11-23-2010

DEM Development from Ground-Based LiDAR Data: A Method to Remove Non-Surface Objects

M.P. Sharma  
*University of Wyoming; West Virginia University, sharma@uwyo.edu*

Ginger B. Paige  
*University of Wyoming*

Scott N. Miller  
*University of Wyoming*

Follow this and additional works at: [http://repository.uwyo.edu/chemical_facpub](http://repository.uwyo.edu/chemical_facpub)

**Publication Information**


This Article is brought to you for free and open access by the Chemical and Petroleum Engineering at Wyoming Scholars Repository. It has been accepted for inclusion in Chemical and Petroleum Engineering Faculty Publications by an authorized administrator of Wyoming Scholars Repository. For more information, please contact scholcom@uwyo.edu.
DEM Development from Ground-Based LiDAR Data: A Method to Remove Non-Surface Objects

Maneesh Sharma 1, Ginger B. Paige 2,* and Scott N. Miller 2

1 WV GIS Technical Center, Department of Geology & Geography, 330 Brooks Hall, West Virginia University, Morgantown, WV 26505 USA; E-Mail: maneesh.april@gmail.com
2 Department of Renewable Resources, University of Wyoming, Laramie, 1000 E. University Ave. WY 82071, USA; E-Mail: snmiller@uwyo.edu

* Author to whom correspondence should be addressed; E-Mail: gpaige@uwyo.edu; Tel.: +1-307-766-2200; Fax: +1-307-766-6403.

Received: 25 September 2010; in revised form: 12 October 2010 / Accepted: 8 November 2010 / Published: 23 November 2010

Abstract: Topography and land cover characteristics can have significant effects on infiltration, runoff, and erosion processes on watersheds. The ability to model the timing and routing of surface water and erosion is affected by the resolution of the digital elevation model (DEM). High resolution ground-based Light Detecting and Ranging (LiDAR) technology can be used to collect detailed topographic and land cover characteristic data. In this study, a method was developed to remove vegetation from ground-based LiDAR data to create high resolution DEMs. Research was conducted on intensively studied rainfall–runoff plots on the USDA-ARS Walnut Gulch Experimental Watershed in Southeast Arizona. LiDAR data were used to generate 1 cm resolution digital surface models (DSM) for 5 plots. DSMs created directly from LiDAR data contain non-surface objects such as vegetation cover. A vegetation removal method was developed which used a slope threshold and a focal mean filter method to remove vegetation and create bare earth DEMs. The method was validated on a synthetic plot, where rocks and vegetation were added incrementally. Results of the validation showed a vertical error of ±7.5 mm in the final DEM.

Keywords: topography; rainfall-runoff plots; vegetation removal; DSM; bare earth model
1. Introduction

Distributed process based hydrologic and erosion models use digital elevation models (DEMs) to discretize watersheds into contributing areas and route surface water across the terrain using overland flow planes and channels [1]. High resolution DEMs improve the ability to differentiate and model overland and concentrated flow paths on the landscape. Remote sensing satellite or airborne technologies such as SAR (Synthetic Aperture Radar), INSAR (Interferometric Synthetic Aperture Radar) and airborne LiDAR (Light Detecting and Ranging) have been used to scan the earth surface to develop digital elevation models (DEMs) used in hydrologic modeling, urban planning and forestry [2,3]. Airborne LiDAR has become an established technology for surveying large areas in relatively short time [4]. However, the captured data from airborne LiDAR often contain vegetation and other non-surface objects along with the surface topographic data. Various algorithms and methods have been developed to remove the vegetation and non-surface objects from the data and develop bare earth DEMs from DSM (digital surface model) [5-7].

Airborne LiDAR can collect point data at a range of resolutions depending on the applications needs and can produce very high (1 m) resolution terrain models. Some projects, such as detailed hydrologic and erosion studies on plot and hillslope scales require even finer resolution data and survey methods such as a Total Station or ground-based LiDAR can be used. Wick [8] demonstrated that topographic data from Total Station could be collected at up to 5 to 10 cm resolution with angular accuracy of 5–10 mm and distance accuracy of 5 mm. For sub-cm resolution data collection, recently developed ground-based LiDAR are being used. Data collected using ground-based LiDAR are not only higher in resolution but also easier and faster to capture as compared to the Total Station [9,10]. Ground-based LiDAR scanners target an area of interest and can produce scans at a range of resolutions depending on the distance to target and the project needs. Each scan must include 3 identifiable points if more than one shot is taken. GPS locations are collected to place the data in real world coordinates. In contrast, a Total Station collects surface elevation data at required individual topographic locations dependent on the field protocol. It is important to note that ground-based LiDAR data, though easy to collect, often require post processing in which multiple shots of the same topography need to be aligned and vegetation and other non-surface objects need to be removed to develop DEMs.

