Combining FIA Plot Data with Topographic Variables: Are Precise Locations Needed?

Stephen P. Prisley[1], Huei-Jin Wang[2], Philip J. Radtke[3], and John Coulston[4]

Abstract: *Plot data from the USFS FIA program could be combined with terrain variables to attempt to explain how terrain characteristics influence forest growth, species composition, productivity, fire behavior, wildlife habitat, and other phenomena. While some types of analyses using FIA data have been shown to be insensitive to precision of plot locations, it has been suggested that terrain-based models may require the use of precise plot coordinates. This study compares results obtained from a variety of terrain-based analyses conducted in the Blue Ridge of North Carolina using both precise and perturbed (fuzzed and swapped) FIA plot locations, and documents differences between field-estimated slope and aspect and GIS-derived slope and aspect. Digital elevation model (DEM) data were used to derive simple topographic parameters such as elevation, slope percent, azimuth of aspect, terrain curvature, flow accumulation, slope position, and compound topographic index. These values were then compared in a pairwise fashion for plots using precise and perturbed coordinates. Correlations between precise and perturbed plot locations ranged from r = -0.006 to r = 0.383, except for precise versus perturbed plot elevations where r = 0.929. Second, a simple, terrain-based forest site quality index (FSQI) was calculated for the each plot. This index defines site quality classes for forest productivity based on azimuth of aspect, slope percent, and slope position. FSQI classifications were compared for precise and perturbed plot coordinates; at best only 40% of plots resulted in the same productivity class (out of 5). Finally, field-obtained estimates of slope and aspect were compared with GIS-derived estimates from precisely-located plots to assess their level of agreement. Correlations between field-measured and GIS-derived values were r = 0.6 for slope and r = 0.4 for aspect. Results of these experiments indicate that perturbed plot locations may not be suitable for such fine-scale applications.*

Keywords: terrain, slope, aspect, elevation, site quality, productivity

Introduction

Forest scientists have known for decades that strong linkages exist between forest productivity and site conditions such as topography. Many efforts have attempted to quantify these linkages so as to be better able to model and predict

---

[1] Associate Professor, Dept. of Forestry, Virginia Tech, Blacksburg, VA 24060; prisley@vt.edu
[2] Graduate Research Assistant, Dept. of Forestry, Virginia Tech, Blacksburg, VA 24060; wanghj@vt.edu
[3] Associate Professor, Dept. of Forestry, Virginia Tech, Blacksburg, VA 24060; pradtke@vt.edu
[4] Head, Methods and Techniques Research, USDA Forest Service, Southern Research Station FIA, Knoxville, TN 37919. jcoulston@fs.fed.us

In: McWilliams, Will; Moisen, Gretchen; Czaplewski, Ray, comps. 2009. 2008 Forest Inventory and Analysis (FIA) Symposium; October 21-23, 2008: Park City, UT. Proc. RMRS-P-56CD. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. 1 CD.

forest site productivity for assessment and management (Davis and Goetz 1990, Bolstad et al. 1998, Franklin 1995). The availability of advanced spatial analysis software and consistent, reliable, national coverage of digital elevation model (DEM) data has enhanced our ability to characterize and quantify topographic conditions at locations where productivity estimates are available. It is natural, then, to look to FIA data as source for consistent estimates of forest productivity across a large geographic area.

A current research effort at Virginia Tech is attempting to evaluate and augment the southern variant of the Forest Vegetation Simulator (FVS) for southeastern mixed forests. As part of this project, productivity data from FIA plots are being related to topographic conditions which are thought to be drivers of forest growth. However, comparison of productivity measures from FIA plots and topographic conditions at the plot locations must be conducted at a relatively fine spatial scale, on the order of 10 to 30 meters, the resolution of the most widely-used DEM data from US Geological Survey.

For a variety of compelling reasons, publicly available FIA data do not report actual plot coordinates. A mechanism called “fuzzing” adds a random error (up to about 1.6 km) to the plot location, and a subset of plots are “swapped” with other plot locations (LaPoint 2005; Guldin et al. 2006). This process of fuzzing and swapping are referred to here collectively as perturbing (McRoberts et al. 2005). Several authors have investigated the reliability of results obtained from perturbed plot locations relative to actual (hereinafter referred to as “precise”) plot locations.

