Mapping river bathymetries: Evaluating topobathymetric LiDAR survey

Daniele Tonina,1,2* James A. McKean,2 Rohan M. Benjankar,3 C. Wayne Wright,4 Jaime R. Goode,5 Qiuwen Chen,6 William J. Reeder,1 Richard A. Carmichael1 and Michael R. Edmondson7

1 Center for Ecohydraulics Research, University of Idaho, 322 E. Front Street, suite 340, Boise, Idaho 83702, USA
2 US Forest Service, Rocky Mountain Research Station, 322 E. Front Street, suite 401, Boise, Idaho 83702, USA
3 Civil Engineering Department, Southern Illinois University Edwardsville, Edwardsville, Illinois 62025
4 US Geological Survey, Remote Sensing, Florida Integrated Science Center, St. Petersburg, FL 33701, USA
5 Mathematics & Physical Sciences Department and Environmental Studies, The College of Idaho, 2112 Cleveland Blvd Caldwell ID 83605, USA
6 Center for Eco-Environmental Research, Nanjing Hydraulic Research Institute, Nanjing 210098, China
7 Idaho Office of Species Conservation, 304 N. 8th Street, Suite 149, Boise, ID 83702, USA

Received 26 May 2018; Revised 10 September 2018; Accepted 13 September 2018

*Correspondence to: Daniele Tonina, University of Idaho, Center for Ecohydraulics Research, 322 E. Front Street, suite 340; Boise, Idaho 83702, USA. E-mail: dtonina@uidaho.edu

ABSTRACT: Advances in topobathymetric LiDARs could enable rapid surveys at sub-meter resolution over entire stream networks. This is the first step to improving our knowledge of riverine systems, both their morphology and role in ecosystems. The Experimental Advanced Airborne Research LiDAR B (EAARL-B) system is one such topobathymetric sensor, capable of mapping both terrestrial and aquatic systems. Whereas the original EAARL was developed to survey littoral areas, the new version, EAARL-B, was also designed for riverine systems but has yet to be tested. Thus, we evaluated the ability of EAARL-B to map bathymetry and floodplain topography at sub-meter resolution in a mid-size gravel-bed river. We coupled the EAARL-B survey with highly accurate field surveys (0.03 m vertical accuracy and approximately 0.6 by 0.6 m resolution) of three morphologically distinct reaches, approximately 200 m long 15 m wide, of the Lemhi River (Idaho, USA). Both point-to-point and raster-to-raster comparisons between ground and EAARL-B surveyed elevations show that differences (ground minus EAARL-B surveyed elevations) over the entire submerged topography are small (root mean square error, RMSE, and median absolute error, M, of 0.11 m), and large differences (RMSE, between 0.15 and 0.38 m and similar M) are mainly present in areas with abrupt elevation changes and covered by dense overhanging vegetation. RMSEs are as low as 0.03 m over paved smooth surfaces, 0.07 m in submerged, gradually varying topography, and as large as 0.24 m along banks with and without dense, tall vegetation. EAARL-B performance is chiefly limited by point density in areas with strong elevation gradients and by LiDAR footprint size (0.2 m) in areas with topographic features of similar size as the LiDAR footprint.

© 2018 John Wiley & Sons, Ltd.

KEYWORDS: topobathymetric LiDAR; streambed bathymetry; performance of green LiDAR

Introduction

Investigations of morphology, habitat conditions and ecosystem function of streams depend on accurate maps of the bathymetry made with a resolution and extent appropriate for studying abiotic and biotic processes (Carbonneau et al., 2012). Stream morphology is shaped by sediment transport, which in turn is a function of the local (0.1 m scale) interaction between stream flow and streamed sediment (Wheaton et al., 2010; Maturana et al., 2013). The performance of multidimensional, numerical, hydrodynamic models in riverine systems directly depends on bathymetric accuracy and resolution (Lane et al., 2005; Pasternack et al., 2006; Tonina and Jorde, 2013), as does water and solute exchange between streams and streamed sediments (Tonina et al., 2015; Benjankar et al., 2016). Stream habitat monitoring (White et al., 2002), modeling (Pasternack et al., 2006; McDonald et al., 2010; Kammel et al., 2016), and management (Le Pichon et al., 2006) require a survey resolution comparable with the scale of organism sizes, their mobility and the entire extent of the river domain used throughout their full life cycle. Furthermore, highly mobile species, such as salmonids, may travel long distances, forcing us to map not just stream reaches or sections, but whole river networks (Fausch et al., 2002; Isaak et al., 2007). Thus, to support these diverse analyses and improve our understanding of riverine systems and processes, bathymetry surveys ideally must meet quite demanding survey criteria, including the ability to map entire stream segments (up to 104 m) with sub-meter spatial resolution and decimeter or even centimeter-scale vertical accuracy.

