JOURNAL OF THE AMERICAN WATER RESOURCES ASSOCIATION
VOL. 38, NO. 1 AMERICAN WATER RESOURCES ASSOCIATION FEBRUARY 2002

EVALUATION OF LIGHT DETECTION AND RANGING (LIDAR) FOR MEASURING RIVER CORRIDOR TOPOGRAPHY[1]

Zachary H. Bowen and Robert G. Waltermire[2]

ABSTRACT: LIDAR is relatively new in the commercial market for remote sensing of topography and it is difficult to find objective reporting on the accuracy of LIDAR measurements in an applied context. Accuracy specifications for LIDAR data in published evaluations range from 1 to 2 m root mean square error (RMSEx,y) and 15 to 20 cm RMSEz. Most of these estimates are based on measurements over relatively flat, homogeneous terrain. This study evaluated the accuracy of one LIDAR data set over a range of terrain types in a western river corridor. Elevation errors based on measurements over all terrain types were larger (RMSEz equals 43 cm) than values typically reported. This result is largely attributable to horizontal positioning limitations (1 to 2 m RMSEx,y) in areas with variable terrain and large topographic relief. Cross-sectional profiles indicated algorithms that were effective for removing vegetation in relatively flat terrain were less effective near the active channel where dense vegetation was found in a narrow band along a low terrace. LIDAR provides relatively accurate data at densities (50,000 to 100,000 points per km^2) not feasible with other survey technologies. Other options for projects requiring higher accuracy include low-altitude aerial photography and intensive ground surveying.
(KEY TERMS: LIDAR – remote sensing; instream flow studies; aquatic ecosystems; watershed management; Geographic Information Systems; surveying.)

INTRODUCTION

During the last decade, the use of remote sensing technology to study aquatic and riparian habitats has become common. Satellite or aircraft mounted instruments collect data used for mapping water temperatures (Wentz *et al.,* 2000), chlorophyll-a concentrations (Avard *et al.,* 2000), terrestrial and underwater topography (Lillycrop *et al.,* 1996), hydrologic data (Bobba et al. 1992), and sampling fish abundances (Lo *et al.,* 2000).

Among the most common uses of satellite and airborne data are defining ground topography and land use for areas ranging from watersheds (Herlihy *et al.,* 1998) to short river reaches (Covington and Hubert, 2000). Topographic data are an important component of river and watershed studies at a variety of scales. Maps of topography are useful for studies of sediment movement and channel change at the river and reach level (Kondolf and Larson, 1995). Similar topographic information is used in impact assessment and instream flow studies that include study sections ranging from hundreds of meters to many kilometers long. As part of impact assessment studies, river corridor topography is typically required to predict and evaluate the potential effects of changes in river stage and streamflow on habitat for aquatic and riparian organisms. Specific data resolution and accuracy requirements vary depending on the extent and purpose of a particular study: watershed mapping is often accomplished using 1:24,000 scale maps while studies of specific river reaches might require 1:5,000 data. Regardless of scale, the use of sophisticated models and spatial analysis tools for aquatic habitat assessment will become more common as demands on fresh water supplies increase, dams come up for relicensing, and the cost of using state-of-the-art technology decreases.

One of the newer technologies commercially available for collecting remotely sensed topography data is LIght Detection And Ranging (LIDAR) – also referred to as airborne laser mapping (Measures, 1991). LIDAR data collection uses a fast-firing laser (typically 4 to 10 kHz) mounted in a small aircraft to measure distances to the surface of the earth based on the round trip travel time for the laser pulse. Distances

[1]Paper No. 01054 of the *Journal of the American Water Resources Association.* **Discussions are open until October 1, 2002.**
[2]Respectively, Fishery Biologist and Computer Specialist, U.S. Geological Survey, 4512 McMurry Avenue, Fort Collins, Colorado 80525-3400 (E-Mail/Bowen: zack_bowen@usgs.gov.1

are initially referenced to the position of the survey aircraft. The position of the survey aircraft is determined by an on-board global positioning system (GPS) receiver. The GPS receiver in the aircraft receives position data which are post-processed and corrected based on concurrently collected position information from a GPS base station receiver located over a known benchmark. The attitude (i.e., pitch, yaw, and roll) of the aircraft is recorded by an inertial measurement unit (IMU) throughout the survey. Distance data, position information, and data from the IMU are used in conjunction with a geoid model (roughly a projection of sea level) to determine longitude (x), latitude (y), and elevation (z) for survey points on the ground. Point measurement density depends on laser frequency, aircraft speed, altitude, ground terrain, and vegetation; 50,000 to 100,000 survey points per square kilometer with spacing between points of 1 to 5 meters are typical (Flood and Glutelius, 1997). Contractor specifications typically state that ground positions and elevations from LIDAR data have root mean square error (RMSE) of 1 to 2 m for x, y positions and 15 to 20 cm RMSE for elevations. These values are consistent with published comparisons of LIDAR with ground-based GPS data from an airfield runway, the Greenland ice sheet (Krabill et al. 1995), and the Assateague Island National Seashore beach (Krabill *et al.*, 2000).

