Estimating vegetation biomass and cover across large plots in shrub and grass dominated drylands using terrestrial lidar and machine learning

Anderson, Kyle E.[1], Glenn, Nancy F. [2]*, Spaete, Lucas P.[2], Shinneman, Douglas J.[3], Pilliod, David S. [3], Arkle, Robert S. [3], McIlroy, Susan K. [3], Derryberry, DeWayne R. [4]

[1]At Present:
The Great Basin Institute
6640 Lockheed Drive
Redding, CA 96002
Work performed at:
Idaho State University
Department of Geosciences
Pocatello, Idaho 83209
kyle.erick.anderson@gmail.com

[2]Boise Center Aerospace Laboratory
Boise State University Department of Geosciences
1910 University Drive
Boise, ID 83725-1535
lucasspaete@boisestate.edu

[3]U.S. Geological Survey, Forest and Rangeland Ecosystem Science Center
970 Lusk Street
Boise, ID 83706
dshinneman@usgs.gov, dpilliod@usgs.gov, rarkle@usgs.gov, smcilroy@usgs.gov

[4]Idaho State University Department of Math
921 S. 8th Ave., Stop 8085
Pocatello, ID 83209-8085
derrdewa@isu.edu

*corresponding author
nancyglenn@boisestate.edu; 208.426.2933

Abstract

Terrestrial laser scanning (TLS) has been shown to enable an efficient, precise, and non-destructive inventory of vegetation structure at ranges up to hundreds of meters. We developed a method that leverages TLS collections with machine learning techniques to model and map canopy cover and biomass of several classes of short-stature vegetation across large plots. We collected high-definition TLS scans of 26 1-ha plots in desert grasslands and big sagebrush shrublands in southwest Idaho, USA. We used the Random Forests machine learning algorithm to develop decision tree models predicting the biomass and canopy cover of several vegetation classes from statistical descriptors of the aboveground heights of TLS points. Manual measurements of vegetation characteristics collected within each plot served as training and validation data. Models based on five or fewer TLS descriptors of vegetation heights were developed to predict the canopy cover fraction of shrubs (R^2 = 0.77, RMSE = 7%), annual grasses (R^2 = 0.70, RMSE = 21%), perennial grasses (R^2 = 0.36, RMSE = 12%), forbs (R^2 = 0.52, RMSE = 6%), bare earth or litter (R^2 = 0.49, RMSE = 19%), and the biomass of shrubs (R^2 = 0.71, RMSE = 175 g) and herbaceous vegetation (R^2 = 0.61, RMSE = 99 g) (all values reported are out-of-bag). Our models explained much of the variability between predictions and manual measurements, and yet we expect that future applications could produce even better results by reducing some of the methodological sources of error that we encountered. Our work demonstrates how TLS can be used efficiently to extend manual measurement of vegetation characteristics from small to large plots in grasslands and shrublands, with potential application to other similarly structured ecosystems. Our method shows that vegetation structural characteristics can be modeled without classifying and delineating individual plants, a challenging and time-consuming step common in previous methods applying TLS to vegetation inventory. Improving application of TLS to studies of shrub-steppe ecosystems will serve immediate management needs by enhancing vegetation inventories, environmental modeling studies, and the ability to train broader datasets collected from air and space.

Keywords: rangelands; carbon; point cloud; lidar; biomass; classification; land cover; remote sensing; machine learning; vegetation type; Structure from Motion (SfM)

1. Introduction

Sagebrush steppe, a shrub- and bunchgrass-dominated biome occupying 47 million hectares of semiarid rangelands in the western United States (Bukowski and Baker 2013), is rapidly being degraded, fragmented, and lost. Many factors contribute to the loss of sagebrush steppe ecosystems, but the greatest driver is the “grass-fire cycle” (D'Antonio and Vitousek 1992) where wildfires promote invasion by nonnative grasses and forbs, which in turn increase the rate and severity of future fires. In many cases, the new regime of frequent wildfire causes the total replacement of sagebrush ecosystems by a new steady state of nonnative pyric grassland (Knick 1999, Balch et al 2013). Deleterious impacts of this shift include increased wildfire hazard and reduced soil retention, forage quality, and biodiversity (Brooks et al. 2004, Rowland et al. 2011, Balch et al. 2013, Ripplinger et al. 2015). One example of the urgent threat to the sagebrush biome is the rapidly changing composition of the 195,000 ha Morley Nelson Snake River Birds of Prey National Conservation Area (NCA) in southwest Idaho, where only about a third of the area is occupied by native shrub communities due to the effects of numerous recent fires (USDI BLM 2008). Improved methods to conserve and restore sagebrush steppe ecosystems are an urgent topic of research (e.g. Pyke et al. 2015). The need for accurate, scalable, and practical methods of vegetation inventory is common to a variety of sagebrush management inquiries, including habitat monitoring, wildfire risk evaluation and behavior modeling, and vegetation treatment evaluation. Hand-measured metrics, such as transect or frame-based measurements of biomass and structure (i.e., cover, density, height), have historically filled this role and remain the most common source of data in sagebrush habitat inventories (e.g. Reiner et al. 2010). Biomass and cover are indicators of productivity and related ecological processes, as well as management processes such as fuel treatments and grazing resources in sagebrush steppe ecosystems (e.g. Davies and Bates, 2010, Pyke et al., 2014). Manual sampling methods of biomass and cover provide precise measurements, but necessitate collections that are highly localized and logistically difficult across vast, remote, and heterogeneous shrubland landscapes. Airborne and spaceborne optical remote sensing provide broad and continuous datasets which are useful for classifying dryland vegetation (e.g. Homer et al. 2012) although most do not collect the necessary structural information to estimate aboveground biomass. The use of airborne laser scanning (ALS) to remotely sense dryland vegetation structure has also been widely developed (e.g. Streutker and Glenn 2006, Mitchell et al. 2011), although ALS sensing encounters difficulty accurately sampling the full structure of low biomass herbaceous plants (e.g. Glenn et al. 2016, Li et al. 2017).

Terrestrial laser scanning (TLS) provides a data source intermediate between precise and localized manual measurements and spatially extensive, coarser measurements from aerial and satellite platforms. Often consisting of a rotating scanner mounted on an elevated platform, TLS instruments enable speedy collection of point clouds representing the 3-D position of the surfaces and objects in the scanner's field-of-view, including herbaceous vegetation. The instrumentation error of TLS measurements is usually negligible, and very high density collections (centimeter to a few centimeters resolution) at ranges up to hundreds of meters are often possible at little logistical expense (Shan and Toth, 2008, Vosselman and Maas, 2010). Although a TLS instrument samples its full field-of-view up to a specified range, it is usually unable to sample objects or surfaces which are behind another object from the instrument's perspective, causing "shadows" of space without points (aka occlusion) (Cifuentes et al., 2014). A common technique is to collect and combine point clouds from several positions around a target area, raising the probability that any given space is in the field-of-view of at least one scanning location (e.g. Cooper et al., 2017, Van der Zande 2008, Wilkes et al., 2017). However, achieving complete sampling coverage of surfaces across large vegetated sites may be impractical when using a field-portable tripod base for the scanner. An approach to mitigate vegetation-caused occlusion is to elevate the instrument (e.g. using vehicle-mounted masts or high points in terrain), so shadowing in the point cloud occurs mostly beneath objects and topmost surfaces are sampled consistently (as with ALS, e.g. Vierling et al. 2013). Systemic irregularities (usually minor) in TLS point cloud density and the positional precision of samples also occur where a collection encompasses a variety of ranges, because the beam diameter of laser pulses (or spot size) and the Euclidean distance between points both increase exponentially with range from the scanner. A review of current TLS technology, workflows, and applications related to the discussion above are provided in Telling et al. (2017). Simple structural traits of plants (such as height) may be measured directly using TLS point clouds, while other ecologically important traits may be predicted by proxy measurements. When scanning targets at consistent ranges, TLS measurement of targets' reflectance at the laser wavelength ("intensity") has been shown to be a useful spectral sensor (e.g. Seielstad et al. 2011, Olsoy et al. 2014b).

