COMPARISON OF LIDAR- AND PHOTOINTERPRETATION-BASED ESTIMATES OF CANOPY COVER

Demetrios Gatziolis

ABSTRACT

An evaluation of the agreement between photointerpretation- and LiDAR-based estimates of canopy cover was performed using 397 90 x 90 m reference areas in Oregon. It was determined that at low canopy cover levels LiDAR estimates tend to exceed those from photointerpretation and that this tendency reverses at high canopy cover levels. Characteristics of the airborne imagery used, and, to a lesser extent, the density of the sampling point pattern employed and the occasional photointerpretation error inflated estimate discrepancies. Where available, LiDAR data can potentially be used to quantify the magnitude of error embedded in estimates of canopy cover obtained via photointerpretation.

INTRODUCTION

Forest canopy cover is an important ecological indicator that is known to affect, among many other phenomena, near-ground solar radiation (Zou and others, 2007), tree regeneration (Stancioiu and O'Hara, 2006), and wildlife habitat (Ganey and others, 2008). It also plays a key role in estimating forest stand attributes from remotely sensed data (Jennings, 1999). The importance of canopy cover for national forest inventory operations has increased since the Food and Agriculture Organization (FAO, 2000) established the 10 percent canopy cover threshold as the universal criterion defining forest land. Prompted by this development, the Forest Inventory and Analysis (FIA) Program of the U.S. Forest Service has recently decided to adopt canopy cover as forest land determinant and it is now participating in an effort designed to model canopy cover across the conterminous U.S. Model predictions are based on Thematic Mapper imagery and ancillary data and will be organized in raster layers. A 5-year updating schedule is envisioned. Canopy cover estimates serving as training data for model development are obtained by manual photointerpretation (PI) of high-resolution airborne imagery.

The term 'canopy cover' adopted by FIA follows the definition suggested by Avery and Burkart (1994) according to whom it is the percent forest area occupied by the vertical projection of tree crowns. In this definition, tree crowns are considered opaque or solid objects and it is implied that canopy cover estimates obtained in the field should only involve observations performed in the exact vertical direction. Dot count (Rautiainen and others, 2005), line intercept (Gregoire and Valentine, 2007) and moosehorn (Fiala and others, 2006) sampling techniques meet this requirement; hemispherical photography (Korhonen and Heikkinen, 2009), a popular alternative, does not, but, reportedly, the effects of the oblique angle view can be minimized by photograph post-processing. All these approaches for field estimation of canopy cover are logistically infeasible for a project with national scope. Estimates based on remotely sensed data are perhaps the only plausible alternative.

Spectral imagery acquired by airborne or satellite platforms conducive to unbiased estimation of canopy cover should have sufficiently fine spatial resolution that allows the identification of individual tree crowns or crown clusters and the delineation of between-crown openings (gaps), and narrow field of view centered at nadir (Korpela, 2004). Where the latter requirement is not met, trees depicted in high-resolution imagery exhibit substantial 'layover' or radial displacement of their crown tops relative to their bases that is intensified as the distance from the image's nadir point increases. This displacement leads to partial obstruction of portions of a tree's crown or of nearby canopy gaps, either by the tree in question or by its neighbors. Consequently, the minimum size of canopy gaps that can be reliably identified in such imagery increases with the distance from the nadir point, ultimately leading to bias in the estimation of canopy cover. Solar illumination and terrain conditions can inflate the bias.

High-density Light Detection and Ranging (LiDAR) data are far less susceptible to bias in part because they are independent of solar illumination and terrain conditions but primarily due to the fact that laser pulses are capable of penetrating tree crowns. LiDAR instruments emit short pulses of light propagated as a narrow beam towards illuminated objects and record the amount of energy that is backscattered to the sensor and the length of time that has elapsed. By processing this information the laser instrument identifies points, also known as returns or

Demetrios Gatziolis, PNW Research Station, dgatziolis@fs.fed.us.

echoes, precisely georeferenced in space, that correspond to the locus of the backscattering. Pulses illuminating hard objects (bare ground, building roofs, etc.) generate a single return. Objects that are not solid, for example tall vegetation, typically generate more returns along the pulse's propagation trajectory. LiDAR data over forested landscapes comprise large sets of returns known as the return 'clouds' that represent sampling of terrain and vegetation materials. Therefore, unlike spectral imagery, LiDAR pulses can sample the portion of a tree crown, even its lower components, positioned away from the flight line of the airborne platform.

This study compares canopy cover estimates obtained via photointerpretation to those assessed from corresponding high-density LiDAR data across a variety of topographic, physiographic, and forest management conditions in Oregon.