Based on the reported accuracy, data density, and relatively low cost per square mile ($500 to $1500), many federal and state agencies are using or considering LIDAR for river corridor mapping projects. LIDAR may substantially improve resolution and coverage for river studies compared to ground-based surveying and requires less time and effort than traditional photogrammetric methods. However, published comparisons of LIDAR data with ground based surveys are limited to evaluations over smooth and homogeneous terrain (runways, beaches, and ice sheets) where limitations of the technology may not be apparent. This study compares one LIDAR data set to ground survey GPS data collected in a river corridor over various types of terrain. Our goal is to provide a practical evaluation of the applicability and accuracy of LIDAR for monitoring, impact assessment, and instream flow studies in river corridors. Our specific objectives are to: (1) evaluate the accuracy of LIDAR data collected over different types of terrain typical of river studies; (2) identify what types of terrain were associated with the largest measurement errors; and (3) identify any method-specific limitations associated with LIDAR data when used for mapping rivers.

METHODS

Study Area

The study area is located in the northeast corner of Utah near the town of Jensen (Uintah County), and includes a 1.5 km section of the Green River adjacent to Dinosaur National Monument. The region receives about 21 cm of average annual precipitation. Surface soils are mostly alluvial clay, sand, and cobble with some rock outcrops. Vegetation is predominantly sagebrush (*Artemisia tridentata*) with salt cedar (*Tamarix ramosissima*) and Russian olive (*Elaegnus angustifolia*) common on islands and bars near the active channel. Elevations in the study area range from 1446 to 1550 m msl.

LIDAR Data Collection

LIDAR data were collected during October 23 and 24, 1999, by a contractor using a proprietary airborne laser mapping system from a small, fixed-wing aircraft. Distancing was accomplished by scanning a laser (4 kHz pulse rate) across the flight path of the aircraft. Aircraft position was measured using an onboard dual frequency GPS receiver. During the airborne survey, a second dual frequency receiver collected measurements over a known benchmark. Laser distances and differentially corrected GPS data were post-processed in conjunction with data from the IMU and a geoid model (GEOID 96) to calculate ground positions and elevations. During final processing a program was used to identify and remove laser returns from tops of vegetation. All LIDAR data processing was completed by the contractor using proprietary algorithms and programs. Hand editing of the final data file was not performed. Imagery was also collected during the airborne survey using a 2000 x 2000 pixel digital camera. The digital images were coregistered and georeferenced to the laser data.

Ground Survey

To evaluate the accuracy of the LIDAR data in different types of terrain, 232 elevations were measured on February 14, 2000, at x, y locations from the LIDAR data set (Figure 1). It was not possible to measure directly the precision of x,y positions from the LIDAR data set. Measurement points were selected prior to field work based on examination of the digital aerial photos to represent the range of terrain types available in the study area (Table 1). Sample size was