TLS has been demonstrated to efficiently replace manual sampling of a variety of common metrics in forested ecosystems, including tree stem count, basal area, biomass, height, location, leaf area index, plant area index, spatial vegetation density, and canopy gap fraction (Henning and Radtke 2006, Yao et al. 2011, Zhao et al. 2011, Zhao et al. 2012, Calders et al. 2014, Richardson 2014). Many of the applications of TLS to shrubland vegetation have studied individual plants, including mapping 2-cm scale shrub structure for fire behavior modeling (Adams 2014), modeling green and woody biomass of shrubs (Olsoy et al. 2014b), measuring shrub volume and limb surface area (Kałuża et al. 2012), and measuring shrub leaf surface area (Loudermilk et al. 2009). Uses of TLS to sample shrubland environments throughout plots have included ranged (<50 m) sensing and biomass estimation of shrubs (Greaves et al. 2015), local estimations of shrub and herbaceous fuelbed volume (Loudermilk et al. 2009, Rowell et al., 2016), identification of individual shrubs across plots and measurement of height and crown area (Vierling et al. 2013), estimation of wildlife visibility through shrub cover (Olsoy et al. 2015), estimation of grass biomass (Cooper et al., 2017), and modeling vegetation density profiles in short- and mixed-height shrublands (Ashcroft et al. 2014). Measurements made using TLS in sagebrush shrubland plots have also been strongly correlated with ALS measurements, showing that TLS collections may be used to “scale up” training data to broader remotely sensed datasets (Li et al. 2015).

A growing body of work has applied machine learning algorithms to classify vegetation and model structural traits using ALS point cloud data either alone (e.g. Li et al. 2017) or in combination with spectral datasets (e.g. García et al. 2011). Machine learning approaches to predictive modeling provide efficient analysis of “wide” data (datasets with many potential predictor variables), and often yield stronger models than can be derived using simple regression methods. Machine learning models commonly report error measures which are “out-of-bag” (aggregated from independent cross-validations internal to the modeling algorithm). For example, the Random Forests algorithm assembles a predictive model as the aggregation of a multitude of decision trees, each of which retains an independent 37% of the dataset for validation. The resulting model’s reported out-of-bag R^2 and root mean square error (RMSE) are aggregations of the errors measured in each of the many trees. The error measurements collected into out-of-bag errors only ever use validation data that has not been used for training, and often provide more accurate measures of model strength than simple cross-validation tests (Breiman 1996, 2001a).

In this study we demonstrate a workflow using TLS to predict biomass and canopy cover of different functional groups of sagebrush-steppe plants across large plots. We use machine learning (Random Forests (RF)) to leverage the information richness of TLS collections by discovering strong relationships between statistical descriptors of point cloud distributions represented by 2D pixels and manually collected measurements of biomass and structure. The main objective of this research is to develop a straightforward method for quantifying biomass and cover in the sagebrush-steppe across large plots (1-ha). The research question we aim to address is, to what extent can canopy cover and biomass of different functional groups of the sagebrush-steppe be quantified without individual classification of plants in TLS point clouds?

After creation, models of predicted features can be applied to whole 1-ha TLS datasets (both the

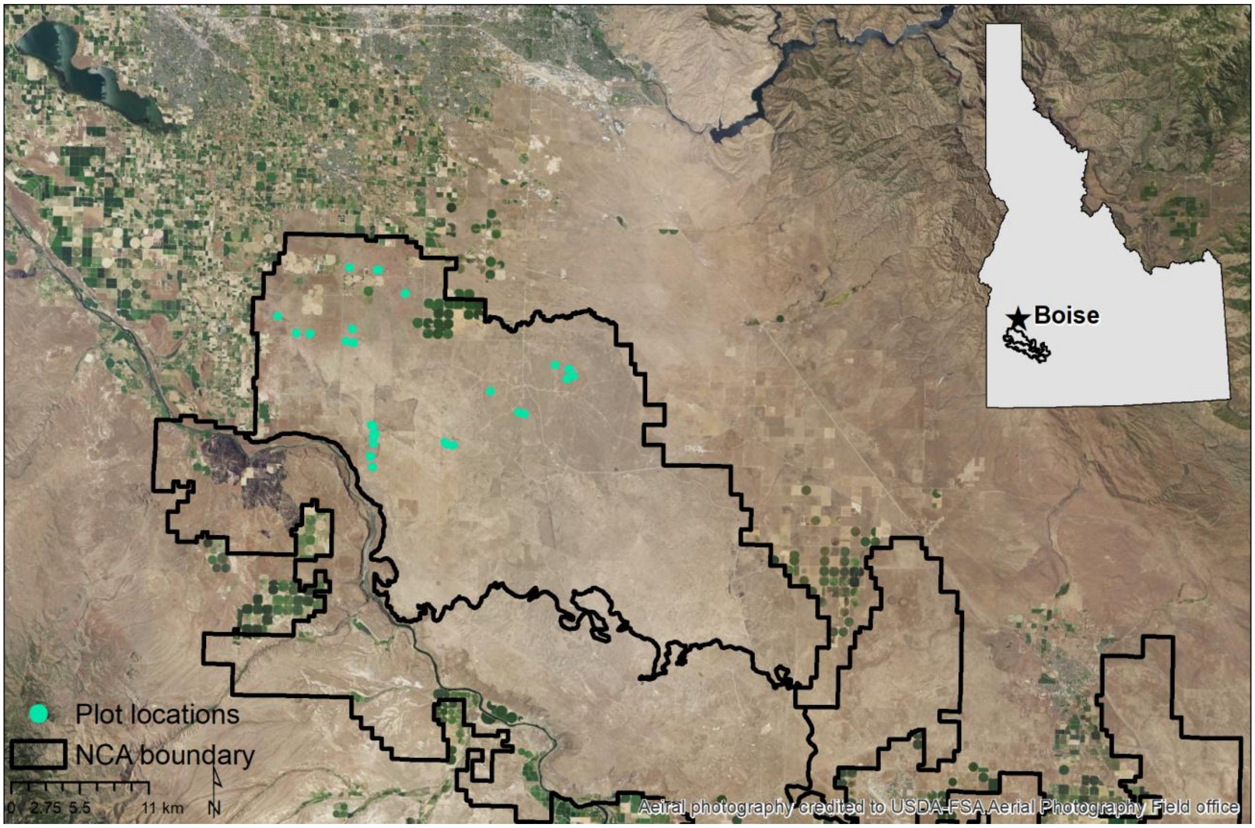
1-ha plots used to develop the model, and new plots in the same study area). Our method does not take any steps to explicitly delineate or classify plants, a challenging task in many lidar-based inventory methods. The models we developed had good predictive power overall, despite imperfect TLS point clouds (collected at oblique angles across plots which included dense shrublands) and some known errors in spatiotemporal matching of TLS and manual collections. Our experience proves this method as an efficient, scalable, and resilient workflow to model shrub-steppe vegetation traits across large areas.

2. Materials and Methods

2.1 Study area

The study was located within the Morley Nelson Snake River Birds of Prey National Conservation Area, which encompasses approximately 242,773 ha of the Snake River Plain Ecoregion in southwestern Idaho, USA (Fig. 1). The mean annual precipitation at the NCA is 24 cm and the average minimum and maximum annual temperatures are 3.5 ° C and 18.0 ° C, respectively, for the period 1980-2010 (PRISM, 2015). Surface geology includes loess windblown soils interspersed by basalt outcrops. The native flora is composed of sparse bunchgrasses (e.g., *Poa secunda, Elymus elymoides*) and an open canopy of shrubs (i.e., < 50% cover) generally less than 1.5 m tall, underlain by biological soil crust. Big sagebrush (*Artemisia tridentata, primarily ssp. wyomingensis*) is the regionally dominant shrub. Frequent wildfire in the NCA, especially in the last 30 years, has created a patchwork of native shrubland communities and degraded areas dominated by short-stature native perennials and non-native annual grasses (predominantly, *Bromus tectorum*) and forbs. Many degraded areas have been seeded with native and non-native perennial grasses, resulting in sparsely distributed, relatively tall (i.e., 30-50 cm) bunchgrasses (USDI BLM 2008).