DEMs prepared from LiDAR data are typically more detailed than those obtained by traditional methods such as satellite data and aerial photography. However, LiDAR data also contain non-surface objects such as vegetation cover. The removal of non-surface objects to develop DEMs from LiDAR data DSMs can be difficult and time consuming. LiDAR can capture data in a wave or in discrete returns. Methods developed for discrete return airborne LiDAR data often use multiple returns to remove vegetation and non-surface objects from DSMs. The majority of bare earth DEMs from LiDAR have been developed by searching for the lowest points in a neighborhood by using morphological filters using curvature method and identifying them as “bare earth”. The curvature method makes use of multiple returns for the same place in identifying the curvature of trees or other non-surface objects and surface topography. Multiple studies have demonstrated the utility of this approach to remove non-surface objects and create bare earth DEMs [5-7,11-14]. In contrast, ground-based LiDAR uses only a single return from a single location; the loss of multiple returns can render some of the airborne LiDAR processing methods ineffective. A variety of ground-based LiDAR instruments (including the
Optech brand model used in this project) provide either first or last discrete return for a given scan. The last return pulse is often considered the best estimate of the ground and is often assumed to be the “bald earth”. However, even the last return will not necessarily be the ground as vegetation canopy may totally obscure the ground [15,16]. In addition, at small scales the utility of first and last return acquisition can be lost when the scanning device is stationary and the returns yield identical results do to the relative size of the footprint. The ability to develop bald earth DEMs from discrete return scanning systems is a challenge because the discrete point cloud data need to be classified and filtered [16]. Approaches that exploit the differences in first and last returns for airborne data can prove ineffectual in processing ground-based scans at small scales; thus there is a clear need for a functional method or approach to remove vegetation and non-surface objects from ground-based LiDAR DSMs acquired with ground-based LiDAR systems.

In this study, a method was developed to remove non-surface objects from high-resolution ground-based LiDAR DSMs to develop bare earth DEMs for small plots using single returns due to the lack in difference between first and last returns. The method was developed and tested using ground-based LiDAR data from rainfall simulator plots installed for hydrologic and erosion studies. The method was validated on a synthetic laboratory plot on which the placement of the surface objects (rocks and vegetation) was controlled.

2. Methods

2.1. Study Area

The field study was conducted on the Kendall 112 sub-watershed in the USDA-ARS Walnut Gulch Experimental Watershed (WGEW), Tombstone, Arizona (31°43′N, 110°41′W) (Figure 1). WGEW is located within the San Pedro River basin, in a semi arid brush-grassland complex in the transition zone between Sonoran and Chihuahuan deserts [15]. The study area is a grass-dominated, small 1.91-ha watershed with an average slope of 9.4%. The vegetation in Kendall 112 watershed is representative of Southwestern rangelands dominated by bunch grasses with associated shrubs. Intensive study plots

**Figure 1.** Location of Kendall 112 watershed in USDA ARS WGEW.
were installed in the watershed for rainfall, runoff and erosion research. These rainfall simulator plots are 2 m by 6 m in size and serve as the test bed for development of bare earth DEMs from ground-based LiDAR. Surface topographic data were collected using a ground-based LiDAR instrument on five of the 2 m × 6 m rainfall simulator plots installed in the sub-watershed [18].

2.2. Ground-Based LiDAR Unit and Data Acquisition

An Optech ILRIS 3-D ground-based LiDAR laser scanner [19] was used to collect high-resolution terrain (surface and vegetation) data on each of the study plots. The technology is based on the time of flight principle [9]. The LiDAR unit consists of a range measurement system and a deflection for the laser beam. The time for the light to travel out to the target and back to the LiDAR scanner is used to determine the “range” to the target. The deflection system directs a laser beam in the scanning direction. The accuracy of distance measurements depends on the intensity of the reflected laser light, which in turn depends on the reflectivity of the object surface. The reflectivity depends on the angle of incidence and surface properties [9]. The Optech system is a pulsed system with a vertical accuracy of 0.3 cm. The field of view is 40° (horizontal) × 40° (vertical). The unit has a measuring rate of 20 kHz. The laser wavelength used for scanning is 1,500 nm. The laser beam class (IEC 600825-1) is Class I. Beam divergence is 0.00974°, and minimum spot step in X and Y axis is 0.00115°. The raw range accuracy of point cloud data is 7 mm at 100 m and raw positional accuracy is 8 mm at 100 m. The instrument can acquire one of two signal returns, first or last. Differences in the first and last return can be attributed to a result of the pulse passing through very sparse vegetation or other substance that might cause multiple returns; dense vegetation, rocks, soil or other hard objects do not generate a second return. Preliminary efforts yielded no difference in the response surface when the first and last returns were compared; thus we were unable to exploit the option of using both first and last return. Our approach was to use the last return to remove any confusion resulting from the occasional pulse traveling through light vegetation and to maintain a consistent approach throughout the study.