McRoberts and others (2005) discussed the impacts of plot location perturbation on model-based and design-based estimation procedures. They note the effects of perturbed plot locations decrease as the size of the sampling unit increases and the spatial autocorrelation increases. The authors propose a variety of ways in which the FIA program may help avoid modeling problems with perturbed plot locations, such as providing a variety of model-based maps of estimates of forest attributes that users could access via the Internet.

Coulston and others (2006a) evaluated biomass estimates derived from kriging and residual kriging at FIA plots in Minnesota. They noted no difference between kriged estimates of biomass from the perturbed and precise plot locations. It should be noted, however, that the only variable obtained from a disparate spatial dataset was leaf area index (LAI), which came from 1-km MODIS imagery. Therefore, it is unlikely that the independent LAI variable showed much difference between perturbed and unperturbed locations.

In another sample application, Guldin and others (2006) computed inventory parameters such as forest area, numbers of live and growing stock trees, and volume from circular woodsheds of varying radii, using perturbed and precise plot locations. The only variation in the analysis, therefore, would be which plots fell into or out of the compact circular woodsheds because of location perturbation.

The authors reported that differences were trivial. A biomass prediction case study involving additional spatial layers indicated that the model developed from perturbed coordinates was no worse than the low performance from the model from precise coordinates ($R^2 = 0.43$). The authors concluded that perturbed FIA data can be used with confidence for similar applications.

In a subsequent article, Coulston and others (2006b) used simulation to develop spatial layers of different resolutions and levels of spatial autocorrelation. These layers were then used in kriging and linear regression models in which the dependent variables came from FIA plots, and models were compared between precise and perturbed locations. For kriging, no differences were noted. For regression, the authors noted that perturbed locations affected model R^2, and that the affects were most evident at finer resolutions and in datasets with lower levels of spatial autocorrelation. Furthermore, differences were most pronounced in models that had higher initial R^2. The authors suggest that regression modeling is only appropriate with very coarse resolution datasets (1-2 km).

The research underway at Virginia Tech differs substantially from most of the applications reported above. In our efforts, forest productivity estimates from FIA data are being linked to terrain characteristics which may change dramatically over distances that are very small relative to the scale at which FIA plots are perturbed. Our situation is most similar to the fine-scale, low autocorrelation linear regression scenarios reported by Coulston and others (2006b) in which perturbed plot locations resulted in substantially poorer model performance. Therefore, we hypothesize that the terrain characteristics extracted from GIS layers at perturbed plot locations will differ substantially from the conditions present at precise plot locations. Such differences would likely prevent adequate prediction of productivity from terrain conditions.

In order to conduct this research, the Virginia Tech Department of Forestry entered into a Memorandum of Understanding with the US Forest Service as part of the Privacy Policy Study Group in 2005. This agreement provided limited access to precise plot coordinates in a closely regulated setting. As part of the agreement, Virginia Tech is required to report differences that would be obtained from perturbed versus precise plot locations. This paper reports our evaluation procedure and documents the results.

Methods

Study Area

This study was conducted using data from the mountain FIA unit in western North Carolina (Figure 1). This area comprises 21 counties covering approximately 17,870 square kilometers. Most of the study area is in the Blue Ridge province of the southern Appalachian Mountains. Elevations in the study area range from 266 to 2033 m.

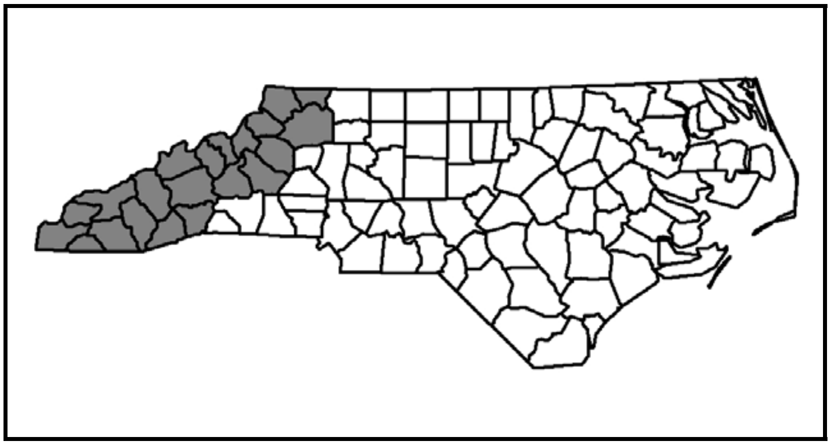


Figure 1: Map of the study area, the mountain FIA unit in western North Carolina.

