USING STATISTICAL METHODS TO CORRECT LIDAR INTENSITIES FROM GEOLOGICAL OUTCROPS

Reuben Reyes and Jerome A. Bellian
Jackson School of Geosciences, Bureau of Economic Geology
University of Texas
Austin, TX U.S.A.
Reuben.Reyes@beg.utexas.edu
Jerome.Bellian@beg.utexas.edu

Erwin W. Adams
12 Shell International Exploration and Production B.V.
Kessler Park 1, 2288 GS Rijswijk, The Netherlands
Erwin.Adams@shell.com

ABSTRACT

Laser intensity reflection strength from ground-based lidar instruments are affected by several natural factors. External factors like humidity, barometric pressure, and airborne dust, can change the reflection intensity of lidar even over relatively short distances (less than 1 km). Instead of applying variable corrections for each external factor, we used statistical methods to correct the laser intensities. Three discrete attributes were applied to separate and isolate point intensity populations. These are distance, angle of incidence and surface roughness. Histograms and correction tables were used to reduce the effect these variables had on resultant laser intensity values. A more consistent intensity distribution was achieved that better related to the rock properties being mapped. This method was used on over one hundred individual scans totaling more than 300 million individual laser observation points. Working statistical methods were coded in software to read and process lidar scans in a native binary format. This reduced the number of processing steps, kept the data in its original format, and minimized disc space allocation. Using this application along with monochromatic lidar can be used to detect any attribute receptive to the laser frequency used by the lidar instrument.

Key words: lidar, statistics, intensity, reflection.

INTRODUCTION

Because reflection intensities normally degrade as a function of distance, software is included with lidar scanners to correct for this degradation. The software works well under optimal conditions, but can not accommodate for all external environmental factors. External factors like humidity, altitude, barometric pressure, and airborne dust, can change the reflection and intensity in lidar scans. Instead of applying a different correction for each of the different external factors, we decided to use a statistical correction of intensities. The method we used created histograms from the raw intensity data. The histogram we created used two separate bins, one for the number of points at a set distance and one for the summation of the intensity values of each point at that distance. If each bin had enough points, it could be represented by a Gaussian curve of the intensity values at a set distance, which we then averaged to get the median value. Originally, the scans we used had intensity values of 0 to 255, so the half intensity was an output value of 127. The median value that we calculated for each bin was subtracted from this original output value (127). The subtraction shifts the input intensities to the middle of the original output values. The histogram bins now had a value that was added to correct for the drop in intensity for that distance. The bins became an index to correct for intensities as a function of distance. The raw data was then reread and the distance of each point was indexed to shift the intensity to the median output value. A multiplier was also used to spread the intensities across the output range.

