Lidar Basics for Mapping Applications

By James D. Giglierano

Iowa Geological Survey
Iowa Department of Natural Resources
109 Trowbridge Hall
Iowa City, IA 52242-1319
Telephone: (319) 335-1594
Fax: (319) 335-2754
e-mail: jgiglierano@igsb.uiowa.edu

INTRODUCTION

The purpose of this article is to help newcomers understand the basics of lidar data collection and processing, especially non-engineering, mapping specialists such as geologists, soils scientists, and those interested in land cover characterization. Many states in the U.S. are embarking on large-scale lidar acquisitions. This will make lidar elevation and other derived products widely available to many different audiences. To make full use of this new source of information, mappers must have or acquire some knowledge of the lidar data collection and handling procedures, and have the capability to convert the vendor supplied files into useful products. In some cases, mappers will do the processing themselves; in others, they will opt to have the processing performed by a vendor or third party. Another case may be that lidar derived topographic data supplied by a local government entity will have no metadata. In this case, the user will have to make some educated guesses as to the type of processing that may have been performed on the data.

HOW LIDAR DATA ARE COLLECTED AND REPRESENTED

The term **LiDAR** is an acronym for Light Detection and Ranging. Light Detection and Ranging basically consists of a laser rangefinder that operates in some form of airborne platform (helicopter, plane, or satellite). The rangefinder takes repeated measurements of the distance from the platform to the ground. The position and elevation of the platform is precisely known by way of airborne GPS along with ground control, so the elevation of the ground surface can be calculated by subtracting the laser rangefinder distance from the height of the platform. Compensation must be made for the tilt and pitch of the airborne platform by way of gyroscopes and accelerometers in the aircraft's inertial measurement unit. A good technical overview of lidar scanning technology is provided by Wehr and Lohr (1999).

Lidar systems record thousands of highly accurate distance measurements every second (newer systems operate at frequencies up to 150 kHz; older systems 30-80 kHz) and create a very dense coverage of elevations over a wide area in a short amount of time. Because lidar is an active sensor that supplies its own light source, it can be used at night and, thus, avoid routine air traffic, or it can be flown under some types of high cloud conditions. Most lidar systems record multiple surface reflections, or "returns," from a single laser pulse. When a laser pulse encounters vegetation, power lines, or buildings, multiple returns can be recorded. The first return will represent the elevation near the top of the object. The second and third returns may represent trunks and branches within a tree, or understory vegetation. Hopefully, the last return recorded by the sensor will be the remaining laser energy reflected off the ground surface, though at times, the tree will block all the energy from reaching the ground. These multiple returns can be used to determine the height of trees or power lines, or give indications of forest structure (crown height, understory density, etc.). Figure 1 shows a single 2 x 2 km tile that consists of 3.3 million first return lidar points.

Another feature of an airborne lidar system is the use of mirrors or other technology to point the laser beam to either side of the aircraft as it moves along its path. Depending on the scanning mechanism, the lidar scans can have a side-to-side, zigzag, sinusoidal, or wavy pattern. While the laser itself pulses many thousands of times per second, the scanning mechanism usually moves from side-to-side at around 20-40 cycles per second. This scanning, combined with the forward motion of the aircraft, produces millions of elevations in a short distance and time. The field of view or angle the scan makes from side-to-side can be adjusted by the operator, but is usually set at 30 to 40 degrees. This creates a swath of around 1 kilometer wide or less. Adjacent swaths overlap from 15 to 30% so that no data gaps are left between flight lines.

The spacing of lidar points on the ground, called "postings," is a function of the laser pulse frequency, scan

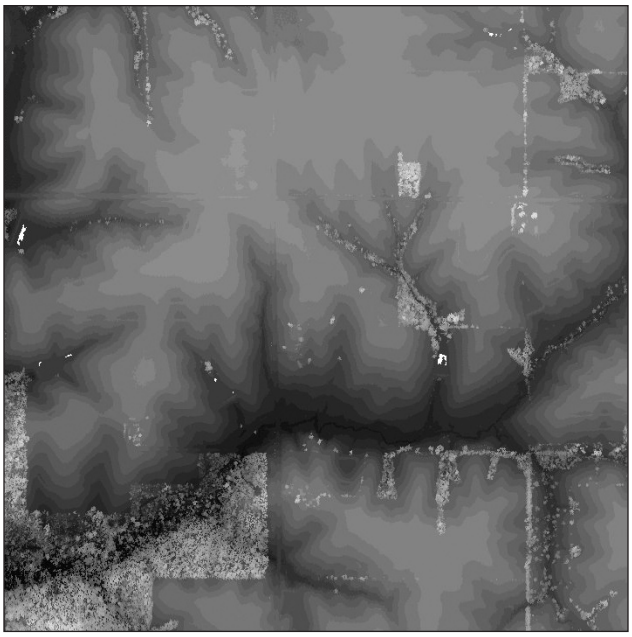

Figure 1. Gray scale image consisting of 3.3 million lidar first return points or “postings.” First returns indicate the tops of trees and buildings as well as bare ground in open areas. White areas are data voids where no returns were recorded, usually due to non-reflecting water surfaces.