METHODS

The 31500 km^2, 75 km wide study area extends from the coastal mountains of Oregon, across the Willamette Valley and the Cascades, eastward to the nearly the Idaho border (Figure 1a), and it is sometimes known as the Oregon transect. It is one of the five pilot study areas selected for the national canopy cover project undertaken by FIA. Forests on the coastal mountains and the western half of the Cascades typically present with high canopy cover which is progressively reduced in the eastern part of the Cascades until the open forests of eastern Oregon are reached. Within the study area, 397 reference areas, each covering 90 x 90 m and centered on FIA plot locations were identified as contained in high-density LiDAR acquisitions in the 2008 – 2010 period. These reference areas will be henceforth mentioned as 'plots.' In each plot, a regularly-spaced 105 point grid was superimposed on 1-m airborne National Agriculture Imagery Program (NAIP) data acquired in 2009 (Figure 1b). Using the NAIP imagery as reference, experienced photointerpreters labeled each of the 105 points in each plot either as belonging either on a tree crown or background objects. Estimates of plot canopy cover were obtained as the ratio of tree points to the total.

To obtain the LiDAR-based estimates of plot canopy cover, the elevation value of each return was first converted to above-ground height by using a digital elevation model (DEM) also generated from the LiDAR data. All returns with height equal to or larger than a threshold were labeled as trees and the remaining ones as background returns. Three height thresholds (1, 2, and 3 m) were considered. Subsequently, raster representations of tree and background return frequencies were computed. Raster cells containing at least one return labeled tree were assigned a value of 1 while cells with only background returns were assigned a value of 0. Cells with no returns were assigned a 'nodata' value and were excluded from further consideration. To ensure that the frequency of nodata cells, and therefore their effect on the canopy cover estimates, is minimized, the resolution of the raster frequency representation was set to the mean laser (footprint) spacing between spatially adjacent pulses. The plot estimates of canopy cover were calculated as the ratio of the value 1 cells to the sum of value 1 and 0 cells. This method for computing canopy cover estimates from laser data was evaluated using precise delineations of tree crowns detailed in Gatziolis and others (2010) and was found to not deviate by more than 3 percent from the field estimates, at least where the density of the LiDAR data exceeded 8 returns per square meter.

To account for registration discrepancies between the LiDAR and NAIP data, all returns on and in the vicinity of a plot were jittered 200 times in two dimensions by using random azimuths and distances drawn from a -5 to 5 m uniform distribution. The magnitude of the jittering was determined by measuring the mean adjustment required to achieve spatial registration by ocular means. The mean LiDAR-based plot canopy cover was finally calculated from the 200 plot-jittering instances.

RESULTS AND DISCUSSION

The scatterplot of PI- vs. LiDAR-based canopy cover indicates that at low cover levels, PI tends to produce lower estimates (Figure 2) than LiDAR. At high canopy cover levels this tendency reverses. A second-order polynomial regression of PI on LiDAR estimates exhibits coefficient of determination $R^2 = 0.787$ with the regression fitted line crossing the 1:1 one at canopy cover of approximately 35 percent. This is in part because at very low canopy cover levels, trees in the landscape can be considered rare events that are not sampled adequately by the point pattern used. At high canopy cover, it is the openings or gaps within the crowns that are rare and undersampled. Given the 1 m resolution of the NAIP imagery, the horizontal footprint of either a small tree or canopy opening would have to exceed 4 m^2, twice the square of the resolution, before it can be identified clearly. To both comply with the minimum identifiable object size requirement and avoid bias due to undersampling of rare events, the density of the point pattern would have to increase by at least an order of magnitude above the present level, an option which is logistically infeasible.

Figure 2 features 3 plots with LiDAR estimates higher than 60 percent and corresponding PI estimates lower than 30 percent and another 3 plots with LiDAR estimates lower than 35 percent and PI estimates higher than 75 percent.

These plots and many others have been examined carefully using various ancillary data in an effort to identify the source of such large discrepancies between the estimates. It was determined that for all six plots photointerpretation error was responsible for the discrepancies. Among the most challenging plots ranked those with uniform hardwood tree crowns mistaken for grass or brush and those covered with snow at the time of the NAIP acquisition. Two other plots with large estimate discrepancy had sustained insect infestation, a condition not anticipated by the LiDAR-based canopy cover estimation procedure which lead to overestimation. Smaller discrepancies were attributed to poor imagery quality, such as hazy conditions and lack of sufficient contrast.

In addition to the undersampling of openings, the PI overestimation of canopy cover compared to LiDAR was attributed to oblique NAIP imagery. In plots or stands with canopy cover higher than 50 percent, or even lower but with trees growing in clusters, the effects of imagery obliqueness are more pronounced. While only a small percentage of pulses had viewing angle greater than 10 degrees, for about 1/3rd of the study area the effective view angle of the NAIP imagery exceeded that angular threshold. In the presence of tall vegetation, steep terrain and fairly low sun elevation angle, conditions that are actually the norm rather than the exception in much of the Pacific Northwest, crown openings are partially or completely obstructed from view. Unless the airborne imagery is acquired with long focal length lens, its information content may not be compatible with unbiased estimation of canopy cover regardless of the diligence and skill of the photointerpreter or the sampling intensity.

