An Evaluation of Lidar-derived Elevation and Terrain Slope in Leaf-off Conditions

Michael E. Hodgson, John Jensen, George Raber, Jason Tullis, Bruce A. Davis, Gary Thompson, and Karen Schuckman

Abstract

The effects of land cover and surface slope on lidar-derived elevation data were examined for a watershed in the piedmont of North Carolina. Lidar data were collected over the study area in a winter (leaf-off) overflight. Survey-grade elevation points (1,225) for six different land cover classes were used as reference points. Root mean squared error (RMSE) for land cover classes ranged from 14.5 cm to 36.1 cm. Land cover with taller canopy vegetation exhibited the largest errors. The largest mean error (36.1 cm RMSE) was in the scrub-shrub cover class. Over the small slope range (0° to 10°) in this study area, there was little evidence for an increase in elevation *error with increased slopes. However, for low grass land cover, elevation errors do increase in a consistent manner with increasing slope.* Slope errors *increased with increasing surface slope, under-predicting true slope on surface slopes >2°. On average, the lidar-derived elevation under-predicted true elevation regardless of land cover category. The under-prediction was significant, and ranged up to −23.6 cm under pine land cover.*

Introduction

Large-scale digital elevation models (DEMs) are widely used in research, education, and management of public resources. Absolute elevation is required for mapping floodplains or conducting visibility studies while surface form (e.g., slope and aspect) is used for hydrologic applications. Prior to the late 1940s, topographic maps were created from field surveys and "artistic sketching" of contour lines (Hodgson and Alexander, 1990). Stereoscopic aerial photography for topographic mapping was developed in the 1940s and became the primary source for large-scale mapping (Jensen, 2000). Interferometric Synthetic Aperture Radar (IFSAR) was developed in the 1960s for medium-scale mapping and more recently, for large-scale mapping. Lidar (Light Detection and Ranging) was developed in the 1980s and is rapidly becoming the preferred method of creating DEMs by counties, states, and some governmental agencies.

While the use of lidar for mapping vegetation characteristics is becoming more frequent (Means, *et al.*, 2000), mapping terrain elevation has been the primary focus of most lidar collections. It is generally known that the leaves and branches can reduce the effectiveness of obtaining elevation information from lidar data (Hodgson *et al.*, 2003). The sensitivity of lidar data collection to system parameters (e.g., height above ground level (AGL), laser instrument) and geography (e.g., relative ruggedness, land cover, seasonality) has not been rigorously researched. Even fewer studies have attempted to assess the accuracy or sensitivity of mapping topographic *slope* with lidar-data sources. Therefore, this study focused on mapping elevation and topographic slope in the winter (leaf-off) season for a mid-latitude geographic area. The research questions examined in this study include:

1) What is the elevation accuracy of a lidar-derived DEM in *Leaf-Off* Conditions?
2) What is the slope accuracy of a lidar-derived DEM in *Leaf-Off* Conditions?
3) Was there a tendency to over- or under-predict elevation or slope?, and
4) Does the absolute elevation and/or surface form accuracy vary or covary with land cover and/or slope?

To test these hypotheses, lidar data were collected in leaf-off conditions over a watershed in the piedmont of North Carolina (Figure 1). The leaf-off lidar data were processed and lidar returns labeled using a combination of automated and manual interpretation approaches. The North Carolina Geodetic Survey collected 1,225 *x, y, z* reference data points in the form of transects within the watershed. Land cover information was collected for each of the 1,225 survey points along the transects by direct field observation. Statistical analyses were performed by comparing the reference data elevations and slopes with the lidar-derived elevation and slope data. The effects of land cover and slope were assessed using (a) signed elevation error, (b) absolute elevation error, and (c) signed slope error.

Background

Overviews of the lidar sensor system for terrain mapping may be found in Jensen (2000) and Fowler (2001) and

M.E. Hodgson and J.R. Jensen are with the Department of Geography, University of South Carolina, Columbia, SC 29208 (hodgsonm@sc.edu).

Jason Tullis is with the Department of Geosciences, University of Arkansas, Fayetteville, AR 72701 (jatullis@uark.edu).

G. Raber is with the Department of Geography, University of Southern Mississippi, Hattiesburg, MS 39406 (george.raber@usm.edu).

B.A. Davis is with the Applications Engineering Division, Earth Science Applications Directorate, NASA Stennis Space Center (bruce.a.davis@nasa.gov).

G. Thompson is with the North Carolina Geodetic Survey, Raleigh, NC 27699 (gary.thompson@ncmail.net).

K. Schuckman is with Earthdata Technologies, LLC, Frederick, MD 21703 (kschuckman@earthdata.com).

Photogrammetric Engineering & Remote Sensing
Vol. 71, No. 7, July 2005, pp. 817–823.