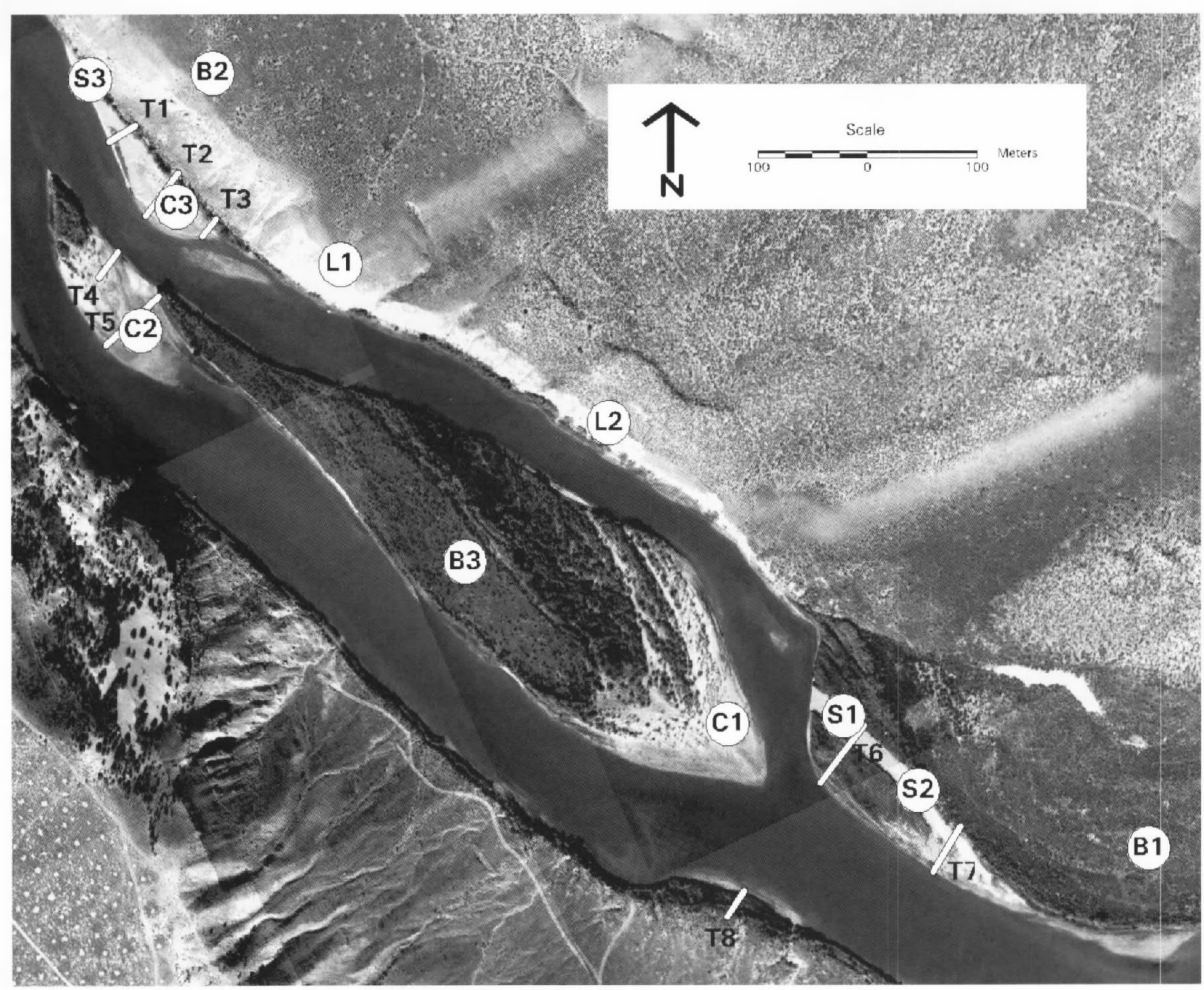

Figure 1. Map of LIDAR Evaluation Study Area Along the Green River in Northeast Utah Showing 11 Data Collection Locations for Point Comparisons (circles) and Eight Transect Locations T1 Through T8 for Cross-Sectional Profiles (white lines). Point measurement circles are labeled S for sand, B for brush, C for cobble, and L for slope. The river flows east to west.

Table 1. Descriptions for Terrain Types, Total Number of Samples, and Number of Samples Per Sampling Location in a Comparison of 232 LIDAR and Ground GPS Elevations Along the Green River Valley in Northeast Utah.

Terrain	Description	Total n / Number of Samples Per Location
Brush	Regularly distributed sagebrush plants 0.3 to 1.0 m in height, 0.3 to 1.3 m in diameter growing in sandy soils; little topographic relief	77 / 32, 12, 33
Sand	Sand bar and side channel filled with fine sand; very little topographic relief	51 / 20, 12, 19
Cobble	Round cobble 10 to 25 cm in diameter with variable embeddedness by fine sand and silt; little topographic relief	92 / 34, 33, 25
Slope	Rock outcrops in sandy soils; large topographic relief; slopes of 45 degrees and larger	12 / 10, 2

not consistent among the different terrain types because some photo-based terrain classifications were changed based on ground observations and some pres-elected measurement points (e.g., steep slope with loose rock) were not accessible. We did not choose and measure alternate points with the same terrain type because the digital photo and full set of x,y positions from the LIDAR data set were not available at the field site. The slope terrain type was measured in two geographic areas. For sand, cobble, and brush terrain types, points were measured in three geographic areas per type to reduce the influence of any unmeasured systematic bias in the LIDAR data. Each point was located and measured using a Trimble 4800 GPS receiver receiving real-time differential corrections from another Trimble 4800 receiver located over a known benchmark at the study area. The LIDAR and ground surveys were conducted using two different benchmarks from a three-point survey network established for this project. Based on previous surveys and checks conducted throughout this study, ground measurement precisions using real time corrections were typically ± 3 cm (x, y, z).

The effectiveness of LIDAR data for capturing near channel topography was evaluated along cross sectional profiles. Ground survey GPS data collected during November 1 through 10, 1999, were used to define topography along cross sections perpendicular to the main river channel. These surveys measured channel geometry between the base flow water surface elevation and the top of bank in the active channel. Cross section length and the number of ground survey points used to define the cross section varied with channel geometry and complexity. A Geographic Information System (GIS – ArcInfo V8.0.1, ESRI 2000) was used to overlay the LIDAR data on a map containing the cross section profiles. Cross section profiles with eight or more LIDAR points falling within 3 m on either side were identified and used to compare topography between LIDAR and ground GPS surveys.

