earth classification. A canopy height model (CHM) was created by subtracting the bare earth point elevations from the predicted canopy points. The CHM was interpolated into a raster surface using a Delaunay triangulation method. The output canopy height raster has a cell size of 0.5 m.

Data Segmentation

Variance Classification

Classification of the forest vertical structure is performed using a focal filter analysis (10 m circular, height variance) of the normalized CHM to generate a raster data set representing the variance of tree height. Utilizing all of the aboveground laser altimetry returns in the analysis produces an assessment of the vertical distribution of pulses in a variety of stand types. The filter size (10 m) is selected to encompass individual tree and clumped tree crowns across the range of expected tree crown widths (based on field measurement) in order to determine variance of height in the landscape. The height variance raster cell size is set at 0.5 m to match the cell size of the CHM. Two structure classes are identified, single strata and multi-strata, based on mean variance values for plots in the study area when compared to the modality of the vertical profiles of laser altimetry points in the coincident area.

Evaluation of the vertical canopy profiles for low variance plots demonstrates that low height variance corresponds with uni-modal stand structure. In the mono-dominant ponderosa pine ecosystem of the Black Hills, uni-modality equates to even-aged stands that are either densely stocked or that have been thinned but are still even-aged. Multi-strata stands are well represented by the high variance plots. These plots are characteristic of mature structure types (e.g., large ponderosa pine with regenerating ponderosa pine understory or white spruce dominated stands that are significantly bi-modal or tri-modal in character).

Two primary structure classes were determined based on the correlation of modality with coincident variance of height values. Raster cells representing height variance between 15.0 and 1.5 were selected to characterize single strata stands. For multi-strata stands, height variance raster cells between 100.0 and 15.0 are selected. An accuracy assessment is conducted by comparing vertical height profiles derived from field data with coincident laser altimetry measurements, in order to test for modality and to verify structure classification (Figure 2)

Percent Canopy Cover Classification

Percent canopy cover (PCC) is derived by calculating the proportion of canopy returns to the total number of returns for a given local area (Chen *et al.*, 2004; Riaño *et al.*, 2003). Crookston and Stage (1999) calculate percent canopy cover as a function of trees per acre and the projected crown area of a tree. Rowell (2005) found a significant relationship between basal area and PCC, demonstrating that PCC is an appropriate metric for evaluating the spatial density of trees within a given area. The high correlation results from the tight relationship between stem diameter at breast height (DBH) and tree crown width projected over a specific area, in the case of basal area, one hectare. For the Black Hills data set, PCC ranges that describe the low-to-high PCC gradient were selected where the low PCC class ranges from 0.00 to 0.270, moderate PCC ranges from 0.271 to 0.550, and high

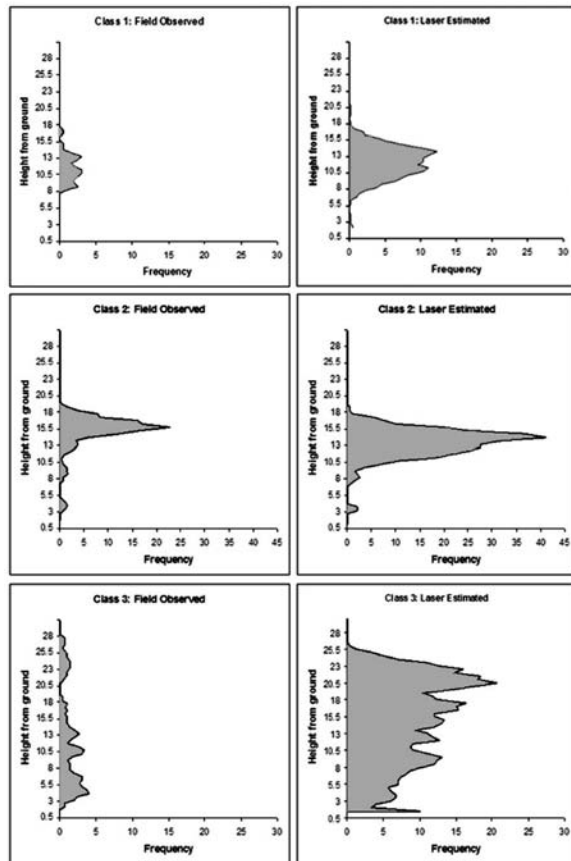


Figure 2. Comparisons of the field modeled canopy height model for class 1 (low PCC), class 2 (low height variance; moderate PCC), and class 3 (high height variance; high PCC) with coincident laser altimetry derived canopy height profiles.

PCC ranges from 0.551 to 1.00. For the purposes of this study, we utilize the low PCC classification, characteristic of open savannah stands in the BHEF, and the high PCC class, characteristic of mixed species and mature stands.

Subset of CHM to a Combined Variance/PCC Classification
The CHM derived from the laser altimetry data was subset according to three primary classes based on the combination of height variance and PCC. These classes are: (a) low PCC, any height variance; (b) moderate PCC, low height variance; and (c) high PCC, high height variance. Classes are selected based upon natural breaks caused by the variety of stand structures present at the BHEF. The CHM is then processed using six permutations (see the following section) to identify stems using the variable-size window analysis, with each run subsetted by the variance/PCC classification.