A variety of survey methods have recently been developed to map stream bathymetry and near-channel environments
(McKean et al., 2009c; Pasternack and Senter, 2011; Walther et al., 2011; Carbonneau et al., 2012). Airborne near-infrared (NIR) LiDAR (Cavalli et al., 2008) has been suggested for qualitative mapping of the main morphological features in shallow systems (water depths less than 0.5 m) because of the inability of the infrared beam to adequately penetrate water, such that to map the bed topography with NIR LiDAR the channel should be completely dry. Real-time kinematic (RTK) differential global position systems (DGPS) have been used for very high resolution surveying of subaerial topography of floodplains and other areas around streams. However, for surveying submerged topography, RTK DGPS is limited to wadable reaches for traditional data collection with surveyors carrying the instruments or with instruments installed on vehicles (Brasington et al., 2000). Its application by installing it on floating devices, coupling with an echo-sounder can overcome the limitation to survey non-wadable sections of streams (Kammel et al., 2016). Multi- and hyper-spectral image analyses (Marcus et al., 2003; Legleiter et al., 2009) have also been used to map water depths (rather than submerged topography), but their success depends on accurately defining local water optical properties and linking those to changes in water depth and turbidity of stream water. These passive optical surveys have generally been limited to short reaches (Legleiter et al., 2015). Multibeam sonar (Conner and Tonina, 2014) has been shown to be very effective in large, navigable streams and rivers where sufficient flow depth is available for power boats. Standard photogrammetric techniques have also shown promising results in defining submerged topography, but as with hyper-spectral image analysis, this approach requires site-specific, light-dependent relationship between depth and water color, which could depend on water turbidity and be influenced by overhanging vegetation, substrate type, and water surface characteristics (Carbonneau et al., 2006).

Following the success of near-infrared LiDAR in mapping terrestrial systems, topobathymetric airborne LiDARs, including the NASA Experimental Advanced Airborne Research LiDAR (EAARL) (McKean et al., 2009b), the Optech Scanning Hydrographic Operational Airborne LiDAR System (SHOALS) (Hilldale and Raff, 2007), the Coastal Zone Mapping and Imaging Lidar (CZML) (Tuell et al., 2010), the Laser Airborne Depth Sounder (LADS) Mk II and Mk 3 (Parker and Sinclair, 2012), the Riegl VQ-820-G (Mandlburger et al., 2011), the Optech Aquarius (Fernandez-Diaz et al., 2014), the Teledyne Optech Titan TopoBathy Sensor (Fernandez-Diaz et al., 2016) and Leica Aquatic Hydrography AB (AHAB) Chiroptera II and HawkEye II and III (Quadros, 2013), have recently been applied to stream systems (McKean et al., 2009a). All of these systems use a green wavelength laser, which has very low absorbance in water, rather than the near-infrared to map submerged topography (McKean et al., 2009a).

Whereas topobathymetric LiDAR performance in littoral systems is well documented and has become an established technique (Nayegandhi et al., 2006, 2009), this approach has not been systematically tested (McKean et al., 2009c, 2014; Mandlburger et al., 2015; Pan et al., 2015) and accepted in riverine systems (Kinzel et al., 2013; Legleiter et al., 2015). Nonetheless, recent investigations showed the advantages of bathymetric LiDAR in ecohydraulics applications that aim to understand the interaction between biotic and abiotic processes in lotic systems (Tonina and McKean, 2010; Tonina et al., 2011; McKean and Tonina, 2013). Previous studies on topobathymetric LiDAR, and specifically the original EAARL systems (McKean et al., 2009c, 2014), highlighted the main problems for application in riverine systems: (1) low resolution of the survey (0.3 points/m²); and (2) difficulty mapping deep pools and banks. A new EAARL system, EAARL-B (Wright et al., 2016), has recently been designed and developed for mapping both riverine and littoral systems to address these limitations. Similar to EARRL, EAARL-B flies at an operating altitude of 300 m and has a 240 m swath width and covers 40–80 km²/h. It uses a small 0.2-m footprint 400 μJ green laser, with 30 000 pulses per second, each 1.2 ns long and a 1 ns digitizer. The new system splits the light beam into three beamlets that are staggered to increase the survey resolution reaching a nominal density of 1 pt/m² with one flight line. Additionally, a fourth dedicated receiver merges the three beamlets to increase the energy of the returning waveform to improve water penetration in deep pools. This allows a wide range of water depths from 0 to 44 m with very clear water (1.5 secchi depth). The LiDAR is co-registered with a color-infrared (CIR) and color digital cameras, and each camera has an 80% image overlap. The new sensor potentially could become the first of a new generation of topobathymetric instruments capable of mapping entire river systems at the resolution needed for most ecohydraulics studies.

Here, we systematically test the performance of the EAARL-B system, which mapped the entire Lemhi River (Idaho, USA approximately 100 km stream length) in October 2013, in mapping the bathymetry of three morphologically distinct reaches of this stream by comparing results against very accurate sub-meter resolution (1.67 points/m²) RTK DGPS ground surveys. Previous studies used coarse ground survey (0.08 points/m²) to compare with EAARL data (0.3 points/m²) (McKean et al., 2009b, 2009c, 2014), a resolution that is not adequate for evaluating its ability to map microhabitat. These previous works also indicated that the EAARL system was not able to map banks because of the very low density in narrow areas (typically 1–2 m wide) where topographical gradients are strong (from 30 degree to vertical banks). These early works did not systematically analyze the performance of the LiDAR system for each morphological feature of a stream reach. Thus, we analyze the performance of the LiDAR in each one of a series of zones with uniform morphology and vegetation (e.g., banks, riffles, pools, runs, and tall versus short vegetation) to investigate the effects of topographic complexity and vegetation on survey accuracy. Our accuracy assessment includes both point measurements (which was not possible in previous work) and comparison of rasters created from point clouds.