The LiDAR scanner was used to collect surface data in the form of a non-spatially referenced 3-dimensional point cloud for each of the five plots. The short wavelength is useful for ensuring accurate distance to any surface object in the laser beam’s path. This creates a shadow effect behind the vegetation where no surface data can be recorded. To account for the shadow effect and develop a complete DSM of the entire plot, LiDAR scans of each plot were taken from multiple directions and then “stitched” together to form an integrated 3-D dataset of each plot using 3-dimensional data processing software [20]. Common control points were established on each of the plots and were included in all scans for the plot. These common points were used to align and merge the point cloud data images. Images were aligned according to their best-fit alignment and comparison using a threshold of 0.01 standard deviation. Data in areas of overlap were thinned and composite images were merged to produce a single point cloud DSM for each plot.

The resulting DSM for each plot was then subjected to the vegetation removal process to produce a bare earth DEM. This is a multi-step process (Figure 2) executed using GIS tools. After creating a DSM, the first step was to create a slope map from the DSM using standard GIS algorithms [19]. The second step was to determine a slope threshold for each plot that could be used to identify and remove areas of vegetation. The vegetation “patches” on each plot have higher local slopes than the adjacent
surface topography. The slope threshold range was based on an evaluation of the 5 study plots in Kendall 112 watershed and an additional set of 5 shrub-dominated plots in WGEW and was determined by comparing the slope map and photographs for the respective plots. The vegetated areas were “removed” from the slope map and converted to null values. A limitation of using a simple slope threshold is that the tops of the vegetation demonstrate low slopes but at higher elevation levels than the surrounding surface topography. The “leftover” vegetation areas were identified by checking the elevation in deleted areas as compared to the adjacent non-deleted areas and comparing the DSM with a photograph of the respective plot.

**Figure 2.** Flow diagram showing steps used to remove vegetation from DSM.
The final step in the vegetation removal process was to fill null value areas to create a bare earth DEM. A $5 \times 5$ focal mean filter (where each pixel is $1 \text{ cm}^2$) was used for this process [21-23]. It was specified that only cells with values (i.e., pixels with elevation values) be used to find the focal mean of the target cell. This means that if a NoDATA (i.e., pixel with a null value) value existed within a neighborhood of the focal mean filter, then the NoDATA value would be ignored [21-23]. This approach minimizes the smoothing effect on the DEM; the NoData values are filled with mean cell value of the neighborhood. The filter was run multiple times in an iterative process from the edge of the patch until all of the NoDATA values within the DSM were filled and a bare earth DEM was created. To avoid excessive smoothing of the data, once a pixel was filled with a value, it could change.

2.3. Validation

A synthetic experimental plot was created in a controlled environment to validate the vegetation removal method. The synthetic 2 m by 4 m plot was prepared in an indoor facility where rock and vegetation were manually manipulated to create pre- and post-vegetation surfaces. The validation test plot was scanned at 5 mm resolution following protocols identical to those used on the field plots. The first scans were taken of the soil surface with rock cover but without vegetation to create a “true” bare earth surface model. It was important to include rocks in the bare earth scans as they can have a significant impact on microtopography and flow dynamics at the plot scale. The objective was to evaluate the ability of the method to remove vegetation with minimal changes to the soil and rock surface of the plot. Vegetation was added to the plot before the next scans to create the DSM (Figure 3). As with the field plots, multiple scans were taken of the plot both without and with vegetation and merged together to develop integrated 3-dimensional bare earth DEM and DSM, respectively. The vegetation areas in the DSM were removed using the newly developed vegetation removal method (Figure 2) to create a final DEM.