0099-1112/05/7107–0817/$3.00/0

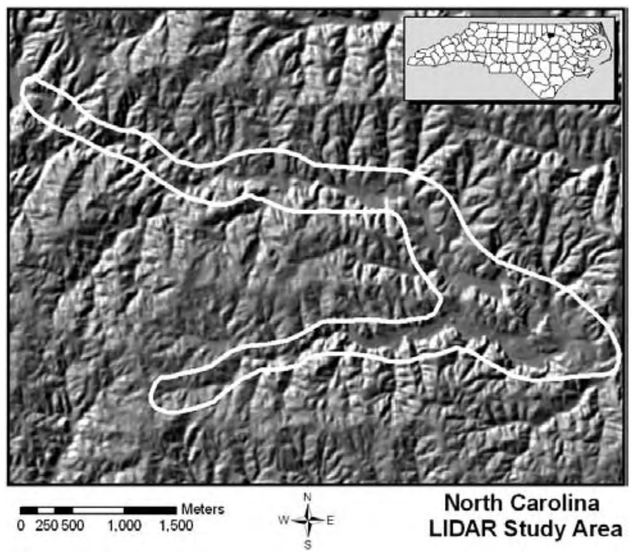

Figure 1. The study area in the piedmont of North Carolina. The underlying shaded relief model was created from USGS Level 2, 30 × 30 m DEMs.

numerous research articles. Here, only the important elements in the lidar collection and processing approach are discussed as they relate to elevation accuracy and surface form accuracy.

The accuracy of the DEM derived from lidar postings (and derived products like surface slope) is affected by several factor groups: sensor, aircraft platform, navigation, lidar point processing, and geographic environment (Table 1). Products derived from the DEM such as surface slope/aspect and drainage channels are also influenced by the same factors affecting elevation accuracy. It is possible, however, for the derived surface slope to be accurate while the elevation data themselves are less accurate; as long as the surface form represented by the lidar observations is internally consistent. For example, a consistent over-prediction in elevation for the lidar postings will not have an impact on surface slope or aspect.

The dominant factors of the lidar system that affect elevation accuracy are reviewed in the next two sections. Results of published studies documenting the terrain model accuracy are presented.

Lidar System

A lidar system for terrain mapping is typically composed of the aircraft platform, sensor, Inertial Measurement Unit (IMU), Inertial Navigation System (INS), and global positioning system (GPS) control. For most mapping applications, the lidar sensor is mounted in a fixed-wing aircraft or helicopter. The GPS and INS (or inertial navigation unit (INU) is some cases) are used to determine the position of the sensor and pointing direction of the laser. The *x, y*, and *z* position of the object the laser pulse is reflected from is determined by modeling the round-trip return time and position/pointing direction of the laser. The absolute accuracy of the GPS approach, INS, pulse length (distance between the start and stop of a single laser pulse), and the footprint size of the laser projected on the ground interact to set the technological limits of lidar data collection. Krabill *et al.*, (2002) provides an empirical approach to evaluating the Airborne Topographic Mapper (ATM) laser system. He documented the combined effects of the GPS, INS, and scan angle on elevation and horizontal accuracy. Errors in the factors can result in appreciable errors at large scan angles (e.g., 59 cm horizontal and 4 cm vertical).

A large number of laser pulses must be emitted to obtain a large number of lidar "returns." Commercial lidar systems today use lasers with a pulse rate of 10,000 to about 70,000 pulses per second (Box, personal communication, 2003). Flying the aircraft at lower altitudes and at a reduced forward speed will also result in a greater number of pulses per unit area.

To obtain a large number of returns from the actual "ground" surface, small laser pulse footprints and multiple laser returns are needed. The footprint is defined by the divergence of the laser and the flying altitude. The laser divergence (0.2 *mr* to 0.33 *mr*) is set while the flying height may be altered (Box, personal communication, 2003). Footprints common for large-scale terrain mapping are between 24 cm to 60 cm. The footprint diameter varies, however, depending on beam divergence and flying heights. From this set of lidar returns, "ground" returns are identified through automated and manual methods of "point cloud" analyses.

The emitted pulse of from 6 to 12 *n*s (approximately 2 m to 4 m in length for the typical lidar sensor) interacts with numerous features encountered in the landscape (Box, personal communication, 2003). Some of the energy reflects from airborne objects (e.g., birds, wires) while most of the energy reflects off vegetation, anthropogenic features, such as buildings, or the "ground." The reflected energy is observed by the sensor as a waveform of varying intensity. Automated methods are used to determine what wave peak is indicative of an object and thus, a "return."

The last identified return pulse may be the best estimate of the ground (often assumed to be the "bald earth"). However, even the last return may not be the ground as vegetation canopy may totally obscure the ground. In addition, the last return may be reflected from an anthropogenically created object (e.g., a building). The goal is to identify these non-ground points using a combination of automated

Table 1. Important Concepts Related to Lidar Surface Accuracy Variations (those examined in this study are in shown in **bold**)

Data Collection				
Sensor	Aircraft	Navigation	Lidar Point Processing	Geography
Pulse Length	Altitude	GPS constellation	Return Identification (e.g., waveform)	Seasonality
Pulse Rate	Forward Speed	Inertial Navigation System (e.g., roll)	Automated Labeling Algorithm	**Land Cover**
Wavelength		Multipathing	Human Classification	**Terrain Slope**
Divergence		Tropospheric Propagation	Interpolation Algorithm	
Scan Angle		GPS Reference Station(s)		

procedures and manual editing through visual and statistical processing of the lidar multiple returns.

