"Bare earth" structure-from-motion data: Evaluating color-based point classification and fine-scale topography

R. Sare[1] and G. E. Hilley[1]

[1]Department of Geological Sciences, Stanford University, Stanford, CA, USA.

Corresponding author: Robert Sare (rmsare@stanford.edu)

Key points

- Machine learning using point colors improves iterative methods for lidar ground point classification
- Low vegetation is successfully filtered in lidar and structure-from-motion (SFM) data of forested landscapes, exposing topography
- Elevation, slope, and curvature can be derived from filtered SFM data with more landscape detail, but less continuity, than bare earth lidar

Abstract

Classification of ground points is a critical step in producing digital elevation models of the Earth's surface for studying landscape processes and geomorphic or ecological change. This paper describes a new algorithm for ground point classification and assesses the relative accuracy of the resulting ground surface in filtered light detection and ranging (lidar) and structure-from-motion (SFM) datasets. This color-enhanced multiscale curvature classification algorithm (MCCRGB) extends a popular lidar classification method (MCC) by introducing classification updates that distinguish vegetation and ground points by color. Multispectral lidar and SFM data imaging a subalpine volcanic tree kill are used to evaluate both methods. We find that color-based classification updates remove tree fall, low canopy, and brush, often requiring fewer iterations on large, ultra-high-density datasets. SFM data capture rills, small channels, and tree fall not visible in the lidar data. "Bare-earth" datasets from each method are internally consistent (mean vertical differences: -0.01 to 0.08 m) and validation at a set of 165 checkpoints shows a mean vertical difference of 0.46 m (standard deviation: 2.21 m) with the SFM ground points. The methods produce consistent topographic derivatives from each data source, including digital elevation models, slope, and profile curvature. While SFM derivatives are more variable and less continuous, the filtered products may be useful for geomorphic mapping and analysis, including mapping of microtopography or measuring landscape change in challenging, forested settings.

1 Introduction

Bare earth lidar data have transformed geomorphic and tectonic research over the last two decades by providing a high-resolution (< 1-m pixel), high-accuracy (< 20 cm) depiction of Earth's surface (Roering et al., 1999; Dietrich et al., 2003; Perron et al., 2009; Zielke et al., 2010; Kirby and Whipple, 2012; Perron and Royden, 2013; Johnstone and Hilley, 2015; Passalacqua et al., 2015; St Clair, Moon, et al., 2015) and changes in Earth's surface over time (Nissen et al., 2012; Scott et al., 2019; Diedrichs et al., 2019; DiBiase and Lamb, 2020). Recent advances in unmanned aerial vehicle (UAV) navigation and computer vision methods promise to image Earth's surface at even higher resolution at lower cost. However, the photogrammetric methods that underpin these point clouds do not penetrate above-ground features as do lidar laser pulses, which limits the use of photogrammetric point clouds in geomorphic and ecological

applications. Nonetheless, bare ground may be visible in these photogrammetric point clouds, and so point-classification algorithms that discriminate these bare-ground points from overlying vegetation would greatly increase the utility of these data for geomorphic and tectonic applications.

The general need to classify ground and vegetation points in both lidar and photogrammetric point clouds has motivated the development of iterative algorithms using curvature or morphological filtering to separate ground and non-ground points based on their relative height (Haugerud and Harding, 2001; Zhang et al., 2003; Evans and Hudak, 2007). More recently, the popular CANUPO algorithm combined multi-scale geometric features such as local planarity with semi-supervised classification based on user-defined class boundaries for multi-class point classification (e.g., bare rock, river gravels, vegetation) (Brodu and Lague, 2012). Other recent work has incorporated color or multispectral information into classification (Ekhtari et al., 2018; Tan et al., 2018). In geomorphology and ecology, these methods are used to differentiate tree crowns and canopy from smoothly-varying topography in lidar data, which often contains multiple returns where laser pulses penetrate vegetation. Classified points are used to interpolate bare earth representations of topography. Misclassification of low vegetation and vegetation on slopes are common failure modes of iterative methods where vegetation height is similar to natural topographic variability (Montealegre et al., 2015). This can complicate geomorphic analysis by introducing non-topographic roughness into bare-earth digital elevation models (DEMs).

The proliferation of photogrammetric point clouds from new modalities requires automated filtering methods that scale to the very high point densities of these datasets and accurately classify the low vegetation they image. Structure-from-motion (SFM) datasets are frequently used in studies of fine-scale topography in geomorphology (Westoby et al., 2012; Bemis et al., 2014; Johnson et al., 2014). Many studies of topographic change use pre- and post-event lidar data (*e.g.*, Nissen et al., 2012; Scott et al., 2019; Diederichs et al., 2019; DiBiase and Lamb, 2020), and recent work has used repeat SFM surveys to measure coastal erosion (James and Robson, 2012; Cook and Dietze, 2019) and landslide deformation (Carey et al., 2019; Pickering et al., 2019). As UAVs and other photogrammetric platforms (*e.g.*, satellite stereo imaging) are applied to more challenging settings, delineating the ground surface in the presence of

vegetation may become an important prerequisite to mapping microtopography or measuring landscape change.

While software such as ENVI Lidar and Terrascan include proprietary filtering algorithms, open-source alternatives provide widely available and reproducible methods for producing useful derivatives from point cloud data. This work provides an automated open-source and scalable method of classifying photogrammetric point clouds, by building on an established ground classification method, multiscale curvature classification (MCC; Evans and Hudak, 2007) to exploit color attributes of photogrammetric data. We validate the accuracy of the new method, called MCCRGB, on a false color, multispectral lidar point cloud and an SFM point cloud of Horseshoe Lake, a subalpine evergreen forest and tree kill area at Mammoth Mountain, California, USA. Classification results are compared to independent classification of ground points performed by the lidar vendor. We find that the new method requires fewer iterations than MCC filtering and reclassifies many low non-ground points that are difficult to filter using only relative height.

We assess the relative accuracy of the ground points from each method. To ensure that the datasets could be vertically compared, SFM point clouds were constructed with ground control points registered to the vendor-classified lidar ground surface. Filtered ground point clouds are vertically consistent, with sub-meter vertical differences between lidar and SFM that are similar to reported differences in previous work that compares SFM to ground-based lidar scanning (Westoby et al., 2012; James et al., 2017). Results are also evaluated at a set of ground checkpoints surveyed with moderate horizontal precision (~0.48 m). While the vertical precision of these checkpoints is poor and the corresponding vertical differences are large (mean: 0.45-0.46 m for both methods), we find that MCC and MCCRGB produce ground point clouds with sub-meter vertical differences when checkpoints are vertically registered to the lidar ground surface.

To evaluate the utility of filtered data for geomorphic mapping and analysis, bare earth DEMs were produced from the ground points for each method and data source. The SFM DEMs capture much of the topographic detail of the lidar DEMs, and include rills, boulders, small channels, and tree fall not visible in the bare earth lidar. Derivatives such as slope and profile curvature from each data source are also consistent (RMSEs of 0.01 and 0.001 m^{-1}). Flow

paths and accumulation areas were determined by routing flow over each DEM. The lidar and SFM flow paths are similar in plan form but differences in tributary flow and junction position lead to discrepancies in accumulation area up to 20% along the largest channel in the area. Finally, SFM DEMs have data gaps where no ground points are available, and we find larger discrepancies between elevation, slope, and curvature in low-point-density areas. Where point density is reasonably high, our results suggest that MCCRGB can efficiently produce ground-only point clouds and derivatives suitable for analyzing fine-scale topography. Classified data from SFM and other modalities will likely enable new discoveries in geomorphology, much as bare earth lidar data continues to advance our understanding of the earth's surface.

2 Methods

2.1 MCC

The MCC method and related techniques iteratively discard non-ground points based on their relative height above an interpolated surface (Haugerud and Harding, 2001; Evans and Hudak, 2007). The algorithm requires two parameters, a scale domain (s) and curvature tolerance (c; hereafter referred to as "height tolerance"). The scale domain determines the resolution of the interpolated surface at each iteration and the height tolerance defines the relative height cutoff for non-ground points. In its original implementation, MCC iterates over three scale domains of $0.5s, s,$ and $1.5s$ using a default height tolerance of $c = 0.3$ m.

At each iteration, a candidate ground surface is interpolated and points exceeding a height tolerance are classified as high, non-ground points. All remaining points are not classified. The non-ground points are discarded and the procedure is repeated until a stopping condition is reached. The algorithm continues to the next scale domain, interpolating a comparison surface and discarding non-ground points until it converges to a final set of points that are classified as ground. The reader is referred to Evans and Hudak, 2007 for a detailed description.

