Precise FIA Plot Registration Using Field and Dense LIDAR Data

Demetrios Gatziolis[1]

Abstract.—Precise registration of forest inventory and analysis (FIA) plots is a prerequisite for an effective fusion of field data with ancillary spatial information, which is an approach commonly employed in the mapping of various forest parameters. Although the adoption of Global Positioning System technology has improved the precision of plot coordinates obtained during field visits, in many circumstances, the coordinate uncertainty of plot coordinates remains substantial. Because Light Detection and Ranging (LIDAR), an emerging remote sensing technology, in its discrete return form delivers three-dimensional data with centimeter precision, FIA field data registered to LIDAR data would yield exceedingly precise plot coordinates. In this study, a fully automated, three-dimensional variant of the world-view method is used to match individual trees, identified by processing dense (> 4 returns/square meter) LIDAR data, to trees tallied in 45 FIA-like plots, thus retrieving the coordinates of the plot center. Results indicate that this method yields precise plot registration in stands exhibiting local heterogeneity in forest structure, performs better in coniferous rather than deciduous stands, and is robust against the tolerances embedded in tree location and height measurements obtained in the field.

Introduction

Meaningful integration of forest inventory and analysis (FIA) plot information obtained in field visits with ancillary, spatially distributed information organized in Geographic Information System layers or imagery acquired via remote sensing requires precise plot registration. Many popular modeling approaches that use FIA information to map forest inventory parameters across the landscape (McRoberts 2006, Ohmann and Gregory 2002, Schroeder *et al.* 2006) operate on the implicit assumption that the coordinates (i.e., registration) of FIA plots are known with precision that exceeds (is smaller than) half the size of the plot or the resolution of the coarser ancillary data source (Heuvelink 1998). Using information from plots with registration that does not meet this precision standard likely reduces the fidelity of the mapping product by a magnitude that is a function of the registration error and the local spatial continuity of the mapped parameter (Goodchild *et al.* 1993).

Historically, FIA plot registration information was collected and maintained to primarily facilitate plot visitation. Plot registration was performed in the office by overlaying analog aerial photographs with maps, usually topoquads or digital raster graphics, on which the location of the plot was pin-marked. This process was laborious and prone to registration errors, especially in mountainous terrain. Attempts to substitute reference maps with satellite or high-resolution airborne multispectral imagery resulted in little or no gain in precision or efficiency; the resolution of satellite imagery was typically too coarse to allow fine tuning of the aerial photographs on the imagery and high-resolution airborne imagery was often poorly registered. The introduction and continuous upgrade of Global Positioning System (GPS) technology, used by FIA field crews since the mid-1990s, have improved plot coordinate precision but, in many instances, registration error has remained substantial, often exceeding the spatial extent of an FIA plot.

Unlike most maps and satellite or airborne imagery, data acquired by using Light Detection and Ranging (LIDAR), an emerging remote sensing technology, has no registration issues. In its discrete form, LIDAR data consist of a collection of points arranged in three dimensions. The points, or 'returns' as

[1] Research Forester, U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station, Forest Inventory and Analysis, Portland Forestry Sciences Lab, 620 SW Main Street, Suite 400, Portland, OR 97205. Phone: 503–808–2038. E-mail: dgatziolis@fs.fed.us.