Lidar Point Labeling for Elevation Mapping

The process typically used by lidar data vendors to classify and remove non-ground points is a combination of an *automatic algorithm* and a *human operator's manual efforts* to classify points in difficult areas. Vendors use their own unique mix of automation and manual intervention (Petzold *et al.*, 1999). Some federal agencies collecting lidar (e.g., NOAA) do not perform point-labeling or attempt to remove returns from vegetation.

While each vendor has developed their own approach to labeling points, most use a two-stage process. Since it is assumed that the final return from all returns associated with a single pulse is the only point that *could* be a ground return, only the final return from each laser pulse is entered into the candidate ground return set. Using a moving neighborhood window, "ground" points are selected based on their elevation with respect to their neighbors. The second step involves a human analyst who visualizes the points in three-dimensions and/or with ancillary data (e.g., orthophotography). Previously labeled "ground" points may be deleted or new "ground" points added. This second step is very labor intensive, but may dramatically improve the quality of the final product.

Some research efforts have specifically focused on this point labeling process. Cobby *et al.* (2001) developed an automated segmentation approach for lidar data so that the point-labeling algorithm varied by land cover category (Petzold *et al.*, 1999). Raber *et al.* (2002) found that vertical accuracy could be improved in lidar derived DEMs if the lidar data was segmented into general land cover classes prior to processing. Despite this effort and a few other research studies on the lidar intensity data, little work has been based on accurate empirical data. Published studies on lidar elevation accuracy seldom include documentation of the point labeling process and the parameter values used (most vendors regard this as proprietary.)

Lidar-derived Elevation and Slope Accuracy

There are few published rigorous studies evaluating the accuracy of airborne lidar data. Lidar companies have conducted "in-house" assessments of the *elevation accuracy*. A few states and counties have performed other rigorous assessments. In a controlled study over the Greenland ice sheet, Krabill *et al.* (2002) predicted an 8.6 cm RMSE theoretical error and documented a 5.4 cm RMSE error. The functional relationship between posting density of "ground" returns and percent canopy closure was found to be strongly linear in a study area predominately covered by pine and some mixed deciduous canopy (Cowen *et al.*, 2000). Although some previous research has established a variability in lidar accuracy by land cover, there is not a substantial body of research that documents expected accuracy estimates for specific land cover categories.

One of the largest lidar-mapping efforts in the world is taking place in the State of North Carolina for FEMA's floodplain mapping program (North Carolina, 2002). Lidar data (4.5 m nominal posting) were collected, processed, converted to a TIN, and then to a 6.1 m (20 feet) and 15.2 m (50 feet) cell size DEM. The target accuracy for their data is 20 cm (RMSE) for coastal counties (composed largely of "flat" terrain) and 25 cm for inland counties (composed largely of rolling or hilly terrain). At least 20 reference points are collected in each of five land cover categories: grass, weeds/crop, scrub, forest, and built-up. A variation on the reporting of accuracy is the "95 percent RMSE calculation" report. In this 95 percent report, the state removes 5 percent of observations in the accuracy assessment that have the highest errors. For the 41 counties studied in the first phase of their program the overall accuracy based on the 95 percent RMSE calculation was 15.15 cm.

Using similar land cover categories as the North Carolina effort, elevation accuracy for leaf-off lidar data was assessed in the piedmont of South Carolina piedmont (Hodgson and Bresnahan, 2004). This study quantified the contribution of error from the lidar system, interpolation algorithm, terrain slope, land cover, and reference data. RMSEs between 17.2 cm and 25.9 cm were found.

Hodgson *et al.* (2003) analyzed North Carolina lidar data collected during *leaf-on* conditions and compared the results to IFSAR-derived DEMs and USGS Level 1 and 2 DEMs. Elevation error for the lidar-derived elevation data varied by land cover category and ranged from 33 cm (low grass) to 153 cm (scrub/shrub). Errors in low grass and high grass were much smaller than those in the more heavily vegetated canopies except for pine forests. Elevation error was not correlated with increasing slope except for the scrub-shrub land cover. However, the slopes only ranged up to 10°. Kraus and Pfeifer (1998) found a 57 cm error in a forested environment (the Vienna woods). These authors also documented a +20 cm systematic over-prediction in elevations.

In a coastal mudslide study area, Adams and Chandler (2002) measured lidar-derived elevation accuracy as 26 cm (RMSE). The lidar-derived elevation data were less sensitive to *terrain slope* than a DEM derived from digital photogrammetry. The authors also found a "slight bias" in the lidar data i.e., a tendency to under-predict elevation. Bowen and Waltermine (2002) assessed the accuracy of lidar derived elevation data and how accuracy varied by topography. Using data collected in a western river corridor, they found an overall accuracy of 43 cm (RMSE). Reference data in the form of transects were used to identify a weakness of vegetation removal algorithms in variable terrain compared with flat terrain.