2.2 MCCRGB

Each iteration of the MCCRGB algorithm is divided into two stages: MCC classification and an update step (Figure 1). First, a relative-height-based classification is performed for a chosen

scale domain as in an iteration of MCC. In the update step, a support vector machine (SVM) classifier is trained using color attributes with the current ground/non-ground classification as training labels. Then, each ground point's class is predicted by the classifier and re-assigned to the non-ground class if they are classified as non-ground by the SVM. The remaining ground points are retained and the procedure is repeated for the current scale domain until a convergence criterion is met. The algorithm repeats this procedure through the remaining scale domains.

The default behavior of MCCRGB is to perform the classification update only for the first scale domain, but classification updates can be performed at any user-defined scales and height tolerances. These are intended to capture the vegetation points in a given height range; for example, the suggested MCC tolerance would yield a classifier trained on non-ground points above 0.3 m of relative height. Additional height ranges could be specified to classify vegetation points below this height tolerance (*e.g.*, lower trees or bushes).

The features used by the SVM classifier are derived from point color attributes and do not incorporate spatial information or relative height. A color index, the normalized green-red difference index (NGRDVI), is calculated as

$$NGRDVI = \frac{Green - Red}{Green + Red} \tag{1}$$

Large positive values of NGRDVI emphasize green points, similar to the normalized difference vegetation index (NDVI) and other indices derived from multispectral or optical imagery (Tucker, 1979; Tan et al., 2018). The 8-bit input (R, G, B) colors are transformed into a CIE-Lab color space of (L, a, b) values to capture perceptual color contrasts. L (lightness) varies from 0 (black) to 100 (white), a from -128 (green) to 128 (red), and b from -128 (blue) to 128 (yellow). This transformation uses the `rgb2lab` function of scikit-image with default illuminant D65 and observer aperture angle of 2° (van der Walt et al., 2014).

To reduce the effects of shadowing, the lightness is ignored, and each point has a corresponding feature vector

$$x_i = [a_i, b_i, NGRDVI_i]^\top \quad (2)$$

In order to delineate a non-linear class boundary, these feature vectors are transformed to a high-dimensional feature space using an approximation to a radial basis function kernel (RBF) before training the SVM (Rahimi and Recht, 2008; 2009). The same transformation is applied in the prediction step. Point labels are assigned as $y_i = 1$, or ground, if the point falls below the height tolerance at the current scale, and 0, or non-ground, otherwise. The classification update step modifies only the labels of ground points, changing the value from 1 to 0 if they are predicted to be in the non-ground class based on their color features. The values of the radial parameter γ of the kernel, the dimensionality of the transformed features, and the regularization parameter α of the SVM were selected using inspection of site SW data (Figure 2) classified using different parameter combinations, as SFM point labels were not available for cross-validation (Table S3).

The classification problem of the MCCRGB update step is to distinguish vegetation points from a poorly-defined set of putative ground points as defined by the intermediate point labels. This is a semi-supervised classification problem or outlier detection problem, as opposed to supervised binary classification, which requires well-defined ground/non-ground classes and a large set of labelled points. In contrast to CANUPO, which is also a semi-supervised method, MCCRGB does not require users to provide labeled points for training. However, it is restricted to a limited number of vegetation classes, while CANUPO can classify multiple texturally distinct point classes, which is especially useful for terrestrial laser scanning data (Brodu and Lague, 2012; Lague et al., 2013).

2.3 Validation

We assessed elevation accuracy with a point-to-point comparison using both field-surveyed checkpoints and randomly selected vendor ground points (Table 1). The elevation of each checkpoint and its nearest horizontal neighbor in the target dataset were compared. The nearest neighbor search was performed by indexing the (x, y) positions of the target dataset in a two-dimensional tree and querying this tree for the checkpoint's nearest neighbor. Randomly selected lidar points classified as ground by the vendor were also used for validation, chosen in

areas of at least 10 SFM points m^{-2} (Figure S7), with vertical differences calculated in the same way.

2.4 Topographic derivatives

Digital elevation models (DEMs) were produced from ground-only point clouds at 1 m resolution by inverse distance weighted interpolation using vendor and MCCRGB classifications (PDAL Contributors, 2018). Topographic slope was calculated using the second order centered finite difference method (Zevenberge and Thorne, 1987) and local profile curvature was calculated in the direction of maximum gradient using a second-degree polynomial fit to elevations in a 5 m square window (Wood, 1996).

For channel network analysis, these products were void-filled with the priority flood algorithm (Barnes et al., 2014). Flow directions and accumulation (watershed) areas were calculated using the D8 algorithm (O'Callaghan and Mark, 1984). In order to compare accumulation area magnitudes, the lidar DEM was clipped to the extent of the SFM DEM prior to flow routing. We report "partial" areas from accumulating flow over the small SFM survey area (0.67 km^2), not the full lidar survey area (9.1 km^2). A channel network was extracted from each DEM by applying a partial area threshold of 30 m^2 (Figure S9).

The distance between flow paths from each DEM was measured by a local pixel-to-pixel search. For each source pixel in the SFM DEM, a corresponding target pixel was identified in the lidar DEM in a 20 x 20 m square window. The target pixel was selected as the pixel having the smallest difference in accumulation area with the source pixel. The local difference between channel networks was calculated as the ℓ_2 distance between source and target pixels. This is intended to capture the adjustment required to match flow paths in the SFM DEM to those in the lidar DEM.

3 Data

3.1 Study site

Horseshoe Lake (HSL, Figure 2) is a subalpine lake south of Mammoth Mountain, California where diffusive release of carbon dioxide (CO_2) has created a volcanic tree kill. Coniferous trees and tree fall comprise the majority of vegetation in the area. The area has been subject to monitoring of atmospheric and soil CO_2 flux surveys by the USGS/California Volcano Observatory since 1993 following the formation of tree kills during seismic unrest in 1989 (Sorey et al., 1998; Werner et al., 2014; Lewicki et al., 2019). This uses a network of 179 flux measurement stations; 165 of these are used as checkpoints in this study. The tree kill area is approximately 0.5 km^2 (Farrar et al., 1999) with elevations between 2730 and 2800 m. Slopes range from flat, bare ground near Horseshoe Lake to slope greater than 25° on the northern and western areas of the tree kill. The topography includes the steep south flank of Mammoth Mountain, moderate gravel- and sand-covered slopes of the central tree kill, and intermittently-active channels and rills draining into Horseshoe Lake.

3.2 Datasets

3.2.1 Lidar data

A multispectral lidar dataset was collected in an airborne survey of HSL by the National Center for Airborne Laser Mapping (NCALM) on September 22, 2018 (Figure 2a). The survey used an Optech Titan multispectral lidar system operating three lasers, two at shortwave infrared wavelengths of 1550 nm (channel 1) and 1064 nm (channel 2), and one green laser with a wavelength of 532 nm (channel 3). Datasets were delivered in Universal Transverse Mercator projected coordinates (NAD83(2011) / UTM Zone 11 N) with ellipsoidal elevations (GRS80). Point classification was performed by NCALM with Terrascan followed by manual verification and cleaning of ground points (hereafter referred to as “vendor ground points”). The reported accuracy is 0.023 m root-mean-square error (RMSE) relative to survey ground control points (GCPs). The point cloud consists of three scan lines, one for each channel, imaging the 9.4 km^2 survey area with an average point density of 8.57 points m^{-2} (NCALM, 2019).

3.2.2 SFM data

An unmanned aerial vehicle survey of HSL was conducted in October 2019 comprising 841 nadir and oblique photographs acquired to mitigate doming error (James and Robson, 2014).

Color point clouds were derived from each survey by SFM photogrammetry using Agisoft Metashape software (Agisoft, 2019). The HSL survey area included nine GCPs surveyed using a Trimble ProXH Global Positioning System (GPS) receiver, which provides sub-meter accuracy (<0.30 m) via differential GPS. The average GCP horizontal (vertical) precision was 0.38 m (0.43 m) following correction against Continuously Operating Reference Station data from reference station P630 (Table S1). A set of 165 GPS checkpoints was surveyed by the same procedure in November 2019 with an average horizontal (vertical) precision of 0.48 m (0.76 m) (Table S2). While no major topographic change occurred during the year between lidar and SFM acquisitions, a hazardous fuels reduction project began in 2019, resulting in tree clearing throughout the survey area and excavation at the borrow pit (BP, Figure 2c).

3.2.3 Data processing

The datasets were augmented before analysis. First, the raw lidar data were reprocessed into false-color point clouds. Each lidar point in the channel 1 scan lines was assigned a color value by a nearest neighbor search. The intensities from the closest two channel 2 and 3 returns were assigned to each point in the channel 1 scan lines, and the 16-bit intensity values were rescaled to 8-bit color depth.