frequency, and flight height (Baltsavias, 1999). While there is usually a nominal posting spacing specified in a lidar project, actual data points have variable spacing that are smaller and larger than the specified spacing. Mappers need to be aware of these effects when viewing final products that were derived from the raw data. The second aspect is that, because the laser scans from side to side, it interacts with the ground in different ways, depending on the angle of incidence. Lidar pulses at the edge of a scan will strike the sides of buildings, whereas pulses at the center of a scan will only strike the roof tops. Likewise, pulses at the edges of scans will pass through trees at an angle. Sometimes this will create “shadows” on the other side where no lidar passes through. In addition, less energy will return to the lidar receiver as it reflects away from the aircraft. This is evident in the intensity images created from the intensity values for each return: one can see overall darkening of the intensity at the edges of swaths. Edges of swaths appear darker than the returns at centers of swaths.

HOW LIDAR POINTS ARE PROCESSED INTO TINS AND DEMS

In the Spring of 2005, the Iowa Department of Natural Resources (DNR) and others, had lidar with a nominal resolution of 1 meter collected by a commercial vendor over the Lake Darling watershed located in Washington County, Iowa. The vendor’s lidar system collected a first and last return from each lidar pulse. From the first and last returns, a so-called “bare earth” return was created using a proprietary classification algorithm developed by the vendor. These classification systems try to sort out non-bare earth returns (tree tops, buildings, power lines, automobiles) from bare earth returns. To distinguish bare earth in forested areas, differences in elevation between the first and last returns, relative changes in elevation, and slope were used. Intensity data were used to identify vegetation and man-made materials. The lidar data for the Lake Darling watershed were collected in April with mostly leaf-off conditions. There are some data voids in forested areas due to non-penetration of the laser through tree canopy, but these areas are generally less than 10 meters across and are easily filled in by interpolation. Leaf-on conditions and tall crops, such as corn, do not allow easy penetration of the laser beam to the ground and should be avoided.

Lidar data for the Lake Darling area were supplied by the vendor in ASCII text format, consisting of 2 x 2 kilometer tiles with x and y coordinates, z elevations, and intensity values. With a nominal 1 meter posting spacing, some tiles had up to 3.3 million points. Postings near the center of the flight lines were close to the nominal 1 meter spacing (Figure 2), while toward the ends of scans, the points converge with the start of the next scan (Figure 3). While some scans converge, others diverge. Where the scans converge, the points can be less than half of the nominal spacing, and likewise, where they diverge, they can be twice the nominal spacing. Because some points can be as close as 0.5 meters, the tiles were initially interpolated to create grids with 0.5 meter resolution, with the idea that no data points should be merged or averaged with any other points. There is a tendency among some users to create grids with resolutions of 3, 5, and even 10 meters to save storage space or reduce the volume of data to process. We desired to create the grids as close as possible to the native resolution of the lidar data to evaluate their full potential to represent the smallest surface features.

To make digital elevation models (DEM) from the tiles, the Surfer 8 software (http://www.goldensoftware.com/products/surfer/surfer.shtml) was used. This software first creates a triangulated irregular network (TIN) before it interpolates the points into a raster DEM; however, once the DEM tiles were initially put together into mosaics, it became obvious that there were noticeable gaps between each tile. To remedy this, a C program was created to sort through the ASCII text files of the adjacent tiles and find points within a 3 meter buffer of the edge of the tile to be processed. Then the tiles were reprocessed adding the 3 meter buffers. When these raster tiles were merged together into a mosaic, the gaps were almost completely eliminated. Leica Imagine (http://gi.leica-geosystems.

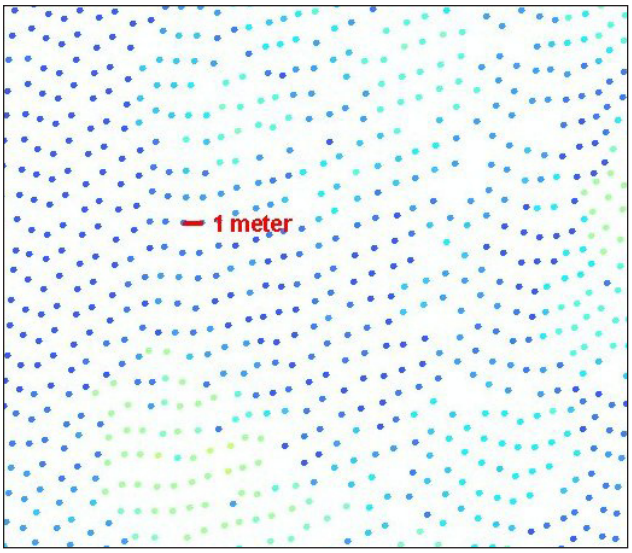

Figure 2. Data near the middle of a lidar flight line. Posting spacing is around one meter at the center of back and forth scans.