A few other studies have focused on the effects of lidar post-processing methods on elevation accuracy. Lloyd and Atkinson (2002) found that kriging with a trend model may be more accurate when the density of lidar postings decreases. In a study on lidar point-labeling research, a 17 cm (RMSE) accuracy was observed in grass and cereal crop land cover (Cobby *et al.*, 2001). Elevation accuracy from the lidar data decreased in a densely wooded environment.

Early work in topographic mapping resulted in the long-standing Koppe's formula for estimating the RMSE elevation error caused by covariation of horizontal error and slope (Maling, 1989). For a constant error in the horizontal location of an observation, the elevation error will increase as slope of the surface increases (Figure 2a). Fundamentally, Koppe's formula is based on the tangent of the surface slope (α) and observed horizontal displacement:

$$\text{Elevation Error} = \text{Tan}(\alpha) \times \text{Horizontal Displacement} \tag{1}$$

For example, a constant surface slope of 20° and a 100 cm typical horizontal displacement of lidar points may result in an elevation error of up to 36 cm.

Little published work to date has focused on the evaluation of terrain slope/orientation accuracy derived from lidar data or any remotely sensed data source. Bolstad and Stowe (1994) and Chang and Tsai (1991) previously noted that modeled slope error in USGS DEMs (from photogrammetric sources) increased with increasing slope in the actual surface.

Fundamental to the problem of mapping elevation and subsequently terrain slope/aspect is collecting enough lidar returns from the ground. Reliably identifying "ground"

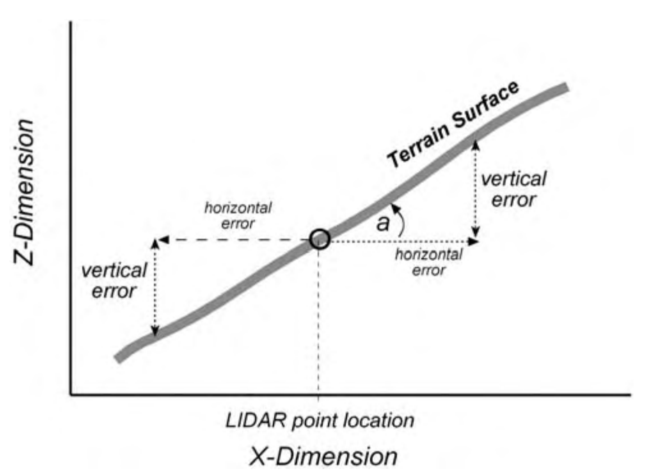

Figure 2. Relationship between elevation error, surface slope, and horizontal error.

returns is determined, in part, by the automated and manual methods used in the point-labeling process. By increasing the pulse density (i.e., the number of pulses emitted from the lidar sensor per unit space) there is a greater likelihood that more pulses will pass through openings in vegetated canopies and reach the ground. As noted earlier, the pulse density may be increased through manipulation of the aircraft forward speed, altitude, and the use of lasers with higher pulse rates. Less understood is the size of the ground projected laser "footprint" on obtaining "ground" returns. There is little empirical work to determine what the expected percentage of "ground" returns from a lidar collection would be in leaf-off versus leaf-on conditions or even in variable canopy (Cowen *et al.*, 2000). The authors are unaware of any empirical assessment of the percentage of ground returns as a function of land cover class.

Methodology

Study Area

The study area (45.65 km^2) in northeastern North Carolina is located in an area of gently rolling terrain (referred to as the *piedmont*) and includes a portion of the Swift and Red Bud Creek watersheds (Figure 1). It consists primarily of farmland, pine plantation, and hardwood forests. Numerous forested areas have been clear-cut. Forest adjacent to rivers and creeks are usually dense. The major agricultural crops are soybeans and tobacco. Elevation ranged from 44 m to 136 m above mean sea level.

Data Sources

Airborne lidar data were collected by Earthdata, Inc. as part of a floodplain mapping effort. The data were collected at 3,657 m (12,000 feet) AGL with a 4.5 m nominal post spacing. Two dates were required to collect the leaf-off lidar data: 28 January and 01 February 2001. The number of emitted lidar *pulses* (i.e., not returns) for the leaf-off data were, on average, one pulse for every 20.48 m^2. Calibration of lidar overflights was conducted using a nearby airport. Block adjustment between flight lines was not performed.

The lidar data were processed using automated methods for identifying ground returns (i.e., using Terrasolid® software) and using manual/visible interpretation methods for identifying ground returns. Specific parameter settings for Terrasolid® were not provided by the vendor. Digital orthoimagery was used as a "backdrop" to aide in the interpretation for both approaches. An average posting density of the identified "ground" returns was one point every 31.1 m^2. Of the 2,218,079 pulses only 66 percent were classified as "ground." A triangulated irregular network (TIN) retaining *all* lidar "ground" returns was created. For each reference point (i.e., the field surveyed points), a TIN-linear interpolated value was computed.