Additionally, the GCPs were vertically referenced to the vendor-defined lidar ground surface. The nearest neighbor of each point was determined using the horizontal coordinates of the surveyed points and lidar points, and each GCP was vertically translated to the elevation of the nearest lidar ground point as defined by the vendor classification. The translations varied from -1.9 m to 1.3 m (Table S1). This registration was performed to mitigate the effect of relatively low vertical accuracy of GCPs and enable relative accuracy assessment between SFM and lidar data.

4 Results

4.1 Point classification

Three sites at HSL were selected to test point cloud filtering: two tree kill sites capturing live, transitional, and dead tree cover (site SW) and a forested slope (site NW), and the borrow pit

(site BP) to the northeast of the tree kill (Figure 2). They contain 46,792 lidar points (4,067,195 SFM), 238,933 (20,902,531) points, and 62,487 (2,873,948) points.

Both lidar and SFM datasets were processed with the MCC method with a scale domain of 1 m and height tolerances of 0.3 m and 0.03 m. A comparison of MCCRGB with baseline MCC classification using a height tolerance of 0.3 m is shown in Figures 3 and 4. The MCC method removes the majority of live and dead standing trees and HSL sites SW (Figures 3c and 4c) and NW (Figures S1c and S2c). Tree fall and vegetative cover imaged in the SFM data are classified as ground points by MCC at site SW (Figure 3b and c) and a small bush is misclassified as ground at site NW (Figure S2c). A small number of points were reclassified at sparsely-vegetated site BP (Figures S3 and S4). Over the entire tree kill area, green lower canopy and tree fall was removed by MCCRGB but classified as ground by MCC (Figure 5). Finally, testing (not shown) indicated that both methods are more sensitive to height tolerance than interpolation scale, with tolerances 0.01 m and less resulting in misclassification of topography, with patches in flat areas and contour artifacts on sloping terrain.

4.2 Ground surface accuracy assessment

Ground point clouds were validated against checkpoints and lidar point elevations (Table 1). The mean and median vertical errors at the checkpoints are similar for each method and dataset, ranging from 0.35 m (MCCRGB applied to lidar) to 0.46 m (MCC and MCCRGB applied to SFM data) with higher standard deviations and RMSEs of 2.08 m to 2.26 m. The minimum and maximum checkpoint errors are also similar, with a range of -15.72 m to 8.24 m, likely due to low GPS vertical accuracy. Only a small number of outlier points have high-magnitude vertical errors against the surveyed checkpoint elevations (Figure S6a-d), primarily near the edge of the survey area. These points are on steep slopes and under tree canopy, conditions that degrade the accuracy of the GPS positions.

Comparison with vendor ground points yields smaller average vertical errors because these points are vertically consistent with both the lidar and SFM data. Both methods yield low mean errors (-0.01 m to 0.06 m) and vertical RMSE (0.07 m to 0.63 m) on each dataset compared to vendor ground points. The minimum and maximum errors are larger for the SFM data (~-1 to 8 m) than for the lidar data (~-1 to 0.6 m) in each case. Inspection of the point clouds shows that

this reflects topographic and ecological change at the tree kill between 2018 (lidar) and 2019 (SFM). The vertical differences at the locations of the checkpoints are also relatively small, with the largest differences near the tree kill boundary (Figure S6f-h). Over the entire survey area, cloud-to-cloud distances between the vendor ground points and MCCRGB SFM ground points are within 1 m (Figure S8). In all SFM-lidar comparisons, tree clearing and excavation created vertical changes in excess of 1 m (Figure S6a-d; Figure S8b).

4.3 Slope and curvature

The lidar and MCCRGB SFM DEMs both capture the distribution of slope and curvature at HSL (Figure 6). The SFM slope field is rougher than lidar slope in the western section of the tree kill; there, fallen trees are apparent in the SFM dataset, but not the lower-density lidar data (Figure 6b). The steepest slopes occur to the north of BP on the southern flank of Mammoth Mountain. The SFM slope field is variable on that forested slope due to data gaps not present in the lidar DEM, which is interpolated from ground returns in the area. Overall, the two slope fields are consistent with high variability; the residual RMSE is 0.011 and standard deviation is 0.103 (Table 2). The lidar-derived profile curvature field is more continuous and less variable than the SFM equivalent (Figure 6c-d). The SFM curvature field has large gaps due to the 5 m window used for the curvature calculation. However, the two curvature fields are consistent with an overall residual RMSE of 0.001 and standard deviation of 0.036 (Table 2).

4.4 Accumulation area

The HSL survey area contains three channels, one of which attains accumulation areas greater than 1 km^2 over its catchment in the full lidar survey area. This channel is similar in plan view and accumulation area in both DEMs (Figure 7). The lidar- and SFM-derived channels are within seven meters of each other along the flow path (Figure 7d) and differ by up to 1000 m^2 in partial accumulation area (Figure 7c, e). Near the head of this channel in the survey area, the accumulation areas differ by a small amount. Flow routing discrepancies result in less flow accumulation in the lower part of the channel in the SFM dataset beyond 250 m of channel length as less flow is delivered from the northern slope (Figure 7e). However, this difference is a relatively small proportion of total accumulation area at this position (Figure 7f). Finally, the flow

paths diverge in the low-relief borrow pit area, resulting in different channel outlets in the two datasets (Figure 7a, b).

5 Discussion

5.1 Point classification

5.1.1 Removal of multiple returns

Intermediate returns are discarded aggressively by the classification updates of MCCRGB (Figure 8). Whereas MCC gradually removes intermediate and first returns with each iteration, MCCRGB reclassified 24,851 ground points as non-ground in a single update step at site SW (iteration 0, Figure 8b). This led to a total of 88,528 non-ground points being discarded following the update, a 41% increase over the initial non-ground points classified by MCC, which removed only 60,204 points out of 238,933 points in the lidar point cloud (Figure 8a). As a result, MCCRGB required only eight iterations in this case and subsequent MCC spline interpolations were slightly more efficient, using fewer points for nearest-neighbor averaging. Finally, the similarity between the MCC and MCCRGB results suggests that the new method quickly discards many points that would eventually be classified as non-ground by MCC.

5.1.2 Which points are reclassified by color?

Recent work on point classification of SFM data for geomorphic applications has focused on specifying a difference index cutoff value for vegetation points (Tan et al., 2018). By contrast, the MCCRGB method imposes no minimum NGRDVI. Instead, the classifier is sensitive to the distinctive greenness of the non-ground training examples, reclassifying points with higher NGRDVI values (Figure 9). The classification approach removes the need to specify an NGRDVI cutoff, which may be useful at sites where there is not a multimodal distribution of NGRDVI, as discussed in Tan et al., 2018. However, in areas with built structures or less color contrast, the use of a color-based classifier may be inappropriate and a morphological filter could yield better results (*e.g.*, Zhang et al., 2003).

5.2 Topographic derivatives

Real topography and data artifacts contribute to slope and curvature variability. Small-scale features like rills, boulders, and fallen trees are better represented in high density SFM data, but SFM cannot image many points below tree canopy or other cover. Consequently, the mean and range of slope and curvature differences between SFM and lidar DEMs is higher in low-point-density areas (Figure 10). Many of these are data gaps and areas where ground is obscured by canopy in the SFM data. Lidar derivatives are less variable in these areas because ground returns are available. However, the lower point density and resolution of lidar products may fail to capture fine-scale topography even where the ground is imaged. Roughness and data gaps also contribute to small differences in flow routing that yield moderate differences in accumulation when aggregated along a channel's length (Figure 7).

6 Conclusions and future work

6.1 Applications in geomorphology

Bare earth lidar data have underpinned many new research directions and discoveries in geomorphology over the past two decades. In terms of critical measurements, geomorphic mapping with lidar has quantified the slip history of major faults, including segments of the San Andreas Fault (Zielke et al., 2010). Lidar has also enabled fundamental advances in geomorphology by providing elevation data at scales relevant to geomorphic processes (Dietrich et al., 2003). These include the study of nonlinear hillslope creep (Roering et al., 1999), lithologic influences on hillslope creep (Johnstone and Hilley, 2015), the characteristic spacing of ridges and valleys (Perron et al., 2009), topographic stresses in the critical zone (St Clair, Moon, et al., 2015), and many applications of the stream-power equation in fluvial geomorphology (e.g., Kirby and Whipple, 2012; Perron and Royden, 2013). Finally, lidar differencing has recently revealed hazards-related landscape change, including near-fault and anelastic earthquake deformation (Scott et al., 2019; Diedrichs et al., 2019) and debris flows following wildfires associated with dry sediment loading (DiBiase and Lamb, 2020). Classified photogrammetric data have the potential to drive discoveries in geomorphology in a similar way.