Variable-size Window Analysis

We adapted an algorithm developed by Popescu *et al.* (2002) where a variable-sized window is moved across the laser altimetry data set searching for local maximum (LM) heights. The LM filter uses an allometric relationship between observed tree height and CW to estimate a search area from the laser altimetry heights (e.g., higher points will produce a larger CW or search radius). For this study, we use site specific predictions for CW based on tree height for pon-

TABLE 1. BLACK HILLS SPECIFIC CROWN WIDTH MODELS USED TO PREDICT VARIABLE WINDOW SIZE, WHERE H IS THE FIELD MEASURED HEIGHT IN THE BHEF

Crown Width Model		R ²
[1] Linear	$CW = 0.2036(H) + 0.3564$	0.480
[2] Polynomial	$CW = 0.014(H)^2 - 0.1538(H) + 2.3563$	0.594
[3] Exponential	$CW = 1.1921e^{0.0634(H)}$	0.454

derosa pine, and compared the results from three models representing linear, polynomial, and exponential functions (Table 1). The window size is determined by rounding the value of the expected CW and dividing by the resolution of the laser altimetry height raster image to generate a square window. The window size cannot be less than 3.0 pixels in width or greater than 15 pixels in width. A maximum height within the search window is determined, and this height is then registered as a treetop. An integer stem location raster dataset is generated from the algorithm representing a laser altimetry derived stem map.

Smoothing/Non-smoothing of Canopy Height Model

In this study, we use a kernel smoothing array of 3×3 pixels for each of the three crown width model runs of the variable window-size algorithm, as well as no smoothing for each model run. Comparisons are made between effectiveness of smoothing or non-smoothing for each classification type. We subset the stem location rasters by class and depending on whether a class shows significantly stronger results using a smoothed or unsmoothed CHM, the stem locations are then merged into a compilation best fit stem location raster.

Results

Variance Classification

Laser altimetry derived canopy height profiles correspond with the variance classification (Plate 1) on the basis of modality. Canopy height profiles for class 1, plots represented by low PCC, are uni-modal with a low frequency of returns from the canopy. Comparisons with field data confirm that these plots are either within even-aged stands that have been extensively thinned (<100 trees hectare⁻¹) or that they are within thinned stands where tree density is low, but trees are large and even-aged (Figure 3, row A). In areas characterized as low variance and moderate PCC (class 2), tree density is highest ($>1,000$ trees hectare⁻¹) with even-aged stand structures that are uni-modal in distribution (Figure 3, row B). Class 3, high PCC stands with high height variance, can be characterized as stands with multi-age structure (mature overstory trees with intermediate strata of young trees) (Figure 3, row C). These stands generally have low to moderate tree density (<300 trees hectare⁻¹) and can also be of mixed-species composition. Class 3 is primarily bi-modal in distribution or in some cases tri-modal distribution depending on species composition. Plots that are characterized as class 1 averaged 25 stems (standard deviation = 12.66), plots characterized as class 2 had an average stem count of 125 stems (standard deviation = 45.00), and class 3 had an average stem count of 38 stems (standard deviation = 23.59). Within class 2 two plots fell near the class 1 to 2 break. These two plots, denoted as RMRS31 and RMRS34, are primarily characterized as class 2 with islands of class 1 within their boundaries.

Comparison of Stem Estimates

Stem counts (Plate 2) from the variable-window analysis were compared with stem counts from field data. Comparisons