Reference Data

An initial set of *in situ* survey reference points (1,195) were collected along 23 transects and additional scattered points (i.e., pavement) throughout the study area. This initial set of 23 transects were collected for evaluating the accuracy of leaf-on lidar data in the Hodgson *et al.* (2003) study. For this new study, 30 additional points were collected over pavement. The additional points over pavement were not collected along a systematic transect as the previous collection and thus, could not be used for evaluating the slope of the surface as with the other 1,195 survey points. Real-time kinematic and static GPS surveying was used to establish the endpoints for each transect and conventional total-station based surveying was used to determine the elevation of the intermediate points. The North Carolina Geodetic Survey (NCGS) and the North Carolina Department of Transportation conducted all surveying. All reference points were transformed to the North Carolina High Accuracy Reference Network (HARN NAD83/95) for 1995. Transects across stream corridors ranged from 100 m to 840 m with a 6.88 m mean distance between surveyed points.

Field crews visited the study area and characterized the land cover at the survey point locations. The greatest percentage of reference points were under forested canopy as this land cover dominated the stream corridors (Table 2). Reference points on pavement, low and high grass were not obscured by tree cover. There were no large canopy breaks in the forested areas where we collected reference data.

Errors in modeled *surface slope* were computed and analyzed using the slope along survey transects. First, the reference slope was computed along each survey transect using the surveyed *x, y, z* data. The observed slope from the different lidar datasets were computed using the Z values interpolated from the TIN. The transect slope was, in effect, slope in 1.5 dimensions as it was along a one-dimensional linear feature but included elevations as another dimension. Slope at each reference point i was

TABLE 2. FREQUENCY OF REFERENCE OBSERVATIONS AND DISTANCE (M) TO NEAREST LIDAR POINT

Statistic	Pavement	Low Grass	High Grass	Scrub/Shrub	Pine	Decid.	Mixed	$Pr > F^a$
Frequency	30	137	266	177	112	281	222	
Mean Slope		1.6°	1.8°	1.8°	2.3°	3.3°	2.5°	.000
Mean Distance to Nearest Point	1.9	2.0	2.2	4.2	2.9	2.4	3.5	

computed from the average slope of the absolute slope along the survey transects:

$$\text{Slope}_i = \frac{|\text{Transect Slope}_1| + |\text{Transect Slope}_2|}{2} \quad (2)$$

This methodology did not consider whether the slope was locally increasing or decreasing as would be necessary in a solar insolation study, but only considers absolute slope as in a rainfall runoff application. Observations representing pavement land cover were not included in any analysis involving surface slope.

The mean reference slope at each survey point was derived by land cover category (Table 2). Both an F-test and a *t*-test were used to determine if these means were significantly different from one another and if some means were in homogenous groups. Overall, there was a significant variation in mean slope of the reference data by land cover category. Terrain slope under deciduous and mixed canopy was significantly steeper than other land covers. The other land covers (i.e., low grass, high grass, scrub-shrub, and pine forest) did not exhibit significantly different mean slopes.

Because steeper slopes were associated with certain land cover categories, separate analyses were conducted to investigate; (a) if elevation error increased with increasing slope while holding land cover constant, and (b) if slope error varied by land cover category while holding slope class constant. These tests were performed by separating the reference points into 2°-width slope classes, from 0° to 8°. As the frequency of observations in the highest slope class (8° – 10°) were so few ($n = 16$), this slope class was removed from analyses with terrain slope.

Observed error was always computed by subtracting the *in situ* values from the lidar-derived values computed as:

$$\textit{Elevation Error}_i = \text{Elevation from LIDAR}_i - \text{Elevation from Reference}_i \quad (3)$$

$$\textit{Slope Error}_i = \text{Slope from LIDAR}_i - \text{Slope from Reference}_i \quad (4)$$

Negative values in elevation or slope error would indicate the surface derived from the lidar data under-predicted elevation or slope, respectively.

Results

Distance to Nearest Ground Return

The average distance to a "ground" return varied by land cover category (Table 2). The mean distance to nearest return was significantly greater under scrub-shrub or mixed forests than the other categories. The complex vegetation structure in a scrub-shrub environment had a large effect on the return signal and was more problematic for the vegetation removal method than other classes. Not surprisingly, the density of ground returns in pavement, low grass, and high grass was great. The density of ground returns under deciduous canopy was greater than under pine or a mixed forest.

Elevation Error and Land Cover

Mean error in elevation ranged from 14.5 cm to 36.1 cm (RMSE) by land cover category (Table 3). The error of survey points under scrub/shrub land cover was the highest (36.1 cm RMSE) and was significantly different from all other land cover categories. The error observed in the other forested categories (pine, deciduous, and mixed) were not significantly different from one another. The observed errors in the low grass (14.5 cm) and high grass (16.3 cm) categories were the lowest and not significantly different from one another. Elevation errors on pavement, however, were significantly larger than the low and high grass categories.