Classified SFM data may be useful in many key areas. They enable measurements of microtopography, such as rilling, deposit or bed roughness at boulder to cobble scales, and

small scale variations in slope that affect near-surface fluid and gas transport. Results of this study also suggest that these data could be used in larger-scale fluvial geomorphology, including channel extraction, slope-area analysis, or steepness/chi analysis, if the survey captures an entire catchment. Change detection around tree cover is another intriguing application. While many studies have focused on large-scale deformation, classified SFM data could be used to measure change in challenging settings, such as small earthquake surface ruptures, coseismic cracking, and debris flows or accelerated erosion in forested catchment areas upstream of fire scars. Most significantly, UAV-based surveys are inexpensive and rapidly deployed, and so abrupt landscape change might be captured under a broader range of circumstances than airborne lidar surveys. Finally, the resolution of space-based optical imaging platforms has increased over the last decade. As the resolution of these images increases, point clouds derived from these coarser-resolution data might be classified using MCCRGB in a similar way to SFM point clouds, which might allow sparse (> 1-m spacing) point clouds to be derived over large, moderately vegetated areas from space.

6.2 Conclusions

Point color and other spectral characteristics provide important information for point classification. We developed a machine learning method that exploits the color of vegetation points to refine iterative, curvature-based point classification. Testing on co-located multispectral lidar and structure-from-motion datasets shows that the ground surfaces produced by this method and standard multiscale curvature classification are consistent. Furthermore, color classification updates remove low vegetation and intermediate returns that can challenge height-based classification in sloping or complex terrain. Finally, we found that both methods produce consistent estimates of topographic slope, curvature, and accumulation area, although SFM derivatives are more variable. Both MCC and MCCRGB produce data gaps in SFM data where the ground is occluded by vegetation. However, if point density is high, MCCRGB can be a more efficient method for filtering large datasets. Color-based classification could enable a wide range of geomorphic studies that were previously limited to arid environments or areas imaged by lidar by filtering novel drone and satellite datasets.

Acknowledgments

We thank Stephen DeLong, Chelsea Scott, and Sam Thiele for insightful discussions; Jennifer Lewicki for sharing survey data for checkpoints; and Aaron Steelquist for field assistance. Lidar data acquisition and processing completed by the National Center for Airborne Laser Mapping. NCALM funding provided by NSF's Division of Earth Sciences, Instrumentation and Facilities Program. EAR-1339015. This work is based on data services provided by the OpenTopography Facility under NSF Award Numbers 1833703, 1833643, and 1833632. Code is available at https://doi.org/10.5281/zenodo.3519582 and data is available on OpenTopography.

References

AgiSoft Metashape Professional (Version 1.6.0) (Software). (2019). Retrieved from http://www.agisoft.com/downloads/installer/. Accessed 11 November 2019.

Barnes, R., Lehman, C., and Mulla, D., (2014). Priority-flood: An optimal depression-filling and accumulation-labeling algorithm for digital elevation models. *Computers & Geosciences*, 62, pp.117–127. https://doi.org/10.1016/j.cageo.2013.04.024

Bemis, S.P., Micklethwaite, S., Turner, D., James, M.R., Akciz, S., Thiele, S.T. and Bangash, H.A., (2014). Ground-based and UAV-based photogrammetry: A multi-scale, high-resolution mapping tool for structural geology and paleoseismology. *Journal of Structural Geology*, *69*, pp.163-178. https://doi.org/10.1016/j.jsg.2014.10.007

Brodu, N. and Lague, D., (2012). 3D terrestrial lidar data classification of complex natural scenes using a multi-scale dimensionality criterion: Applications in geomorphology. ISPRS Journal of Photogrammetry and Remote Sensing, 68, pp.121-134. https://doi.org/10.1016/j.isprsjprs.2012.01.006

Carey, J.A., Pinter, N., Pickering, A.J., Prentice, C.S. and Delong, S.B., (2019). Analysis of Landslide Kinematics Using Multi-temporal Unmanned Aerial Vehicle Imagery, La Honda, California. *Environmental & Engineering Geoscience*, pp.1-17. https://doi.org/10.2113/EEG-2228

Cook, K. L., & Dietze, M. (2019). Short Communication: A simple workflow for robust low-cost UAV-derived change detection without ground control points. *Earth Surface Dynamics*, *7*(4), 1009–1017. https://doi.org/10.5194/esurf-7-1009-2019

DiBiase, R. A., & Lamb, M. P. (2020). Dry sediment loading of headwater channels fuels post-wildfire debris flows in bedrock landscapes. Geology, 48(2), 189-193. https://doi.org/10.1130/G46847.1

Diederichs, A., Nissen, E.K., Lajoie, L.J., Langridge, R.M., Malireddi, S.R., Clark, K.J., Hamling, I.J. and Tagliasacchi, A. (2019). Unusual kinematics of the Papatea fault (2016 Kaikōura

earthquake) suggest anelastic rupture. Science Advances, 5(10), eaax5703. https://doi.org/10.1126/sciadv.aax5703

Dietrich, W. E., Bellugi, D. G., Sklar, L. S., Stock, J. D., Heimsath, A. M., & Roering, J. J. (2003). Geomorphic transport laws for predicting landscape form and dynamics. Geophysical Monograph-American Geophysical Union, 135, 103-132. https://doi.org/10.1029/135GM09

Ekhtari, N., Glennie, C. and Fernandez-Diaz, J.C. (2018). Classification of airborne multispectral lidar point clouds for land cover mapping. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *11*(6), pp.2068-2078. https://doi.org/10.1109/JSTARS.2018.2835483

Evans, J. S., and Hudak, A. T. (2007). A Multiscale Curvature Algorithm for Classifying Discrete Return LiDAR in Forested Environments. *IEEE Transactions on Geoscience and Remote Sensing: A Publication of the IEEE Geoscience and Remote Sensing Society*, *45*(4), 1029–1038. https://doi.org/10.1109/TGRS.2006.890412

Farrar, C. D., Neil, J. M., and Howle, J. F. (1999). Magmatic carbon dioxide emissions at Mammoth Mountain, California. *US Geol. Surv. Open-File Rep*, *98*, 4217. Retrieved from https://pubs.usgs.gov/wri/wri98-4217/

Haugerud, R. A., and Harding, D. J. (2001). Some algorithms for virtual deforestation (VDF) of LIDAR topographic survey data. *International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences*, *34*(3/W4), 211–218.

James, M.R. and Robson, S., 2012. Straightforward reconstruction of 3D surfaces and topography with a camera: Accuracy and geoscience application. *Journal of Geophysical Research: Earth Surface*, *117*(F3). https://doi.org/10.1029/2011JF002289

James, M.R. and Robson, S., 2014. Mitigating systematic error in topographic models derived from UAV and ground-based image networks. *Earth Surface Processes and Landforms*, *39*(10), pp.1413-1420. https://doi.org/10.1002/esp.3609

James, M. R., Robson, S., & Smith, M. W. (2017). 3-D uncertainty-based topographic change detection with structure-from-motion photogrammetry: precision maps for ground control and directly georeferenced surveys: 3-D uncertainty-based change detection for SfM surveys. *Earth Surface Processes and Landforms*, *42*(12), 1769–1788. https://doi.org/10.1002/esp.4125

Johnson, K., Nissen, E., Saripalli, S., Arrowsmith, J.R., McGarey, P., Scharer, K., Williams, P. and Blisniuk, K., 2014. Rapid mapping of ultrafine fault zone topography with structure from motion. *Geosphere*, *10*(5), pp.969-986. https://doi.org/10.1130/GES01017.1

Johnstone, S. A., & Hilley, G. E. (2015). Lithologic control on the form of soil-mantled hillslopes. Geology, 43(1), 83-86. https://doi.org/10.1130/G36052.1

Kirby, E., & Whipple, K. X. (2012). Expression of active tectonics in erosional landscapes. Journal of Structural Geology, 44, 54-75. https://doi.org/10.1016/j.jsg.2012.07.009

Lague, D., Brodu, N. and Leroux, J., 2013. Accurate 3D comparison of complex topography with terrestrial laser scanner: Application to the Rangitikei canyon (NZ). *ISPRS journal of photogrammetry and remote sensing*, *82*, pp.10-26. https://doi.org/10.1016/j.isprsjprs.2013.04.009

Lewicki, J. L., Evans, W. C., Montgomery-Brown, E. K., Mangan, M. T., King, J. C., & Hunt, A. G. (2019). Rate of Magma Supply Beneath Mammoth Mountain, California, Based on Helium Isotopes and CO 2 Emissions. Geophysical Research Letters, 46(9), 4636–4644. https://doi.org/10.1029/2019GL082487