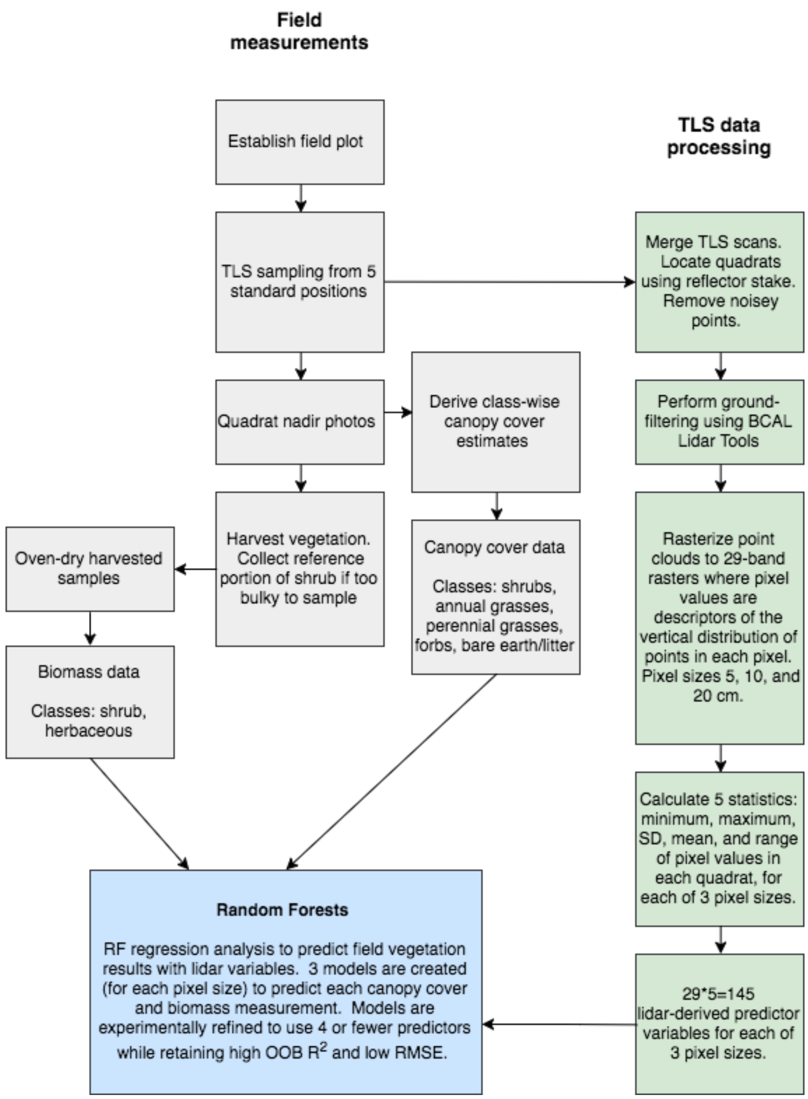
Figure 1. The NCA study area and location of plots with manual and TLS vegetation sampling. The background image is a National Agriculture Imagery Program (NAIP) true-color image.



2.2 Data collection

Our workflow for data collection and processing is detailed in Fig. 2 and described in detail below. We performed all TLS sampling between 15 May and 14 June 2013. By this date grasses and forbs were mostly senescent, but structurally intact. We used a stratified random sampling approach to locate twenty-six 1-ha plots, measuring 100 m by 100 m, for manual and TLS vegetation sampling throughout the western NCA. The sites spanned a gradient of plant community compositions, including intact shrublands, areas dominated by non-native grasses, and seeded sites containing taller perennial bunchgrass species. The plots were split evenly among sites dominated by shrubs and grasses (n=13 each). In each 1-ha plot, vegetation characteristics were collected manually in nine 1-m^2 quadrats spaced 25 m apart in a 3 by 3 grid centered on the plot (Fig. 3 and 4). This resulted in a total of 234 1-m^2 quadrats across 26 plots where paired manual and TLS sampling was performed.

Figure 2. Workflow of data collection, data processing, and Random Forests analysis resulting in predictive models for each cover and biomass class.



We deployed elevated disc reflectors at each plot corner to provide control points for coregistration and georegistration of TLS scans. A small reflector on a tall stake was placed at the center point of each quadrat to precisely mark its location in the TLS point cloud (after TLS collection, the stake was replaced by a surveyor's flag to mark the center point for manual vegetation sampling). The sides of square 1-m^2 quadrats were aligned with cardinal directions. We performed the TLS collection using a Riegl VZ-1000 near-infrared (1550 nm) scanner mounted on a 2-m tripod. At a range of 100 m, this instrument has a reported standard deviation of error of 8 mm and a beam diameter of 30 mm (corresponding to a beam divergence of 0.3 mrad) (Riegl, Austria). Single-return scans were performed with 0.02° of separation between pulses. Plots were scanned from five positions, once from the approximate midpoint of each side (using 180° scans) and once from the approximate plot center (using 360° scans). Our experimental setup took approximately 1-2 hours to collect five scans at each 1-ha plot. Slight leeway in scanner location allowed for adaptation to reduce occlusion in each scan (Fig. 3). After scanning was complete, the quadrat stake reflectors were replaced with surveyor's flags.

Figure 3. Examples of 1-ha plot layout and TLS-derived data. Black circles show scanning positions while black squares show locations of 1-m^2 manual sampling quadrats (enlarged for visibility). Coloring shows the maximum aboveground height of TLS points in 5-cm pixels in plots which are seeded with bunchgrasses (A), shrub-dominated (B), and native and non-native annual grass-dominated (C). Pixels occluded from sampling appear as white.

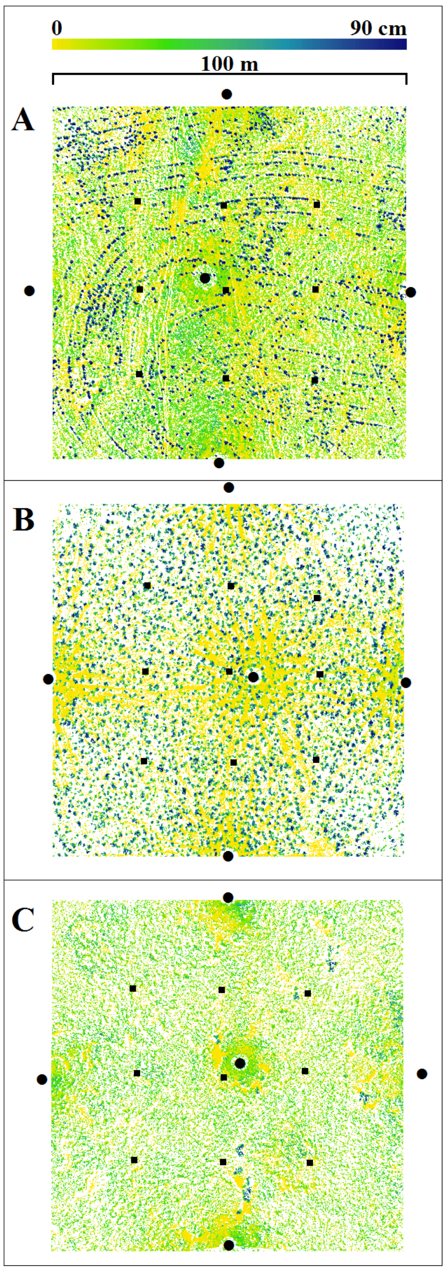
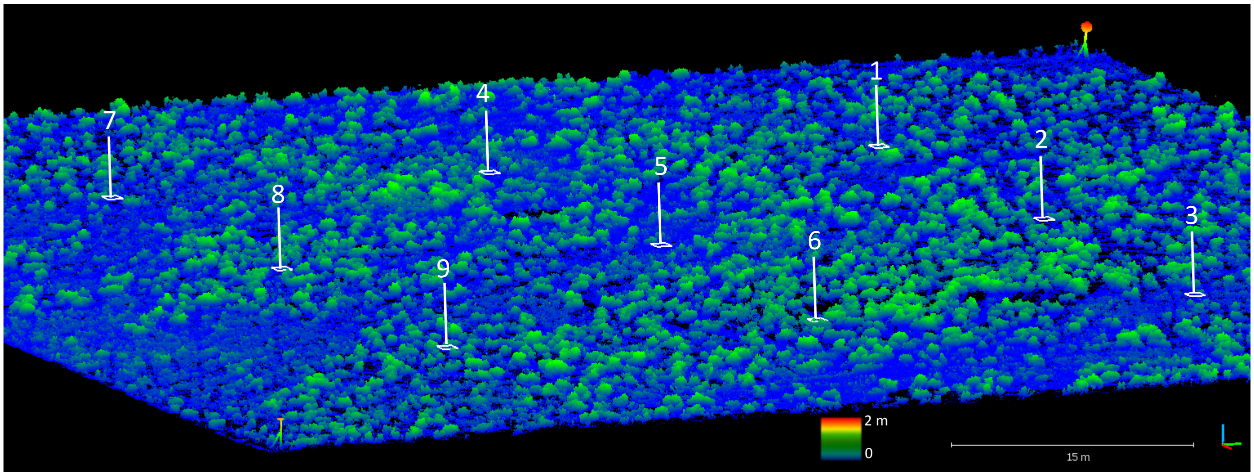


Figure 4. An example shrub-dominated plot with quadrats. White lines and numbers show quadrat location/layout. Elevated disc reflectors are shown at each corner. Quadrat 2 was omitted from the analysis due to occlusion (n=136 lidar returns).