**Methodology**

**Study area**

The Lemhi River Basin (3260 km²) is located near the Idaho-Montana border with an elevation range between 1585 and 2745 m (Figure 1). Annual precipitation ranges between 230 and 1016 mm and hydrology is snowmelt dominated, with 70% of precipitation falling as snow during the winter months between November and April. The Lemhi River is a gravel-bed stream with a bankfull width ranging between 8 and 20 m. Its minimum, average, and maximum daily mean discharges measured at the confluence with the Salmon River are 0.02, 7.11, 73.91 m³/s, respectively (Borden, 2014).

We selected three geomorphologically different channel reaches as test sites for quantifying the accuracy of the EAARL-B survey in different stream morphologies. Reach 1 has an average reach slope of 0.53% and includes a long bend with a deep pool (1.5 m), riffles and runs. It is approximately 160 m long and 10 m wide (Figure 2) with substrate dominated by gravel (2–64 mm) and cobbles (64–256 mm). Reach 2 is a morphologically complex reach, which includes pools (~1.6 m deep), riffles, runs, and vegetated point bars (Figure 3). It is
approximately 235 m long and 10 m wide with bed slope of 0.75%. Its substrate varies from fine (sand <2 mm) to very coarse (boulder >256 mm) sediment. Reach 3 is a 110 m long and 15 m wide simple, straight and plane-bed reach with 0.75% bed slope (Figure 4). It contains engineered and armored banks with a 45° slope. The substrate is dominated by cobbles and boulders. We also surveyed the surface of a paved road adjacent to reach 3 to test the EAARL-B system accuracy without the influence of on-site error sources, such as abrupt elevation changes, woody debris, vegetation, type of reflective substrate, and water depth commonly found in complex riverine environments.

Topographic and bathymetric surveys

We performed a high-resolution sub-meter (1.67 points/m²) topographic and bathymetric survey (hereafter ground-survey) during late summer (August 2013) low, wadable flows with a RTK DGPS (Leica GPS System 500) with an accuracy of 0.01 m horizontally and 0.03 m vertically. Each point was set to reach a global error of less than 0.03 m before being recorded; in some cases such as below vegetation, the accuracy was decreased to 0.06 m in order to record the point. Overall, 37 480, 21 727 and 20 966 topographical points were collected at reach 1, 2 and 3, respectively. Our ground-survey included the following features: channel bathymetry, point bars, islands, bank, top of bank, water edge, water surface elevation, adjacent floodplain and a section of a paved road. The ground-survey data were separated into four main geomorphic areas: channel, banks, floodplain, and paved road. Each geomorphic area was further divided into homogenous morphological sub-areas, including runs, pools, vertical and sloped banks, and areas with dense tall and short vegetation to quantify the capability of the EAARL-B sensor in different contexts (Table I). Some portions of floodplain near gently sloping banks in reach 1 and 2 were characterized by intense livestock movements, which left deep approximately 0.2 m closely-spaced (0.2 m), irregular holes of 0.3 m diameter. We ground surveyed them to see if the EAARL-B system could characterize this feature.

EAARL-B data were collected in a 3-day mission in October of the same year as the ground survey, during clear water conditions. Discharge (measured at US GS gage 13305000 LEMHI RIVER NR LEMHI ID, which is 1 km downstream of site 3) remained constant around 2.8 m³/s till the end of September (during ground survey) and then increased to 6.8 m³/s (during EAARL-B flight) at the beginning of October as some diversion dams were closed. High flows typically are around 25 m³/s at...
this location. Similarly, the Idaho Department of Water Resources gage stations at Cotton Lane (just upstream of sites 1 and 2) and at McFarland campground (upstream of site 3 but downstream of site 1 and 2) show a similar increase in discharge but do not report any floods or significant flow events between the two surveys (Figure 5). We believe that this change in discharge may have caused some limited change in microtopography, e.g. limited scour and deposition but no major elevation changing which can affect the overall topography at the meter scale. During the flights, water depths ranged from 0 m along gently sloping banks to 1.6 m in a deep pool. Both ground- and EAARL-B surveyed point coordinates were in Universal Transverse Mercator Zone 12 North for their horizontal location and their elevation was in ellipsoidal heights referenced with the IGS08 vertical datum.

EAARL-B data analysis

All data were processed in the Airborne Lidar Processing System (ALPS) using both the bathymetric and bare earth modes (Nayegandhi et al., 2006, 2009; Nagle and Wright, 2016). Both modes interrogate the full waveform of each laser reflection and attempt to identify energy returns from the
ground surface. The bathymetric mode also identifies the air–water interface, calculates the refraction of each laser pulse at that interface, and reduces the velocity of the laser energy during that portion of the travel path that is under water.

Following ALPS processing, the bare earth and bathymetric data from all three days were combined to increase the data density and eliminate any gaps in coverage. EAARL-B and ground surveyed points were overlaid over Google Earth aerial images. We checked the images and ground points co-registration by testing the match between the road center line of the aerial images and the ground points surveyed along the road centerline. This verification showed very good match, which allowed us to hand digitize the delineation of (a) each bank above the water level, (b) the bed of the channel and all other submerged areas, and (c) the floodplain exposed above the waterline. Bare earth data were extracted from all areas above the water and bathymetric data from all submerged areas. All data were first analyzed to separate between ground and non-ground points with LAS Tools (Isenburg, 2012) (where LAS is binary file format for efficient exchange of 3-dimensional point cloud data sets among users). Successively, the data set was filtered with the above ground level analysis tool of Quick Terrain Modeler (QTM, Applied Imagery) to remove laser returns from above the ground surface. The final filtered point cloud had a density of 1 point/m² that was much higher than the original EAARL (0.3 points/m²) and was used to create 1 m and 0.5 m DEMs with the Natural Neighborhood interpolation method (Sibson, 1981).