Bias in Elevation Error

The mean signed errors for each land cover class was tested (using a *t*-test) for a difference from 0.0 (Table 3). The means for all land cover categories except high grass, were significantly different than 0.0. Since all the mean signed errors were negative, there was a tendency for the dataset to consistently under-predict elevation, regardless of land cover class.

Elevation Error and Slope Angle

The expectation was that elevation error would increase as surface slope increases based on fundamental research in topographic mapping (Maling, 1989). Since elevation error was previously found to vary by land cover category, a one-way ANOVA test between mean absolute errors for slope categories was conducted, controlling for land cover category (Table 4). Surprisingly, only a few land cover categories exhibited a significant relationship between elevation error and terrain slope. Only for lidar elevations collected in low grass did elevation error increase as terrain slope increased. None of the other land cover categories exhibited a statistically significant relationship with terrain slope class.

The lack of a consistent relationship between land cover, slope class, and elevation error was not anticipated. Since elevation error was significantly related to slope under low grass, it is possible that the strong effect of vegetation cover for the other land cover categories dominated other sources of elevation error. For these relatively small slopes (0°-8° range) land cover appears to play a much larger role in the introduction of elevation error for leaf-off conditions as found by Hodgson *et al.* (2003) in leaf-on conditions.

Slope Error and Slope Angle

The mean absolute slope error for each reference point was computed using the observed slope of the transect segments as reference values. Neighboring segments were averaged to compute surface slope at the reference points (Equation 2). As an entire set, mean absolute slope error was smaller for the 0° to 2° slope class than the other three slope classes (Table 5). For the lowest slope class (0° to 2°) the mean error was 0.44°. Slope errors for all higher categories ranged from 0.70° to 0.81°.

Another interpretation of the slope errors is to compute a proportional error: the mean error for a slope class divided

Table 3. Terrain Model Error (cm) by Land Cover Class

Statistic	Pavement	Low Grass	High Grass	Scrub/Shrub	Pine	Decid.	Mixed	$Pr > F$[a]
Mean Absolute Error	19.9	11.1	12.2	26.2	24.5	19.9	20.1	*.000*
Mean Signed Error	−11.4	−6.7	−1.50	−15.7	−23.6	−16.0	−11.4	
RMSE	22.6	14.5	16.3	36.1	27.6	27.3	24.3	
$Pr > T$ for signed error = 0.0	*.004*	*.000*	*.134*	*.000*	*.000*	*.000*	*.000*	

[a]F-test for significant difference among land cover categories.

Table 4. Mean *Absolute* Elevation Error by Land Cover Class and Slope Class

	Slope Class	Land Cover Class: Low Grass	High Grass	Scrub/ Shrub	Pine[a]	Decid.	Mixed
Leaf-Off	$0 - 2^0$	9.8	12.3	28.4	25.2	18.8	20.6
	$2 - 4^0$	11.4	12.2	25.1	22.8	23.0	21.0
	$4 - 6^0$	18.8	9.7	15.7		13.3	17.4
	$6 - 8^0$	25.4	15.7	16.9		21.1	18.1
Pr > F		*.000*	*.560*	*.237*	*.371*	*.083*	*.747*

[a]Mean values for the larger slope classes are not reported as the frequency of observations \was less than 5.

Table 5. Mean Error (degrees) in Slope by Reference Slope Class

Statistic	Reference Slope Class (degrees): 0° – 2°	2° – 4°	4° – 6°	6° – 8°	*Pr > F*
Frequency	766	238	98	55	
Mean Absolute Error	.44°	.74°	.81°	.70°	.000
Mean Signed Error	.10°	−0.20°	−0.26°	−0.35°	.000

by the midpoint of the slope class. For example, the lidar-derived slope data showed a 0.44° to 1.00° or error rate of 44 percent of the observed slope angles. Mean error at the highest slope class (6° to 8°) was larger but not as proportionately large (only 10 percent).

As a set, slope was under-predicted as actual terrain slope increased (Table 5). The mean signed slope errors, however were relatively small, ranging from + 0.10° (0° to 2° class) to −0.35° (6° to 8° class). If this trend continued to even greater slopes, the error rates would be less than 10 percent but would be under-predicted.

Separate univariate linear models were computed between mean signed error and slope class and land cover class. Pine land cover was not included in this analysis as the frequency of observations in the higher land cover categories was low. The results indicate mean signed slope error varies by slope class when controlling for land cover (Table 6). Slope was over-predicted for the lowest slope category and under-predicted for the higher categories. This relationship was statistically significant for the high grass, scrub-shrub, mixed land cover, and deciduous. Slope errors in the mixed forests were not monotonically related to reference slope.

In summary, mapping terrain slope[1] using leaf-off lidar data increasingly under-predicted terrain slope for larger slopes (i.e., between 2° and 8°) and over-predicted for very low slopes. While the errors were not very large (e.g., typically less than 1°) the slopes in this study are relatively small.