Terrain Data

Digital elevation data were obtained from the US Geological Survey "seamless" web site (http://seamless.usgs.gov) and consisted of 10m grid cells containing elevation in meters. From this dataset, additional terrain datasets were developed using functions in ArcGIS 9.2 software. These additional layers included slope percent, azimuth of aspect, terrain shape index, terrain curvature, and flow accumulation. Slope percent was computed with the ArcGIS slope tool, which applies the Horn (1981) algorithm using elevations at eight adjacent cells. Azimuth of aspect and terrain curvature were calculated using the default ArcGIS algorithms. Because of the difficulties in dealing with circular data, azimuth of aspect was also transformed using a cosine transformation. Terrain shape index (TSI), an indicator of local landform convexity or concavity (McNab 1989), was computed as the difference between the elevation at a cell and the average elevation of a 35m circular neighborhood. High positive values represent areas of convexity such as ridges, and more negative values represent concave landforms such as coves or drainages. Flow accumulation is a hydrological-based indicator of landscape position. The flow accumulation algorithm computes the number of grid cells whose runoff would flow eventually through a given cell. Low numbers represent areas near ridge crests and high numbers represent valleys, bottoms, and drainageways.

Two additional terrain descriptors were also derived: slope position and forest site quality index (FSQI). To quantify the relative positions of points on the landscape, an indicator of slope position was required. Slope position was defined as the percentage of the flow path distance from a stream to a ridge. To obtain this metric, the flow accumulation layer was used to define streams, which were then masked from the analysis. The flow length tool in ArcGIS was then used to determine both the uphill and downhill flow length for each cell. Uphill flow length is the number of cells along the flow path uphill to a ridge. Downhill flow length is the number of cells along the flow path downhill to a stream. Slope position for a grid cell is then calculated as the downhill flow length at that cell divided by the sum of uphill and downhill flow lengths. The resulting value for slope position (0 -1.0) is then reclassified into codes for six landscape positions: summit, shoulder, backslope, footslope, terrace, and floodplain (Cotton et al. 2008).

Forest site quality index is an ordinal value integrating slope, aspect and the slope position class defined above (Meiners et al. 1984). The FSQI values have been shown to be correlated with upland oak site index, and are being applied in ongoing forest productivity research. FSQI scores were developed using the slope, aspect, and slope position class layers developed for western North Carolina. Scores were categorized into five classes corresponding to site index ranges (Cotton et al. 2008).

FIA Data

FIA plot data from cycle 7 (ca. 2002) were obtained from the Southern Research Station for North Carolina. Both published (perturbed) and actual (precise) coordinates for each plot were obtained. A total of 1,022 plots were used in this analysis.

Analysis Approach

A spatial dataset of perturbed and precise plots was created from the FIA tables containing coordinates of plot locations. All data were projected to the UTM Zone 17, NAD 83 coordinate system and plots were overlaid with the raster terrain datasets, extracting cell values for each plot location from each layer. Scatterplots and correlation estimates were produced to compare terrain values for perturbed versus precise locations. Categorical FSQI site index classes for perturbed and precise plot locations were compared in a contingency table.

Results

Perturbed versus Precise Coordinates

The influence of perturbed locations on terrain-based models was examined by comparing topographic variables derived from precise versus perturbed coordinates. Despite the trivial contribution to bias in topographic GIS derivatives, perturbed locations did influence the precision of derivatives (Figure 2 and Table 1). Dispersion of differences between the two plot locations had a wide range and its corresponding standard errors were far from zero. Examples of observed discrepancies between precise and perturbed plots illustrate these differences: (1) aspect obtained from a precise coordinate was toward north with azimuth 7°, but that obtained from a perturbed coordinate was toward south with azimuth 175°; (2) a precise location has a steep slope of 99%, but its perturbed location has a relatively flat slope of 2%. Maps of locations with similar terrain discrepancies are depicted in Figure 3 (without actual plot locations). The strength of topographic correlation was weak or none (*r* = 0 – 0.383) between perturbed and exact locations, except the elevation variable (*r* = 0.929).