Montealegre, A. L., Lamelas, M. T., & de la Riva, J. (2015). A Comparison of Open-Source LiDAR Filtering Algorithms in a Mediterranean Forest Environment. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *8*(8), 4072–4085. https://doi.org/10.1109/JSTARS.2015.2436974

National Center for Airborne Laser Mapping (2019). Point Classification for Hazard Applications: Mammoth Lakes, CA. Collected by National Center for Airborne Laser Mapping (NCALM). Distributed by OpenTopography. https://doi.org/10.5069/G9TT4P3S

Nissen, E., Krishnan, A.K., Arrowsmith, J.R. and Saripalli, S., (2012). Three-dimensional surface displacements and rotations from differencing pre-and post-earthquake LiDAR point clouds. *Geophysical Research Letters*, *39*(16). https://doi.org/10.1029/2012GL052460

O'Callaghan, J.F. and Mark, D.M., (1984). The extraction of drainage networks from digital elevation data. *Computer vision, graphics, and image processing*, 28(3), pp.323-344. https://doi.org/10.1016/S0734-189X(84)80011-0

Passalacqua, P., Belmont, P., Staley, D.M., Simley, J.D., Arrowsmith, J.R., Bode, C.A., Crosby, C., DeLong, S.B., Glenn, N.F., Kelly, S.A. and Lague, D., (2015). Analyzing high resolution topography for advancing the understanding of mass and energy transfer through landscapes: A review. *Earth-Science Reviews*, *148*, pp.174-193. https://doi.org/10.1016/j.earscirev.2015.05.012

PDAL Contributors, (2018). PDAL Point Data Abstraction Library. doi:10.5281/zenodo.2556738

Perron, J. T., & Royden, L. (2013). An integral approach to bedrock river profile analysis. Earth Surface Processes and Landforms, 38(6), 570-576. https://doi.org/10.1002/esp.3302

Pickering, A.J., Prentice, C.S. and DeLong, S.B., (2019). *Landscape change associated with the upper Scenic Drive landslide, La Honda, California, January 10–June 28, 2017* (No. 2019-1024). US Geological Survey. https://doi.org/10.3133/ofr20191024

Rahimi, A., & Recht, B. (2008). Random Features for Large-Scale Kernel Machines. In J. C. Platt, D. Koller, Y. Singer, & S. T. Roweis (Eds.), *Advances in Neural Information Processing Systems 20* (pp. 1177–1184).

Rahimi, A., & Recht, B. (2009). Weighted Sums of Random Kitchen Sinks: Replacing minimization with randomization in learning. In D. Koller, D. Schuurmans, Y. Bengio, & L. Bottou (Eds.), *Advances in Neural Information Processing Systems 21* (pp. 1313–1320).

Roering, J. J., Kirchner, J. W., & Dietrich, W. E. (1999). Evidence for nonlinear, diffusive sediment transport on hillslopes and implications for landscape morphology. Water Resources Research, 35(3), 853-870. https://doi.org/10.1029/1998WR900090

Sare, R. and Hilley, G. E. (2019). pymccrgb: Color- and curvature-based classification of multispectral point clouds in Python. *Journal of Open Source Software*, 4(42), 1777, https://doi.org/10.21105/joss.01777.

Scott, C.P., Arrowsmith, J.R., Nissen, E., Lajoie, L., Maruyama, T. and Chiba, T., 2018. The M7 2016 Kumamoto, Japan, Earthquake: 3-D Deformation Along the Fault and Within the Damage Zone Constrained From Differential Lidar Topography. *Journal of Geophysical Research: Solid Earth*, *123*(7), pp.6138-6155. https://doi.org/10.1029/2019GL082202

Sorey, M. L., Evans, W. C., Kennedy, B. M., Farrar, C. D., Hainsworth, L. J., & Hausback, B. (1998). Carbon dioxide and helium emissions from a reservoir of magmatic gas beneath Mammoth Mountain, California. *Journal of Geophysical Research*, *103*(B7), 15303–15323. https://doi.org/10.1029/98JB01389

St Clair, J., Moon, S., Holbrook, W. S., Perron, J. T., Riebe, C. S., Martel, S. J., ... & Richter, D. D. (2015). Geophysical imaging reveals topographic stress control of bedrock weathering. Science, 350(6260), 534-538. https://doi.org/10.1126/science.aab2210

Tan, Y., Wang, S., Xu, B., & Zhang, J. (2018). An improved progressive morphological filter for UAV-based photogrammetric point clouds in river bank monitoring. *ISPRS Journal of Photogrammetry and Remote Sensing: Official Publication of the International Society for Photogrammetry and Remote Sensing* , *146*, 421–429. https://doi.org/10.1016/j.isprsjprs.2018.10.013

Tucker, C.J., (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote sensing of Environment*, *8*(2), pp.127-150. https://doi.org/10.1016/0034-4257(79)90013-0

U.S. Geological Survey and California Geological Survey (2006). Quaternary fault and fold database for the United States, accessed 1 May 2019 at http//earthquake.usgs.gov/hazards/qfaults/.

van der Walt, S., Schönberger, J. L., Nunez-Iglesias, J., Boulogne, F., Warner, J. D., Yager, N., et al. (2014). scikit-image: image processing in Python. *PeerJ*, *2*, e453. https://doi.org/10.7717/peerj.453

Werner, C., Bergfeld, D., Farrar, C.D., Doukas, M.P., Kelly, P.J. and Kern, C., (2014). Decadal-scale variability of diffuse CO2 emissions and seismicity revealed from long-term monitoring (1995–2013) at Mammoth Mountain, California, USA. *Journal of Volcanology and Geothermal Research*, *289*, pp.51-63. https://doi.org/10.1016/j.jvolgeores.2014.10.020

Westoby, M.J., Brasington, J., Glasser, N.F., Hambrey, M.J. and Reynolds, J.M., (2012). 'Structure-from-Motion' photogrammetry: A low-cost, effective tool for geoscience applications. *Geomorphology*, *179*, pp.300-314. https://doi.org/10.1016/j.geomorph.2012.08.021

Wood, J. (1996). The geomorphological characterisation of digital elevation models (Doctoral dissertation, University of Leicester).

Zevenbergen, L. W., & Thorne, C. R. (1987). Quantitative analysis of land surface topography. Earth surface processes and landforms, 12(1), 47-56. https://doi.org/10.1002/esp.3290120107

Zhang, K., Chen, S. C., Whitman, D., Shyu, M. L., Yan, J., & Zhang, C. (2003). A progressive morphological filter for removing nonground measurements from airborne LIDAR data. *IEEE transactions on geoscience and remote sensing*, *41*(4), 872-882. https://doi.org/10.1109/TGRS.2003.810682

Zielke, O., Arrowsmith, J. R., Ludwig, L. G., & Akçiz, S. O. (2010). Slip in the 1857 and earlier large earthquakes along the Carrizo Plain, San Andreas fault. Science, 327(5969), 1119-1122. https://doi.org/10.1126/science.1182781

Figures

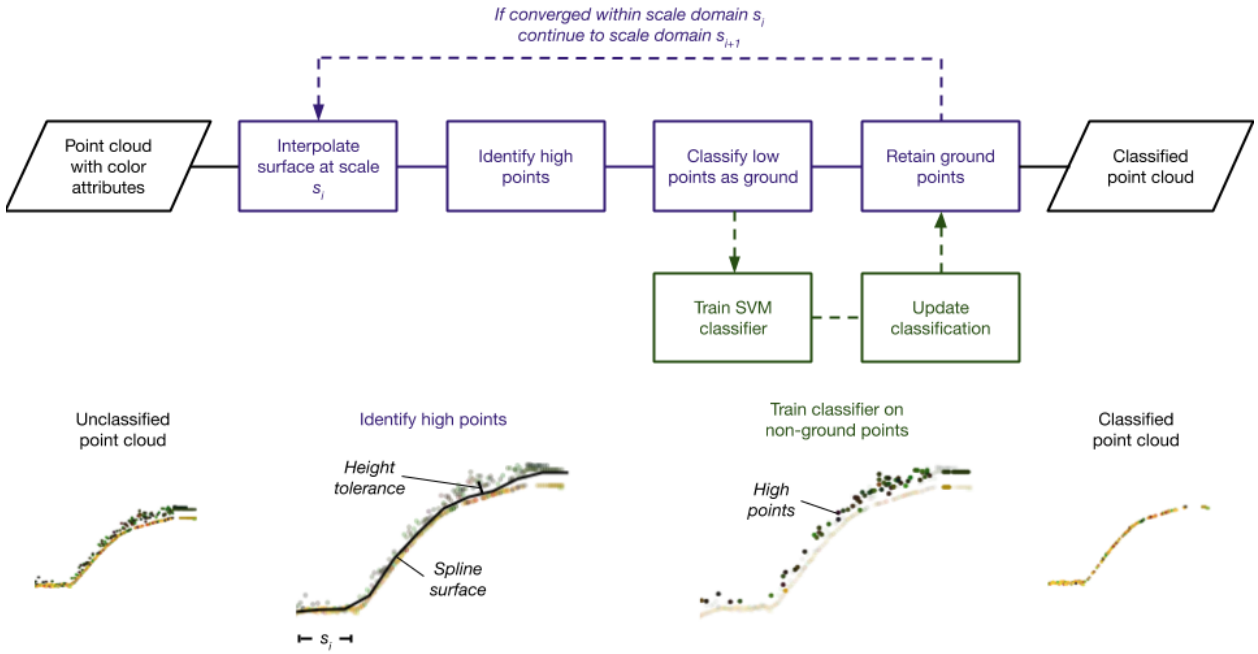

Figure 1. MCCRGB method. Purple steps show the MCC algorithm. Green steps are MCCRGB reclassification of ground points. SVM: Support vector machine.