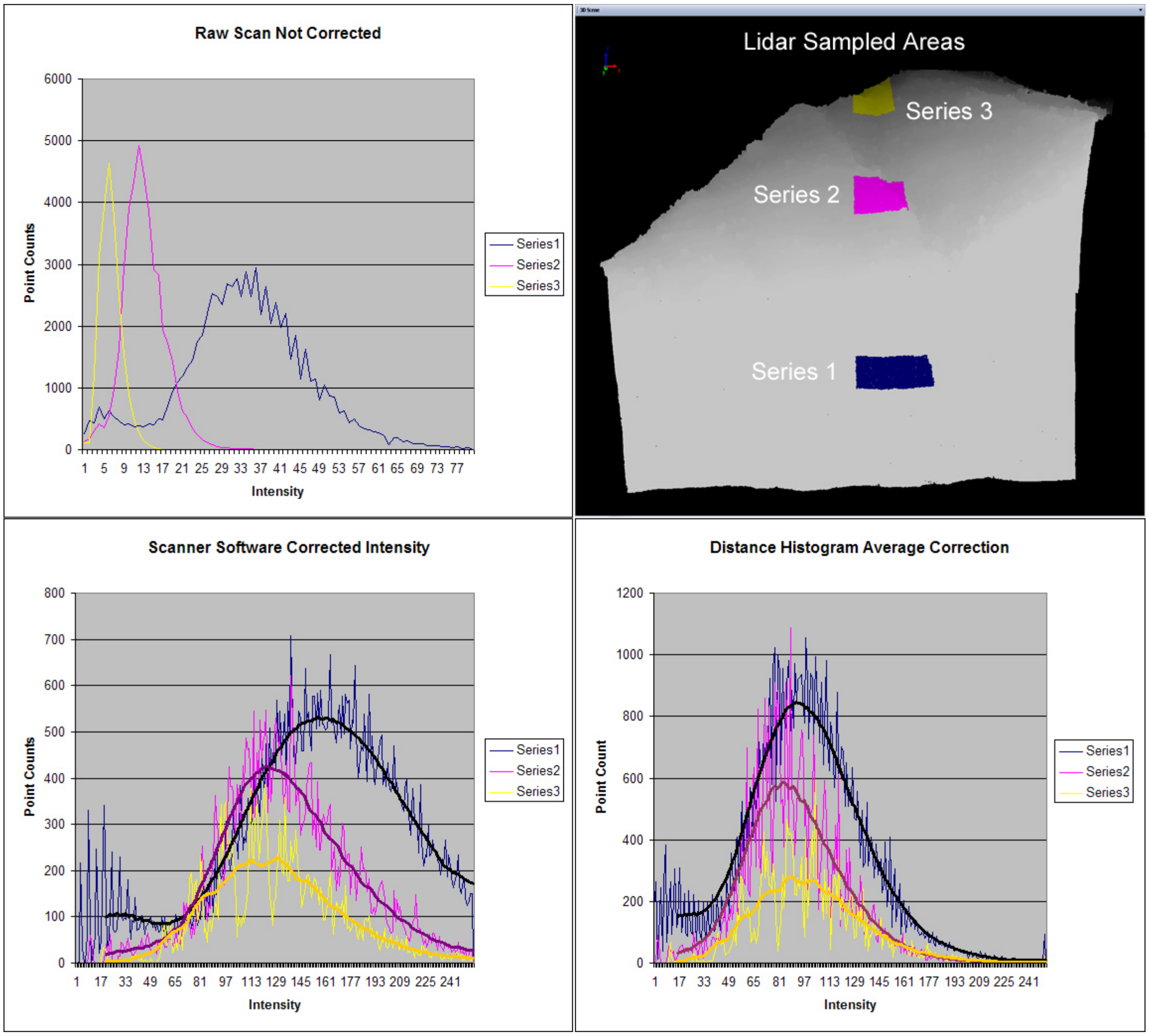


Figure 1. Histograms of distance series sampled areas from outcrop point cloud.

The lidar scans of outcrops had an even distribution of intensities at near and far distances. This produced an adequate population for each bin of distances. If this were not the case, using this method would still work, but would be skewed to the median. In some outcrops having large areas with little variation in intensity, using this method might produce more detail, but may not represent true intensity values for the entire outcrop.

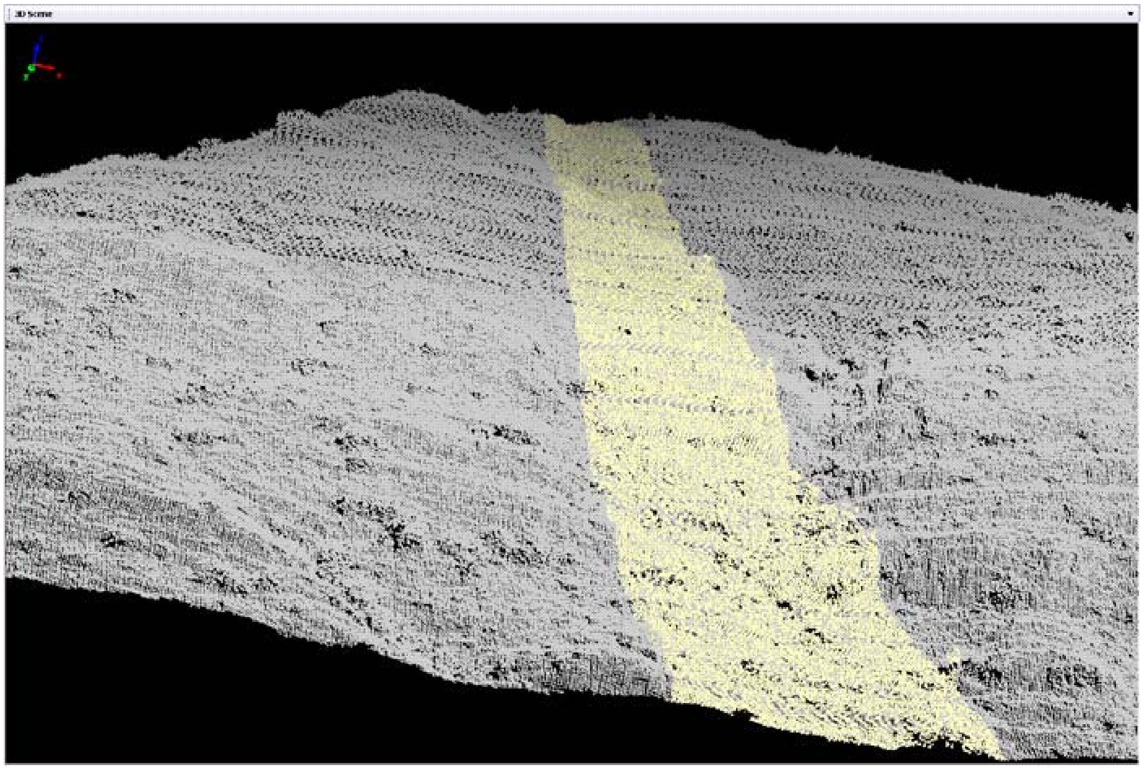


Figure 2. Outcrop point cloud band of points used to create "Intensity average over Distance".