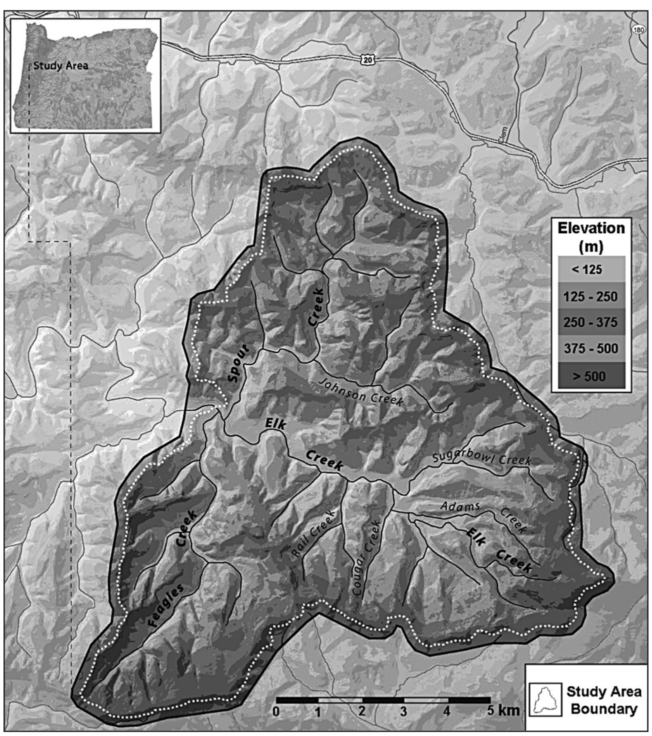
they are some times referred to, represent the locus of where narrow beams of infrared light, emitted from the airborne platform, are intercepted by the target that the beams illuminate. Because the three-dimensional location of each LIDAR point is recorded with centimeter precision, FIA field data registered to LIDAR data would yield exceedingly precise plot coordinates. This article introduces a novel, fully automated method, which uses tree data collected in a typical FIA plot and dense LIDAR data to retrieve the plot center coordinates. The method is evaluated using data from 45 FIA-like plots, established in a 9,500-hectare study area, located at the coastal forests of Western Oregon.

Methods

Study Area

The study area extends over about 9,500 hectares of the upper part of the Big Elk watershed located in Lincoln County, OR (fig. 1). More than 90 percent of the area is forestland, 47 percent of which is privately owned and under very intensive, timber-oriented management. The State of Oregon owns 1,550 hectares of land and 3,850 hectares are part of the Siuslaw National Forest where management has been limited to occasional precommercial thinning (very little of which occurred after 1984). Elevation in the study area ranges from 66 m to 1,123 m above sea level, and terrain is characterized by steep slopes that often exceed 100 percent. The forestland contains Douglas fir, bigleaf maple, and red alder, with the hardwoods dominating the buffer zones around the drainage network.

Figure 1.—*Map of the study area located at the coastal range of the State of Oregon.*