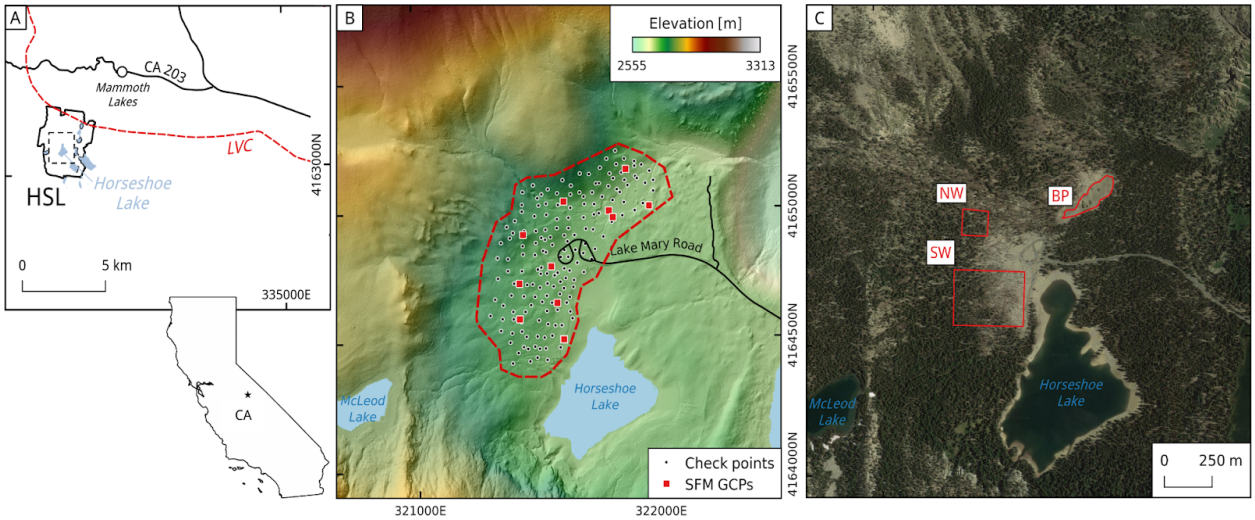

Figure 2. Horseshoe Lake tree kill. a) Map with extent of panels b-c (dashed black line), lidar dataset (solid black line), and Long Valley Caldera (LVC; dashed red line), b) Bare earth shaded relief map with SFM survey area (dashed red line), ground control points (red squares), and checkpoints (black circles). c) Satellite image showing differences in tree cover between test sites (red polygons). Lidar acquired 9/22/2018, imagery acquired 7/26/2016, copyright 2019 Google, DigitalGlobe. NAD83/UTM Zone 11 N. NW: Northwest treekill, SW: Southwest treekill, BP: Borrow pit.

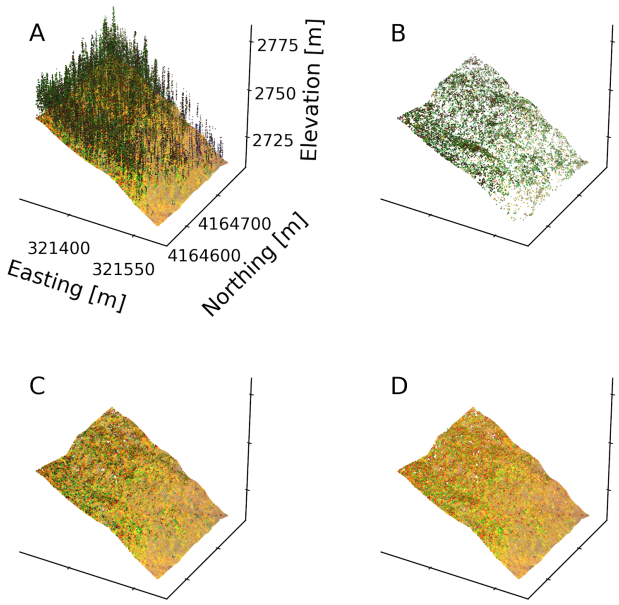

Figure 3. Results from lidar data at site SW. a) Full point cloud, b) Ground points that are reclassified as non-ground by MCCRGB, c) MCC ground points, d) MCCRGB ground points.

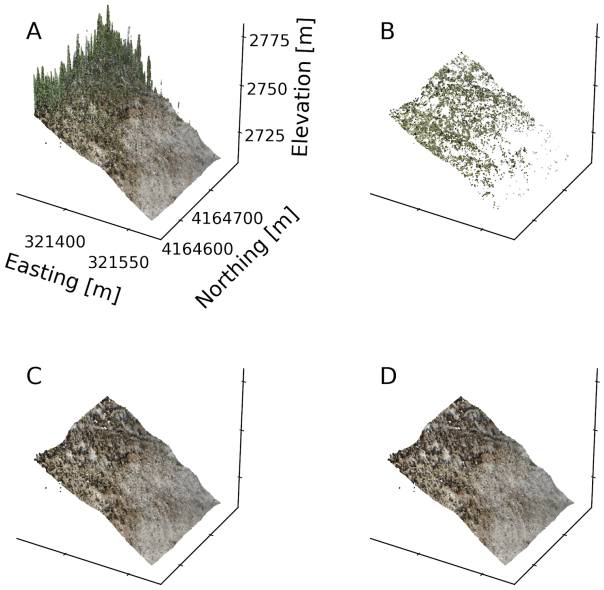

Figure 4. Results from SFM data at site SW. a) Full point cloud, b) Ground points that are reclassified as non-ground by MCCRGB, c) MCC ground points, d) MCCRGB ground points.