Discussion

Lidar point density (ground returns) for heavily vegetated categories (i.e., scrub-shrub, pine, deciduous, and mixed) were significantly lower than the other three categories. It is not known if the higher elevation error in these categories results from point density alone, return labeling, or other

[1]The mapping unit size is considered 12 m as the slope at reference points was averaged from the two adjacent transect segments. The average distance for transect segments was approximately 6 m.

Table 6. Mean *Signed* Slope Error (degrees) by Land Cover Class and Slope Class

	Slope Class	Land Cover Class: Low Grass	High Grass	Scrub/ Shrub	Pine[a]	Decid.	Mixed
Leaf-Off	$0° - 2^0$	.04	.11	.05		.15	.11
	$0° - 4^0$	−.09	−.23	−.36		−.04	−.40
	$4° - 6^0$	.00	−.78	−.18		.06	−.28
	$6° - 8^0$	−.30	−.70	−.45		−.24	−.23
Pr > F		*.091*	*.000*	*.017*		*.252*	*.252*
Linearity Test		*.091*	*.000*	*.008*		*.050*	*.001**

[a]There were not enough pine points in the higher slopes classes to statistically evaluate the relationship between reference slope and slope error.
*Indicates the deviation from linearity is significant; thus, the results of the test are inconclusive.

sources. If larger errors in heavily vegetated land cover is primarily from point density then increasing the lidar posting density (e.g., through higher pulse rates and lower AGLs) may obviate this problem.

The pattern of large RMSEs for heavily vegetated land cover is consistent with some previous findings in leaf-off conditions (e.g., Kraus and Pfeifer, 1998) and leaf-on conditions (Hodgson *et al.*, 2003). However, other studies in leaf-off conditions found elevation error under pine cover to exhibit error as low as pavement, low and high grass land cover (e.g., Hodgson and Bresnahan, 2004). The findings that scrub-shrub and deciduous land cover were the highest appears to be consistent with all other studies regardless of leaf conditions (e.g., Hodgson *et al.*, 2003; Hodgson and Bresnahan, 2004).

The lidar-derived elevations in this study were consistently under-predicted regardless of land cover. An under-prediction in elevations has been noted by Adams and Chandler (2002) and Hodgson and Bresnahan (2004), but this was not consistent for all land cover categories. Kraus and Pfeifer (1998), however, noted a +20 cm over-prediction in their Vienna woods. The consistent under-prediction in this study could be just a characteristic of this specific dataset. Several error sources might explain this bias. For instance, a vertical error in the GPS or INS approach (Table 1) could cause such a negative shift in the elevations (Kraus and Pfeifer, 1998; Krabill *et al.*, 2002; Maas, 2002). The surprisingly large under-prediction for reference points on pavement (−11.4 cm) is a pattern noticed by North Carolina Flood Plain Mapping staff in many of their datasets.

There was little evidence that elevation error increased with increasing surface slope across the low slopes (from 0° to 8°) typical of flood prone areas. Only the low grass category exhibited the expected monotonic increase in error as slope increased. It is likely that the relatively low slopes in this study were not large enough to introduce strong observable error greater than the error introduced from other sources (e.g., point density, return labeling). Theoretically, this range in slopes would, in the worst case of 8° slopes only introduce vertical errors of 14 cm (assuming a 100 cm horizontal error).

Surface slope angles were slightly over-predicted on low slopes (0° to 2°) but almost always under-predicted on higher slopes (>2°). The relatively high variability of elevations resulting from laser returns on near-surface objects would explain the over-prediction on low slopes. Additional research should explore the effects of interpolation algorithms and slope algorithms for modeling surface slope from lidar data. Krabill *et al.* (2002), for example, advocates spatial averaging of lidar returns to minimize error. This study used reference slope as the average slope of adjacent segments along survey transects where the transect segments were approximately 6 m in length. Modeled slope was derived from linear interpolation of elevations in a TIN model. Modeling surface slope may be effected by the reference data, lidar posting-density, interpolation algorithm, land cover, and mapping unit (e.g., grid cell size).

Reporting the accuracy of a lidar-derived terrain dataset without reporting the collection conditions or geography of the study area is problematic. The results of this study indicate the elevation accuracy of lidar data vary by *land cover* type. Where summary statistics are reported (e.g., overall RMSE) the distribution of reference points across different land cover categories and slope classes should also be reported. If the distribution of observations is not equal to the area within each land cover category, the overall RMSE will not be representative of the accuracy of the overall study area.

References

Adams, J.C., and J.H. Chandler, 2002. Evaluation of LIDAR and medium scale photogrammetry for detection soft-cliff coastal change, *The Photogrammetric Record*. 17(99):405–418.

Bolstad, P.V., and T. Stowe, 1994. An evaluation of DEM accuracy: Elevation, slope, and aspect, *Photogrammetric Engineering & Remote Sensing*, 60(11):1327–1332.