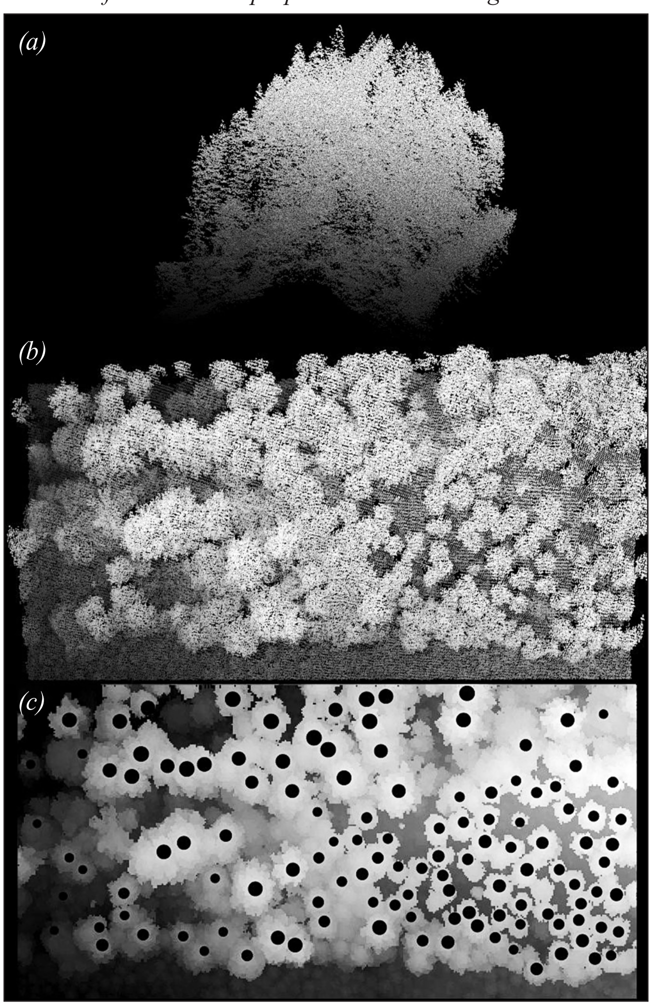
Vegetation and LIDAR Data

Forty-five FIA–like plots were established in the study area during the summer of 2005, stratified in classes of cover type (conifers, hardwoods, and mixed), tree size, and stand density. To ensure that in each plot an adequate number of trees would be present, the radius of the plot was expanded from the standard 7.32 m to 15 m. Plot centers were georeferenced by using a real time kinematic (RTK) GPS instrument. Compared to local benchmarks, the RTK instrument yielded a two-dimensional root mean square error of 2.6 cm. All trees with diameter at breast height (d.b.h.) exceeding 12.7 cm or of dominant or codominant status regardless of d.b.h. were tallied in each plot. Sketch maps depicting the presence, type, and height of understory vegetation (if any) were produced. LIDAR data over the study area was acquired at leaf-on conditions in July 2005 using an aircraft-mounted Optech 3100 system from an average height of 1,000 above ground along flight lines with 50 percent sidelap. The LIDAR instrument, at 71 kHz pulse rate, captured a 20^{o} scan width (10^{o} from nadir), averaged 9.81 points per square meter, and had spot spacing of 32 cm and laser footprint diameter of 33 cm.

Data Transcription to Point Patterns

Matching plot and LIDAR data required conversion to a common form. This conversion was accomplished by computing or identifying the tops of individual trees in the plot and on the LIDAR data sets. Using the distance and azimuth from the plot center and the tree height recorded by the field crew, the tree-dimensional coordinates of tops for dominant or codominant trees were computed. The origin of the coordinate system coincided with the plot center. It was assumed that all treetops were positioned vertically above corresponding tree bases, unless, as stated in the FIA field protocol, substantial tree leaning mandated recording of the leaning angle and computing an offset. The set of points representing individual treetops computed from field data is henceforth mentioned as the "plot point pattern." Identifying individual treetops in the LIDAR data set was a far more complex exercise. It involved processing the LIDAR data to generate ground and vegetation raster surfaces of 1 m resolution and then querying the vegetation surface to detect local maxima believed to correspond to the tops of individual tree crowns (fig. 2). The height of each identified tree was computed as the distance between the vegetation and ground surfaces at the two-dimensional location of the treetop. Gatziolis (2006) details the algorithms employed in the processing of the LIDAR data. The set of points representing individual treetops in real-world coordinates computed from the LIDAR data for a square area of approximately two hectares containing the plot is referred to as the 'stand point pattern' (fig. 3).

Figure 2.—*(a) Perspective and (b) nadir view of raw LIDAR data for one of the stands hosting a plot. (c) Raster representing the elevation of the vegetation surface derived from the LIDAR data. Darker/lighter tones correspond to lower or higher elevation. The dark circles indicate individual treetops identified by processing the vegetation surface raster. The size of the circles is proportional to tree height.*



Point Pattern Matching

Initial attempts to match plot and stand point patterns using Euclidean geometry, heuristic decision rules, and brute computing force proved inadequate and inefficient. It was realized that a pattern matching method should perform adequately in stands with few trees as well as stands with numerous trees and be robust against errors in point locations introduced during the data transcription stage. Common sources of such errors include the various tolerances in the measurement of tree location and height adopted by FIA to promote data collection efficiency, omission or commission of trees identified using LIDAR data, uncertainty in LIDAR-derived tree heights, and the ever-present discrepancies in two dimensions between treetop and base. After some additional experimentation, a modified version of the 'world-view' method (Murtagh 1992) was adopted (fig. 4).

The original world-view method retrieves the coordinates for a group of stars on an image acquired with a telescope by using reference star maps. After converting all the stars on the image to points, the method selects a star located near the centroid of the group and counts the number of stars present in each one-degree sector radiating outwards, or viewed, from the

Figure 3.—*Three-dimensional arrangement of (a) a stand point pattern derived from LIDAR data and (b) a plot point pattern derived from field data. Both patterns depict tops of individual trees. The vertical lines are included to enhance visualization. Scale is variable between and within each point pattern.*