Overlays of the sampling point pattern with the NAIP imagery questioned the choice of regularity in the former for several plots examined. The arrangement of points in the pattern yields a 9-m distance between a point and its immediate neighbors. This point spacing is a multiple of the planting distance for many commercial forests in the Pacific Northwest. Although certainly not an issue in ‘natural’ forest stands, systematic sampling can have unintended implications where sample points happen to consistently lay on crowns or canopy openings. Alternatively, random sampling point pattern could perhaps be employed in regions with substantial component of commercial forests.

Modifying the object height threshold that separates trees from background objects was found to have a small overall effect on the agreement between PI and LiDAR estimates of canopy cover. For height threshold equal to 1 m, 2 m, and 3 m the root mean square discrepancy between the two types of estimates was 14.98, 15.20, and 15.92 percent respectively. For 21 plots, 5.3 percent of the total, increasing the threshold from 1 to 3 m resulted in a more that 4 percent change in estimate discrepancy. It should be noted that within the study area there was hardly any portions with non-tree vegetation of mean height larger than each of the thresholds specified and for the majority of the open forests in eastern Oregon there is little or no understory. Such conditions facilitate precise estimation of canopy cover, at least for the LiDAR-based approach. In different biomes and dominant cover types, the distinction between tree and non-tree vegetation might be less clear. While LiDAR data do describe the vertical structure of vegetation, we are yet to see in literature methodologies and applications capable of accurately and consistently discerning bushes and brush from tree overstory. In such conditions, how well the height threshold selected represents the vegetation profile will likely determine the accuracy of the estimates obtained.

Assuming that the LiDAR-based estimates of canopy cover are either unbiased or, if not, only marginally biased, the results of this study suggest that the PI-based estimates contain substantial bias at least for plots with low or high true canopy cover. Considering that the primary objective for the PI effort is to support the national canopy cover project, it should be concerning that the bias in the PI estimates will propagate through the modeling function and likely bias the outputs. Given the model structure types considered for the national project, it is unlikely that one can assess *a priori* the effect of the bias in the input to any bias in the output. Perhaps the only viable option is to repeat the modeling effort once with PI estimates as input and once with their LiDAR equivalent and compare the outputs, at least in regions where high-density LiDAR data is available. Such a comparison could lead to useful insights towards methodological improvement in the PI process and in the structure of models employed for future implementations of the national canopy cover project.

CONCLUSION

An evaluation of the agreement between PI- and LiDAR-based estimates of canopy cover was performed using a large number of plots across a variety of vegetation and topographic conditions. The evaluation indicates that the agreement between estimates relates to the value of canopy cover. There is sufficient evidence to suggest that the PI approach tends to underestimate low and to overestimate high canopy cover. In addition to bias, PI estimates appear to be imprecise as well, in part because of the characteristics of the airborne imagery used. The magnitude of the bias can be quantified where high-density LiDAR data is available. Additional investigations are needed to determine if bias removal or reduction can be achieved.

LITERATURE CITED

Avery, T.E.; Burkart, H.E. 1994. Forest Measurements, 4th edition, McGraw-Hill, New York, USA, 331 pp.

Fiala A.C.S.; Garman, S.L.; Gray, A.N. 2006. Comparison of five canopy cover estimation techniques in the western Oregon Cascades. Forest Ecology and Management, 232:188–197.

Food and Agriculture Organization. 2000. On definitions of forest and forest change. Forest Resources Assessment Program. Working paper 33. Rome, Italy. 15p. URL: ftp://ftp.fao.org/docrep/fao/006/ad665e/ad665e00.pdf (last accessed December 15, 2010).

Gatziolis, D.; Fried, J.S.; Monleon, V. 2010. Challenges to estimating tree-height via LiDAR in closed-canopy forests: a parable from western Oregon. Forest Science, 56(2):139-155.

Gregoire, T.; Valentine, H. 2007. Sampling strategies for natural resources and the environment. Applied Environmental Statistics, Volume 3. Chapman & Hall/CRC, London, UK. 474p.

Ganey, J.L.; Cassidy, R.H.; Block, W.M. 2008. Estimating canopy cover in forest stands used by Mexican spotted owls: Do stand-exam routines provide estimates copmarte to field-based techniques? Res. Paper RMRS-RP-72WWW, US Forest Service, Rockey Mountain Research Station, Fort Collins, Colorado. 8p.