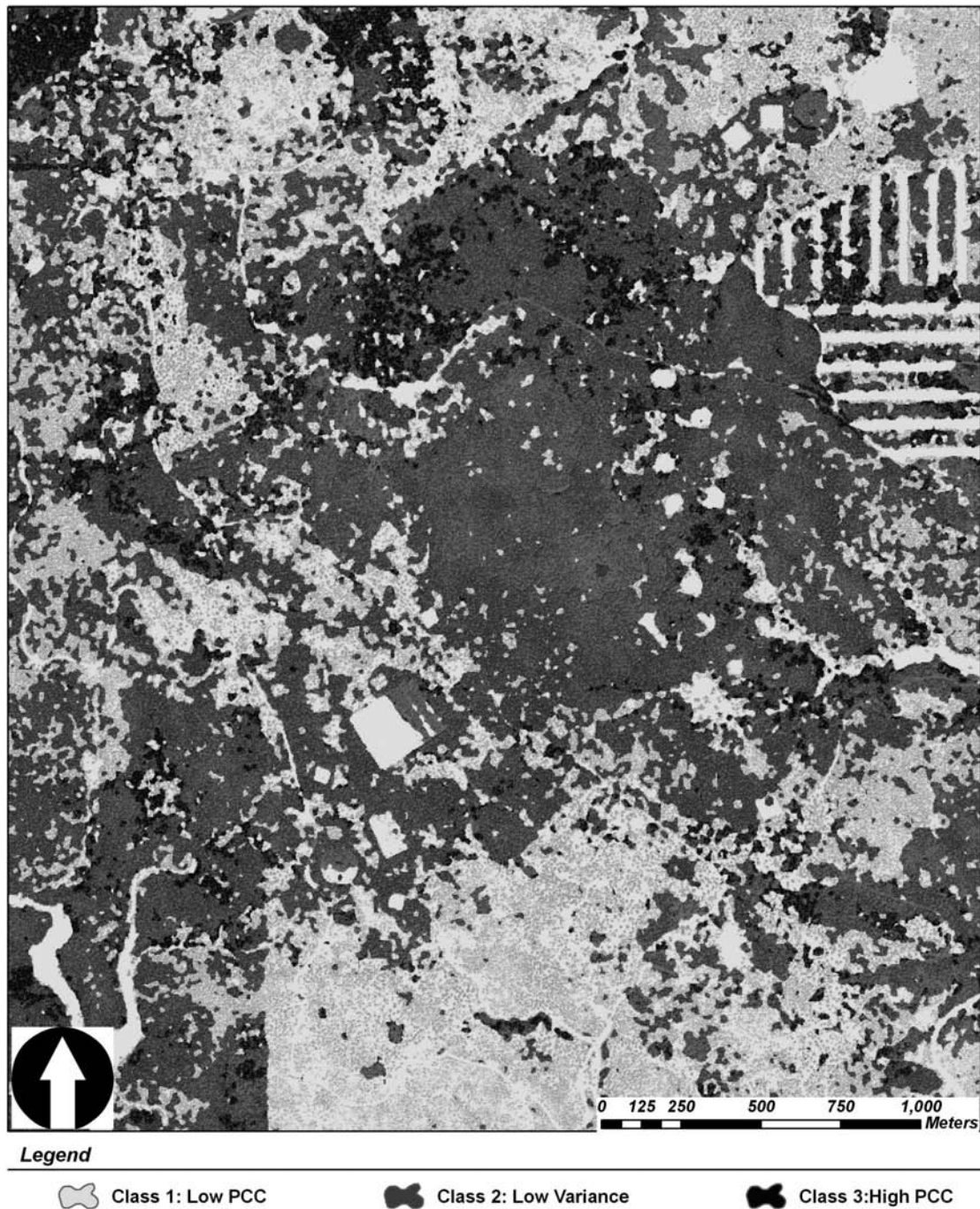
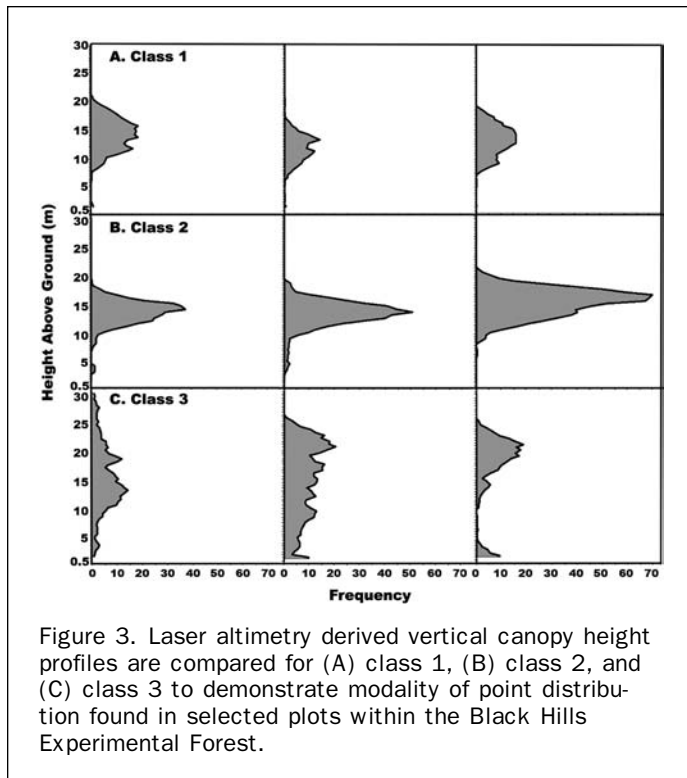


Plate 1. A map depicting the spatial distribution of height variance/percent canopy cover classes in the central Black Hills Experimental Forest.

were conducted to understand the types of trees (e.g., dominant, co-dominant, intermediate) that are identifiable using the variable-size window analysis. A baseline dataset is created using the variable-size window algorithm that incorporates a general crown width model derived from the observed tree height/crown width relationship for 397 trees (Table 1, Equation 1) with no smoothing. Comparison of laser-based stem detection and field observed tree counts

(Table 2) for all trees within each plot using the baseline model reflects mixed results with respect to accuracy of stem prediction ($R^2 = 0.24$, RMSE = 107.54 stems). When the field observed stem counts were limited to only dominant and co-dominant populations, the relationship improved slightly ($R^2 = 0.42$, RMSE = 115.22 stems). A second model run using the same crown width prediction coupled with a 3×3 smoothing kernel is performed to evaluate changes in prediction



($R^2 = 0.85$, RMSE = 46.02 stems). Using combinations of a smoothed CHM for class 1, no-smoothing of the CHM for class 2, and smoothing of the CHM for class 3, dominant and co-dominant stem counts are well characterized with relatively low error ($R^2 = 0.88$, RMSE = 28.59 stems, Figure 4). The comparisons indicate that a saturation effect occurs in plots with stem densities exceeding 200 stems plot^{-1} (900 m^2). In these stand types, characterized as dense (“dog hair”) even-aged ponderosa pine, model performance breaks down when predicting stem locations for dominant and co-dominant trees (Figure 5). The two distinct trends in Figure 6 clearly demonstrate that, in managed stands (e.g., low density, thinned), without CHM smoothing there is a potential for overestimating stem counts. Stems are more effectively located in class 2 when no smoothing occurs ($R^2 = 0.64$, RMSE = 1.83 stems) versus using smoothing ($R^2 = 0.56$, RMSE = 5.36 stems). This finding is also supported by Parker *et al.* (2005), who found that smoothing reduced the estimated number of stems acre^{-1} when compared with estimates from unsmoothed data.