Manual vegetation sampling at the scale of this study was made possible through collaboration with a larger, multi-year project being performed by the U.S. Geological Survey (Shinneman et al. 2015). The U.S.G.S. sampled vegetation characteristics in the field approximately 10 days following TLS sampling. At each quadrat, a nadir photograph centered on the plot was collected from a height of 2 m, imaging an area approximately 1 x 1.5 m. The topmost plant species (or lack thereof) were identified at 100 gridded sample points in each photograph using Samplepoint Software (see Pilliod and Arkle 2013), providing an estimate of the canopy cover of each species across the photo. Species-level data were aggregated to represent canopy cover of the following classes: bare earth/litter, annual grasses, perennial grasses, forbs, and shrubs. Aboveground vegetation within or overhanging each 1-m^2 quadrat was harvested and categorized as shrub or herbaceous. Where shrubs were too bulky to be harvested efficiently, a portion was collected for reference, and the number of equivalent-weight portions remaining was estimated for the quadrat. Samples were oven-dried at 65° C for at least 72 hours and then weighed.

2.3 Processing

We subsampled the TLS point clouds representing the quadrats to a minimum spacing of 1 cm between points using an octree filter. Points representing quadrat marker reflectors and other spurious (“noise”) points were manually removed. Using the BCAL Lidar Tools software (http://bcal.boisestate.edu/tools/lidar), ground filtering (classification of points as ground or vegetation) was performed using an iterative grid-based filtering approach that has been widely applied in shrub-steppe ecosystems (e.g. Streutker and Glenn 2006). The same software was then used to calculate 29 statistical descriptors of the vertical distribution of aboveground TLS points (Table 1), storing this information in 29-band raster files. The BCAL Lidar Tools exploit the rich information about height distributions in 3D point clouds by creating point statistics directly from the point cloud and reporting those in a 2D pixel representation. Each point cloud was used to create three 29-band rasters, each using a different pixel size (5, 10, and 20 -cm) to calculate descriptors of point distribution. Considering only pixels containing TLS points, we calculated the minimum, maximum, mean, range, and standard deviation of each of the 29 descriptors listed in Table 1 for each quadrat. As an example, the minimum, maximum, mean, range, and standard deviation of the 50th percentile of all height points within each pixel were calculated at the quadrat level. Calculating five statistics about each of 29 descriptors yielded a total of 145 statistics about point cloud distributions in each quadrat. Hereafter we refer to these 145 statistics as predictors.

2.4 Quadrat quality control

Through the data review process, we identified 28 quadrats as unfit to include in our analysis.

Twenty-two of these were discarded due to mistakes (mainly in quadrat placement) stemming from errors in communication between TLS and manual field sampling teams. One quadrat was discarded due to a rare ground filtering error that was identified in a cursory inspection of the classified TLS point cloud. In addition, we set a minimum threshold of 150 TLS returns (after subsampling to 1-cm spacing) to include a quadrat in the modeling. This threshold was set to exclude quadrats where occlusion prevented collection of any meaningful structural data (see Fig. 4 for an example). Five quadrats were discarded using this criterion. After removing these 28 quadrats, the remaining 206 were used for further analysis.

Table 1. Descriptors calculated from the TLS point cloud distribution within each pixel. The minimum, maximum, mean, range, and standard deviation of each of the descriptors (n=29) within the bounds of each quadrat were used as predictor variables. All points with a modeled height greater than 0 were classified as vegetation.

Descriptor
Minimum height
5th percentile height
10th percentile height
25th percentile height
50th percentile height
75th percentile height
90th percentile height
95th percentile height
Maximum height
Mean height
Standard deviation of heights
Range of heights
Interquartile range of heights
Kurtosis of heights
Skewness of heights
Variance of heights
Coefficient of variation of heights
Mean absolute deviation from mean height (AAD) = mean(|height - mean height|)
Median absolute deviation from median height (MAD) = 1.4826 x median(|height - median height|)
Texture of heights (standard deviation of heights between 5 cm and 15 cm)
Canopy relief ratio of height points = (mean height - min height)/(max height - min height)
Percent of returns modeled as ground
Percent of heights between 0 and 1 m tall
Percent of heights between 1 and 2.5 m tall
Count of vegetation returns
Count of ground returns
Count of all returns
Ratio of vegetation returns to ground returns
Ratio of vegetation returns to total returns

2.5 Random Forests analysis

We used Random Forests to leverage the detailed structural information collected by TLS, predict diverse vegetation traits using a common method, and apply an automated heuristic approach to analyze datasets which are complicated by varying scan angles, varying point densities, and patchy occlusion. RF regression (implemented with Salford Predictive Modeler Software Suite version 7, Salford Systems, San Diego, CA) was used to develop models predicting field-measured canopy cover and biomass of vegetation functional groups using the TLS-derived 145 predictors. In each model we found the bulk of predictors to be of low influence, and the inclusion of most actually decreased model performance in the testing datasets. We derived models using an automated forward selection procedure, which creates a 1-predictor model using the strongest solitary predictor, a 2-predictor model by identifying the second predictor which yields the strongest model in combination with the first, and so on. We also tested automated backward selection (iterative removal of the least important predictor) and manual trial-and-error procedures of model derivation, but found that forward selection discovered superior models in every case. For each of the three sets of predictors created using 5, 10, and 20 -cm pixels, we collected the first five models of each vegetation feature that were produced by forward selection. From among these we selected the model of each feature with the highest R^2 and lowest RMSE that used five or fewer predictors. All R^2 and RMSE values used and reported are out-of-bag.

Spatial autocorrelation between field observations was considered given that the quadrat observations within a plot were close together and could possibly exhibit autocorrelation. We tested for spatial autocorrelation between field observations by taking the residuals from the RF model and running a one-way ANOVA with the 26 plots as the treatments. If autocorrelation was present, the residuals from any plot would tend to be mostly positive, or mostly negative. If there was no autocorrelation the residuals would have random variation around a mean of zero. Using this method, we found no evidence of autocorrelation.

3. Results and Discussion

3.1 *Field Canopy Cover and Biomass*

The distribution of field-measured biomass and fractional canopy cover were highly non-normal, with most biomass and cover estimates clustering near the low or high ranges of measurements. Likewise, the standard deviation of measurements approached or exceeded the mean measurement of each variable (Table 2). For example, the mean shrub, bare earth/litter, and annual grass cover was 8%, 41%, and 35%, whereas their corresponding standard deviation was 14%, 27%, and 38%, respectively.

Table 2. Statistics describing the manual measurements of cover and biomass (n = 206). Minimum values were all 0. Columns 25th & 75th are percentiles and SD is standard deviation.

Feature	25th	Median	75th	Max	Mean	SD
Shrub cover (%)	0	0	8	61	8	14
Bare earth/litter cover (%)	13	43	61	94	41	27
Annual grass cover (%)	0	13	73	100	35	38
Perennial grass cover (%)	1	7	21	70	13	15
Forb cover (%)	0	0	3	68	4	9
Shrub biomass (g)	0	0	18	2476	106	322
Herbaceous biomass (g)	57	97	180	1193	146	158

3.2 *Predicted Canopy Cover and Biomass*

Five of seven Random Forest models achieved out-of-bag R^2 > 0.5 correlation with manual measurements. Descriptors calculated using a 5-cm pixel size yielded the strongest predictors of forb cover (R^2 = 0.52, RMSE = 6%) and herbaceous biomass (R^2 = 0.61, RMSE= 99 g) (Tables 3 & 4). A 10-cm pixel size yielded the strongest predictors of shrub cover (R^2 = 0.77, RMSE = 7%), annual grass cover (R^2 = 0.70, RMSE = 21%), perennial grass cover (R^2 = 0.36, RMSE = 12%), bare earth/litter cover (R^2 = 0.49, RMSE = 19%), and shrub biomass (R^2 = 0.71, RMSE = 175 g) (Tables 3 & 4). A 20-cm pixel size did not yield the strongest predictors of any feature. The precision of our model predictions ranged between 46% and 165% of mean manual measurements (by comparing the lowest RMSE values from Tables 3 & 4 with mean manual measurements in Table 2). For example, the prediction of the bare earth/litter cover class had the lowest RMSE in comparison to the mean of the manual measurements (46%), and the RMSE of the annual grass cover was 60% of the mean manual measurement. These classes were also the dominant cover classes in the field, as measured by mean percent cover data (41% and 35%, respectively, Table 2). In comparison, our model predictions which had high RMSE values (e.g. forb cover and shrub biomass) corresponded to classes that had low vegetation percent cover in our field plots.