Data Analysis

The statistical error measurements, Root Mean Squared Error (RMSE), Absolute Mean Error (ABSE), Mean Error (ME), and Maximum error (MAXE) were used to compare the LIDAR elevations to the ground GPS elevations.

$$RMSEz = \sqrt{\frac{\sum_{i=1}^{n}\left(Zlidar_i - Zgps_i\right)^2}{n}}$$

$$ABSEz = \frac{\sum_{i=1}^{n}\left|Zlidar_i - Zgps_i\right|}{n}$$

$$MEz = \frac{\sum_{i=1}^{n}\left(Zlidar_i - Zgps_i\right)}{n}$$

$$MAXEz = MAX_{i=1}^{n}\left|Zlidar_i - Zgps_i\right|$$

The root mean square error and absolute mean error are measures of the deviation of the LIDAR elevations from the ground GPS elevations. Overestimations and underestimations are accounted for as errors in these measures. According to the Federal Geographic Data Committee National Standard for Spatial Data Accuracy (1998), Accuracyz = 1.960 x RMSEz. The mean error describes the overall bias of the LIDAR data to overestimate or underestimate elevations compared to the ground GPS data. In this error measurement, overestimations and underestimations cancel each other out. The maximum error describes the largest point of deviation of the LIDAR elevations compared to the ground GPS elevations. This value, along with the mean error describes the range of deviations.

To correct for any systematic bias resulting from setup, calibration, or measurement errors in either the LIDAR data or ground GPS data, LIDAR elevations were block corrected by subtracting the mean error between LIDAR and ground GPS measurements (-44 cm) from the original LIDAR data set. Errors and error statistics were computed using the block corrected LIDAR elevations. Normality of the error distributions by terrain type was assessed using the Komolgorov-Smirnov test. A Kruskal-Wallis rank test was used to test the hypothesis that error magnitude was the same among different terrain types. Data from LIDAR and ground GPS measurements along cross sectional profiles were compared graphically. The digital orthophotos were used to provide context and possible explanations for observed differences in LIDAR and ground GPS measurements of elevation.

RESULTS

Error statistics computed using all 232 elevation measurements indicated differences between the original LIDAR and ground GPS measurements larger than the 15 to 20 cm RMSEz accuracy specifications found in most published results and advertising literature (Table 2). Block correcting the LIDAR elevations resulted in a 30 percent reduction in RMSEz and greater than 50 percent reduction in absolute mean error (Table 2). Even with these corrections, the RMSEz was still about two times higher than commonly advertised accuracy specifications. The large maximum error relative to the mean error indicated that large outliers were present in the LIDAR data.

TABLE 2. Root Mean Square Error (RMSE), Absolute Mean Error (ABSE), Mean Error (ME), and Maximum Error (MAXE) Obtained by Comparing LIDAR and Ground GPS Measurements of Elevation (n = 232). Error statistics in the center column are based on the original data. Error statistics in the right column were calculated based on block corrected LIDAR elevations.

Error Statistic	Original Data (cm)	Block Corrected Data (cm)
RMSE	62	43
ABSE	56	22
ME	-44	0
MAXE	191	233

Error distributions were not consistent among different terrain types ($p < 0.001$, Figure 2). Root mean square error was largest for the slope terrain type (RMSEz = 111 cm, n = 12) followed by sand (RMSEz = 53 cm, n = 51), cobble (RMSEz = 19 cm, n = 92), and brush (RMSEz = 9 cm, n =77). The largest range of errors (380 cm) was observed in slope terrain and the largest number of outliers (eight) was observed in sand terrain. Error values from the cobble terrain type were normally distributed ($p > 0.05$). Error values for brush, sand, and slope terrain were not normally distributed ($p < 0.05$). The largest error magnitudes were in the slope terrain type and in the sand terrain type near the edge of the channel (Figure 3).

Cross sectional profiles near the active channel also revealed differences between LIDAR and ground GPS elevations (Figure 4). Most of the relatively large differences were where LIDAR elevations were higher than ground GPS elevations. The large LIDAR overestimations typically were observed in areas where riparian vegetation was relatively dense and bank slopes were steep (Figure 1 and Figure 4, panels