### EAARL-B data accuracy analysis

EAARL-B accuracy was quantified by comparing EAARL-B and ground surveys using two methods: raster-to-raster (hereafter raster) and point-to-point (hereafter point) based comparison (Skinner, 2011; Woodget et al., 2015). The raster comparison computes the elevation difference between DEMs for each individual pixel. This method is sometimes referred to as a difference of DEMs (DoD), and might be affected by interpolation errors in areas where survey points are missing or sparser than the raster resolution, and/or by smoothing due to assignment of a single elevation value to an entire cell. The point comparison is made between elevations of closely coincident points in both ground and LiDAR surveys. While avoiding the

| Table 1. Comparison between EAARL and EAARL-B technical specifications with bold specifications highlighting noticeable improvements of EAARL-B over EAARL |
|-----------------|-----------------|-----------------|
| **EAARL**      | **EAARL-B**     |
| Total system weight | 113 kg          | 113 kg          |
| Maximum power requirement | 28 VDC at 26 A  | 28 VDC at 26 A  |
| Nominal survey altitude 300 m (AGL) | 300 m (AGL)    | 300 m (AGL)    |
| Nominal survey flight speed | 50 m/s       | 50 m/s       |
| Inertial navigation system | 200 Hz         | 200 Hz         |
| Precision kinematic | GPS 2 Hz       | GPS 2 Hz       |
| Laser wavelength | 532 nm          | 532 nm          |
| Laser energy | Up to 70 μJ/pulse | 400 μJ        |
| Laser pulse length | 1.2 ns (full width at half-maximum) | 1.2 ns |
| Laser field-of-view | 1.5–2 mrad     | 1.5–2 mrad     |
| Laser beam divergence | 0.5 mrad     | 0.5 mrad     |
| Laser spot diameter | from 300 m AGL 20 cm | from 300 m AGL 15 to 20 cm |
| Laser pulse frequency | 3–10 kHz      | 30 kHz        |
| Raster scan rate | 25/s           | 25/s           |
| Number of Waveform channels | 1            | 4              |
| Digitizer temporal resolution | 1 ns (14.9 cm in air, 11.3 cm in water) | 1 ns |
| Data swath width | from 300 m AGL 240 m | from 300 m AGL 240 m |
| Single pass sample spacing | 2 m cross-track, 2 m average along-track | 1 m cross-track, 1.5 m average along-track spacing |
| Horizontal accuracy | <1 m RMSE | <1 m RMSE |
| Water depth range (~ 1.5 x secchi depth) | 0 to 28 m | 0 to 44 m |
| Co-registered digital cameras | CIR and color | CIR and color |

![Figure 5.](image-url)
interpolation issues of the DoD method, the point-based approach might include errors due to real changes in elevation within the distance between compared points, because points are never at exactly the same location in both the ground and LiDAR surveys. For example, elevations may change rapidly between top and bottom of a boulder or on sloped surfaces like banks or the sides of pools. Additionally, ground-survey data are point measurements, whereas EAARL-B data are spatially averaged elevations within the footprint area, typically a circle of 0.2 m diameter.

In the DoD analysis, we used 0.5 (hereafter sub-meter, SMR) and 1 m (hereafter meter, MR) resolution raster to analyze accuracy at the scale of individual DEM grid cells. These high-resolution cell sizes are typically used in hydrodynamic modeling for quantifying aquatic habitat quality (Gard, 2009; Benjankar et al., 2014; Kammel et al., 2016). Raster surfaces are interpolated with the Natural Neighborhood interpolation method and comparison was restricted to the extent of the ground-survey areas so the comparisons between the EAARL-B and ground-survey raster had similar spatial extents.

The point-based analysis is restricted to ground points that have at least one EAARL-B point within a 0.1 m radius, which coincides with that of the laser footprint. We selected the closest EAARL-B point when more than one was available. The 0.1 m distance threshold minimizes the error due to elevation variability between the two points being compared and still preserves a large sample of points for statistical analysis. Typically, we had only a single point and rarely two points within the 0.1 m radius, which prevented us from studying the statistics of point cloud elevations, such as mean and variance, within the EAARL-B footprints.

The errors between the EAARL-B and ground-survey elevations are reported with their frequency distribution, median absolute error (M, the residual that is half way through the ordered data set), mean error (or bias, B), root mean square error (RMSE), and the correlation coefficient ($R^2$) for the regression of EAARL-B and ground-survey elevations for each morphologic area and sub-area. We calculated errors, residuals, $Res$, as ground-survey elevations, $z$, minus EAARL-B-surveyed elevations,

$$Res_i = z_{g,i} - z_{L,i}$$

(1)

$$M = median(|Res|) = \begin{cases} \frac{1}{2} (Res_{n+1} + Res_1) & \text{if } n \text{ is odd} \\ \frac{1}{2} (Res_1 + Res_{n+1}) & \text{if } n \text{ is even} \end{cases}$$

(2)

$$B = \frac{1}{n} \sum_{i=1}^{n} Res_i$$

(3)

where $n$ is the total number of points, $i$ is the $i$th point and subscripts $L$ and $g$ stand for EAARL-B and ground survey, respectively. Therefore, positive bias indicates a higher elevation value for ground-survey data than for the EAARL-B data. Low magnitudes of RMSE, $M$, and $B$ and high values of $R^2$ indicate good agreement between surveys. Mean error helps to identify systematic bias between the two surveys, whereas the median absolute error is insensitive to the effect of large residuals (outliers). Thus, $M$ is more robust to outliers than RMSE, because RMSE squares the residuals before averaging them giving a higher weight to large errors than $M$. The comparison between $M$ and RMSE values helps to identify the presence of large errors within the residual distribution.