We found that while marginal improvements in model quality were made available by testing several pixel sizes for predictor creation, the benefit was unlikely to be great. Despite a sixteen-fold difference in the area of the pixel sizes we tested, in only one case was the difference in strength between the strongest models produced by each pixel size greater than $R^2 = 0.05$ (shrub biomass, difference of $R^2 = 0.13$) (results provided in Supplementary Material). While we did not find a single pixel size for predictor calculation to be consistently superior, predictors from 20-cm pixels never yielded the strongest model, and predictors from 10-cm pixels produced the best across-the-board performance. Given that differences in the models were low, a 10-cm pixel size can be interpreted to be appropriate for predicting vegetation cover and biomass in our study area, representing a compromise between too fine a resolution (5-cm) that over-represents occlusion and too coarse (20-cm) which generalizes subtle differences in the point cloud. Future studies may also wish to test several pixel sizes to discover which yield the most useful predictors of the local environment.

Our use of RF was straightforward. For each vegetation feature, we used a forward selection procedure to derive models using one to five predictors, for each of the 5, 10, and 20 -cm pixel predictor sets, and selected the model with the highest R^2 from the fifteen produced. This method is not comprehensive, and it is possible somewhat stronger combinations of predictors exist to model some of our targets. We expect our experience to match the common case where the top several competing models of a single feature exhibit similar strengths (even though the predictors they use may differ), minimizing the importance of which specific model is selected (Breiman 2001b). We presented only models using up to five predictors, although using one or two fewer predictors would generally not cost much predictive strength, and allowing one or two more would not cost much parsimony. The combination of predictors used is inconsistent among models, but some predictors were used more commonly than others (Tables 3 & 4). For example, a statistic describing the 50th percentile (median) of point heights in pixels was the top single predictor in every model, except those of shrub biomass.

Table 3. Predictions of percent canopy cover for annual grass, bare earth/litter, forb, perennial grass, and shrub classes using the optimal pixel size to calculate point statistics, as generated by the first 5 predictor sets yielded by forward stepwise selection. Predictors are listed in the order they were added to the predictor set, and resultant models' predictive strength and root mean square error (RMSE, in %) are also listed. Bolded is the model explaining the most variance and with the lowest RMSE. If N=4 is bolded, then the model used the first four predictors; if N=5 is bolded, then the model used all five predictors. Additional results on the remaining pixel sizes are presented in Supplementary Material.

Vegetation	Pixel size	Predictors	N	R^2	RMSE
Annual grass	10	Mean of 50th percentile heights	1	.59	24
		Standard deviation of maximum heights	2	.62	23
		Mean of ratio of vegetation returns to total returns	3	.67	22
		Mean of ratio of vegetation returns to ground returns	4	.69	21
		Minimum of 50th percentile heights	**5**	**.70**	**21**
Bare earth /litter	10	Maximum of 50th percentile heights	1	.38	21
		Standard deviation of interquartile range of heights	2	.45	20
		Standard deviation of ratio of vegetation returns to ground returns	3	.48	20
		Range of percent of vegetation 0 < & <= 1 m high	4	.48	19
		Mean of percent of vegetation 0 < & <= 1 m high	**5**	**.49**	**19**
Forb	5	Minimum of 50th percentile heights	1	.47	6
		Standard deviation of minimum heights	2	.48	6
		Maximum of 50th percentile heights	3	.51	6
		Mean of canopy relief ratio	**4**	**.52**	**6**
		Mean of 50th percentile heights	5	.51	6
Perennial grass	10	Maximum of 50th percentile heights	1	.19	14
		Minimum of coefficient of variation of heights	2	.27	13
		Maximum of 90th percentile heights	3	.32	13
		Minimum of kurtosis of heights	4	.33	13
		Maximum of interquartile range of heights	**5**	**.36**	**12**
Shrub	10	Maximum of 50th percentile heights	1	.66	8
		Standard deviation of 90th percentile heights	2	.72	7
		Mean of 50th percentile heights	3	.76	7
		Range of skewness of heights	4	.76	7
		Minimum of 50th percentile heights	**5**	**.77**	**7**

Table 4. Predictions of biomass for herbaceous and shrub classes using the optimal pixel size to calculate point statistics, as generated by the first 5 predictor sets yielded by forward stepwise selection. Predictors are listed in the order they were added to the predictor set, and resultant models' predictive strength and root mean square error (RMSE, in grams) are also listed. Bolded is the model explaining the most variance and with the lowest RMSE. Both herbaceous and shrub biomass were best predicted using the first four predictors. Additional results on the remaining pixel sizes are presented in Supplementary Material.

Vegetation	Pixel size	Predictors	N	R^2	RMSE
Herbaceous	5	Mean of 50th percentile heights	1	.47	115
		Minimum of mean of heights	2	.57	104
		Mean of count of vegetation returns	3	.60	100
		Minimum of 50th percentile heights	**4**	**.61**	**99**
		Maximum of 50th percentile heights	5	.60	99
Shrub	10	Mean of range of heights	1	.58	209
		Mean of absolute deviation from mean heights	2	.68	183
		Mean of standard deviation of heights	3	.67	185
		Minimum of coefficient of variation of heights	**4**	**.71**	**175**
		Mean of count of returns	5	.69	178

3.3 *Estimates Without Individual Plant Classification*

The class-wise characteristics of vegetation functional groups were predicted without explicit classification and delineation of individual plants or vegetation classes. However, our workflow using pixel statistics to extract information from point clouds yielded models with lower fit to ground truth measurements than those developed using per-plant measures such as crown area or volume (e.g. Vierling et al. 2013, Olsoy et al. 2014a,b, and Greaves et al. 2015). We are unaware of any studies that have attempted to sample large plots (1-ha) in dense shrubland using common oblique scanning from a tripod, and the literature may not represent the difficulty of classifying and delineating small and closely-spaced plants in point clouds where occlusion is pervasive. Automated classification approaches, such as spatial wavelet analysis and eigenvalue separation, have not been demonstrated across point clouds where occlusion is common and the sampled vegetation is small and spatially mixed, as is the common case in TLS collections of desert shrublands. Modeling vegetation characteristics on a per-area, rather than per-plant basis, is especially valuable when complementary manual sampling considers all of the vegetation within (and overhanging) a quadrat, and none of the vegetation extending outside of the quadrat.

There are some disadvantages to avoiding explicit classification in a TLS-based vegetation inventory. The strongest predictive relationships between plant structural indices and traits such as biomass would be expected when a single, complete plant is considered. By contrast, our approach aggregates structural information from predefined grids across 1-m^2 quadrats. This results in measurements that combine information from all plants and plant classes in a quadrat and excludes portions of plants which extend beyond the quadrat's edge. Aggregating structural data from unclassified plants risks confusing a decision tree when different vegetation compositions of quadrats exhibit similar signals (e.g. the aggregated measurements of points representing several tall and narrow bunchgrasses might resemble the measurements of a single tall and stout shrub). We expect the high RMSE of the shrub biomass predictions were partly caused by these challenges. In fairness, we would expect some misidentification of vegetation to occur in any TLS-based workflow due to structural similarity of certain species in different functional groups (e.g. tall forbs, such as tumble mustards and thistles, resemble shrubs and smaller perennial grasses resemble annual grasses).