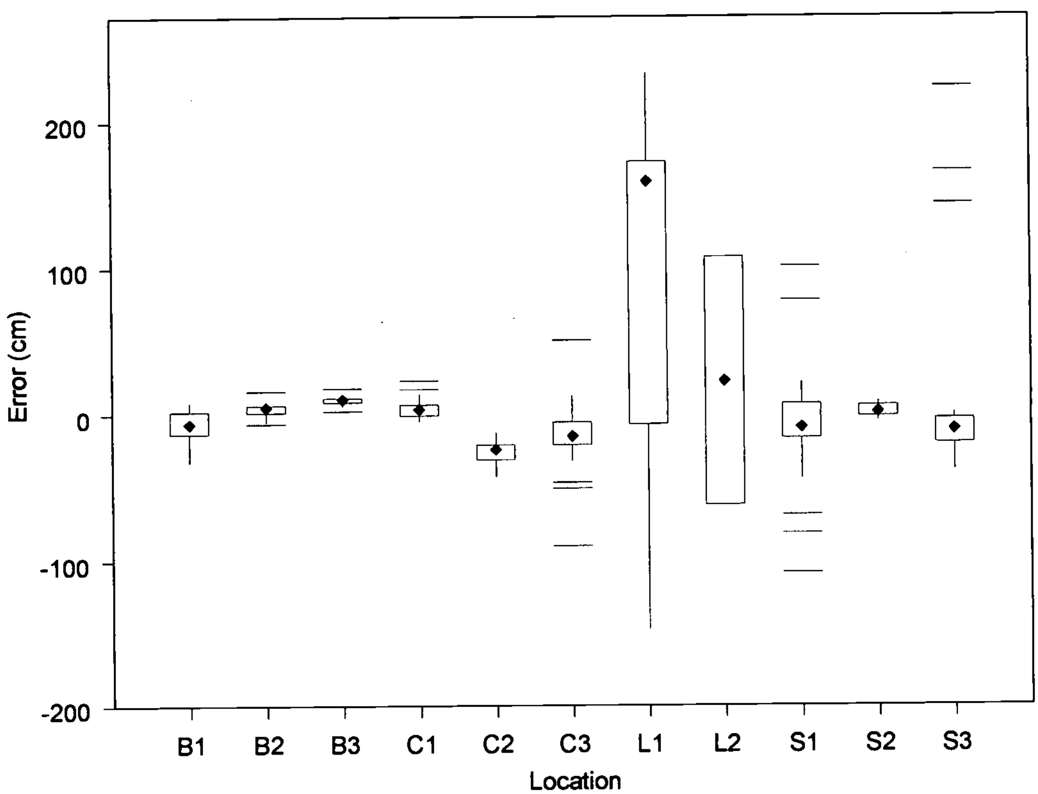

Figure 2. Elevation Error Distribution for Each Measurement Location Comparing LIDAR and Ground GPS Elevations. Diamonds represent the median error, box lengths represent the interquartile range (25th to 75th percentile), vertical lines represent the range excluding outliers, and the horizontal lines represent outliers (values 1.5 box lengths or greater from the 25th or 75th percentiles) (S for sand, B for brush, C for cobble, and L for slope).