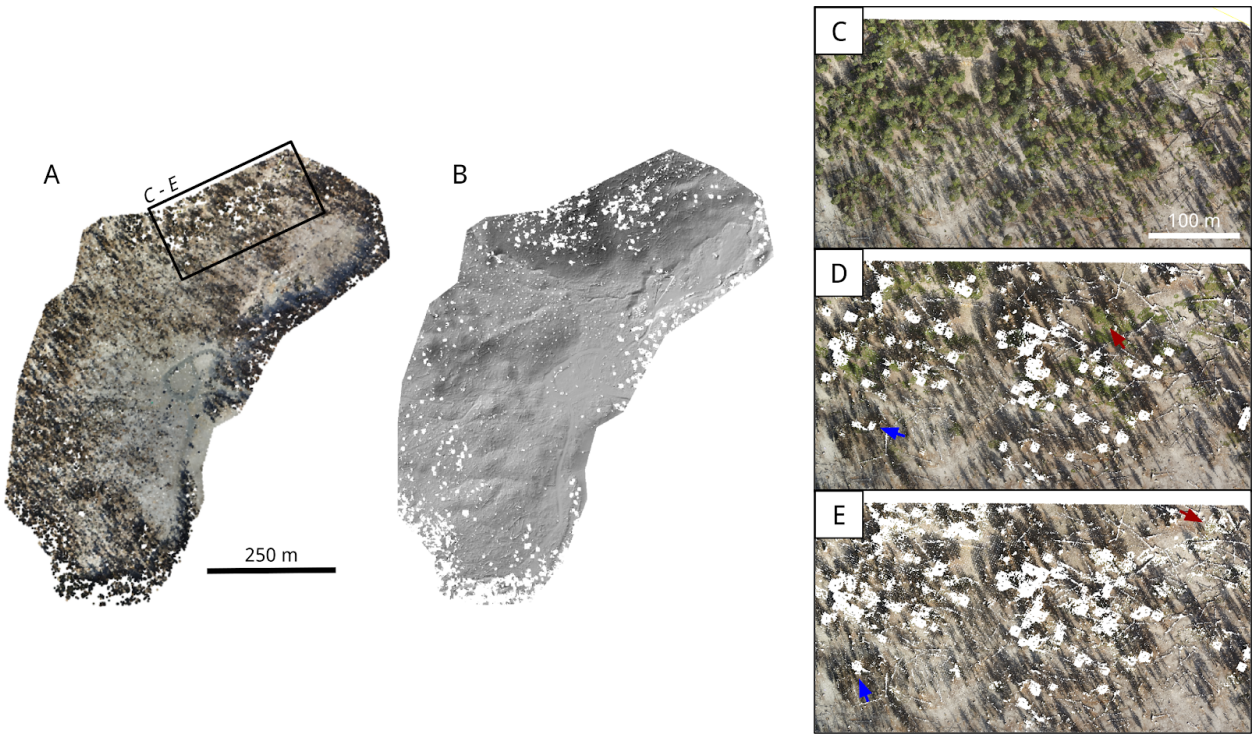

Figure 5. MCCRGB results. a) Ground-only point cloud and b) digital elevation model, 1 m resolution, c) Unfiltered point cloud of slope with tree fall, d), MCC ground point cloud, e) MCCRGB ground point cloud. Blue arrows indicate errors of omission and red arrows errors of commission.

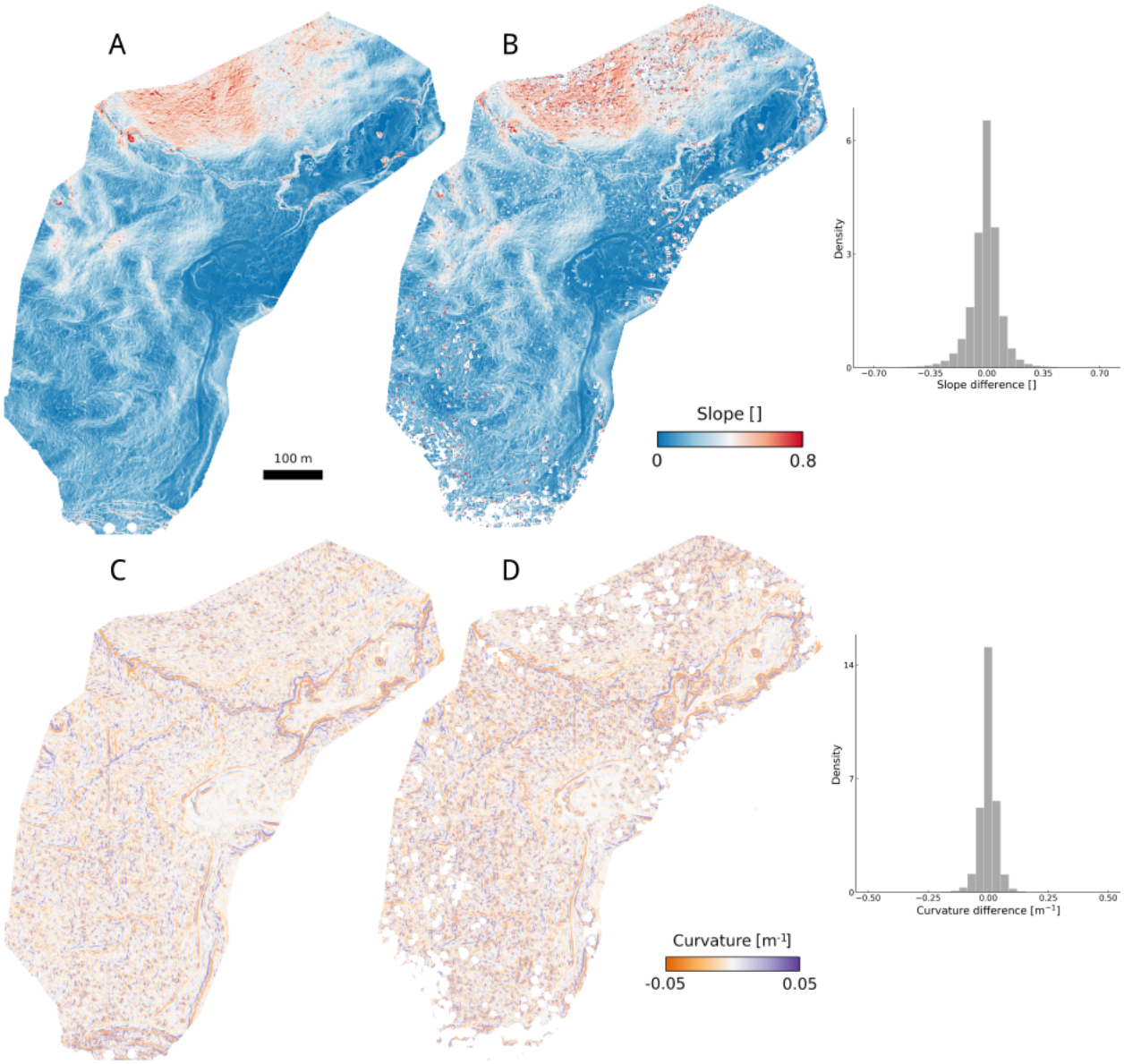

Figure 6. Topographic slope and curvature. Slope derived from a) vendor bare earth DEM and b) SFM MCCRGB ground-only DEM. Profile curvature derived from c) vendor bare earth DEM and d) SFM MCCRGB ground-only DEM using a 5 m window. Insets show histograms of differences (lidar - SFM).

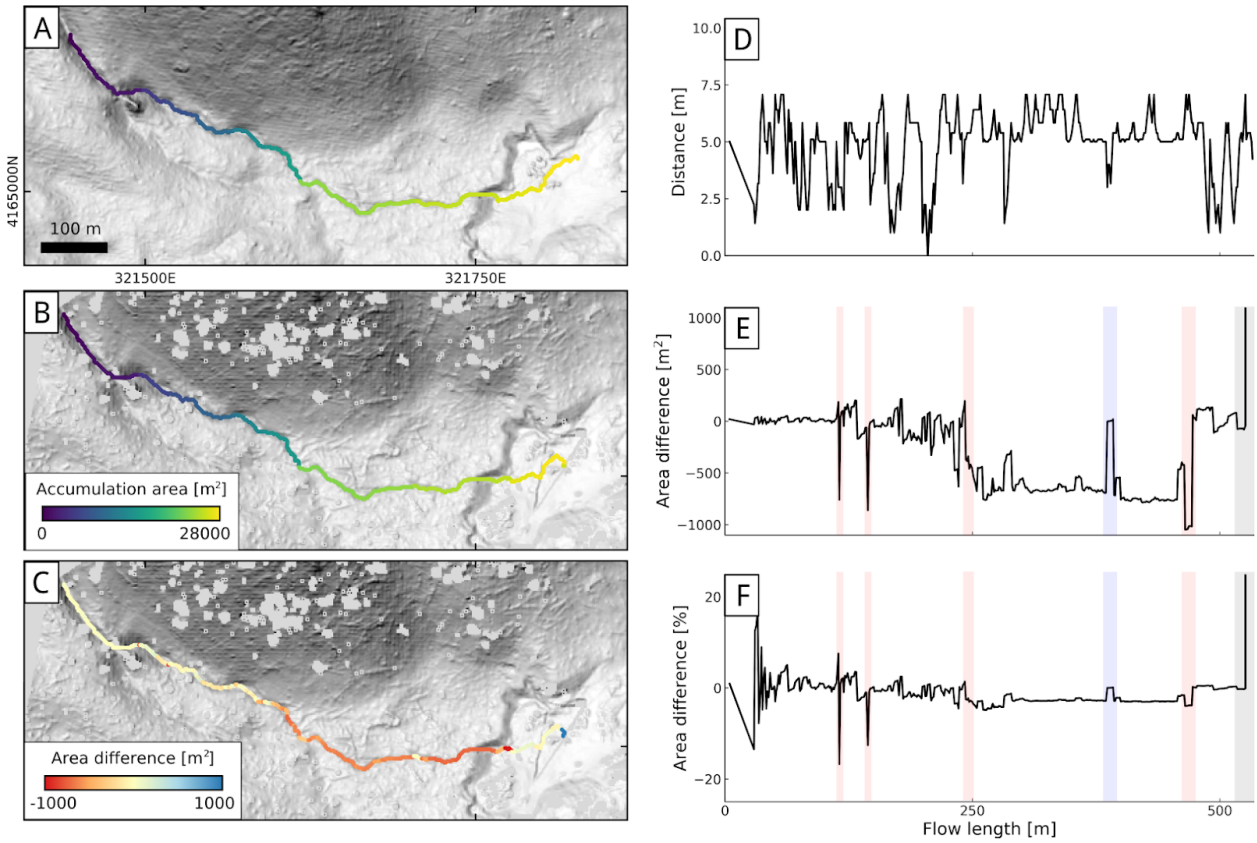

Figure 7. Partial accumulation area from a) lidar DEM and b) SFM DEM for largest channel in survey area. c) Area difference (lidar - SFM). d) Distance between channels, e) Area difference, and f) Relative area difference as percentage of lidar area along flow path in b). Shaded areas highlight effects of small tributaries with greater flow accumulation in lidar DEM (red) or SFM DEM (blue) and large area difference due to flow paths diverging near the outlet (gray).