We report the results in three sections, one for each of three main geomorphic features: channel, banks and floodplain. Additionally, we analyzed the accuracy of the system by surveying a nearby paved road. The channel area covers the streamed section and is fully submerged. The bank area is a narrow band, partially submerged, between the bottom and top of the bank, and thus is where data analysis transitions between the bathymetric and terrestrial algorithms. The floodplain area is a terrestrial environment.

### Results

EAARL-B residuals did not increase significantly with water depth as also observed by others (Fernandez-Diaz et al., 2014; McKean et al., 2014) (Figure 6(a)). The slope and coefficient of correlation ($R^2$) of the regression line fitted to all submerged data between residuals ($Res$) and water depths were 0.18 (which is almost flat) and 0.13, respectively (Figure 5(a)). The large errors were at or near vertical banks, below overhanging vegetation (willow Channel), where a small horizontal difference in surveyed points can cause a large error in elevation (Figure 6(b)).

Point density of the filtered point cloud of EAARL-B survey ranged from 0.49 to 1.26 points/m$^2$ with an averaged point density of 0.9 point/m$^2$ (Table II). The point density did not correlate with any errors (Figure 7), because it was adequate to define macro-bedforms, such as pools, riffles and bars. Surprisingly, high densities near and larger than 1 point/m$^2$ were in areas with large errors, e.g. vertical banks and pools (Table II), conversely low density (0.49 points/m$^2$) was on the paved road and the fine run where errors were small (Figure 7). Pools in the surveyed reaches are narrow and typically near banks, resulting in strong elevation changes in the cross-stream direction. These areas have strong elevation gradients, suggesting that point density needs to be large to capture strong topographic gradients well. This result suggests that a point density of 1 point/m$^2$ is adequate to survey in-channel streambed

![Figure 6](https://example.com/figure6.png)

**Figure 6.** Error residual for sub-meter DEM for submerged location as a function of water depth for the morphologically complex Reach 2 (a) shown for all points grouped in channel and on banks (b) shown for each submorphologic unit. [Colour figure can be viewed at wileyonlinelibrary.com]
features in small streams like the Lemhi. However, higher density is needed in areas characterized by strong elevation changes, like banks and steep pool sides.

### Paved surface

Paved surfaces, like roads, provide an ideal area to test the minimum expected error of EAARL-B because of their gentle change in elevation, lack of submergence (completely dry), uniform reflectivity of the asphalt, and lack of vegetative cover. The largest errors, $M = 0.05$ m, $B = -0.05$ m and RMSE $= 0.04$ m, are from the point analysis (Figure 10). These large errors could be due to the low number of points, only four, within a 0.1 m radius. Conversely, both sub-meter and meter DEM have similar errors: RMSE is $0.03$ m, which is similar to many DGPS surveys, and M and B are even smaller.

## Channel

The overall RMSEs for the channel vary from $0.10$ to $0.15$ m for both raster and point based comparisons (Figures 8, 9 and 10) regardless of substrate size (from sand to boulder, $>256$ mm) and morphological complexity (straight featureless and meandering pool–riffle reaches). Errors are a bit smaller for the sub-meter than meter resolutions and for the raster than for the point comparison. Both bias (mean error) and median absolute error values are always lower than RMSEs and residuals do not vary systematically with depth (Fernandez-Diaz et al., 2014; McKean et al., 2014) or morphological unit. Correlations between EAARL-B and field-survey elevations are strong ($R^2 > 0.9$) regardless of comparison and raster resolution, except for reach 3.

The RMSE in deep pools varies from $0.16$ to $0.19$ m for reach 2 and 1, respectively (Figures 9 and 8). The largest error (RMSE of $0.38$ m) is estimated in a deep pool (reach 2) below
overhanging vegetation adjacent to a bank with tall dense vegetation (Figure 9). This area has the highest bias among all sub-morphological categories and very low correlation values (R²). These errors are mostly due to the sparse and limited ground-survey points that were collected in this area. Conversely, errors are low (e.g. RMSEs of 0.07 and 0.09 m), for both SMR and MR resolutions in the fine shallow run with median grain size of 30 mm. The RMSEs for the coarse steep run with cobble (64–256 mm) and boulder (>256 mm) size substrates are 0.08 m and 0.09 m for SMR and MR. These areas have low bias, however, coarse steep run has good correlation (R² > 0.68), whereas the fine shallow run has very low correlation (R² < 0.32).