Tree Height

Detecting individual stems from a laser altimetry data set is used to refine estimates of tree heights, (a *native* measurement of laser altimetry). Tree stems derived from a laser altimetry data set can be used to refine estimates of tree heights (a *native* measurement from laser altimetry) because the measurements of local maxima provide direct comparisons with field measurements that could not otherwise be easily obtained using all first returns in a canopy point cloud.

Tree height estimates from discrete return laser altimetry are well documented in the literature (e.g., Magnussen *et al.*, 1999; Næsset and Bjerknes, 2001; Pearson *et al.*, 2002). Accuracy of the maximum predicted height is paramount to successful characterization of secondary forest inventory variables (e.g., DBH). Discrete return laser altimetry in coniferous forests tends to underestimate maximum height due to the low probability that laser pulses will consistently reflect enough energy from the tops of trees. Magnussen and Boudewyn (1998) attribute the underestimation as a factor of the laser altimetry returns primarily being composed of “locations below the tree tops.” At the BHEF, a comparison between field observed maximum height and laser predicted maximum tree height showed a significant relationship ($R^2 = 0.94$, RMSE = 6.46 meters). The laser-predicted tree height exhibited the anticipated underestimation of tree height. However, estimated tree heights generally fell near the identity line (1:1) (Figure 7). Whether or not laser pulses strike the apices of trees is, in part, a factor of scanning geometry, i.e., proximity to nadir. The position of a tree relative to nadir clearly affects the laser recorded maximum tree height, and trees sampled at increasingly large look-angles were more likely to provide accurate maximum heights. No discernable relationships were identified between plot stem density and the average difference between the observed maximum tree height and the laser predicted tree height (Table 1). Analysis of plot proximity to nadir (data not shown) suggests that the closer