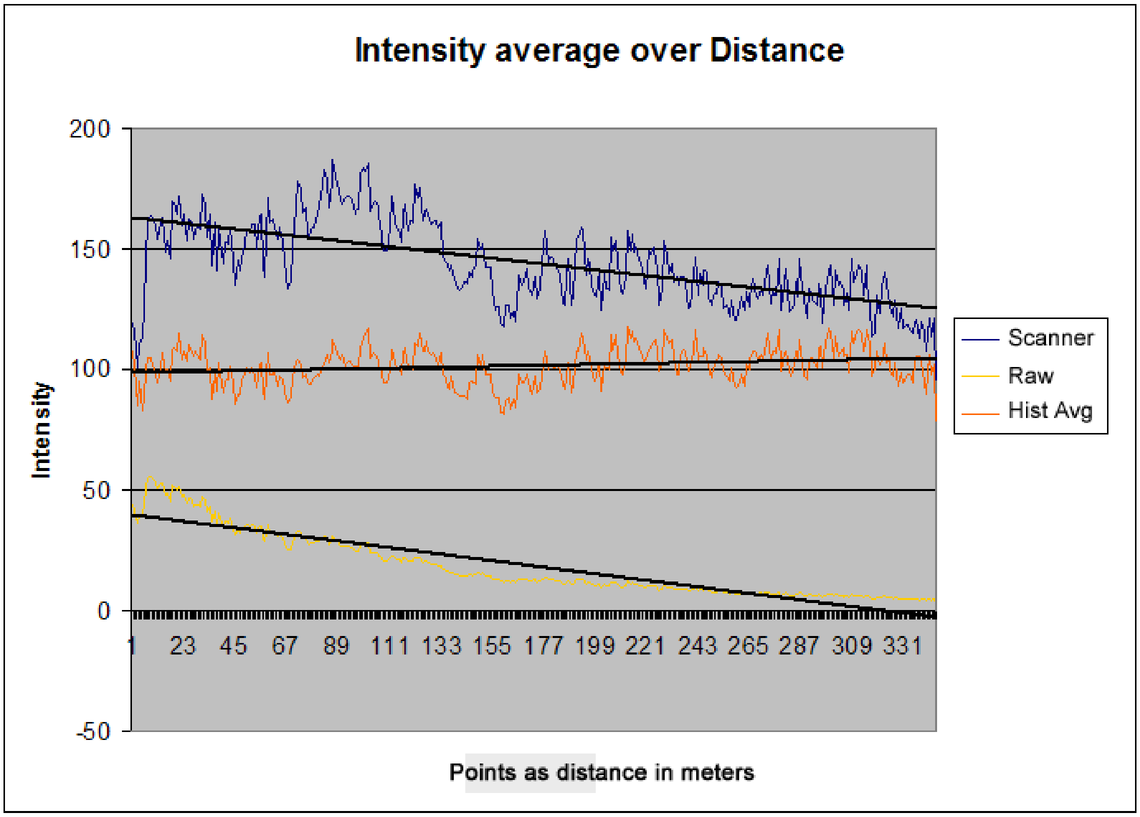


Figure 3. Note that the slope is flat across near and far distances with curve "Hist Avg".

2D HISTOGRAM

This statistical method was accomplished by writing a computer program that would read raw lidar data. Using raw data had several advantages: data was kept in its original form, data was compact, and data was gridded in an array of equal angles. Because the data is in an array, the program calculated distance and geometries quickly. We compared the intensity correction from the original software included with lidar scanner to our statistical corrections, and we found that our statistical method produced better results. Once we corrected the intensity values for distance, we identified other factors that needed intensity corrections. Two other factors that we corrected for were "roughness of surface" and the "angle of incidence" from laser source. A similar statistical method was used to correct the intensity values for "roughness of the surface". The gridded raw data array allowed the program to compare adjacent point geometries and defined how smooth or rough the surface was. This calculated roughness factor was used to create a 2D histogram of distance and roughness. A second 2D histogram using "angle of incidence" and distance was used to correct intensities. Both methods produced better results.

3D HISTOGRAM

The final statistical method used all three correction methods combined into a single 3D histogram. The 3D histogram corrected the intensity values for distance, roughness, and angle of incidence. This final statistical method produced the best results. We tested the results by analyzing high intensity horizons that ran along a curved outcrop surface. Curved outcrop surfaces decrease intensity values due to angle of incidence and roughness. By comparing the results of our methods to the original scan, we found that the corrections produced more consistent, and thus better, intensity values than the original scanner software. The graphs and images below identify the curved outcrop horizon that was corrected using this 3D histogram approach.