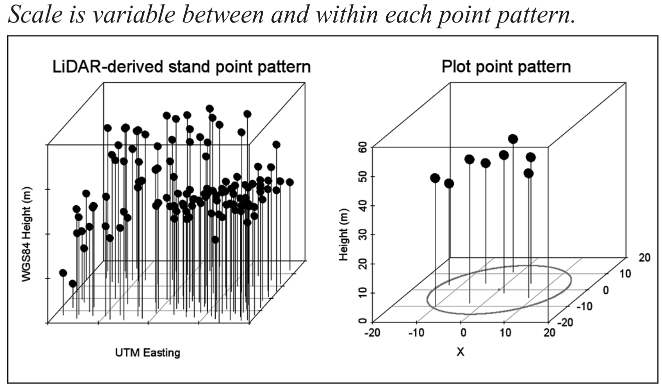
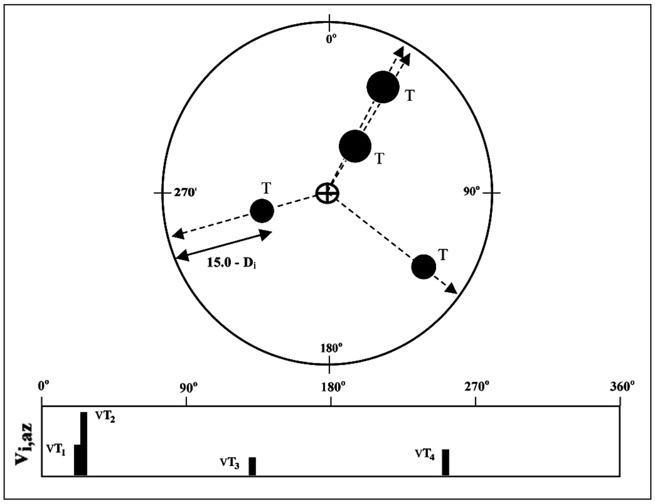


Figure 4.—*Graphical illustration of the modified world-view method for a plot point pattern. The top part features a nadir view of the plot with solid circles representing the location of four (T1 to T4) treetops. Circle size is proportional to tree height. Dash lines point at the azimuth at which a tree is viewed from the plot center. The lower part shows the values of the view vector contributed by each tree in the form of bars.*



central star. The 360-element long (view) vector of star counts is compared with the view vectors computed for all stars in the reference maps present in the general vicinity of the sky the telescope was pointing at the moment the image was acquired. The similarity of view vectors computed for the central star, once by using the image and once by using the reference maps, allows the central star to be identified in the reference maps and thus registers the star group.

Substituting the points representing stars with the points representing treetops provides a functional implementation of the world-view method. Unfortunately, the original form of the method exhibits four notable weaknesses when used in matching plot and stand point patterns. First, because it operates strictly in two dimensions, it ignores valuable tree height information. Second, it assumes that at least one tree is near the plot center. Where this is not the case, the tree labeled as ‘central’ would be at some distance from the plot center, and it would produce an asymmetrical plot view with azimuth-specific scope, longer for view sectors across the center of the plot and shorter in the opposite direction. Third, only the azimuth of a tree relative to the central tree affects the view vector; its distance from the central tree does not. Lastly, no provision exists on how errors of commission or omission should be handled. Apparently such errors do not exist in star reference maps, but are common where individual trees are identified by processing remotely sensed data. To neutralize these weaknesses, the original method was modified as follows: the plot view vector is now computed from the plot center instead of from the location of the tree closest to the plot center; stand views are computed for each vertex of a lattice of user-specified resolution covering the two dimensional extent of the stand; and instead of simply counting the tree(s) in a view sector, each tree in the plot contributes a value *v* to the view sector that corresponds to the tree’s azimuth computed as

$$v_{i,az} = P_R - D_i + \left| \frac{(H_i - \overline{H})}{\overline{H}} \right| * \mathrm{var}(H) \qquad (1)$$

where *az* represents the azimuth of tree *i* in reference to the plot center (view origin), H_i is the tree height, P_R is the plot radius, D_i is the distance of the tree from the plot center, *var(H)* is the variance, and $\overline{H}$ is the mean height of the trees in the plot. Hence, the closer a tree is to the plot center and the more its height differs from the mean height in the plot, the larger its influence on the view vector. The *var(H)* coefficient is, in essence, a scale adjustment between the distance and height components. The value of the view vector ***V*** at *az* is computed as the sum of $v_{i,az}$ for all trees (if any) present at the *az* sector of the view, or

$$V^T = \sum v_{i,1}, \sum v_{i,2}, \ldots, \sum v_{i,360}\ , \forall i \qquad (2)$$