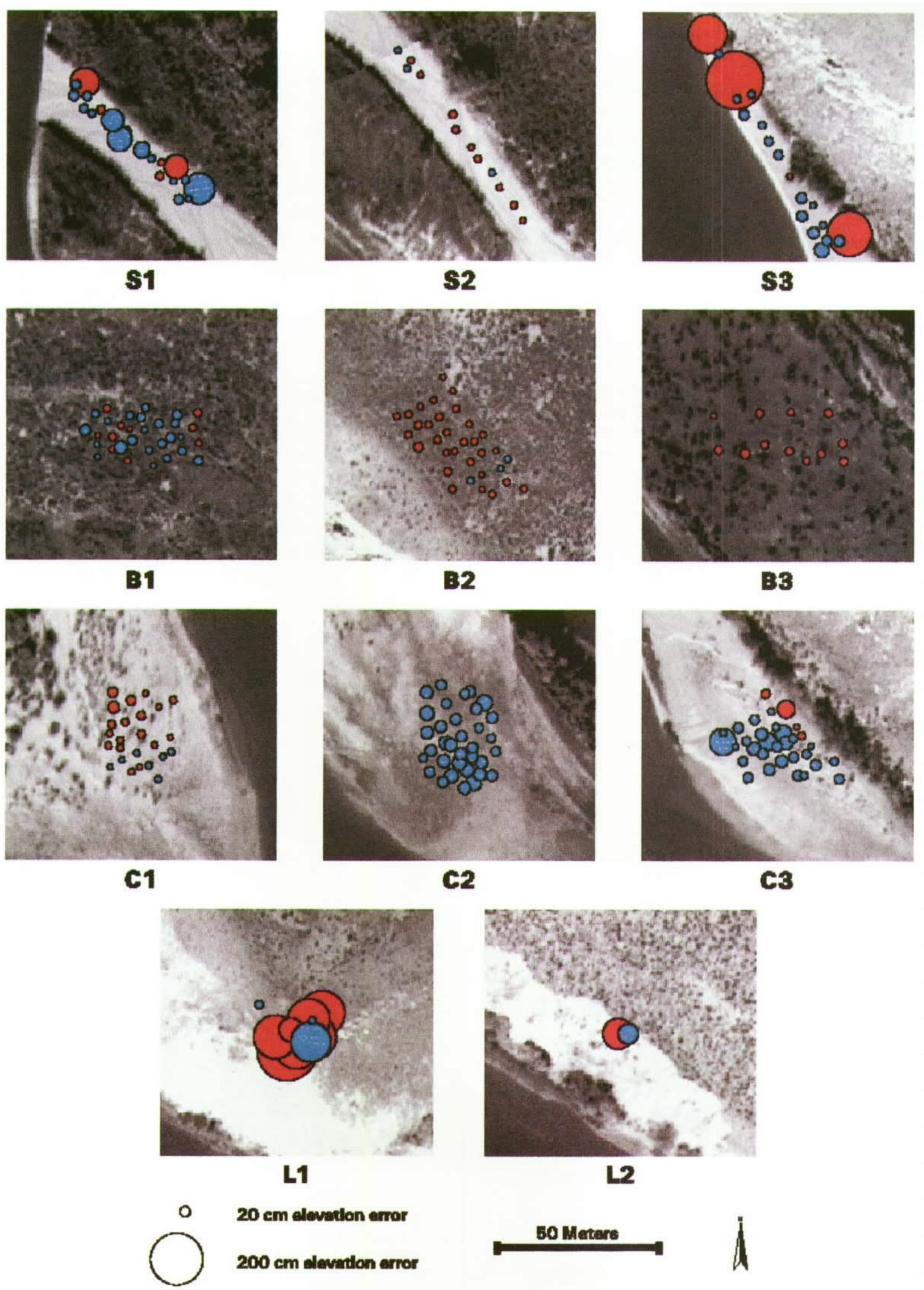

Figure 3. Elevation Errors for Individual Survey Points. Blue indicates LIDAR underestimation and red indicates LIDAR overestimation of ground GPS elevations. Circle size is proportional to error magnitude (S for sand, B for brush, C for cobble, and L for slope).