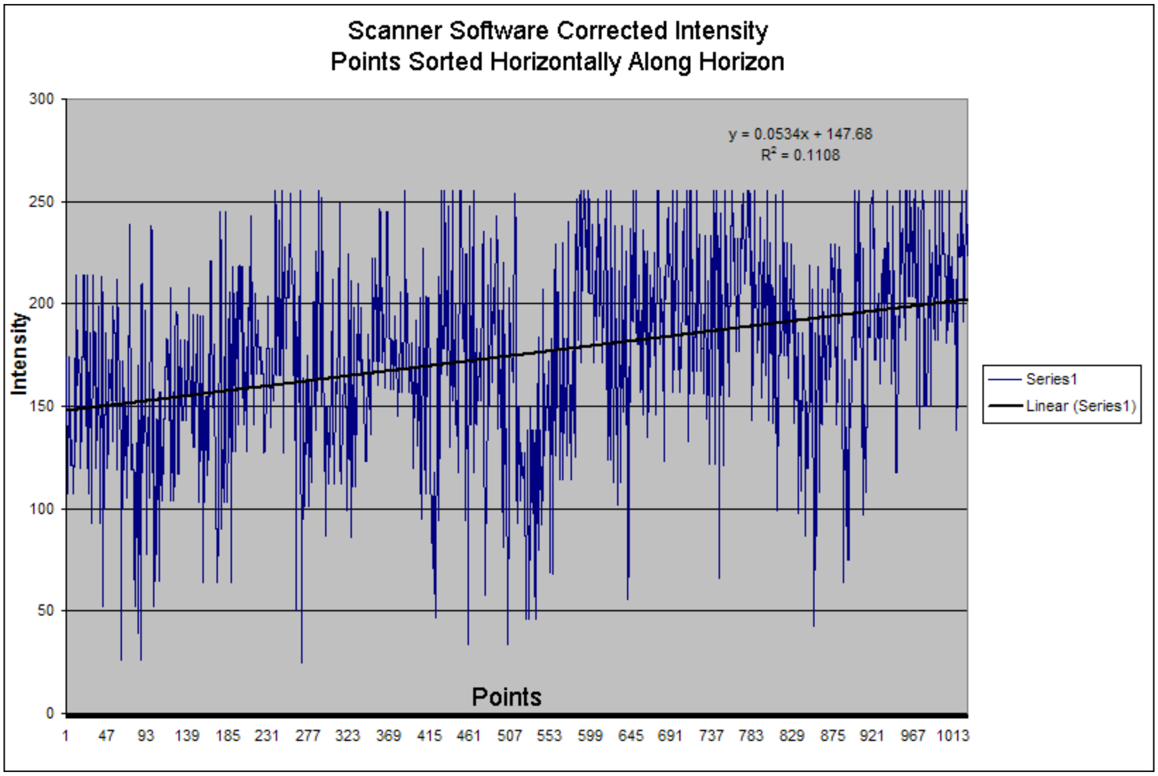


Figure 4. Original scanner correction of curved horizon.

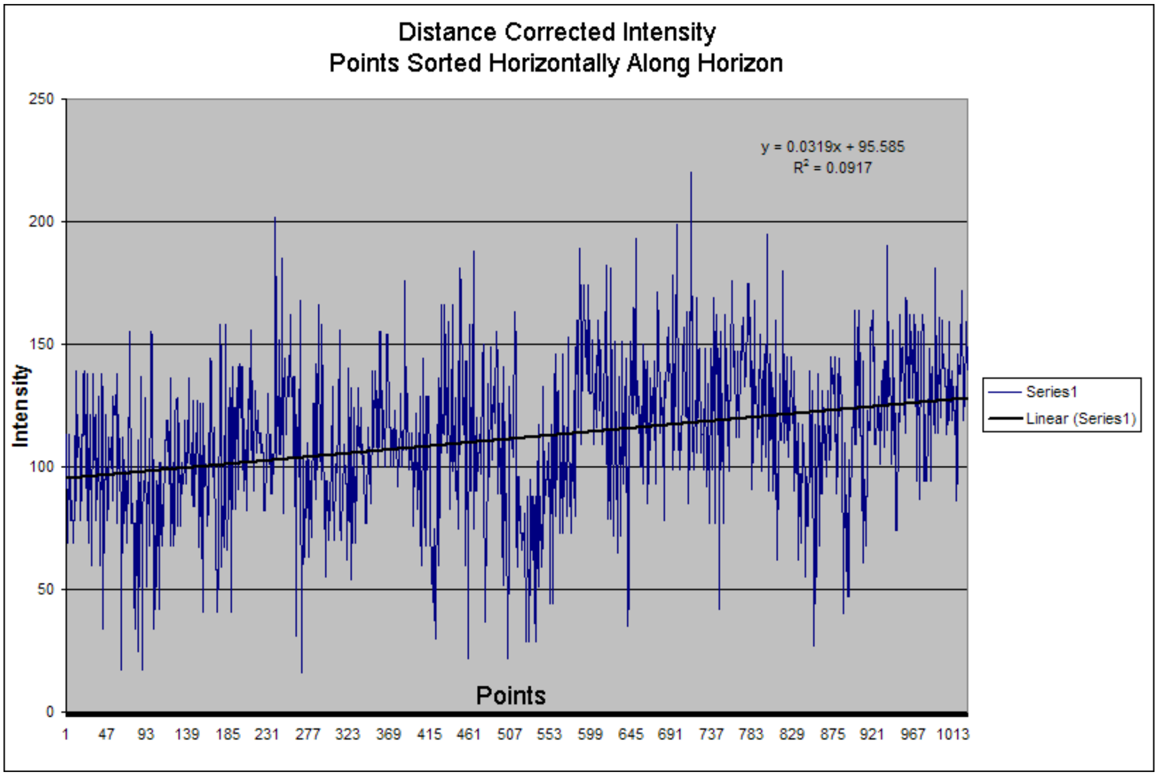


Figure 5. Curved horizon intensity correction using distance only.