**Figure 3.** Photos of the synthetic validation plot with and without vegetation.

A two-step process was used to evaluate the results of the vegetation removal method and validate the final DEM. First, a difference map (DSM minus final DEM) was developed to evaluate the elevation values and locations of the vegetation areas removed from the DSM to create the final DEM. This allowed the direct comparison of the “removed” and filled areas to locations on the plot where
vegetation had been placed. Second, the final DEM was compared to the original bare earth DEM and an error map created using Equation 1 to show the distribution of error in the final DEM.

$$\% \text{ErrorMap} = \frac{(DEM - FinalDEM)}{DEM} \cdot 100$$

(1)

3. Results and Discussion

3.1. Field Plots

The vegetation removal method was carried out on all five of the grass and shrub-dominated plots on the Kendall sub-watershed; results are shown using one of the plots. DSMs were created for each plot at 1 cm resolution. An example of a merged DSM prior to vegetation removal is shown in Figure 4. Vegetated areas are recognizable as abrupt changes in elevation within the DSM, note the presence of large shrubs in the top and left of the plot. Figure 5 shows a slope map (in degrees) of the same plot; again the vegetation areas are identifiable. The vegetation areas show up as high slope areas when compared to the low slope topographic areas.

A slope threshold range of 58°–65° was determined to be appropriate for deletion based on evaluation of the rainfall runoff plots. The slope threshold ranged from 58° to 64° for the grass-shrub dominated plots in Kendall 112 and 59°–65° for the additional shrub-dominated plots that were used to test the threshold. Figure 6 shows a plot from which areas with a slope greater than 60° were removed. As stated earlier, the top of vegetation can show up in DSM as areas of high elevation-low slope areas. Figure 7 shows the plot after running a 5 × 5-cell focal majority filter; it replaced pixels with elevation values (green pixels) in between pixels with null value (blue vegetation patch) with the majority value of the filter. Note that the surface is devoid of aberrant high spots. At this point only data representing the ground remain in the terrain data.

**Figure 4.** DSM created from XYZ point cloud data collected from ground-based LiDAR unit.
**Figure 5.** Slope map of the plot. Slope values were derived from 1 cm DSM of plot. Reddish areas have slopes greater than 60°.

![Slope map of the plot](image)

**Figure 6.** Plot with vegetation areas removed. A slope threshold of 60° was used to delete the vegetation patches. Blue patches in the map are vegetation areas converted to null values.

![Plot with vegetation areas removed](image)

The final step in the vegetation removal process was to fill the deleted areas in the raster image (Figure 7) using a focal mean filter to create the bare earth DEM (Figure 8.) Results show that a maximum of 60 cm elevation, related to grass and shrub vegetation was removed from different areas of the DSM (Table 1). The highest local elevation value in the DSM was 1523.17 m (Figure 3) compared with 1522.63 m in the bare earth DEM (Figure 7). The range of elevation removed from shrub-dominated plots was 40–62 cm and from 5 grass-dominated plots was 23–44 cm. On shrub-dominated plots, the filled areas can be large and therefore require many passes using the focal
mean filter to interpolate the fill values. The result is that the filled pixels closest to the edge of the shrub areas are most likely more accurate than the filled pixels close to the center of a removed patch.

**Figure 7.** Raster image after $5 \times 5$ focal majority filter was run. Blue patches show removed vegetation areas. Inside the blue patches there are few green colored areas as compared to patches in Figure 5.

**Figure 8.** Bare earth DEM at 1 cm resolution. The vegetation was removed from the DSM and deleted areas filled with focal mean filter.
Table 1. Comparison of elevation values between DSM and DEM for the field plot.

<table>
<thead>
<tr>
<th></th>
<th>DSM</th>
<th>DEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum (m)</td>
<td>1521.71</td>
<td>1521.75</td>
</tr>
<tr>
<td>Maximum (m)</td>
<td>1523.17</td>
<td>1522.63</td>
</tr>
<tr>
<td>Mean (m)</td>
<td>1522.18</td>
<td>1522.16</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.25</td>
<td>0.25</td>
</tr>
</tbody>
</table>

3.2. Validation Plot Results

The difference map comparing the DSM with the final DEM for the validation plot is shown in Figure 9. Elevation values differed between the DSM and final DEM from a low of −0.14 m (red) to a high of +0.47 m (dark blue). The light green to greenish yellow color that dominates the DEM indicates areas where there is negligible change in elevation values in both DEMs. Positive values indicate higher relief areas in the DSM as compared to final DEM. Negative values represent areas that have higher elevation in the final DEM. The relative changes in the final DEM are due to the iterative process of the filter.

Figure 9. Difference map showing areas where vegetation was removed from the DSM for the validation plot. The dark blue color indicates areas of higher elevation in the DSM (height of vegetation). The reddish color indicates areas that are relatively higher in the final DEM. Light greenish to yellowish areas indicate no change in elevation values between the two models.