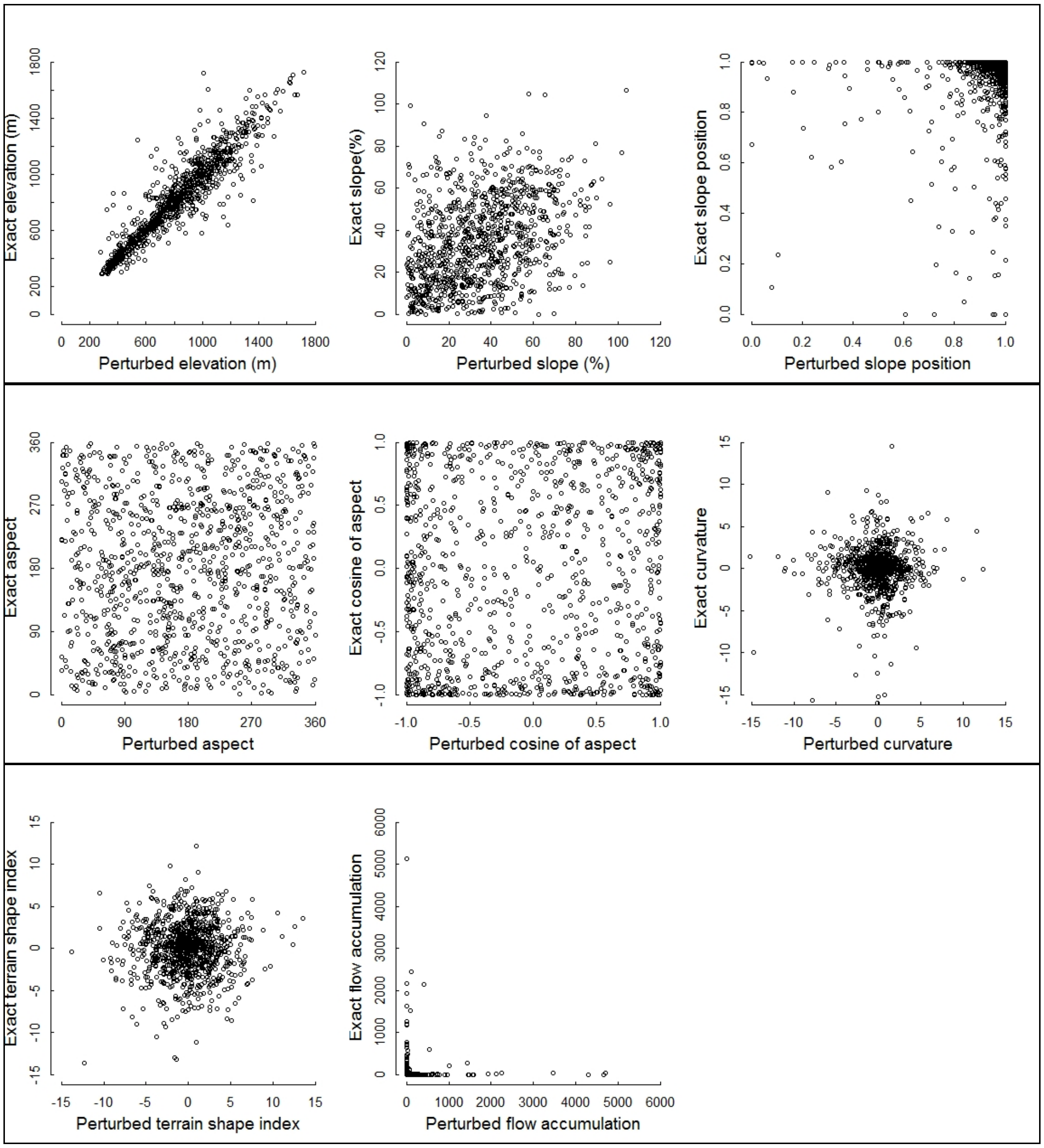


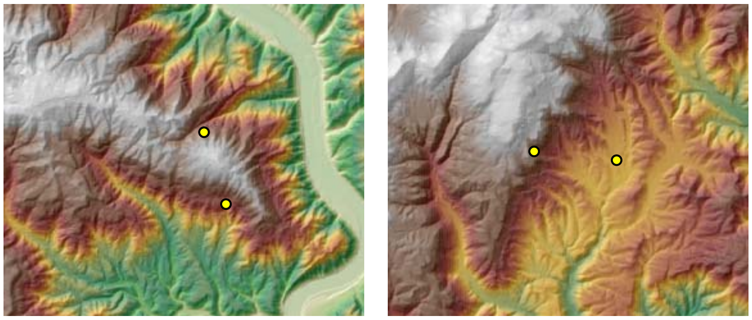
Figure 2: Scatterplots of terrain variables for perturbed and precise coordinates; incuding elevation, slope, slope position, aspect, cosine of aspect, curvature, terrain shape index, and flow accumulation (sample size = 1022).

Forest site quality index (FSQI) consists of scores from 3 to 16. Only 15.4% of plots had the same score for perturbed and precise locations, resulting in a correlation of 16.7%. When aggregated to five classes, FSQI exhibited only a 40% overall agreement (Table 2). In this agreement analysis, no evidence existed to support the assertion that perturbed coordinates could provide sufficient information in evaluating site quality (Table 2). After chance agreement was excluded (using the kappa statistic), the agreement rate was 5%.

Table 1: The performance of perturbed coordinates is evaluated by comparing to GIS-derived terrain values from precise coordinates.

	r	*Bias*	*SE* [a]	$\sqrt{MSE}$ [b]
Elevation (m)	0.929	6.06	112.9	113.06
Slope (%)	0.383	-0.18	23.14	23.14
Slope Position (%)	0.201	~ 0	0.18	0.18
Aspect (˚)	0.044	3.25	139.63	139.67
Cos(aspect)	0.093	0.03	0.95	0.95
Curvature	0.038	0.09	3.53	3.53
Terrain Shape	0.036	0.13	4.37	4.37
Flow Accumulation	-0.006	-49.07	2811.69	2812.12

[a] *SE*: Standard Error
[b] *MSE*: Mean Square Error



Aspect difference

Slope difference