Box, D., 2003. 3Di, Inc., and Eagleview Research, personal communication.

Bowen, Z.H., and R.G. Waltermire, 2002. Evaluation of light detection and ranging (LIDAR) for measuring river corridor topography, *Journal of the American Water Resources Association,* 38(1):33–41.

Cobby, D.M., D.C. Mason, and I.J. Davenport, 2001. Image processing of airborne laser altimetry data for improved river modeling, *ISPRS Journal of Photogrammetry and Remote Sensing*, 56(2):121–138.

Cowen, D.J., J.R. Jensen, C. Hendrix, M.E. Hodgson, and S.R. Schill, 2000. A GIS-assisted rail construction econometric model that incorporates LIDAR data, *Photogrammetric Engineering & Remote Sensing*. 66(11):1323–1326.

Fowler, R., 2001. Topographic Lidar, *Digital Elevation Model Technologies and Applications: The DEM Users Manual*, (D. Maune, editor), American Society for Photogrammetry and Remote Sensing, Bethesda, Maryland, pp. 207–236.

Hodgson, M.E., J.R. Jensen, L. Schmidt, S. Schill, and B. Davis, 2003. An Evaluation of LIDAR- and IFSAR-derived Digital Elevation Models in Leaf-on Conditions with USGS Level 1 and Level 2 DEMs, *Remote Sensing of Environment*, 84:295–308.

Hodgson, M.E., and B.E. Alexander, 1990. Use of Historic Maps in GIS Analyses, *Proceedings of the 1990 ASPRS-ACSM Annual Convention*, 3:109–116.

Hodgson, M.E., and P. Bresnahan, 2004. Accuracy of Airborne Lidar-Derived Elevation: Empirical Assessment and Error Budget, *Photogrammetric Engineering & Remote Sensing*. 70(3):331–339.

Jensen, J.R., M.E. Hodgson, H.E. Mackey, and W. Krabill, 1987. Correlation Between Aircraft MSS and LIDAR Remotely Sensed Data on a Forested Wetland, *Geocarto International*, 2(4):39–54.

Jensen, J.R., 2000. Active and Passive Microwave, and LIDAR Remote Sensing, Chapter 9, *Remote Sensing of the Environment: An Earth Resource Perspective*, Prentice Hall, New Jersey, pp. 285–332.

Kenward, T., and D.P. Lettenmaier, E.F. Wood, and E. Fielding, 2000. Effects of digital elevation model accuracy on hydrologic predictions, *Remote Sensing of Environment*. 74:432–444.

Krabill, W.B., W. Abdalati, E.B. Frederick, S.S. Manizade, C.F. Martin, J.G. Sonntag, R.N. Swift, R.H. Thomas, and J.G. Yungel, 2002. Aircraft laser altimetry measurement of elevation the Greenland ice sheet: technique and accuracy assessment, *Journal of Geodynamics*, 34:357–376.

Kraus, K., and N. Pfeifer, 1998. Determination of terrain models in wooded areas with airborne laser scanner data, *ISPRS Journal of Photogrammetry and Remote Sensing*, 53:193–203.

Lloyd, C.D., and P.M. Atkinson, 2002. Deriving DSMs from LIDAR data with kriging, *International Journal of Remote Sensing*, 23(12):2519–2524.

Maling, D.H., 1989. *Measurements from Maps*, (Pergamon Press, New York).

Maas, H-G., 2002. Methods for measuring height and planimetry discrepancies in airborne laserscanner data, *Photogrammetric Engineering & Remote Sensing*, 68(9):93–940.

Maune, D.F., 1996. Introduction to digital elevation models (DEM), Chapter 6, *Digital Photogrammetry: An Addendum to the Manual of Photogrammetry*, American Society for Photogrammetry and Remote Sensing, Bethesda, Maryland.

Means, J.E., S.A. Acker, B.J. Fitt, M. Renslow, L. Emerson, and C. J. Hendrix, 2000. Predicting forest stand characteristics with airborne scanning LIDAR, *Photogrammetric Engineering & Remote Sensing*. 66(11):1367–1371.

North Carolina Floodplain Mapping Program, 2002. URL: http://www.ncfloodmaps.com, The State of North Carolina, Raleigh, North Carolina (last date accessed: 13 March 2005).

Petzold, B., P. Reiss, and W. Stossel, 1999. Laser scanning – surveying and mapping agencies are using a new technique for the derivation of digital terrain models, *ISPRS Journal of Photogrammetry and Remote Sensing*, 54:95–104.

Raber, G.T., J.R. Jensen, S.R. Schill, and K. Schuckman, 2002. Creation of digital terrain models using an adaptive LIDAR vegetation point removal process, *Photogrammetric Engineering & Remote Sensing*, 68(12):1307–1315.

Thrower, N.J.W. and J.R. Jensen, 1976. The orthophoto and orthophotomap: characteristics, development and application, *The American Cartographer*, 3(1):39–56.

(Received 21 October 2003; accepted 02 April 2004; revised 10 May 2004).