Jennings, S.B.; Brown, N.D.; Sheil, D. 1999. Assessing forest attributes and under-storey illumination: Canopy closure, canopy cover, and other measures. Forestry, 72(1):59-74.

Korhonen, L.; Heikkinen, J. 2009. Automated analysis of in situ canopy images for the estimation of forest canopy cover. Forest Science, 55(4):323-334.

Korpela, I. 2004. Individual tree measurements by means of digital aerial photogrammetry. Silva Fennica Monographs, 3. 93 p.

Rautiainen, M.; Stenberg, P.; Nilson, T. 2005. Estimating canopy cover in Scots pine stands. Silva Fennica, 39(1):137-142.

Stancioiu, P.T.; O'Hara, K.L. 2006. Regeneration growth in different light environments of mixed species, multiaged, mountainous forests of Romania. European Journal of Forest Research, 125: 151–162, DOI 10.1007/s10342-005-0069-3.

Zou, C.B.; Barron-Gafford, G.A.; Breshears, D.D. 2007. Effects of topography and woody plant canopy cover on near-ground solar radiation: Relevant energy inputs for ecohydrology and hydropedology. Geophysical Research Letters, L24S21, doi:10.1029/2007/2007GL031484.

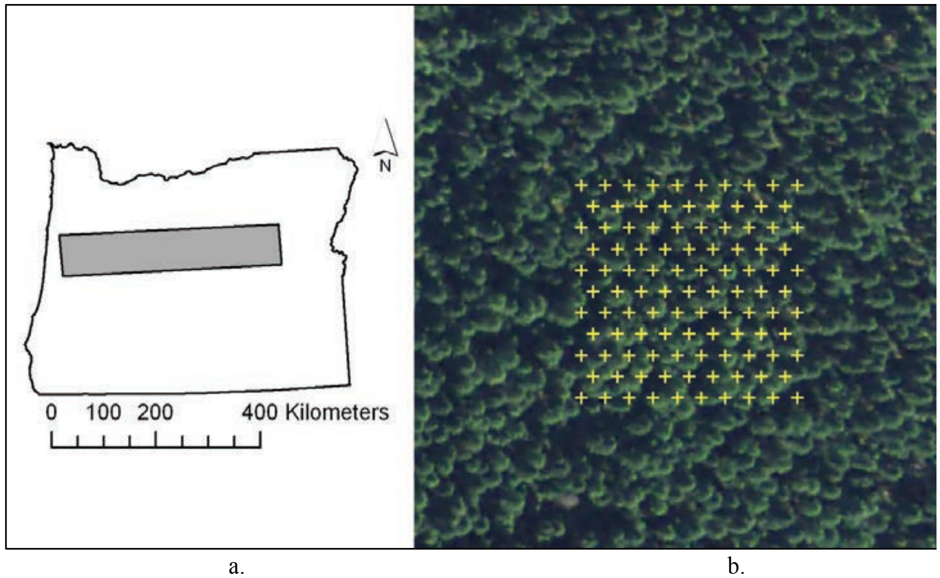

Figure 1—a. Study area (shaded rectangle) and State of Oregon boundary, b. 90 x 90 m sampling point pattern on NAIP panchromatic imagery for a randomly selected location.

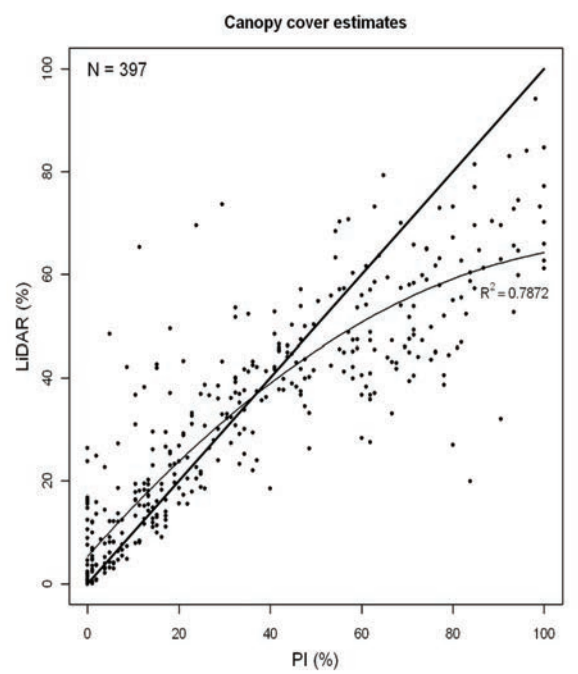

Figure 2—Scatterplot of LiDAR-vs. photointerpretation-based canopy cover estimates with 1:1 (thick) line and second-order regression fit (thin line).