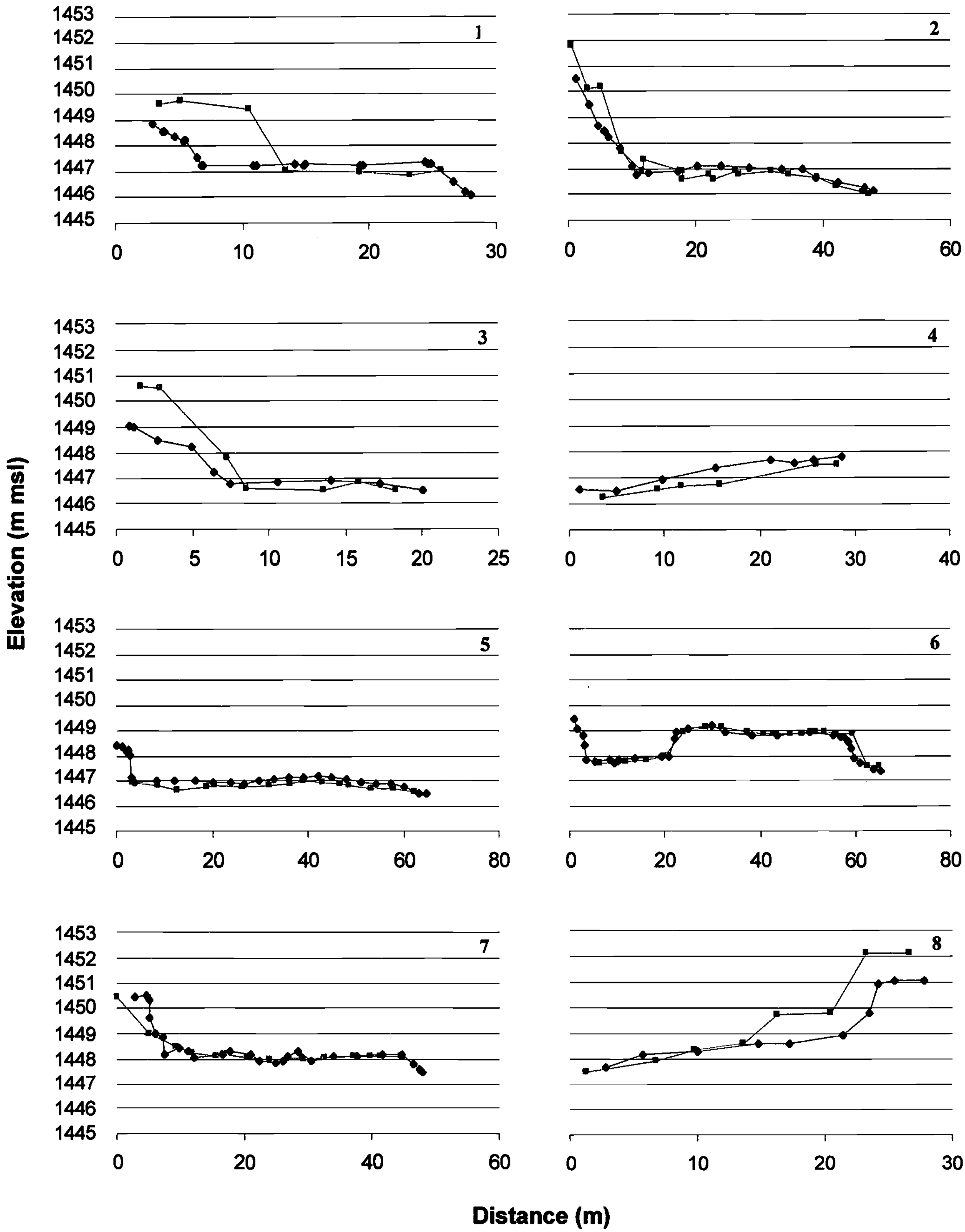

Figure 4. Cross-Sectional Profiles Based on Ground GPS (blue diamonds) and LIDAR (pink squares) for the Eight Transects (T1 through T8) in Figure 1.

1, 2, 3, and 8). Other cross sections in more open terrain showed LIDAR elevations consistently lower than GPS measurements (Figure 1 and Figure 4, panels 4 and 5). Four of the cross sections revealed only slight differences in elevation and profile shape between LIDAR and ground GPS measured elevations (Figure 4, panels 2, 5, 6, and 7).

DISCUSSION

A block correction of the LIDAR data was required to reduce the effects of systematic bias and improve agreement between the LIDAR and ground GPS elevation data. Potential sources of systematic bias include setup and calibration errors in the LIDAR or ground GPS systems, error in the ground GPS network, and errors introduced during data processing. Because vertical control was limited to a single point in both surveys, a uniform block correction based on the bias (mean error) of the LIDAR data was applied. If ground GPS data at precise x, y locations from the LIDAR survey were not available and no correction were applied, the RMSEz would have been 30 percent larger. This finding highlights the importance of collecting at least a minimal set of ground survey validation data as part of a LIDAR project.

Over all terrain types combined, observed errors were larger than those commonly reported in the literature. Most published evaluations of LIDAR accuracy have been conducted over relatively smooth and homogeneous surfaces (e.g., Krabill *et al.*, 1995; 2000). Similarly, calibration and accuracy checks are typically conducted over straight, well surveyed road surfaces. Our comparison of 232 LIDAR elevations to those from ground GPS over a range of terrain types indicated that observed errors can be twice as large (RMSEz = 43 cm) as found in typical LIDAR accuracy specifications. Although larger than typically reported, the observed overall RMSEz (43 cm) was very close to a value from a steep dune profile along the Assateague National Seashore beach (Krabill *et al.*, 2000; standard deviation of errors = 49 cm) indicating this level of error is not unprecedented and is probably a real limitation of the technology in steep, variable terrain.

Error statistics varied widely across terrain types. Values for RMSEz in cobble and brush terrain types (19 and 9 cm, respectively) were consistent with published error statistics and advertised accuracy specifications. In this survey, points collected in cobble and brush terrain types were in areas where there was little topographic relief. The low RMSEz in brush indicated that algorithms used to remove vegetation were effective in these areas. Large errors in steep, rocky terrain and in areas with abrupt elevation changes contributed to the overall RMSEz value (43 cm). The RMSEz for sand (53 cm) was strongly influenced by eight outliers (Figure 2). Measurements responsible for the outliers were in the active channel near the bank (Figure 3). It was not possible to measure directly the precision of x,y positioning from the LIDAR survey with the data available. However, based on proximity to the bank and the magnitude of the LIDAR elevation outliers (> 0.7 m), it appears that the outlier measurements resulted from horizontal positioning error associated with the LIDAR data: elevation was based on a laser-measured distance to the dry channel bed and the corrected GPS position indicated the measurement was on the bank, or vice versa. If outliers in the sand terrain type are discounted, the error distribution is similar to those for cobble or brush (Figure 2). The largest error statistics were observed in the slope terrain type (RMSEz = 111 cm) which was characterized by abrupt changes in elevation and a very irregular surface where small errors in x, y positioning would result in large differences in elevation. A relatively small sample size (n = 12) also contributed to the large RMSEz in slope terrain.

Ground GPS data collected along transects and compared to LIDAR elevations revealed strengths and limitations associated with LIDAR data. Cross-sectional profiles indicated algorithms that were effective for removing vegetation in relatively flat terrain were less effective in sloped terrain near the active channel. Four of the eight cross sections measured showed LIDAR overestimation of elevation resulting from laser returns from tops of riparian vegetation (Figure 4, panels 1, 2, 3, and 8). Overestimates occurred where dense vegetation was found in a 10 to 50 m wide band along a low terrace adjacent to the channel. Because the filtering algorithm must screen on relative changes in elevation over a specified ground distance, it is logical that discrimination between ground and vegetation would be less precise in areas where ground elevation was highly variable over short distances. On at least one cross section profile (Figure 4, panel 7) it appears that a combination of the laser sampling pattern and erroneous filtering resulted in oversmoothing along a relatively variable profile. There was no clear reason why LIDAR underestimated elevations along Transects 4 and 5, extending across a mid-channel bar. LIDAR generally captured the shape of the cross section along Transects 2, 5, 6, and 7.