3.4 *Field Considerations*

Our five-position TLS sampling protocol often provided redundant coverage at excess resolution, but occlusion was still a challenge in plots with high shrub cover. In some cases, TLS sampling of vegetation shorter than the top canopies of shrubs was sparse across most of the plot. Nonetheless, we discarded only 2% of quadrats due to a low number of TLS returns. Partial occlusion was common in the remaining quadrats. Measuring with gridded presence/absence windows, average quadrat sampling coverage of pixels with any number of points in them was 59% (std = 24%) using a 5-cm grid, 80% (std=19%) using a 10-cm grid, or 92% (std =16%) using a 20-cm grid. That we succeeded in developing strong models despite occlusion shows that our methods function well using practical TLS field collections.

Sampling density (returns per m^2) varied widely, depending on occlusion and position relative to the scan position layout. We calculated statistics about the counts of TLS returns per quadrat by quadrat position (center, middle-edge, corner), and the counts of returns in every 1-m^2 grid cell across plots (which we stratified by grass and shrub-dominated plots). In addition to statistics about return counts, we calculated the percentage of quadrats or 1-m^2 grid cells which did not meet the 150 return minimum threshold for modeling. We found that sites adjacent to scan positions (e.g. center quadrats) were commonly sampled with thousands more returns than sites that were further away. The distribution of sampling densities of the total population of 1-m^2 quadrats closely resembles that of the total population of 1-m^2 grid cells, indicating that our quadrat placement protocol was adequate to represent sampling variability throughout plots. Additionally, only 4% of 1-m^2 grid cells were below the 150 point minimum threshold. Taken together, these results confirm that our models can be applied with the reported strength nearly continuously across 1-ha plots (Table 5).

Table 5. For center (n=1 quadrat x 26 plots), middle-edge (n=4 x 26), and corner (n=4 x 26) quadrat positions, and across all 1-m^2 grid cells in grass (n=10,000 grid cells x 13 plots) and shrub-dominated plots (n=10,000 x 13), we report statistics about the counts of TLS returns and the percentage of quadrats or grid cells below the 150 return minimum threshold. We report the same statistics for all quadrats (n=9 quadrats x 26 plots) and all 1-m^2 grid cells (n=10,000 grid cells x 26 plots). Refer to Figures

3 and 4 for quadrat layout.

Region type	Min	25th	Median	75th	Max	Mean	< 150 returns
Center quadrats	1,474	5,708	8,245	12,600	34,770	10,260	0%
Middle-edge quadrats	2	805	1,504	2,782	12,370	1,976	2.9%
Corner quadrats	22	514	930	1,482	6,730	1,213	1.9%
All quadrats	2	666	1,268	2,731	34,770	2,558	2.1%
1-m^2 grid cells in grass plots	0	683	1,216	2,402	70,110	2,534	1.4%
1-m^2 grid cells in shrub plots	0	581	1,312	3,011	103,000	2,936	6.6%
1-m^2 grid cells in all plots	0	639	1,257	2,679	103,000	2,730	4.0%

The greatest source of occlusion in the TLS sampling was dense shrub canopies, which in the highest shrub cover plots blocked sampling of almost half of 5-cm gridded windows within the hectare area. On average, the 5 quadrats discarded due to occlusion had relatively high shrub cover (21%), average bare earth/litter cover (41%), and low annual grass cover (5%), reflecting the general composition of the shrub-dominated plots in which they occurred. That these quadrats contained substantial amounts of both the most and least physically prominent vegetation cover classes (shrub and bare earth/litter) supports our field observation that instances of near-total occlusion within quadrats is mainly a consequence of surrounding vegetation, and not low-lying or impenetrable vegetation within quadrats themselves.

While the algorithm we used for ground classification has been widely tested in similar shrubland environments (e.g. Glenn et al., 2011, Mitchell et al., 2011, Streutker and Glenn, 2006), error due to occlusion and confusion of plants and the ground surface may have resulted in the low R^2 of the model predicting coverage of bare earth/litter. Imperfect ground classification and surface modeling will also result in errors in point cloud measurements of height (e.g. Ashcroft et al. 2014, Fan et al. 2014). Better sampling coverage of 1-ha plots with densely spaced shrubs could be achieved by scanning from more than 5 positions, moving positions from the plot edge inward, or by further elevating the TLS. Although classification errors were a possibility, and we discarded a single quadrat due to a ground filtering issue, classifying vegetation and ground was not a major operational challenge in this workflow.

As a result of pairing this pilot study with a pre-existing campaign to measure and harvest vegetation, spatial matching of point clouds to areas sampled manually was imperfect, potentially yielding imprecise compositional measurements if the quadrat vegetation is not representative of its surroundings. The area considered in photograph samples and canopy cover inventory (1.5 m^2) was larger than the 1-m^2 quadrats considered in the TLS data. We were not able to adjust the area considered in TLS point clouds to match the extent of photos due to inconsistent photo orientation (allowing field technicians to work around their environment).

Small discrepancies in quadrat placement within TLS point clouds (marked with field flags in scans and on the ground) versus the actual manual sampling locations may have introduced some erroneous biomass values to our dataset, especially where a relatively large amount of a quadrat's vegetation has been wrongly included or excluded. Although growth and decomposition of vegetation is slow in our field area, a typical delay of up to two weeks between TLS and manual sampling could also allow for compositional changes (e.g. trampling, grazing, or senescent plants or litter blowing in or out of the plot) in quadrats between collections. The method of harvesting only a representative portion of large shrubs likely caused some imprecision in shrub biomass measurements. Ideally, future studies would conduct TLS and vegetation sampling at the same time using the same field team.

The TLS methodology presented here allows for repeat scanning to monitor changes in vegetation structure on a per-area basis. Once predictive models have been trained to satisfactory strength for a study area, the need for further carefully-coordinated manual sampling is eliminated.

3.5 *Future Studies*

Future studies might also enhance implicit vegetation classification within RF models by calculating additional pixel statistics of high-resolution spectral imagery gathered from airborne or spaceborne platforms. A single band spectral dataset may also be collected by normalizing the intensity of TLS pulse returns to range effects (e.g. Nield et al. 2014, Zhu et al. 2015), but the complex model required to resolve intensity effects of vegetation size, angle, and spectral reflectance, atmospheric conditions, and beam divergence in sagebrush steppe vegetation has not been demonstrated. Structure-from-motion (SfM) derived point clouds from optical imagery of similar precision and density to those used in this study have recently been published (Cooper et al., 2017, Wallace et al., 2017, Olsoy et al., in review). The methods developed herein could be applied to such data with the potential benefits of fewer areas of occlusion with a (near) nadir sampling platform. One should consider that the understory of shrub-dominated plots or other high biomass vegetation near the ground surface may be under-sampled with the use of optical imagery to generate the point clouds (e.g. Wallace et al., 2017). Regardless of platform, measures of occlusion could potentially be used as an inverse measure of vegetation presence or absence across a large plot, with careful consideration of beam divergence with scan range and use of visibility models (e.g. Lin and West, 2016, Murgoitio et al. 2014, Zhao et al. 2012). Future studies should consider the minimum number of field measurements needed for a statistically robust relationship between field data and point cloud statistics.

We would expect future applications of our approach to remove some of the sources of error we listed, allowing even stronger models to be developed. Despite some preventable challenges, our workflow produced models of strength, demonstrating the capability to use a TLS/machine learning approach to extend localized manual vegetation sampling in sagebrush steppe habitats to much larger plots. In sum, our methods are automatable, applicable to a broad range of mixed and short-stature vegetation communities, yield models which are continuous, and provide analysis of “messy” clouds where occlusion is common, plants are small, and vegetation classes are mixed.

4. Conclusions

This study illustrates an efficient and effective method to relate TLS point clouds with groundtruthing data for prediction of cover and biomass of shrubs and grasses at 1-m^2 scale across large plots. Our method of TLS sampling was time efficient and the workflows to calculate predictors from point clouds and generate models can be largely automated. Once generated, a model can be applied to a 1-m^2 grid across the whole plot. A significant strength of our method of calculating TLS predictor variables is that it does not require explicit classification and delineation of vegetation groups being studied—a challenging and time consuming task which may be impossible when vegetation is dense and vegetation classes are spatially mixed. Although our workflow is highly transferrable to point clouds derived from SfM and to similar ecosystems outside of our study area, new models will need to be trained based on data collection methods, specific ecosystem conditions, and considerations of timing due to phenology. Our methods supply a convincing demonstration of the ability of machine learning to exploit the richness of point clouds, generating models of shrub-steppe biomass and cover which are accurate, efficient to develop, and easy to extrapolate as continuous rasters across large plots.