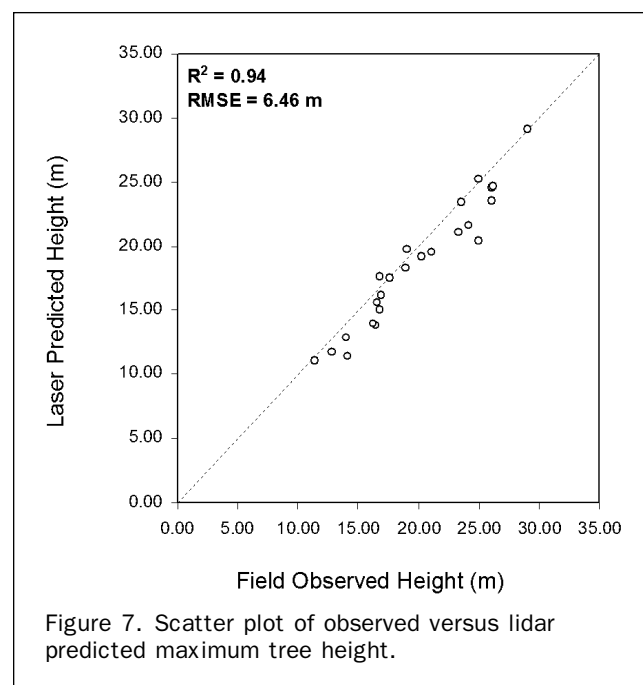
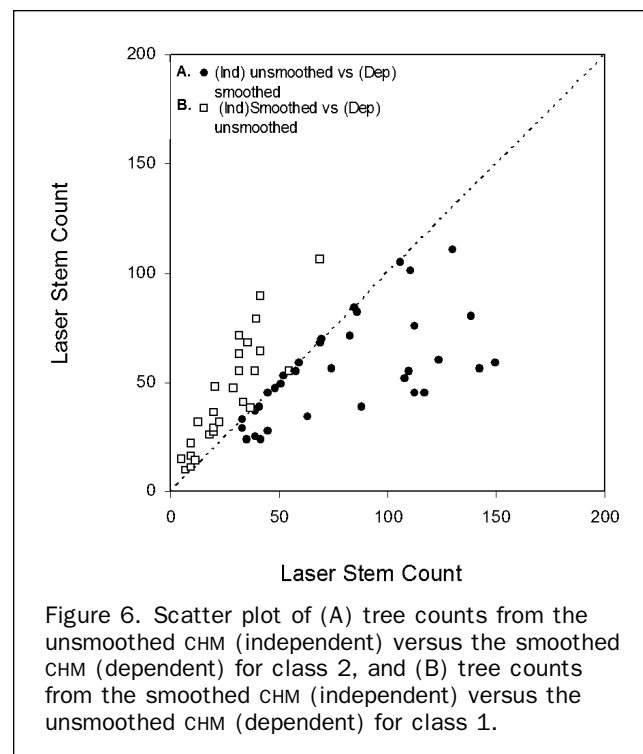
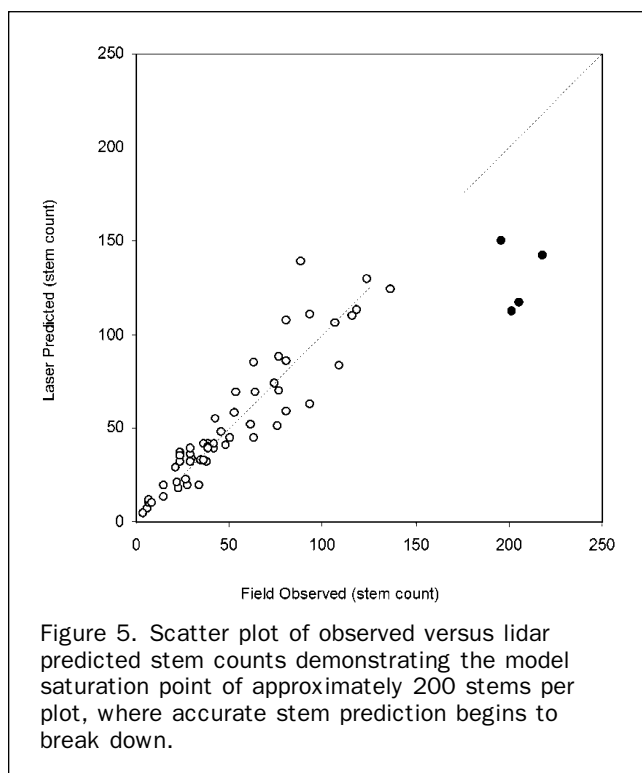
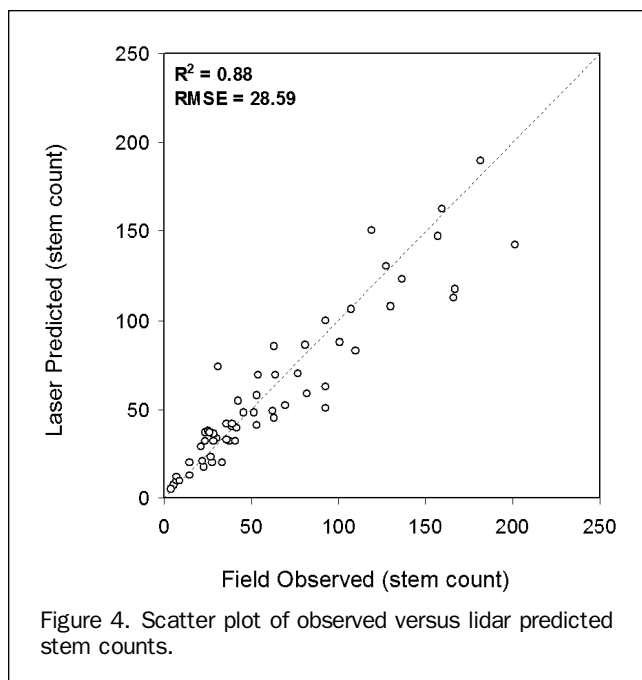
TABLE 2. SUMMARY RESULTS OF SMOOTHING VERSUS NO SMOOTHING OF THE CANOPY HEIGHT MODEL FOR ALL STEMS AND DOMINANT/CO-DOMINANT STEMS, IN ADDITION TO THE COMBINED BEST-FIT OUTPUT

Sample	Model	R^2	RMSE (Stems)
All Stems	No-smoothing	0.24	96.68
	Smoothing	0.61	21.57
Dominant/ Co-Dominant	No-smoothing	0.85	46.45
	Smoothing	0.42	104.70
Dominant/ Co-Dominant	Combined	0.88	28.59

accuracy with smoothed data. Results show that this smoothing procedure improves stem detection using the variable window-size algorithm. Comparison of smoothed CHM stem count predictions with observed stem counts per plot (all trees) reflects a stronger relationship ($R^2 = 0.61$, RMSE = 20.91 stems). One significant outlier is a mixed coniferous/deciduous plot where the field observed stem count (322 stems) differed significantly from the laser altimetry derived stem count (63 stems). Using only dominant and co-dominant populations (Table 3), the relationship between laser altimetry derived stems and observed stems improves further

TABLE 3. RESULTS OF OBSERVED STEMS FOR ALL AND DOMINANT/CO-DOMINANT TREES WITH PREDICTED STEM LOCATIONS

Class	Average Stems			Standard Deviation Stems			RMSE (stems)	
	Observed (All)	Observed (D/CD)	Predicted	Observed (All)	Observed (D/CD)	Predicted	All	D/CD
1	24.36	24.36	26.64	13.08	13.08	15.60	11.40	11.40
2	116.93	98.20	87.95	66.07	47.67	41.78	152.18	49.57
3	64.00	41.20	40.80	23.59	16.15	7.40	51.88	0.89



to nadir the less accurate the laser maximum height estimate, while on average near nadir plots underestimated maximum tree height by a mean error of 2.27 m.

Discussion

Segmentation by Variance/PCC Classification and Effects of Smoothing

Smoothing of the CHM improved stem detection within stands that are characterized as low stem density (class 1).