Figure 3: Examples of results for aspect and slope based on perturbed and precise coordinates. Perturbed and precise plots differed in aspect by approximately 170 degrees; perturbed and precise plots differed in slope by 97%. These maps depict similar terrain conditions but do not indicate real plot locations.

Table 2: The contingency table of forest site quality index (FSQI) for perturbed versus exact coordinates: FSQI is computed based on GIS-derived aspect, slope, and slope position for each plot. 1022 plots are used to evaluate the FSQI agreement between perturbed and actual coordinates. Overall agreement of FSQI classes between perturbed and precise coordinates is 40%. The proportion of FSQI agreement is 0.05 (kappa statistic), after chance agreement is excluded.

		Precise coordinates				
	FSQI	3.0-4.5	4.5-7.5	7.5-10.5	10.5-13.5	13.5-16.0
Perturbed coordinates	3.0-4.5	1	3	1	4	1
	4.5-7.5	0	53	78	31	3
	7.5-10.5	3	84	254	166	12
	10.5-13.5	0	42	152	98	10
	13.5-16.0	0	4	10	6	2

Field-Measured versus GIS-Derived Variables

Field-obtained slope and aspect values were available for 767 FIA plots for comparing the performance between field and GIS-derived slope and aspect using precise coordinates. The linear association between field and GIS-derived variables was stronger than the correlation between perturbed and precise plot locations. For slope, $r = 0.611$ and for aspect $r = 0.405 \sim 0.546$ (Figure 4). Only 53% of plots had GIS-derived slope within 10% of the field measurement (±10% is the MQO for subplot slope). For aspect, only 21% of plots had a GIS-derived value within 10 degrees of the field measurement (±10° is the MQO for aspect).