The percent error map comparing the bare earth DEM with the final DEM for the validation plot showed error in the range of ±5% (Figure 10). Figure 11 shows the percent error map with the
polygons showing rocks and vegetation. Most of the positive error is in the areas from where vegetation was removed and most of the negative error is from the areas where some rocks were removed. As a result the final DEM shows slightly lower elevation in reddish colored areas of the percentage error map and slightly higher elevation in blue colored of the percentage error map. Minimum, maximum, mean and standard deviation of elevation values for both DEMs were calculated (Table 2). There was an arbitrary datum of 3.5 m set for the plot base (Figure 3). The maximum height of rock that was placed in the test plot was 15 cm. The maximum error shown in Figure 10 is 5%. At a maximum error of 5 %, the difference in elevation between final DEM and bare earth DEM was 7.5 mm, calculated using the absolute elevation values (Table 2.) The raw range of accuracy for collection of LiDAR data is 7 mm; therefore, the error in the created bare earth DEM (±7.5 mm) only slightly exceeds the error range of data collection.

**Figure 10.** Percent error map for the final DEM. Dark blue areas indicate an error of 3.5% (4.9 mm) while dark red areas indicate an error of −5% (7.5 mm). Yellow to light blue areas show areas of negligible error.

3.3. **Implications**

We identified a suitable method for vegetation removal (primarily shrubs) at the plot scale. This solution allows very high resolution LiDAR-derived DEMs to be used in fine-scale runoff and erosion research, thus enabling improved understanding of the mechanics of rill and inter-rill erosion. The use of ground-based LiDAR is a non-destructive method for gathering a rich data set that is very suitable for long-term research efforts as it allows for periodic sampling without human disturbance due to sampling. Plots in the research area demonstrate relatively little abrupt changes in topography, which is important as steep slopes and cliffs would introduce confusion and noise to the algorithm we developed. Our approach exploits differences in terrain and vegetation morphology by identifying
vertical structures for removal. Future research is aimed at reducing the need for expert opinion in the algorithm development, specifically the determination of slope thresholds. Slope thresholds identified in this study that determine whether a landscape element will be removed or retained should work for most grass and shrub-dominated rangelands, but additional verification data are required for systems with different vegetation structure such as oak-woodlands. Research is suggested at larger scales, specifically large hillslope to watershed scales to determine how this approach will perform when geomorphic variation increases.

Figure 11. Percent error map for the final DEM showing polygons for known locations of the rocks and vegetation on the validation plot.

Table 2. Comparison of elevation values between the DSM and final DEM for the validation plot.

<table>
<thead>
<tr>
<th></th>
<th>DSM</th>
<th>final DEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum (m)</td>
<td>3.87</td>
<td>3.92</td>
</tr>
<tr>
<td>Maximum (m)</td>
<td>4.78</td>
<td>4.80</td>
</tr>
<tr>
<td>Mean (m)</td>
<td>4.34</td>
<td>4.36</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.22</td>
<td>0.23</td>
</tr>
</tbody>
</table>

4. Conclusions

Ground-based LiDAR can be used to collect very high resolution topographic and terrain data to create high resolution DEMs. High resolution DEMs are used to generate accurate channels and planes needed for hydrologic and erosion simulation models. Ground-based LiDAR allows the collection of high-resolution data over fairly large areas in a reasonable amount of time. However, LiDAR data also contain non-surface objects such as vegetation cover. Though the non-surface data are invaluable for
examining vegetation structure and distribution across the terrain, they impede our ability to easily develop bare earth DEMs. The challenge is to remove vegetation cover from LiDAR data and generate a DEM that accurately represents surface topography. In this study, a method was developed and tested to create high resolution (1 cm) bare earth DEMs from ground-based LiDAR using single (last) return data. The vegetation removal method used a slope threshold and a focal mean filter to remove vegetation and create a bare earth DEM. Validation of the vegetation removal method resulted in an error in the DEM up to ±7.5 mm. The sub-cm error indicates that the vegetation removal method can be used to remove vegetation at plot scale. Although the vegetation removal method has been developed for plot scale, it should also be valid for vegetation removal from upland areas that have topography similar to the rainfall simulator plots. However, this method needs to be expanded to develop additional methods to remove vegetation from terrains that have uneven topography such as complex hillslopes, ravines and river valleys.

Acknowledgements

We gratefully acknowledge the Wyoming NASA Space Grant Consortium for providing partial funding for this project. We also thank the University of Wyoming, Graduate School and Department of Renewable Resources and staff at the USDA-ARS Walnut Gulch Experimental Watershed for their support. Sudhir Raj Shrestha and Abhishek Tilaye are acknowledged for their help during this study. In addition, we thank the reviewers from Remote Sensing; their thorough and thoughtful reviews made the manuscript much stronger.

References


© 2010 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/3.0/).