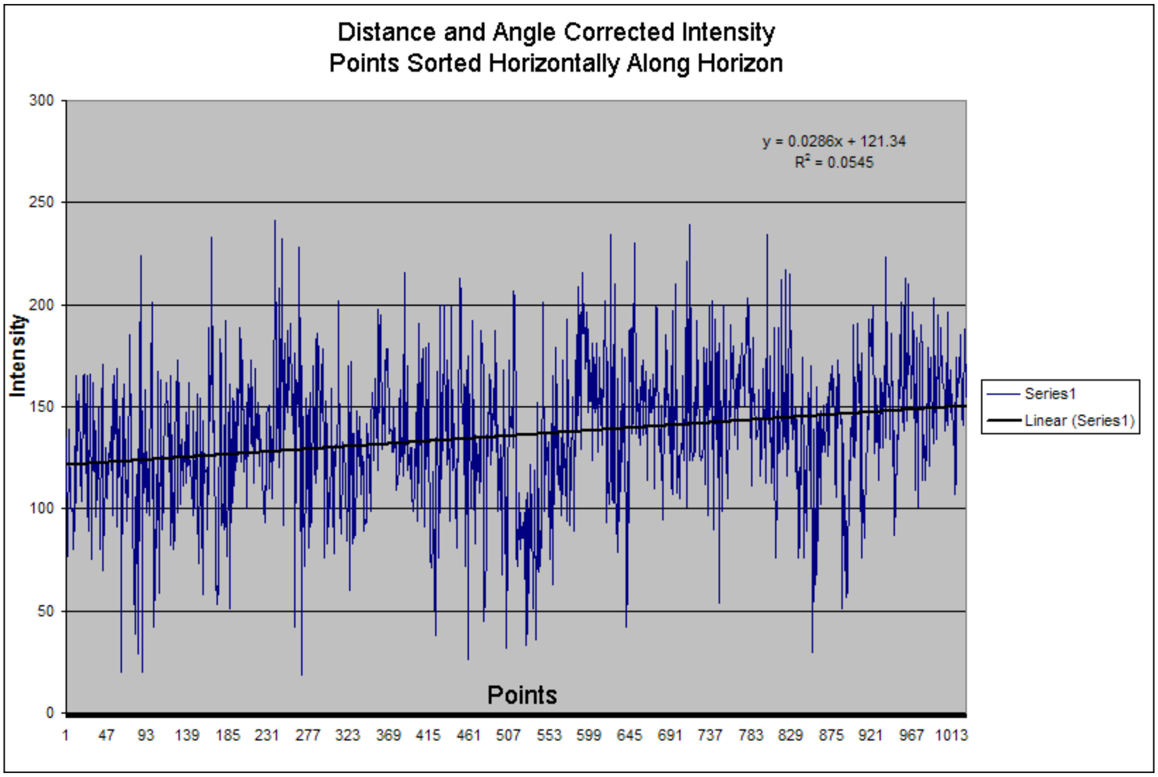


Figure 6. Curved horizon intensity correction using distance and angle of incidence.

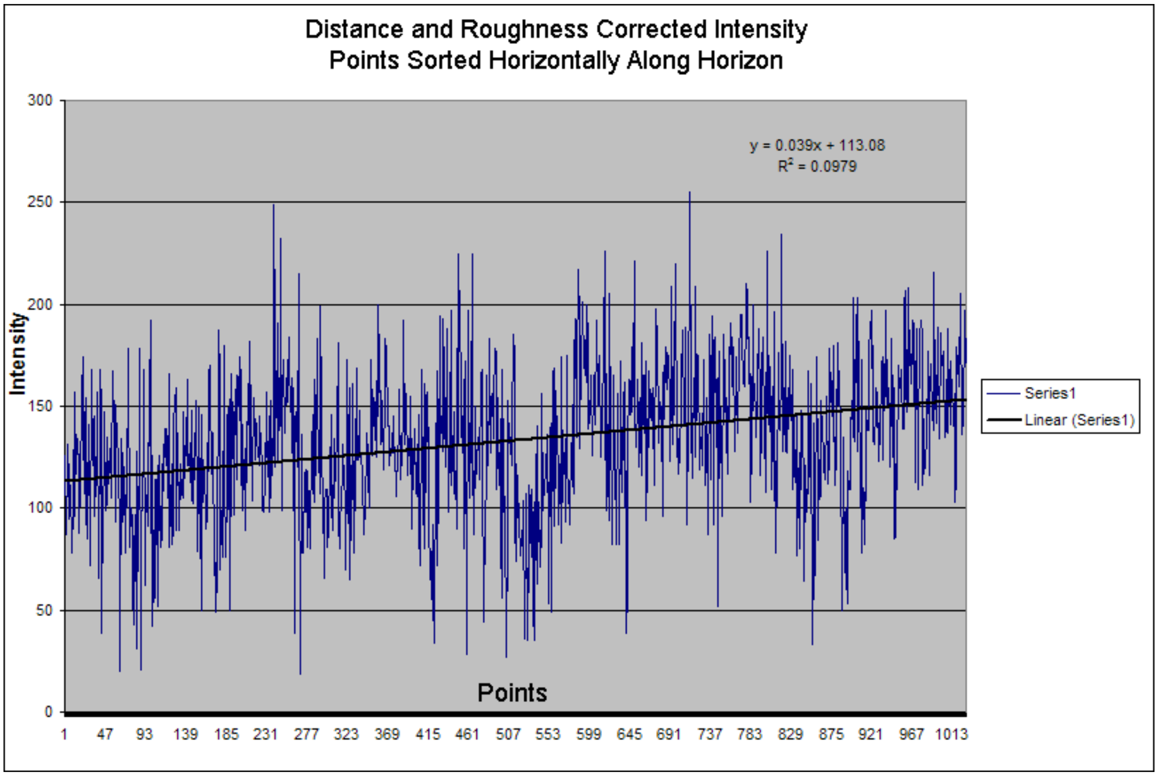


Figure 7. Curved horizon intensity correction using distance and roughness.