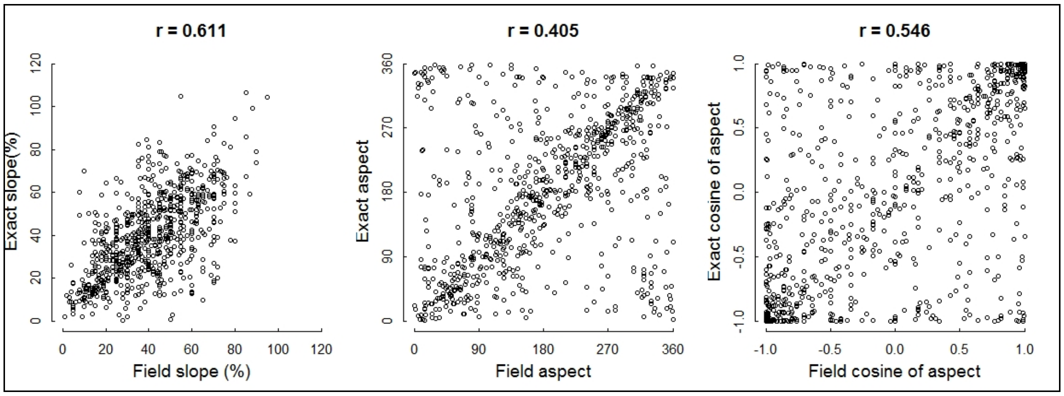


Figure 4: Linear correlations between field-derived and GIS-derived values for slope, aspect, and cosine of aspect (sample size = 767).

Discussion and Conclusions

We know that microscale variation in landform and terrain affects the type and quality of the vegetation in mountainous regions, largely due to temperature and soil moisture limiting to plant growth. Terrain information can be easily derived through GIS procedures because of widely available digital elevation models, and terrain derivatives can improve predictive models (Davis and Goetz 1990). The perturbed coordinate system, however, cannot provide useful information derived from fine-scale digital elevation data. In addition, it is unlikely that FIA can provide all possible terrain variables from field or even GIS measurements: the variety of terrain-related, GIS-derived variables currently used in ecological and hydrological applications is large and growing. Thus, it appears that FIA data in publicly-available form (perturbed locations) are not useful in conjunction with fine-scale spatial applications.

It also appears that the terrain variables that are currently measured in the field (slope and aspect) do not correlate with GIS-derived variables as strongly as we might hope. The algorithms broadly used in GIS-based terrain analysis do not measure the same things as are measured in FIA field procedures. For example, slope is measured in the field using a percent-scale clinometer, observing from the uphill to downhill edges of a subplot. The default GIS approach computes a weighted average gradient in N-S and E-W directions. Therefore, it is not appropriate to interchange field slope/aspect with GIS-derived slope/aspect in either developing or applying predictive models.

Clearly, modeling forest productivity or other phenomena that vary at spatial scales on the order of currently available digital elevation models cannot be adequately performed with the publicly-available FIA plot locations.

Researchers pursuing such endeavors must find other avenues to accomplish this research.

Opportunities for spatial modeling with FIA data with ancillary geospatial data have increased proportional to the availability of geospatial data (e.g. DEMs). These opportunities will continue to increase as more geospatial information at higher resolution becomes available. However, the usefulness of the ancillary data is related to the accuracy of the precise plot locations. Although errors in precise locations were not considered in this analysis, these errors create additional uncertainty for the types of analyses presented here. For example, the FIA program typically uses recreation-grade GPS units to collect coordinates of each inventory plot. These types of GPS units have under-canopy locational accuracies of approximately 7.5 m (Bolstad et al. 2005). We recommend that the accuracy of the precise location should be available to clients who are using these precise coordinates. This information is particularly important when conducting research to develop models based on fine resolution (e.g. 10 m) geospatial data and FIA plot attributes. To keep pace with the increased resolution of ancillary geospatial data, we also recommend that the FIA program adopt new, more accurate, GPS technology as it becomes available and affordable.