CONCLUSIONS

LIDAR is relatively new in the commercial market for remote sensing of topography and it is difficult to find objective reporting on the quality of data sets. This report evaluated the practical accuracy of one LIDAR data set over various terrain types in a western river corridor. LIDAR provides relatively accurate topographic data at densities (50,000 to 100,000 points per km^2) not feasible with other survey technologies. The requirements for ground surveying to control the LIDAR model are minimal. This saves time and money and facilitates surveys over fragile or inaccessible areas. Minimal preflight ground surveying and fast data processing mean that topographic data can be developed quickly – on the order of days or weeks. Accuracy of LIDAR elevation data varies depending on terrain type. Most published estimates of LIDAR accuracy reflect values obtained in relatively flat, homogeneous terrain. Although LIDAR is capable of producing 15 to 20 cm RMSEz elevation data, horizontal positioning limitations (1 to 2 m RMSEx,y) associated with each laser return increase the probability for larger observed elevation errors in areas with variable terrain and large topographic relief. LIDAR elevation data are based on systematic random sampling by a scanning laser. This limits the user's ability to define linear features, such as the top of a steep bank, at a resolution smaller than the distance between sample points (typically 1 to 5 m). However, it is possible to use co-registered digital imagery in conjunction with the LIDAR data to hand digitize elevation breaklines that follow visible linear features. Co-registered digital imagery is also useful for data interpretation and quality control and should be obtained if possible.

(Note: The mention of trade names does not constitute endorsement or recommendation for use by the Federal Government.)

LITERATURE CITED

Avard, M. M., F. R. Schiebe, and J. H. Everitt, 2000. Quantification of chlorophyll in reservoirs of the Little Washita River Watershed Using Airborne Video. Photogrammetric Engineering and Remote Sensing 66:213218.

Bobba, A. G., R. P. Bukata, and J. H. Jerome, 1992. Digitally Processed Satellite Data as a Tool in Detecting Potential Groundwater Flow Systems. Journal of Hydrology 131:25-62.

Covington, J. S. and W. A. Hubert, 2000. Relations Between On-Site and Aerial Measurements of Streamside Features and Cover for Trout in Alluvial Valley Streams in Wyoming. North American Journal of Fisheries Management 20:627-633.

Environmental Systems Research Institute, Inc. 2000. ArcInfo V8.0.1. Redlands, California.

Federal Geographic Data Committee, 1998. Geospatial Positioning Accuracy Standards, Part 3: National Standard for Spatial Data Accuracy. FGDC-STD-007.3-1998. *Available at* http://www.fgdc.gov/standards/documents/standards/accur acy/chapter3.pdf

Flood, M. and B. Glutelius, 1997. Commercial Implications of Topographic Terrain Mapping Using Scanning Airborne Laser Radar. Photogrammetric Engineering and Remote Sensing 63:327-329,363-366.

Herlihy, A. T., J. L. Stoddard, and C. B. Johnson, 1998. The Relationship Between Stream Chemistry and Watershed Land Cover Data in the Mid-Atlantic Region. U.S. Water, Air, and Soil Pollution 105:377-386.

Kondolf, G. M. and M. Larson, 1995. Historical Channel Analysis and Its Application to Riparian and Aquatic Habitat Restoration. Aquatic Conservation: Marine and Freshwater Ecosystems 5:109-126.

Krabill, W. B., R. H. Thomas, C. F. Martin, R. N. Swift, and E. B. Frederick, 1995. Accuracy of Airborne Laser Altimetry Over the Greenland Ice Sheet. International Journal of Remote Sensing 16:1211-1222.

Krabill, W. B. *et al.*, 2000. Airborne Laser Mapping of Assateague National Seashore Beach. Photogrammetric Engineering and Remote Sensing 66:6571.

Lillycrop, W. J., L. E. Parson, J. L. Irish, and M. W. Brooks, 1996. Hydrographic Surveying With an Airborne Lidar Survey System. Proceedings of the Second International Airborne Remote Sensing Conference and Exhibition 1:279-285.

Lo, N. C. H., J. R. Hunter, and J. H. Churnside, 2000. Modeling Statistical Performance of an Airborne Lidar Survey System for Anchovy. Fishery Bulletin 98:264282.

Measures, R. M., 1991. Laser Remote Sensing: Fundamentals and Applications. Krieger Publishing Co., Melbourne, Florida, 524 pp.

Wentz, F. J., C. Gentemann, D. Smith, and D. Chelton, 2000. Satellite Measurements of Sea Surface Temperature Through Clouds. Science 288:847-850.