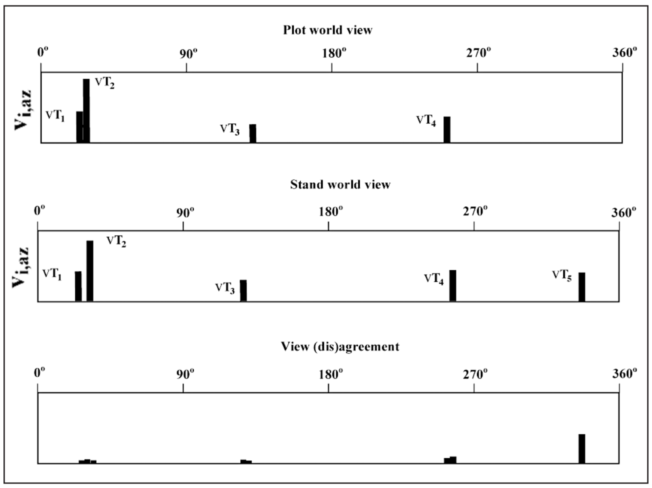
The plot world-view vector is subsequently compared to the stand-view vectors computed at each vertex of the lattice. The comparison entails computing the sum of absolute discrepancies between the azimuth-paired elements of the stand and plot vectors and producing a statistic

$$\Phi = \sum_{az=1}^{360} \left| PLOT \sum v_{i,az} - STAND \sum v_{i,az} \right|, \forall i \qquad (3)$$

that quantifies the agreement between the plot world view and each stand-view instance. The lattice vertex at which Φ is minimized is expected to represent the most likely location of the plot center within the stand. To ensure that the occasional imprecision in tree azimuths recorded in the field does not inflate Φ, the value of each one-degree plot view sector PLOT$v_{i,az}$ is subtracted for each of the five one-degree stand vector values

STAND$v_{i,[az-2,az+2]}$ centered on *az,* and the smallest absolute value is used in the computation of Φ (fig. 5). The performance of the method, in terms of its ability to retrieve the plot center, is evaluated by calculating the two-dimensional Euclidean distance (retrieval precision) between the stand location with the smallest Φ value and the plot center location recorded with the RTK instrument.

Figure 5.—*Graphical illustration of the view vectors computed for the plot (top) and the stand (middle) when viewed from the plot center. The collective height of the bars on the lower portion of the figure represents the amount of disagreement between the views.*



Results and Discussion

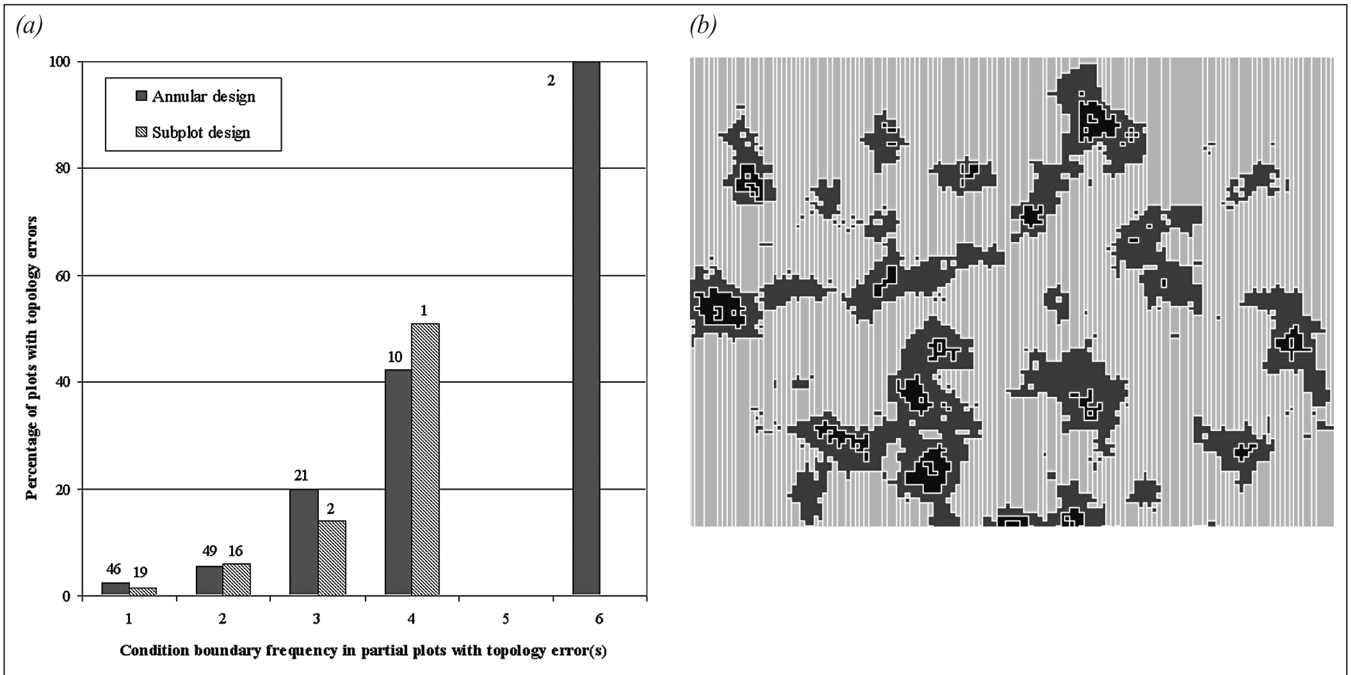
The modified world-view method applied over a stand containing a plot, ideally, will produce one group of distinctly small Φ values clustered spatially around a single stand-wide Φ minimum that corresponds to the plot center. Pending primarily on stand and plot conditions, the method is found to occasionally yield more than one clusters of local Φ minima. Rarely, it produces a nearly uniform distribution of Φ values with numerous weak local Φ minima. The latter case indicates method failure. In the study area, the modified world-view method failed for 4 of the 45 plots. For eight other plots, the retrieval precision exceeded the plot radius (table 1). For the remaining 33 plots, 73 percent of the sample total, retrieval precision was better (smaller) than the plot radius and, hence, superior to the expected registration precision currently available for FIA plots.