As shown in previous studies (McKean et al., 2009c; Skinner, 2011), errors in the channel increase near and at the banks in the raster analysis because of the smoothing of banks as a result of interpolation. This outcome could be attributed to low point density relative to the change in slope and the EAARL-B footprint size (0.2 m diameter), within which elevations are

Figure 7. RMSE averaged among all the sub-areas plotted against the averaged point density within each sub-area. [Colour figure can be viewed at wileyonlinelibrary.com]

Figure 8. Frequency distribution (%) of errors, RMSE, Median absolute error (M), Bias (B), Number of samples (N) (i.e. N refers to the number of cells used in the raster comparison and number of points used for the point comparison) and coefficient of correlation (R²) for Reach 1. Blue, red and green bars represent sub-meter (SMR), meter (MR), and point-by-point (P) comparisons, respectively. [Colour figure can be viewed at wileyonlinelibrary.com]
averaged. The same effect is present in all three reaches (Figure 11). To remove this effect, we trimmed the channel bathymetry to include only the area at least 2 m away from either bank and defined the in-channel sub-area in reach 3 (Table II and Figure 10). All errors for the in-channel sub-area are noticeably lower than any other in-channel sub-area for reach 3 in both comparisons (Figures 8, 9 and 10). This interpretation is further supported by lower errors for the sub-meter resolution than for the meter resolution.

Pools in reach 1 and 2 are very close to the channel banks (Figures 2 and 3), and typically very narrow (2–3 m wide) (Figure 12(b), (c)). The relatively high RMSEs of the pool sub-area highlight the effect of the banks on the raster comparison. This morphological unit also is characterized by the highest bias between 0.13 and 0.17 m in reach 2 and 1, respectively. The point-based comparison error is smaller in reach 1 (RMSE of 0.16 m) than 2 (RMSE of 0.21 m) because reach 1 lacks tall, dense riparian vegetation, which is present along the banks of reach 2. This result shows the impact of willow vegetation. The effect of overhanging riparian vegetation on channel bathymetry is better shown by the comparison between surveys in the channel under the willow (Table II: willow-channel) area (reach 2). This zone has the largest errors and lowest correlation among all geomorphic sub-areas within the channel for the raster comparison (RMSEs of 0.33 and 0.38 m with similar bias and median absolute errors for sub-meter and meter resolutions, respectively) (Figure 9). EAARL-B data were influenced by tall dense willow vegetation, as well as the effects of the bank. Errors are lower for the point comparison, but the number of samples (point comparison) is comparatively low (six points).

The comparison between EAARL-B and measured ground points in the side channel morphology quantifies the ability to detect small channels. This is a small, narrow (less than 5 m wide), and shallow (less than 0.5 m deep), spring-fed reach. All errors are comparable with the overall channel error of reach 1 for raster comparison. The ability of the sensor to detect complex topographical variations is also illustrated in Figure 12, which reports four cross-sections from reach 2. These span from simple and featureless streambed areas (cross-section 1) to a very complex section with a vegetated bar in the middle and a deep pool (cross-section 3).

Bank

All errors (e.g. RMSE ranged between 0.16 and 0.23 m), except bias, were larger for channel banks than for the other geomorphic areas (i.e. channel and floodplain) except those below...
dense overhanging vegetation in all three reaches regardless of the type of comparison (Figures, 8, 9 and 10). Therefore, the largest errors occurred in areas with abrupt elevation changes (Figure 11). Nevertheless, the estimated RMSEs for the channel banks are lower than the values (0.4–0.73 m) reported in previous studies (McKean et al., 2009c; Skinner, 2011), which also observed the smoothing or widening of channel banks. The large errors at channel banks stem from the combined effects of abrupt changes in elevation, EAARL-B geo-location errors, the 0.2 m EAARL-B footprint size and interpolation error in generating a raster from the field- and LiDAR-survey points for the raster comparison (McKean et al., 2009c). Although the point density of EAARL-B points over the surface of channel banks was higher than other locations, it was still not sufficient to capture the strong gradients in the steep banks, which resulted in high RMSEs. Figure 12(b) and (d) show this issue, although the LiDAR tracks the banks very well, because of the strong slope at the vertical bank the elevation error is around 0.2 m.

Figure 10. Frequency distribution (%) of errors, RMSE, Median absolute error (M), Bias (B), Number of samples (N) (i.e., N refers to the number of cells used in the raster comparison and number of points used for the point comparison) and coefficient of correlation ($R^2$) for Reach 3. Blue, red and green bars represent sub-meter (SMR), meter (MR), and point-by-point (P) comparisons, respectively. [Colour figure can be viewed at wileyonlinelibrary.com]

Figure 11. Maps of the error distribution for raster comparison based on sub-meter resolution. Positive errors indicate a higher elevation for the ground-survey data than for the EAARL-B data. [Colour figure can be viewed at wileyonlinelibrary.com]
Errors are lower along sloped than vertical banks (Figure 12(d)). However, even the sloped banks are very narrow areas, whose widths are between 1 and 3 m and could have changes of elevations as high as 1 m near the edge of the channel.

Floodplain

Floodplains are relatively smooth surfaces with minimal abrupt elevation changes and consequently have low errors (e.g. RMSEs ranging from 0.06 to 0.08 m with similar median absolute error and almost zero bias) for the raster comparison (Figures 8, 9 and 10). Point comparison has slightly higher errors (e.g. RMSEs ranging from 0.08 to 0.14 m). Dense, tall vegetation present in reach 2 and 3 contributes to the larger RMSEs for the floodplain in these reaches than reach 1. Both M and RMSEs were less than 0.06 m for the short vegetation, whose height was approximately 0.2 m, and on terraces for the raster comparison and less than 0.1 m for the point-based comparison. These errors are slightly higher for the livestock-stamped floodplain than for the short vegetation floodplain (reach 1 and 2). Livestock-stamped floodplains, present mostly in reach 2, have noticeable, closely-spaced, irregular deep holes, which were captured during the ground-survey, but less effectively by EAARL-B, because these features have dimensions similar to the EAARL-B footprint (McKean et al., 2009c).