Running the algorithm using a smoothed CHM for these stand types significantly reduced the effects of the irregular spatial point distribution from the laser altimetry dataset that influences the spatial-continuity of the CHM pixels. Smoothing of the CHM in open stands represents the most influential component for improving estimates of stem counts. Using the smoothing kernel array reduced multiple stem locations on single trees, specifically reducing the number of stems that fall on the edges of

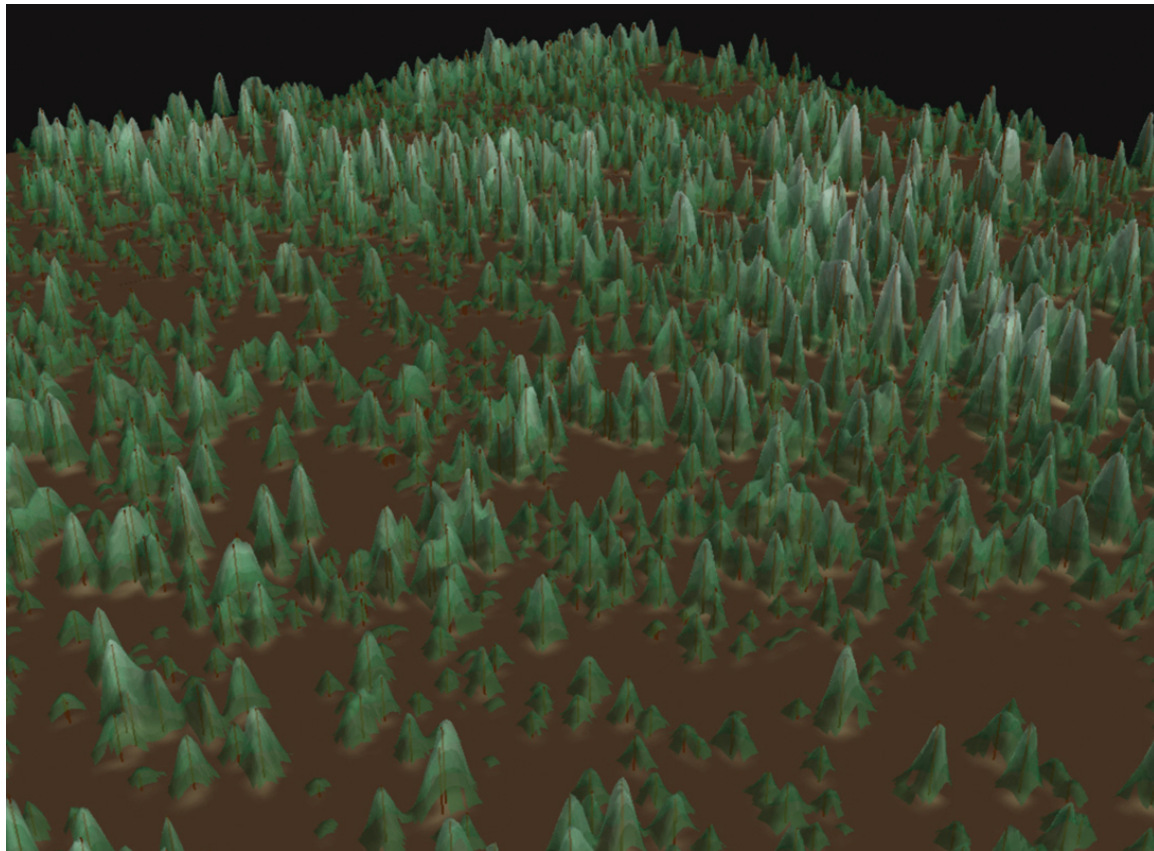


Plate 2. The smoothed canopy height profile draped over modeled stems located by the variable window-size algorithm.

areas denoted as individual trees. All three crown width models predicted similar numbers of stem locations, but smoothing reduced the occurrence of false positives registered as stem locations.

Smoothing the CHM for stands within class 2 generalized the CHM so that stems are under-predicted in most cases. Within class 2 stand types, two distinct management practices occur. One management practice is even-aged shelter thinning, where trees are evenly spaced and have been thinned to maximize growth for timber harvest. The other management type can be characterized as non-thinned control stands, where the stands are even-aged. Class 3 exhibits similar stand structure to class 1; however class 3 is composed of intermediate trees growing below the dominant mature canopy. CHM smoothing within these stand types improves stem estimates within white spruce dominated stands and is problematic in mature ponderosa pine dominated stands. White spruce in the Black Hills typically have dense branches running from the top to base of the tree, offering little potential for seedlings to grow below them. Additionally, many of these stands are comprised of both clumps of trees and isolated single trees. For ponderosa pine dominated stands in class 3, CHM smoothing reduces the effectiveness of detecting intermediate stems that are adjacent to or have a majority of their crown covered by the dominant overstory.

A classification based subset of the CHM, using a smoothed CHM for class 1 and an unsmoothed CHM for model runs in classes 2 and 3, shows improvement regard-

ing estimation of dominant and co-dominant trees. The model is still unable to attain accurate predictions for stems >200 stems plot^{-1} , if the plot has irregularly spaced trees (unmanaged).

Effects of Crown Width Equation and Window Size

The model used to predict crown width from laser altimetry estimated height plays a significant role in how many stems are actually estimated. Detecting stem locations using allometrically derived crown width predictions that are incorporated in the variable-size window algorithm proposed by Popescu *et al.* (2002) becomes problematic in multiple management forest systems like the BHEF. Evaluation of crown width/tree height allometry in the study area demonstrates that utilizing a singular crown width prediction does not accurately encompass the variability apparent in the multi-management forest structure of the Black Hills (Rowell, 2005).