Plot center retrieval precision depends primarily on cover type and stand structure. Precision was higher for coniferous than deciduous stands and higher in the presence of local heterogeneity in tree height and spatial arrangement. The retrieval precision median for the 20 coniferous plots was 2.7 m. Figure 6a shows, in raster form, the spatial arrangement of Φ values computed over a coniferous stand. The portion of the raster in warmer tones depicts a single cluster of cells with small Φ values that envelope the plot center. Although a few other clusters of local minima are present, their values are too large to be considered candidate locations of the plot's center. In deciduous stands, the retrieval precision median was 6.1 m. Although the arrangement and shape of clusters in rasters of Φ values computed over stands is similar for all cover types, the value variability among local Φ minima for deciduous stands is typically lower than in coniferous stands. The lack of a pronounced Φ minimum hinders the process of determining which local minimum is the true location of the plot center among the many present, and often leads to sizeable retrieval errors (fig. 6b). Homogeneous stand conditions, regardless of cover type, cause either retrieval failure or large errors, as it is the case in plots installed in young, unthinned plantations of Douglas fir or high density red alder stands. Thinned Douglas fir plantations, even those having the operation performed on a grid or planting line, have adequate tree height variability to allow precise retrieval, a benefit attributed to the inclusion of the *var(H)* component in equation (1); without the scale adjustment between tree height and distance from the plot center provided by *var(H)*, the location of the plot center in plantations cannot be retrieved.

Table 1.—*Frequency of plots in classes of registration precision and cover type. Registration precision is the distance between the actual plot center and the location identified as plot center by the modified world-view method.*

		Cover type			
		Conifers	Mixed	Deciduous	Total
Registration precision	< 1 m	4	1	—	5
	1–5 m	12	4	2	18
	5–10 m	1	3	4	8
	10–15 m	—	1	1	2
	> 15 m	3	2	3	8
	NA	—	—	4	4
	Total	20	11	14	45

Figure 6.—*One-meter resolution raster of plot center retrieval precision Φ computed by using the modified world view method over (a) coniferous stand and (b) deciduous stand, each hosting a plot. Squares represent the true plot location and circles the location identified by the world-view method.*