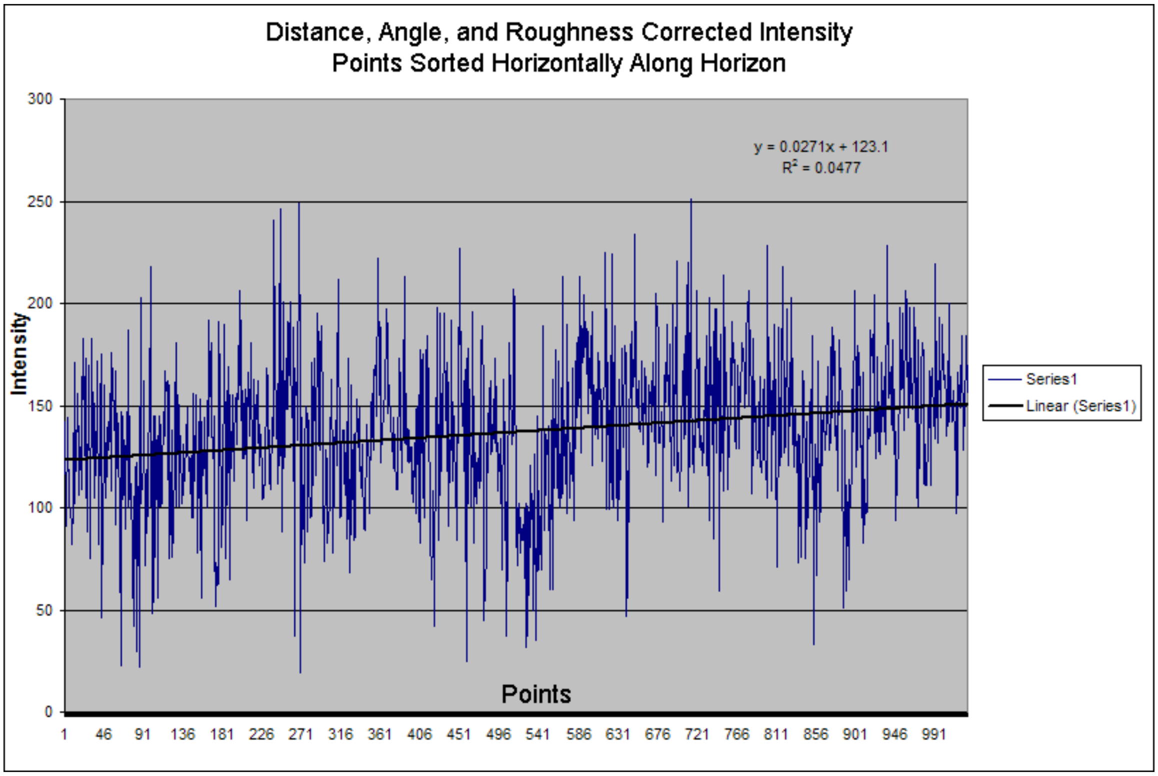


Figure 8. Intensity correction using distance, angle of incidence and roughness.

Type of correction:	Slope
Original scanner correction:	0.0534
Distance only:	0.0319
Distance and angle:	0.0286
Distance and roughness:	0.039
Distance, angle and roughness:	0.0271

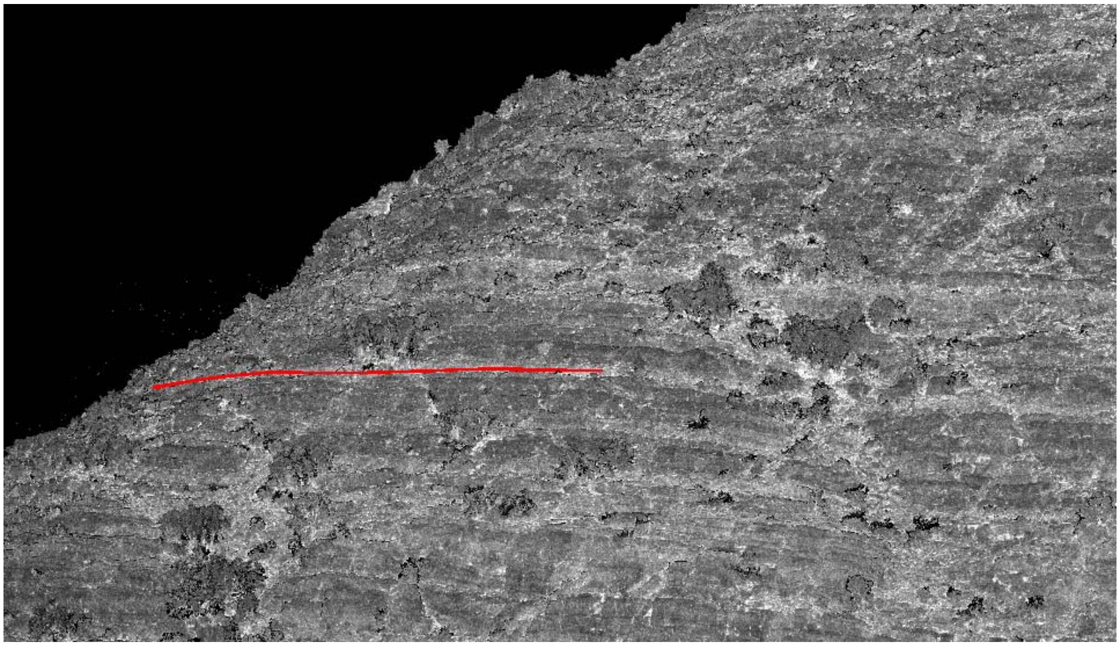


Figure 9. Area on outcrop used for plots from point intensity image.