Discussion

Surprisingly, none of the error metrics used in this study increased with decreasing EAARL-B point density regardless of comparison (point and raster) (Table II). However, errors correlated with morphological units, with large errors in areas with strong changes in elevation such as banks. These areas were also those with the highest EAARL-B point density of 1.26 points/m². This density was not sufficient to map the changes in elevation with EAARL-B because changes occur noticeably within spatial distances shorter than the spacing among EAARL-B points. EAARL-B point density was also high within areas with tall vegetation where RTK DGPS was less effective due to accessibility. This underscores the ability of airborne topobathymetric LiDARs to survey even vegetated areas. Our EAARL-B survey was during autumn when most trees were leafless, which improved bare earth detection but did not help with RTK DGPS survey.

Error analysis over a paved road provides the EAARL-B system accuracy without the influence of error sources, such as abrupt elevation changes, woody debris, type and density of vegetation, type of reflective substrate and water depth commonly found in the complex riverine environment. Many of these sources of error typically are found simultaneously along banks. Additionally the uniform elevation over the paved road minimized the error due to the horizontal uncertainty of EAARL-B surveyed points. Assessing horizontal accuracy is very difficult because land surface lacks benchmark points, which are well-defined and easy to survey by the system, such that it is very difficult to disentangle vertical and horizontal accuracies. Previous work on EAARL suggests this error is less than 1 m (McKean et al., 2009c).

As expected, all error indices are lowest for the road surface at reach 3 among all the morphological sub-areas regardless of comparison methods, whereas channel banks have the largest error indices (Figures 8, 9 and 10). The largest errors are found in areas with tall, dense vegetation, such as willow. EAARL-B signals might have been scattered over this area, or the filtering (specifically for vegetation) algorithm used to detect bare-earth might require improvement (Skinner, 2011). RMSEs are noticeably smaller for point than for raster comparisons for the areas under dense vegetation (Willow-channel), although there are only six ground-survey points, which are within 0.1 m from an EAARL-B survey points (Figure 9). This lack of coverage resulted in erroneous interpolated cells in the raster and in few points for comparison in the point-based analysis.

Ground-survey points are few (compared with EAARL-B survey) in such areas because of dense bushes, which limited the access to those locations. A crew of three people was required to collect ground points in areas with dense, tall vegetation. Two people opened the branches to allow the third person to reach a spot and have clear sky view for recording the location with RTK DGPS. Below dense vegetation, sky view could be limited and thus DGPS position lock might not have been possible and, when possible, might be highly uncertain.

On the floodplain with livestock-stamped ground, some error may have also come from changes due to animal movements or deterioration of those features over time. Although, the landlord told us that no livestock returned to the place between the surveys, big game and other wild animals might have roamed that area. Thus, some of the observed errors may have come from real changes rather than random instrument errors.
Median absolute errors are considerably lower than RMSEs for all geomorphic areas regardless of comparison method. This indicates that some extreme elevation errors (outliers) in the EAARL-B survey may increase the values of RMSE. The good performance of the EAARL-B system also is indicated by strong correlations between EAARL-B and field-survey data ($R^2$>0.9) for the raster comparison in all three reaches for the majority of channel and floodplain areas. As expected, banks have the lowest $R^2$ compared with other geomorphic areas. The low $R^2$ value of 0.32 for the fine shallow run in reach 2 (Figure 9), despite its low M, B, and RMSE values, is due to the narrow range of elevation variability in the area (less than 0.24 m) and close to the RMSE of 0.07 m. These two effects strongly suppress any slope in the regression equation fit to the trend.

Our results suggest that EAARL-B could overcome some of the limitations of previous topobathymetric LiDARs, whose RMSE ranged from 0.14 to 0.52 m (Kinzl et al., 2013; Fernandez-Diaz et al., 2014). These previous works studied the performance of topobathymetric LiDAR in wider and deeper rivers. We argue that the new EAARL-B would work effectively in those systems, as long as there is enough water clarity, because bathymetric changes are typically more gradual in large systems than in small streams like the Lemhi river. Our results for the entire streambed show uncertainty (RMSE=0.11 m) between ground and LiDAR surveys larger than those reported by Mandilburger et al. (2015), who quantified standard deviation (similar to RMSE) of 0.04 m for the river bed. If that section is similar to our fine-shallow run then the results are comparable, suggesting that topobathymetric technology is starting to become a more mature system for examining riverine environments.