A few other factors are collectively responsible for the discrepancy in retrieval precision between coniferous and deciduous plots. Dominant red alder and bigleaf maple trees growing near streams and creeks form nearly continuous canopies with few openings and have large, umbrella-shaped crowns composed of several branches or groups of branches each. In such conditions, even the most sophisticated algorithms that identify individual trees by processing imagery or LIDAR data are known to perform suboptimally and produce many tree omission or commission errors. Unlike deciduous species, the pyramid-shaped crowns of conifers produce comparatively fewer tree identification artifacts. Because the horizontal discrepancy between the location of the center of a tree stump, recorded by FIA, and its crown top, detected by processing LIDAR data, increases monotonically with crown size, the two dimensional distance between treetop and base for most red alder and maple trees exceeds that for Douglas fir. Hence, the plot point pattern extracted from LIDAR data approximates the plot point pattern computed from field data better for conifers than for deciduous species. Similar trends characterize the representation fidelity of the stand point pattern. Ultimately, the amount of noise present in the representation fidelity of the plot and stand point patterns determines whether the plot location can be retrieved and with what precision. Although a portion of the noise present in a plot point pattern originates in field measurement errors, its impact on retrieval precision was found to be minimal. Alternative world views, 20 for each plot, computed using intentionally introduced errors in tree height, stump distance, and azimuth, of magnitude comparable to that typically encountered in FIA data collection, produced results similar to those obtained by using the actual field measurements. This indicates that the accuracy and precision of field measurements is not a factor in the success or failure of the world-view method nor its plot retrieval precision.

Because the expected number of trees in a plot, and therefore the number of points in a plot pattern, is related linearly to the plot area, progressively smaller plot radii would produce more and more sparsely populated view vectors. Φ values computed from only a few non-zero Σv_i are affected by the stochasticity in the positional relationship between tree base and treetop much more than those computed in the presence of numerous non-zero Σv_i's. Smaller plot radii cause a reduction in the specificity of a Φ value to a particular location within the stand. The current FIA plot design comprises four circular subplots of radius nearly half the one used in the study area plots. Although the total area between a standard FIA and a study plot is approximately the same, applying the method to each of the four subplots and identifying the plot center as the spatial average, for example the centroid, of the individually retrieved subplot centers will produce poor plot registration and a substantial increase in the number of method failures. A functional alternative merges all subplot views into a single V^T vector. Computing the merged plot view vector presents no difficulties; calculating the merged stand-view vector for each lattice vertex, however, is a computational challenge because it requires identifying the subset of the stand point pattern located within the four subplots centered on the lattice vertex processed at the time. On-going experimentation with data from a limited number of FIA plots present in the study area indicates that the alternative approach produces retrieval precisions equivalent to those mentioned in this study but with a five-fold increase in computation time. Spatial indexing of the point patterns is expected to reduce the computational intensity of the alternative method.

Conclusion

This study has demonstrated that the majority of plots established in stands scanned with LIDAR technology can be registered via the modified world-view method. The application of the method is expeditious and requires virtually no user input, and therefore no expertise in handling LIDAR data, or advanced computer resources. The occasional method failures or imprecise results occur strictly in stands that are exceedingly homogeneous. The strong spatial continuity and low stand attribute variability that characterizes such stands, however, ensures that information from plots registered poorly by the modified world view method likely will not degrade the output of models that overlay plot and other ancillary information. Registration of actual FIA plots is of fidelity equal to the one obtained by processing continuous, larger plots commonly used elsewhere, only it is computationally more intensive.

Literature Cited

Gatziolis, D. 2006. Evaluation of tree height measurement accuracy using FIA and LIDAR data. Proceedings, Society of American Foresters 2006 national convention. Bethesda, MD: Society of American Foresters. 19 p. [on CD-ROM].

Goodchild, M.F.; Parks, B.O.; Staeyert, L.T. eds. 1993. Environmental modeling with GIS. New York: Oxford University Press.

Heuvelink, G. 1998. Error propagation in environmental monitoring. London: Taylor and Francis. 127 p.

McRoberts, R.E. 2006. A model-based approach to estimating forest area. Remote Sensing of Environment. 103(1): 56-66.

Murtagh, F. 1992. A new approach to point pattern matching. Publications of the Astronomical Society of the Pacific. 104: 301-307.

Ohmann, J.L.; Gregory, M.J. 2002. Predictive mapping of forest composition and structure with direct gradient analysis and nearest neighbor imputation in coastal Oregon. Canadian Journal of Forest Research. 32: 725-741.

Schroeder, T.A.; Cohen, W.B.; Song, C.; Canty, M.J.; Zhiqiang, Y. 2006. Radiometric calibration of LANDSAT data for characterization of early successional forest patterns in Western Oregon. Remote Sensing of Environment. 103(1): 16-26.