There are urgent needs for quick and accurate vegetation measurements that provide ecological and management indicators in the highly imperiled sagebrush steppe and in other dryland ecosystems. Our method of vegetation inventory across large plots has immediate applicability to numerous research and management needs which presently rely on localized manual measurements, including ecological productivity and status, evaluation of wildlife habitat, evaluation of landscape management practices, and fuel load surveys for wildfire risk. TLS-based models of vegetation characteristics may also serve as a stepping stone to train broader-scale datasets collected from air or space (e.g. Li et al. 2015, Greaves et al. 2017). The spatially explicit, realistic, high-resolution vegetation information across large plots may also be an invaluable data source for landscape simulations, such as wildlife habitat use, wildfire behavior, or erosion processes.

Acknowledgements

This work was supported by a Joint Fire Sciences Program grant (Project ID: 11-1-2-30), NOAA's Earth System Research Laboratory (ESRL, Physical Sciences Division) Award NA10OAR4680240, the NSF Idaho EPSCoR Program, and by the National Science Foundation under award number EPS-0814387. We thank Randy Lee at Idaho National Laboratory for use of the TLS, and Dr. Rupesh Shrestha, Dr. Aihua Li, Mr. Peter Olsoy, Mr. Kyle Gochnour, and Mr. Samuel Gould for providing lab and field assistance. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government. This article has been peer reviewed and approved for publication consistent with USGS Fundamental Science Practices (http://pubs.usgs.gov/circ/1367).

Literature Cited

Adams, T. (2014). Using Terrestrial LiDAR to Model Shrubs for Fire Behavior Simulation. Retrieved from: http://scholarworks.umt.edu/etd/4173/ in May 2015.

Ashcroft, M. B., Gollan, J. R., & Ramp, D. (2014). Creating vegetation density profiles for a diverse range of ecological habitats using terrestrial laser scanning. *Methods in Ecology and Evolution*, *5*(3), 263-272.

Balch, J. K., Bradley, B. A., D'Antonio, C. M., & Gómez-Dans, J. (2013). Introduced annual grass increases regional fire activity across the arid western USA (1980–2009). *Global Change Biology*, 19(1), 173-183.

Breiman, L. (1996). Out-of-bag estimation. *Technical report 1996b*. Berkeley, CA: Statistics Department, University of California Berkeley.

Breiman, L. (2001a). Random forests. *Machine learning*, *45*(1), 5-32.

Breiman, L. (2001b). Statistical modeling: The two cultures (with comments and a rejoinder by the author). *Statistical Science*, *16*(3), 199-231.

Brooks, M. L., D'antonio, C. M., Richardson, D. M., Grace, J. B., Keeley, J. E., DiTomaso, J. M., Hobbs, R.J, Pellant, M., & Pyke, D. (2004). Effects of invasive alien plants on fire regimes. *BioScience*, 54(7), 677-688.

Bukowski, B. E., & Baker, W. L. (2013). Historical fire regimes, reconstructed from land-survey data, led to complexity and fluctuation in sagebrush landscapes. *Ecological Applications*, *23*(3), 546-564.

Calders, K., Newnham, G., Burt, A., Murphy, S., Raumonen, P., Herold, M., Culvenor, D., Avitabile, V., Disney, M., Armstrong, J., & Kaasalainen, M. (2014). Nondestructive estimates of above-ground biomass using terrestrial laser scanning. *Methods in Ecology and Evolution, 6*(2), 198-208.

Cifuentes, R., Van Der Zande, D., Farifteh, J., Salas, C., & Coppin, P. (2014). Effects of voxel size and sampling setup on the estimation of forest canopy gap fraction from terrestrial laser scanning data. *Agricultural and Forest Meteorology*, 194, 230-240.

Cooper, S.D., Roy, D.P., Schaaf, C.B. and Paynter, I. (2017). Examination of the Potential of Terrestrial Laser Scanning and Structure-from-Motion Photogrammetry for Rapid Nondestructive Field Measurement of Grass Biomass. *Remote Sensing*, 9(6), p.531.

Davies, K.W., Bates, J.D. (2010). Vegetation Characteristics of Mountain and Wyoming Big Sagebrush Plant Communities in the Northern Great Basin. *Rangeland Ecology and Management*, 63, 461–466.

D'Antonio, C. M. and P. M. Vitousek. (1992). Biological Invasions by Exotic Grasses, the Grass /Fire Cycle, and Global Change. *Annual Review of Ecology and Systematics, 23*, 63-87.

Fan, L., Powrie, W., Smethurst, J., Atkinson, P. M., & Einstein, H. (2014). The effect of short ground vegetation on terrestrial laser scans at a local scale. *ISPRS Journal of Photogrammetry and Remote Sensing*, *95*, 42-52.

García, M., Riaño, D., Chuvieco, E., Salas, J., & Danson, F. M. (2011). Multispectral and LiDAR data fusion for fuel type mapping using Support Vector Machine and decision rules. *Remote Sensing of Environment*, *115*(6), 1369-1379.

Glenn, N.F, Spaete, L., Sankey, T., Derryberry, D., Hardegree, S., Mitchell, J. (2011). Errors in LiDAR derived shrub height and crown area on sloped terrain, Journal of Arid Environments, 75 (4), 377-382.

Glenn, N.F., Neuenschwander, A., Vierling, A.A., Spaete, L.P., Li, A., Shinneman, D., Pilliod, D.S., Arkle, R., McIlroy, S., (2016). Landsat 8 and ICESat-2: Performance and potential synergies for quantifying dryland ecosystem vegetation cover and biomass. *Remote Sensing of Environment, 185: 233-242*.

Greaves, H.E., Vierling, L.A., Eitel, J.U., Boelman, N.T., Magney, T.S., Prager, C.M. and Griffin, K.L. (2017). Applying terrestrial lidar for evaluation and calibration of airborne lidar-derived shrub biomass estimates in Arctic tundra. *Remote Sensing Letters*, 8(2), pp.175-184.

Greaves, H. E., Vierling, L. A., Eitel, J. U., Boelman, N. T., Magney, T. S., Prager, C. M., & Griffin, K. L. (2015). Estimating aboveground biomass and leaf area of low-stature Arctic shrubs with terrestrial LiDAR. *Remote Sensing of Environment*, *164*, 26-35.

Henning, J. G., & Radtke, P. J. (2006). Ground-based laser imaging for assessing three-dimensional forest canopy structure. *Photogrammetric Engineering & Remote Sensing*, *72*(12), 1349-1358.

Homer, C. G., Aldridge, C. L., Meyer, D. K. and Schell, S. J. (2012). 'Multi-scale remote sensing sagebrush characterization with regression trees over Wyoming, USA: Laying a foundation for monitoring'. *International Journal of Applied Earth Observation and Geoinformation*, 14(1), pp. 233-244

Kałuża, T., Tymków, P., & Strzeliński, P. (2012). Use of remote sensing for investigating riparian shrub structures. *Polish Journal of Environmental Studies*, *21*(1).

Knick, S. T. (1999). Requiem for a sagebrush ecosystem? *Northwest Science*, 73(1).

Li, A., Dhakal, S., Glenn, N.F., Spaete, L.P., Shinneman, D.J., Pilliod, D.S., Arkle, R.S., & McIlroy, S.K. (2017). Lidar aboveground vegetation biomass estimates in shrublands: prediction, uncertainties and application to coarser scales, *Remote Sensing*, *9*(9), 903; doi:10.3390/rs9090903.

Li, A., Glenn, N.F., Olsoy, P.J., Mitchell, J.J., Shrestha, R., (2015). Aboveground biomass estimates of sagebrush using terrestrial and airborne LiDAR data in a dryland ecosystem. *Agricultural and Forest Meteorology*, 213, 138-147. doi: 10.1016/j.agrformet.2015.06.005

Lin, Y. and West, G. (2016). Reflecting conifer phenology using mobile terrestrial LiDAR: A case study of Pinus sylvestris growing under the Mediterranean climate in Perth, Australia. *Ecological Indicators*, 70, 1-9.

Loudermilk, E. L., Hiers, J. K., O'Brien, J. J., Mitchell, R. J., Singhania, A., Fernandez, J. C., Cropper, W.P. & Slatton, K. C. (2009). Ground-based LIDAR: a novel approach to quantify fine-scale fuelbed characteristics*. International Journal of Wildland Fire*, *18*(6), 676-685.