Sensitivity analysis conducted between linear, polynomial, and exponential models all showed different levels of accuracy. In all models there is agreement that prediction of crown width for trees between 10.0 and 15.0 m in height are more accurate because there is a significant population of trees within this range affecting the modeled mean value. As tree heights become greater, there is a marked increase in variability regarding the mean crown width prediction. The results demonstrate that the crown width model used to predict the stem location will affect how many tree locations are detected. Smaller variable window sizes are sometimes

representing non-tree locations along edges of dominant and co-dominant trees. In general, changes in stem counts occur on a marginal basis, regardless of the model used. More importantly, the variable window-size algorithm misses numerous potential local maxima because the algorithm is unable to distinguish subtle changes in the CHM. We found that in high density stands there are noticeable regions that are un-represented with regard to stem prediction. In all cases, these areas have low variance of top of canopy height, <1.0 m difference between adjacent local maxima. Popescu *et al.* (2004) revised the LM algorithm by employing a circular window in lieu of using a square window. That study found improvements regarding height estimation from stem locations, but makes no conclusions regarding stem counts. Using a square window analysis does introduce a potential problem as adjacent local maxima (potential stem locations) may be included in the dominant local maximum sphere of influence. Secondly, adjacent areas that are very close in height could be treated as a flat plain regardless of window shape and therefore, no local maximum would be detected with no subsequent stem ID.

Conclusions

This study demonstrates an approach to refine stem identification using laser altimetry as the sole data source. We show that in the Black Hills of South Dakota, laser-based height variance and percent canopy cover can be used to inform decisions regarding which areas would benefit from data smoothing and which are better left unsmoothed. Stands characterized by low PCC are indicative of open savannah ponderosa pine. Thinning and management decisions in these stands generally promote even-aged shelter belt structure. In these stands, smoothing the CHM improves estimates of stem counts. Conversely, stem ID in areas with moderate PCC and low height variance is significantly improved using an unsmoothed CHM at stem densities <2,200 ha⁻¹. The low height variance classes are characterized by even-aged closed canopy stands. In some cases, these areas are dense, unmanaged *dog hair* stands, which prove problematic for predicting stem locations regardless of smoothing. We document a threshold of approximately 2,200 stems ha⁻¹ as the point beyond which acceptable stem count accuracy is no longer obtainable. In the high PCC, high height variance areas, decisions regarding whether to smooth the CHM appear to be species dependent (in the BHEF, spruce versus pine). In the latter stands, utilization of multi-spectral or hyper-spectral imagery may prove more prudent for distinguishing between coniferous species.

In terms of estimating crown width from tree height in the BHEF, we show that the form of the model used to relate the two (linear, polynomial, and exponential) is relatively insignificant to LM algorithm performance. However, it is important to recognize the influence of the tree height mode on the modeled mean height value (e.g., the best crown width estimates are obtained for trees with heights similar to model mode height).

Our study supports previous work that demonstrates the efficacy of using multiple approaches to estimate stems within a single landscape, specifically to refine estimation of tree biophysical parameters. The "Catch-22" is that a forester needs to be quite familiar with a landscape prior to acquiring laser altimetry data in order to take full advantage of what we (collectively) are learning about the technology. It is not yet clear whether results from the Black Hills are portable to other locations or vice versa, but it is our hope that a critical mass of results will soon be obtained that allow algorithms and models to be applied regionally and/or by type, species, structure with little modification.

Future research in the tree detection domain should certainly utilize richer data sets in order to characterize canopies in dense even-aged stands (following the Scandinavian example). Additionally, methods to segment landscapes prior to laser data acquisition could inform mission planning and result in multiple-mode acquisitions tailored to what we think is on the ground. It does seem clear that a one-size-fits-all approach to forest inventory even within relatively small spatial areas may not be the best approach. However, the primary issues with using laser altimetry for forestry still remain the relatively high cost of data for even small acquisitions and the spatial irregularity of data caused by complex topography. These considerations aside, laser altimetry derived stem detection is currently feasible in many forest types, and should lead to better estimates for all subsequent biophysical variables that may be derived from improved height estimates.

Acknowledgments

This work was supported in part by NASA EPSCOR Grant NCC5-588 (L. Vierling, Principal Investigator) and by the Upper Midwest Aerospace Consortium Forest Public Access Resource Center (UMAC Forest PARC). We thank Meghan Calhoun, Rick Clawges, Katy Derr, Beth Hansen, Shane Hansen, Michael Toomey, and Rachel Smith for field assistance; also Maribeth Price, Patrick Zimmerman, and Roger Johnson for helpful discussions regarding the study. We also thank Horizons, Incorporated for the donation of the laser altimetry data used in this study. Ross Nelson, Randolph Wynne, and three anonymous reviewers provided helpful and insightful comments on an earlier version of the manuscript.