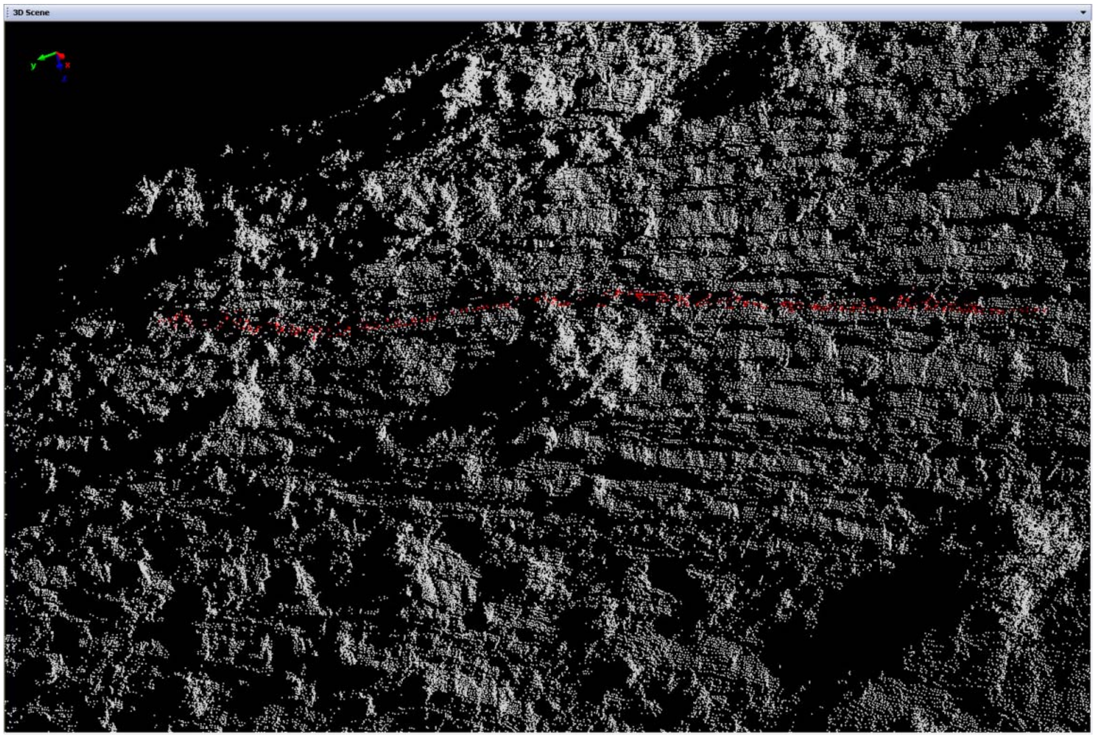


Figure 10. Close up of same area used for plots from point cloud viewer.

One other correction used with multiple lidar scans was to remove high concentration of points where scans overlap. One of the problems that occur with multiple lidar scans is that a bright band, represented by high concentration of points is produced in the area of scan overlap. This area is visually distracting and has an excess number of points. To correct for this, a computer program was written to blend overlapping areas of point cloud data sets. This program takes two overlapping point clouds scans from ground biased lidar and produces a gradient in the overlapping areas of the point clouds. The blend or gradient is accomplished by gradually reducing the number of points across the overlapping areas. The program identifies overlapping areas and then applies a function to reduce the number of points in this area. Before the program is used, the two overlapping data sets must be registered and aligned. This program can not register or align two overlapping data sets. The software used to register the point clouds prior to the blending is a commercial software called Polyworks. After two data sets are registered, each data set is saved as X, Y, Z, and intensity. The program reads both data sets and finds the overlapping areas. The overlapping areas are stored in separate linear arrays. This is needed to insure changing geometries in the overlapping areas are properly blended. Once the arrays are setup, the edges of the overlapping area are defined, and a blend takes place in the form of a non-linear gradient. In the overlapping area a probability function is used to randomly reduce the points from 100% to 0% across one data set and from 0% to 100% across the other. This probability curve is a half cosine function that determines the number and identity of the points removed. Below is an example of the cosine probability curve in the overlapping area.