Our analyses suggest that the EAARL-B sensor effectively detects large morphological structures, including ripples, pools, and runs. However, it may not detect the exact position and elevation of cobble-size structures, whose dimensions are smaller than the LiDAR footprint (0.2 m), but the presence of cobbles results in a rough streambed bathymetry. Thus, the presence of these features is partially captured due to the added variability in bed elevation. Similarly, features like the cattle-stamped holes, whose size is similar to the EAARL-B footprint, were not effectively detected by the LiDAR survey. EAARL-B cannot detect subtle topographic features, with amplitudes within its uncertainty of an average 0.11 m and because of low point density to detect the increase in roughness near those features. This suggests salmonid egg nests, called redds, cannot be surveyed with current bathymetric technology because their size is smaller (Crisp and Carling, 1989; Bjornn and Reiser, 1991) than the detection limit of EAARL-B or any other topobathymetric LiDAR. Thus, EAARL-B surveys can be used to detect and study habitats formed by large bedforms including pools, ripples, runs, bends, and gentle-sloping banks. However, it cannot be used to study micro-habitats formed by single features whose sizes are smaller or similar to the EAARL-B horizontal resolution (meter scale) and vertical elevation uncertainty (0.11 m within the streambed and 0.23 m on the banks), such as large cobbles or logs (Crowder and Díplas, 2000; Tullos et al., 2016).

We argue that these limitations of the EAARL-B are common to any topobathymetric LiDAR and mainly stem from system horizontal resolution and footprint size, which both depend on the flying height from which the instrument is deployed. The accurate detection of small bedforms and vertical banks would require a slower and lower flying system, such as a small drone, with high stability. Advances in this direction have been shown in some recent publications by using an optical sensor and a structure from motion (SfM) technique (Woodget et al., 2015; Carrivick et al., 2016; Dietrich, 2016). Although a helicopter-borne system is potentially an attractive solution, it would not provide the vertical stability necessary for accurate vertical measurements by LiDAR and would be extremely expensive. Besides these limitations, this analysis shows that the EAARL-B system provides bathymetric survey data within the channel similar, if not better, than other survey methods such as multibeam sonars (Conner et al., 2009), which have been used for deep navigable rivers (4–5 m deep), often with low water clarity, where a LiDAR based system may not be an effective surveying tool. Nevertheless, the current study (performance of EAARL-B) was focused on a shallow gravel-bed river, where water turbidity is typically not an issue. Different survey techniques may have different resolutions and accuracies. The latter would restrict the possibility to merge areas covered with different techniques to provide continuous coverage. Because the uncertainty of the EAARL-B (0.11 m) is similar to that of the multibeam survey (0.15 m), these surveys could be merged to provide continuous coverage without losing accuracy. The merged DEM from both techniques could cover the entire riverscape from streams to rivers and potentially provides a continuous map of riverine systems, an almost complete census, as advocated by Pasternack and Senter (2011) and most recently by Kammel et al. (2016). This approach would allow studying the riverine system at a scale and extent needed to improve water management and sustainability of water resources and ecosystems (Rice et al., 2010; Parasiewicz et al., 2013; Brooker, 2016). Bathymetric information would allow simulating stream flow under different water management conditions to predict the impact of human activity on aquatic habitat (Li et al., 2015a, 2015b) and monitoring over time to advance our understanding of river evolution (Brasington et al., 2000; Wheaton et al., 2010; Vericat et al., 2014). Our next investigation will focus on quantifying the cascading effects of the bathymetric errors quantified in this work on hydraulic quantities modeled with numerical modeling.

Conclusions

Our results show that most of the difference between EAARL-B and ground-surveyed topography is associated with EAARL-B point density and footprint size only in areas of strong topographic changes, once the proper algorithm (bathymetric or terrestrial) to analyze the data is selected. EAARL-B accuracy along gradually varying submerged topography (RMSE=0.06 m) is comparable with that for paved terrestrial surfaces (RMSE=0.03 m). Large errors on the order of 0.24 m are measured at vertical and steep banks, where EAARL-B point density was insufficient to characterize these features accurately. The errors reduce to 0.11 m for gently sloping banks. EAARL-B has RMSEs similar to its footprint radius (0.1 m) when the mapped area has topographical features with dimensions similar to the laser footprint. Overall, RMSEs within the channel are around 0.11 m for both complex and plane bed reaches and independent of streambed grain size, from sand to cobbles.

These results suggest that EAARL-B can effectively survey streambeds and their macro-bedform such as pools, ripples, runs, bends, and gentle-sloping banks at an unprecedented resolution and extent. These will provide the opportunity to map riverine systems without the need to sample them or upscale local information from ‘representative reaches’ to the entire watershed. However, its horizontal spatial resolution (meter scale) and vertical uncertainty (centimeter scale) prevent its use to detect local bed features, such as cobbles and redds, and structures with a comparable size (0.2 m) to the sensor footprint. Consequently, the sensor does not provide the resolution...
needed to study micro-habitat characteristics and processes. In our future research, we will quantify how these mapping errors and resolution limitations propagate into numerical hydraulic models to predict meter scale hydraulic quantities such as flow velocity, depth, and shear stress, which are commonly used to study riverine environments.

Acknowledgements—This research was partially supported by the Idaho Office of Species Conservation (grant number LEM023 16), by Bonneville Power Administration (Bonneville Power Administration’s Lemhi River Restoration Project, Project # 2010-072-0), by the US Forest Service (grant number 16-CR-11221634-049) and by the University of Idaho Office of Research (seed grant 2011). We also thank Quantitative Consultants Inc. for logistical support and for facilitating access to land and streams. We thank Mr. Merrill Beyerel for access to his property, which made this research project possible. We thank Frank Gariglio, Ryan Carnie, Roberto Marivela for helping with data collection. The comments from the associate editor, two anonymous reviewers and Dr. Lee Harrison on the earlier version improved the manuscript. Any opinions, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the supporting agencies.

References