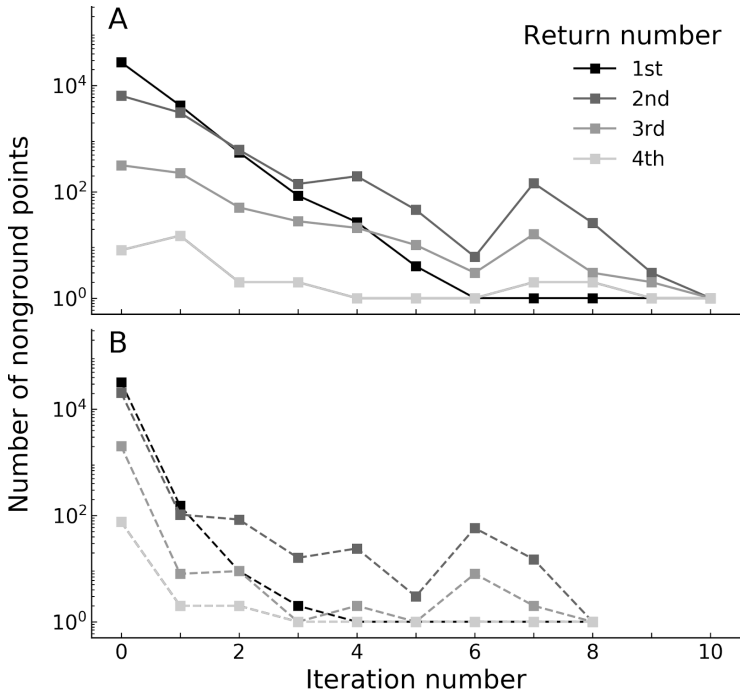

Figure 8. Lidar points classified as non-ground by a) MCC (solid lines) and b) MCCRGB (dashed lines) in each iteration. Note that the MCCRGB update step classifies more intermediate second and third returns as non-ground at first iteration. Based on lidar data at site SW (Figure 3). Only multiple returns shown. Vertical axis in logarithmic scale.

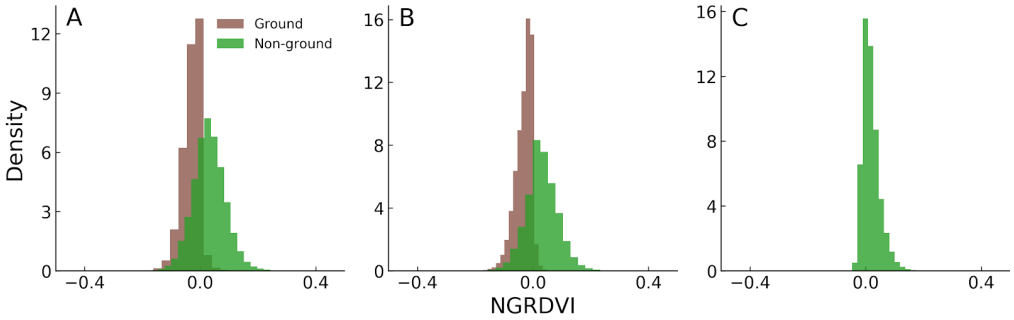

Figure 9. Histograms of NGRDVI from points classified as ground (brown) and non-ground (green). a) MCC only, b) MCCRGB, c) Points updated to non-ground by MCCRGB. Based on SFM data at site SW (Figure 4).

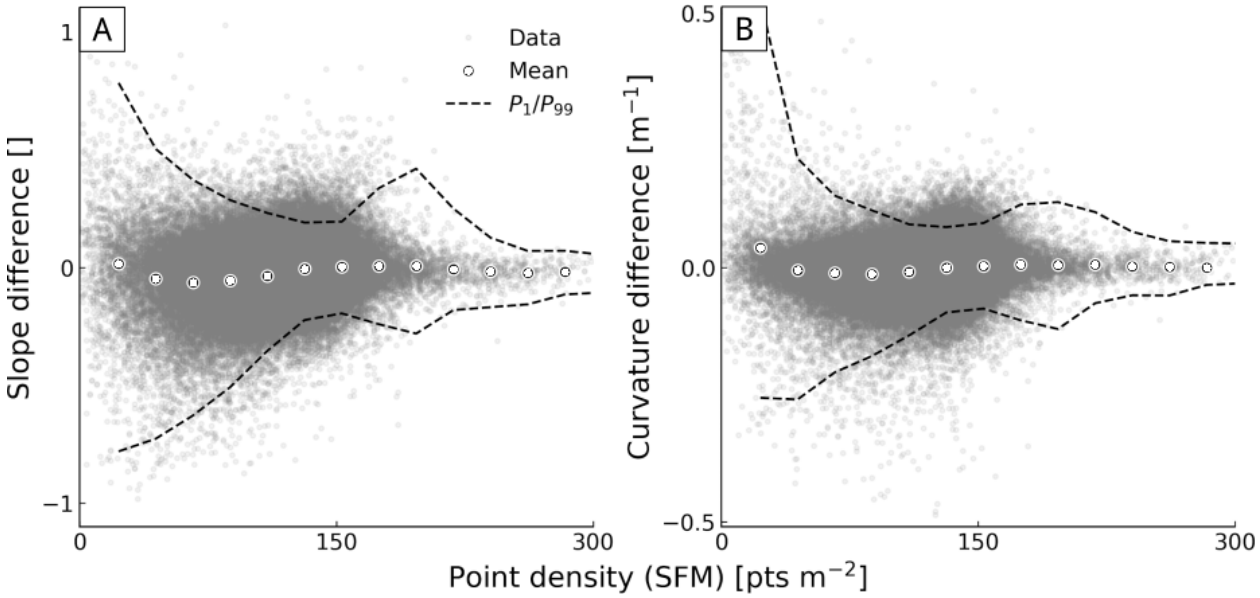

Figure 10. Point density. a) Slope and b) curvature differences as a function of point density. Windowed means (white circles) and 1st/99th percentiles (dashed lines) plotted over data (grey circles).

Table 1. Error statistics for checkpoints at Horseshoe Lake tree kill.

Table 2. Statistics for slope and curvature differences (lidar - SFM) at Horseshoe Lake tree kill. Distributions shown in Figure 6.

Table 1: Error statistics for ground points at Horseshoe Lake tree kill.

Statistic (m)	Surveyed check points				Registered check points				Vendor ground points			
	Lidar + MCC	Lidar + MCCRGB	SFM + MCC	SFM + MCCRGB	Lidar + MCC	Lidar + MCCRGB	SFM + MCC	SFM + MCCRGB	Lidar + MCC	Lidar + MCCRGB	SFM + MCC	SFM + MCCRGB
Minimum	-15.72	-15.72	-15.63	-15.63	-0.38	-0.59	-0.66	-0.64	-0.50	-0.70	-1.32	-1.08
Maximum	7.77	7.40	8.24	8.22	0.33	0.34	5.49	5.48	0.32	0.32	6.39	8.28
Mean	0.36	0.35	0.45	0.46	-0.01	-0.02	0.11	0.09	-0.01	-0.02	0.05	0.08
Median	0.45	0.46	0.47	0.46	0.00	0.00	0.04	0.03	0.00	0.00	0.02	0.03
Standard deviation	2.08	2.08	2.19	2.21	0.08	0.09	0.56	0.55	0.07	0.07	0.44	0.60
RMSE	2.11	2.11	2.23	2.26	0.08	0.10	0.57	0.56	0.07	0.08	0.44	0.61
n	165	165	165	165	165	165	165	165	1000	1000	1000	1000

Table 2: Statistics for slope and curvature differences (Lidar - SFM) at Horseshoe Lake tree kill.		
	Slope difference ( )	**Curvature difference (m-1)**
RMSE	0.011	0.001
Mean absolute error (MAE)	0.066	0.024
Standard devation	0.103	0.036
Minimum	-0.994	-0.644
Maximum	1.501	0.891
n (pixels)	331982	296790