Acknowledgements

This research was supported by the USDA Forest Service, Southern Research Station under Agreement No. SRS 05-CA-11330134-251. The authors gratefully acknowledge the support of the Southern Research Station, specifically Cathryn Greenberg and David Loftis. We also acknowledge valuable assistance from Chad Keyser of the Forest Management Service Center, Bill Burkman and Sam Lambert of the SRS FIA unit, and Claudia Cotton and Spencer Riddile of the Virginia Tech Department of Forestry.

Literature Cited

Bolstad, P., Jenks, A., Berkin, J., Horne, K. 2005. A comparison of autonomous, WAAS, real-time, and post-processed global positioning system (GPS) accuracies in northern forests. Northern Journal of Applied Forestry 22(1). 5-11.

Bolstad, P.V., Swank, W., and J. Vose. 1998. Predicting southern Appalachian overstory vegetation with digital terrain data. Landscape Ecology 13:271-283.

Cotton, C.; Prisley, S.; Fox, T. 2008. Mapping upland hardwood site quality and productivity with GIS and FIA in the Blue Ridge of North Carolina. In: McWilliams, Will; Moisen, Gretchen; Czaplewski, Ray, comps. 2008. Forest Inventory and Analysis (FIA) Symposium 2008; October 21-23, 2008; Park City, UT. Proc. RMRS-P-56CD. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. 1 CD.

Coulston, J.W.; Reams, G.A.; McRoberts, R.E.; Smith, W.B. 2006a. Practical considerations when using perturbed forest inventory plot locations to develop spatial models: a case study. In: McRoberts, R.E.; Reams, G.A.; VanDeusen, P.C.; McWilliams, W.H., eds. 2006. Proceedings of the sixth annual forest inventory and analysis symposium; 2004 September 21-24; Denver, CO. Gen. Tech. Rep. WO-70. Washington, D.C.: U.S. Department of Agriculture, Forest Service. 126 p.

Coulston, J.W.; Ritters, K.H.; McRoberts, R.E.; Reams, G.A.; Smith, W.B. 2006b. True versus perturbed forest inventory plot locations for modeling: a simulation study. Canadian Journal of Forest Research 36: 801-807.

Davis, F.W., Goetz, S., 1990. Modeling vegetation pattern using digital terrain data. Landscape Ecology 4, 69-80.

Franklin, J. 1995. Predictive vegetation mapping: Geographic modelling of biospatial patterns in relation to environmental gradients. Progress in Physical Geography 19:474-499.

Guldin, R.W.; King, S.L.; Scott, C.T. 2006. Vision for the future of FIA: paean to progress, possibilities, and partners. In: McRoberts, R.E.; Reams, G.A.; VanDeusen, P.C.; McWilliams, W.H., eds. 2006. Proceedings of the sixth annual forest inventory and analysis symposium; 2004 September 21-24; Denver, CO. Gen. Tech. Rep. WO-70. Washington, D.C.: U.S. Department of Agriculture, Forest Service. 126 p.

Horn, B.K.P. 1981. Hill shading and the reflectance map. Proc. IEEE 69(1):14-47.

LaPoint, E. 2005. Access and use of FIA data through FIA spatial data services. In: McRoberts, R.E.; Reams, G.A.; VanDeusen, P.C.; McWilliams, W.H., eds. 2005. Proceedings of the fifth annual forest inventory and analysis symposium; 2003 November 18-20; New Orleans, LA. Gen. Tech. Rep. WO-69. Washington, DC: U.S. Department of Agriculture, Forest Service. 222 p.

McNab, W. 1989. Terrain shape index: quantifying the effect of minor landforms on tree height. Forest Science 35:91-104.

McRoberts, R.E.; Holden, G.R.; Nelson, M.D.; Liknes, G.C.; Moser, W.K.; Lister, A.J.; King, S.L.; LaPoint, E.B.; Coulston, J.W.; Smith, W.B.; Reams, G.A. 2005. Estimating and circumventing the effects of perturbing and swapping inventory plot locations. Journal of Forestry 103(6): 275-279.

Meiners, T.M., D.W. Smith, T.L. Sharik, and D.E. Beck. 1984. Soil and plant water stress in an Appalachian oak forest in relation to topography and stand age. Plant and Soil 80:171-179.