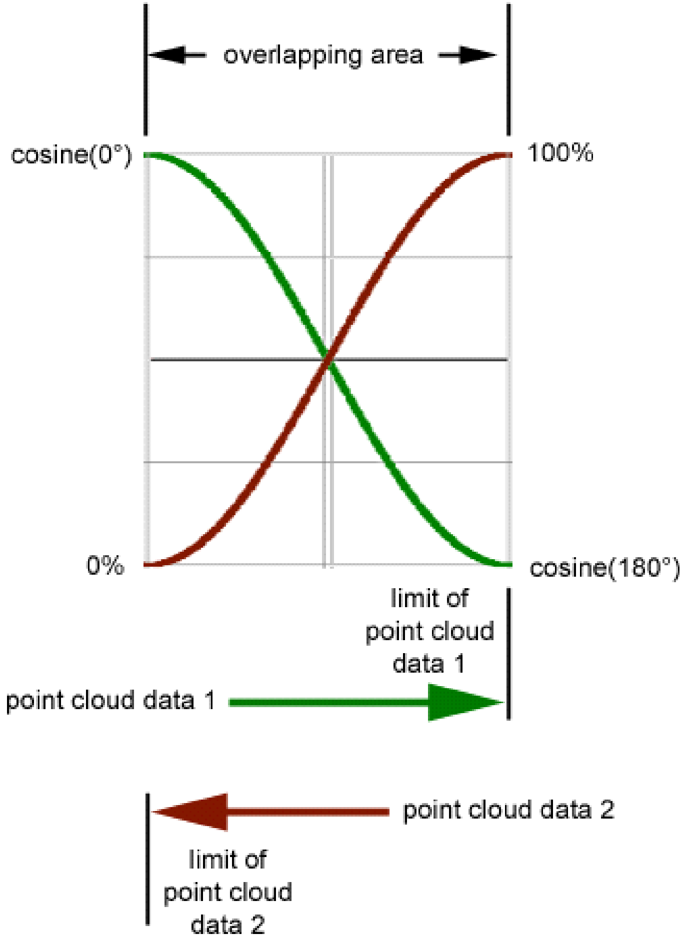
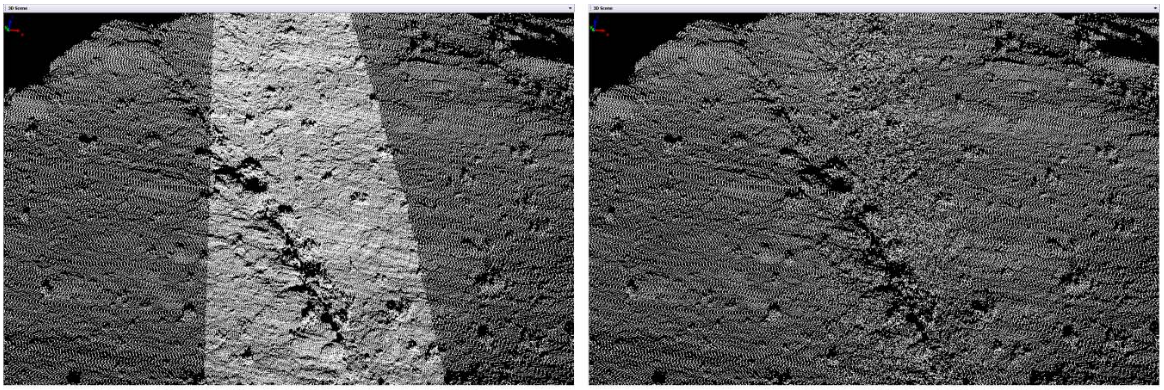


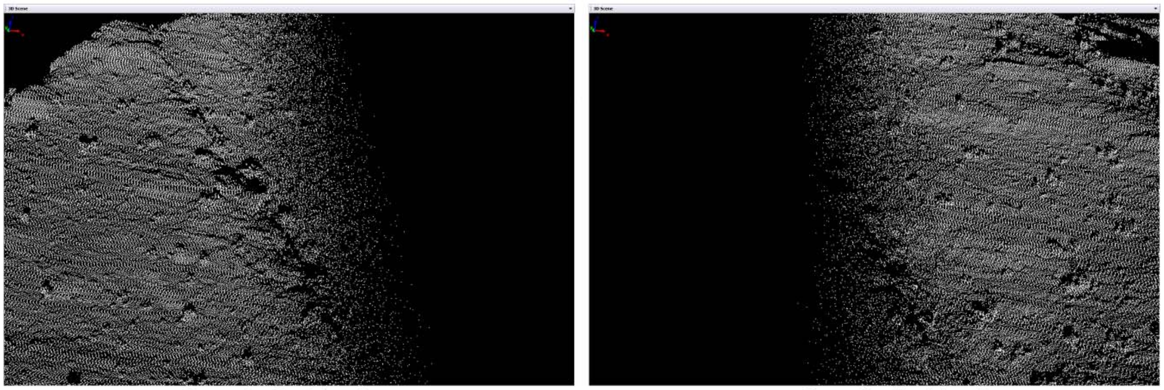
Figure 11. Cosine probability curve.

After the blend takes place the program can output merged point clouds as one data set with X, Y, Z, and intensity or it can output each point cloud data set independently. Figure 12 shows examples of the results before and after blending.



Overlapping lidar scans without blending

Both scans after blending



Scan 1 after blending

Scan 2 after blending

Figure 12. Before and after blending of point clouds.

The program used to blend overlapping point cloud data sets was written in C++. The blending program can not register or align two overlapping data sets. This program is only used to remove banding after two overlapping point clouds are aligned and registered.

REFERENCES

Bellian, J.A., C. Kerans, and D.C. Jennette, 2005. Digital outcrop models: applications of terrestrial scanning lidar technology in stratigraphic modeling, *Journal of Sedimentary Research*, 75, 166-176.

Hasegawa H., 2006. Evaluations of LIDAR reflectance amplitude sensitivity towards land cover conditions, *Bulletin of Geographical Survey Institute*, Vol 53, pp. 43-50, Mar. 2006.

Slob, S., H.R.G.K. Hack, and A.K. Turner, 2002. An approach to automate discontinuity measurements of rock faces using laser scanning techniques, *Proceedings of ISRM International Symposium EUROCK* 2002, Funchal, 2002 Nov. 25-28, pp. 87-94.