Mitchell, J. J., Glenn, N. F., Sankey, T. T., Derryberry, D. R., Anderson, M. O., & Hruska, R. C. (2011). Small-footprint LiDAR estimations of sagebrush canopy characteristics. *Photogrammetric Engineering & Remote Sensing*, *77*(5), 521-530.

Murgoitio, J., Shrestha, R., Glenn, N., Spaete, L. (2014). Airborne LiDAR and terrestrial laser scanning derived vegetation obstruction factors for visibility models. *Transactions in GIS*, 18(1): 147-160.

Nield, J.M., King, J., Jacobs, B., (2014). Detecting surface moisture in aeolian environments using terrestrial laser scanning, Aeolian Research, 12, 9-17, DOI: 10.1016/j.aeolia.2013.10.006

Olsoy, P.J., Mitchell, J.J., Levia, D.F., Clark, P.E., Glenn, N.F., (2016). Estimation of big sagebrush leaf area index with terrestrial laser scanning. *Ecological Indicators*, 61(2), 815-821.

Olsoy, P.J., N Glenn, & Clark, P. (2014a). Estimating sagebrush biomass using terrestrial laser scanning (TLS). *Rangeland Ecology & Management*, *67* (2): 224-228.

Olsoy, P.J., Glenn, N. F., Clark, P. E., & Derryberry, D. R. (2014b). Aboveground total and green biomass of dryland shrub derived from terrestrial laser scanning. *ISPRS Journal of Photogrammetry and Remote Sensing*, *88*, 166-173.

Olsoy, P.J., Forbey, J.S., Rachlow, J.L., Nobler, J.D., Glenn, N.F., & Shipley, L.A. (2015). Fearscapes: Mapping functional properties of cover for prey with terrestrial LiDAR. *BioScience, (January 2015) 65 (1): 74-80. 10.1093/biosci/biu189*

Pilliod, D.S., Arkle, R.S., (2013). Performance of quantitative vegetation sampling methods across gradients of cover in Great Basin plant communities. *Rangeland Ecology and Management*, *66*, 634-647.

PRISM Climate Group, (2015). Oregon State University, http://prism.oregonstate.edu, LT71m dataset created Dec 2015.

Pyke, D. A., Chambers, J. C., Pellant, M., Knick, S. T., Miller, R. F., Beck, J. L., & McIver, J. D. (2015). Restoration handbook for sagebrush steppe ecosystems with emphasis on greater sage-grouse habitat—Part 1. Concepts for understanding and applying restoration (No. 1416). US Geological Survey.

Pyke, D. A., Shaff, S.E., Lindgren, A.I., Schupp, E.W., Doescher, P.S., Chambers, J. C., Burnham, J.S., Huso, M.M. (2014). Region-Wide Ecological Responses of Arid Wyoming Big Sagebrush Communities to Fuel Treatments. *Rangeland Ecology and Management*, 67, 455-467.

Reiner, A. L., Tausch, R. J., & Walker, R. F. (2010). Estimation procedures for understory biomass and fuel loads in sagebrush steppe invaded by woodlands. *Western North American Naturalist*, 312-322.

Richardson, J. J., Moskal, L. M., & Bakker, J. D. (2014). Terrestrial Laser Scanning for Vegetation Sampling. *Sensors*, *14*(11), 20304-20319.

Ripplinger, J., Franklin, J., & Edwards, T. C. (2015). Legacy effects of no-analogue disturbances alter plant community diversity and composition in semi-arid sagebrush steppe. *Journal of Vegetation Science*, 26(5), 923-933.

Rowell, E., Loudermilk, E.L., Seielstad, C. and O'Brien, J.J. (2016). Using simulated 3D surface fuelbeds and terrestrial laser scan data to develop inputs to fire behavior models. Canadian Journal of Remote Sensing, 42(5), pp.443-459.

Seielstad, C., Stonesifer, C., Rowell, E., & Queen, L. (2011). Deriving fuel mass by size class in Douglas-fir (Pseudotsuga menziesii) using terrestrial laser scanning. *Remote Sensing, 3*(8), 1691-1709.

Shan, J., Toth, C.K. (Eds.), Topographic Laser Ranging and Scanning, CRC Press, Boca Raton, FL (2008), pp. 87–127.

Shinneman, D.J., Pilliod, D.S., Arkle, R.S., Glenn, N.F. (2015). Quantifying and predicting fuels and the effects of reduction treatments along successional and invasion gradients in sagebrush habitats: Final Report to the Joint Fire Science Program, p. 1-44.

Streutker, D. R., & Glenn, N. F. (2006). LiDAR measurement of sagebrush steppe vegetation heights. *Remote Sensing of Environment*, *102*(1), 135-145.

Telling, J., Lyda, A., Hartzell, P. and Glennie, C. (2017). Review of Earth science research using terrestrial laser scanning. Earth-Science Reviews.

USDI Bureau of Land Management. (2008). *Snake River Birds of Prey National Conservation Area Proposed Resource Management Plan and Final Environmental Impact Statement, Boise, ID.*

USGS Forest and Range Ecosystem Science Center. (2000). *BRIEFING PAPER: Loss of the sagebrush ecosystem of the Snake River Plain, Boise, ID.*

Van der Zande, D., Jonckheere, I., Stuckens, J., Verstraeten, W. W., & Coppin, P. (2008). Sampling design of ground-based lidar measurements of forest canopy structure and its effect on shadowing. *Canadian Journal of Remote Sensing*, *34*(6), 526-538.

Vierling, L. A., Xu, Y., Eitel, J. U., & Oldow, J. S. (2013). Shrub characterization using terrestrial laser scanning and implications for airborne LiDAR assessment. *Canadian Journal of Remote*

Sensing, *38*(6), 709-722.

Vosselman, G., Maas, H.-G. (Eds.), Airborne and Terrestrial Laser Scanning (first ed.), CRC Press. Boca Raton, FL (2010).

Wallace, L., Hillman, S., Reinke, K. and Hally, B. (2017). Non-destructive estimation of above-ground surface and near-surface biomass using 3D terrestrial remote sensing techniques. *Methods in Ecology and Evolution*.

Wilkes, P., Lau, A., Disney, M., Calders, K., Burt, A., de Tanago, J.G., Bartholomeus, H., Brede, B. and Herold, M. (2017). Data acquisition considerations for Terrestrial Laser Scanning of forest plots. Remote Sensing of Environment, 196, pp.140-153.

Yao, T., Yang, X., Zhao, F., Wang, Z., Zhang, Q., Jupp, D., Lovell, J., Culvenor, D., Newnham, G., Ni-Meister, W., Schaaf, C., Woodcock, C., Wang, J., Li, X., & Strahler, A. (2011). Measuring forest structure and biomass in New England forest stands using Echidna ground-based lidar. *Remote sensing of Environment*, *115*(11), 2965-2974.

Zhao, F., Strahler, A. H., Schaaf, C. L., Yao, T., Yang, X., Wang, Z., Schull, M.A., Román-Colón, M.O., Woodcock, C.E., Olofsson, P., Ni-Meister, W., Jupp, D.L.B., Lovell, J.L., Culvenor, D.S., & Newnham, G. J. (2012). Measuring gap fraction, element clumping index and LAI in Sierra Forest stands using a full-waveform ground-based lidar. *Remote Sensing of Environment*, *125*, 73-79.

Zhao, F., Yang, X., Schull, M. A., Román-Colón, M. O., Yao, T., Wang, Z., Yao, T., Wang, Z., Zhang, Q., Jupp, D.L.B., Lovell, J.L., Culvenor, D.S., Newnham, G.J., Richardson, A.D., Ni-Meister, W., Schaaf, C.L., Woodcock, C.E., & Strahler, A. H. (2011). Measuring effective leaf area index, foliage profile, and stand height in New England forest stands using a full-waveform ground-based lidar. *Remote Sensing of Environment*, *115*(11), 2954-2964.

Zhu, X., Wang, TJ., Darvishzadeh, R., Skidmore, AK., Niemann, KO., (2015). 3D leaf water content mapping using terrestrial laser scanner backscatter intensity with radiometric correction, *ISPRS*

Photogrammetry and Remote Sensing, 110, 14-23, 10.1016/j.isprsjprs.2015.10.001**.**