References

- Chen, X., L.A. Vierling, E. Rowell, and T. DeFelice, 2004. Using lidar and effective LAI data to evaluate IKONOS and Landsat 7 ETM+ vegetation cover estimates in a ponderosa pine forest, *Remote Sensing of Environment*, 91(1):14–26.
- Crookston, N.L., and A.R. Stage, 1999. *Percent Canopy Cover and Stand Structure from the Forest Vegetation Simulator*, United States Department of Agriculture, Forest Service, Rocky Mountain Research Station, General Technical Report, RMRS-GTR-24.
- Gobakken, T., and E. Næsset, 2004. Estimation of diameter and basal area distributions in coniferous forest by means of airborne laser scanner data, *Scandinavian Journal of Forest Research*, 19(6):529–542.
- Haugerud, R.A., and D.J. Harding, 2001. Some algorithms for virtual deforestation (VDF) of LIDAR topographic survey data, *International Archives of Photogrammetry and Remote Sensing*, Vol. XXXIV, Part 3/W4, Commission III, pp. 211–217.
- Holmgren, J., and A. Persson, 2004. Identifying species of individual tree using airborne laser scanner, *Remote Sensing of Environment*, 90(4):415–423.
- Leckie, D., F. Gougeon, D. Hill, R. Quinn, L. Armstrong, and R. Shreenan, 2003. Combined high-density lidar and multispectral imagery for individual tree crown analysis, *Canadian Journal of Remote Sensing*, 29(5):633–649.
- Magnussen, S., and P. Boudewyn, 1998. Derivations of stand heights from airborne laser scanner data with canopy-based quantile estimators, *Canadian Journal of Forest Research*, 28:1016–1031.
- Magnussen, S., F. Gougeon, D. Leckie, M. Wulder, 1999. Predicting tree heights from a combination of LIDAR canopy heights, and digital stem counts, *Proceedings of the IUFRO Conference on Remote Sensing and Monitoring*, 01–03 June, Rogow, Poland.
- Maltamo, M., K. Eerikäinen, J. Pitkänen, J. Hyypä, and M. Vehmas, 2004. Estimation of timber volume and stem density based on scanning laser altimetry and expected tree size distribution functions, *Remote Sensing of Environment*, 90(3):319–330.

- Næsset, E., and K.O. Bjerkness, 2001. Estimating tree heights and number of stems in young forest stands using airborne laser scan data, *Remote Sensing of the Environment*, 78(3):328–340.
- Næsset, E. 2002. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data, *Remote Sensing of the Environment*, 80(1):88–99.
- Næsset, E., 2004. Accuracy of forest inventory using airborne laser-scanning: Evaluating the first Nordic full-scale operational project, *Scandinavian Journal of Forestry*, 19(6):554–557.
- Parker, R.C., and A.L. Mitchel, 2005. Smoothed versus unsmoothed LiDAR in a double-sample forest inventory, *Southern Journal of Applied Forestry*, 29(1):40–47.
- Persson, Å., J. Holmgren, and U. Söderman, 2002. Detecting and measuring individual trees using an airborne laser scanner, *Photogrammetric Engineering & Remote Sensing*, 68(9):925–932.
- Popescu, S.C., R.H. Wynne, and R.E. Nelson, 2002. Estimating plot-level tree heights with lidar: Local filtering with a canopy-height based variable window, *Computers and Electronics in Agriculture*, Elsevier Science, 37(1–3):71–95.
- Popescu, S.C., R.H. Wynne, and R.E. Nelson, 2003. Measuring individual tree crown diameter with lidar and assessing its influence on estimating forest volume and biomass, *Canadian Journal of Remote Sensing*, 29(5):564–577.
- Popescu, S.C., and R.H. Wynne, 2004. Seeing the trees in the forest: Using lidar and multispectral data fusion with local filtering and variable window size for estimating tree height, *Photogrammetric Engineering & Remote Sensing*, 70(5):589–604.
- Popescu, S.C., R.H. Wynne, and J.A. Scrivani, 2004. Fusion of small-footprint lidar and multispectral data to estimate plot-level volume and biomass in deciduous and pine forests in Virginia, USA, *Forest Science*, 50(4):551–565.
- Riaño, D., E. Meier, B. Allgöwer, E. Chuvieco, and S.L. Ustin, 2003. Modeling airborne laser scanning data for spatial generation of critical forest parameters in fire behavior modeling, *Remote Sensing of Environment*, 86(2):177–186.
- Rowell, E., 2005. *Estimating Forest Biophysical Variables from Airborne Laser Altimetry in a Ponderosa Pine Forest*, Masters thesis, South Dakota School of Mines and Technology, Rapid City, South Dakota, 99 p.
- St-Onge, B., J. Jumelet, M. Cobello, and C. Véga, 2004. Measuring individual tree height using a combination of stereophotogrammetry and lidar, *Canadian Journal of Forest Research*, 34(10):2122–2130.
- Wulder, M., K.O. Niemann, and D.G. Goodenough, 2000. Local maximum filtering for the extraction of tree locations and basal area from high spatial resolution imagery, *Remote Sensing of Environment*, 73(1):103–114.
- Wulder, M.A., and D. Seeman, 2003. Forest inventory height update through the integration of lidar data with segmented Landsat imagery, *Canadian Journal of Remote Sensing*, 29(5):536–543.
- Zimble, D.A. D.L. Evans, G.C. Carlson, R.C. Parker, S.C. Grado, and P.D. Gerard, 2003. Characterizing vertical forest structure using small-footprint airborne LiDAR, *Remote Sensing of Environment*, 87(2–3):171–182.