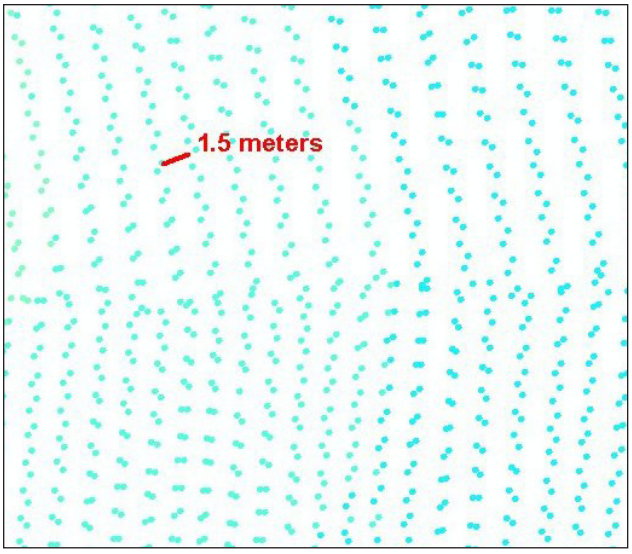

Figure 3. Data along overlap of two adjacent lidar flight lines (top and bottom of figure). The posting spacing is highly variable at edges of flight lines. Some postings are less than one meter apart at the end of one and beginning of the next scan, while the distances between points in different sets of scans are as much as 1.5 meters apart.

com/) was used to mosaic all the tiles into one large raster DEM file. From the DEM, shaded relief images were created and compressed. DEMs and shaded relief images were then easily imported into ArcGIS software (http://www.esri.com/) for display and further analysis.

Field examination of the lidar bare earth shaded relief images was conducted in January, 2006. It was surprising how well lidar shaded the relief images represented the smallest topographic features, including small slope changes of less than half-a-meter, even in forested areas. There were some data voids due to lack of penetration through the dense tree canopy, but there were enough data points to show good definition of incised stream channels, meander scars, and gullies (Figure 4). Man-made features such as road ditches and embankments, terraces, and dams were also well defined. Tillage patterns are evident as regular lineated textures on crop fields parallel to the road grid. These are not scanning artifacts as the individual scans are at a slight angle to the east-west flight lines.

Because the bare earth processing does not remove 100% of the forest artifacts, a distinctive bumpy pattern is left in the bare model, which indicates the presence of forest cover (Figure 5). During field examination, it was noticed that different canopy structures were represented by different patterns in the artifacts. In the tall canopy floodplain forest, most of the bumps were removed, which left a mostly smooth surface, while on side slopes with thick understory or brush cover, the texture on the shaded relief image is rougher in appearance. Interestingly, the bare earth processing removed nearly all of the numerous tree falls in the stream channels, which allows drainage tracing programs to work well when following flow paths downstream. Also, areas with pine trees were very smooth, which indicated nearly complete penetration by the laser beam.

HOW TO USE LIDAR PRODUCTS FOR MAPPING APPLICATIONS

Once the raw lidar point tiles are processed into high-resolution DEMs, other useful mapping products can be derived. The derived shaded relief image previously mentioned (Figure 4) is very useful for visual display and interpretation, and can be combined with colorized elevation images for extra information content. Another useful display product is the slope map, which can be derived from the DEM using the grid processing tools found in almost every GIS package. Usually, a choice can be made whether to calculate the slope rate in degrees or as percent (45 degree slope = 100%). A slope map based on percent can be grouped into slope classes typically used by soil survey mappers (slope class A = 0-2%, B = 2-4%, etc.) and readily compared to soil polygons displayed by slope class (Figure 6). Figure 7 shows the new level of detail available in slope classes derived from lidar data.

In addition to the elevation component of the lidar return, many systems produce an intensity component that indicates the strength of the lidar return. This intensity value is mostly influenced by the reflectance of the material struck by the laser pulse, but is also influenced by the scan angle. (Laser pulses directed away from the airplane at significant angles do not reflect back as much light energy as a pulse directed straight down from the plane). Because most lidar systems use a laser that emits light in

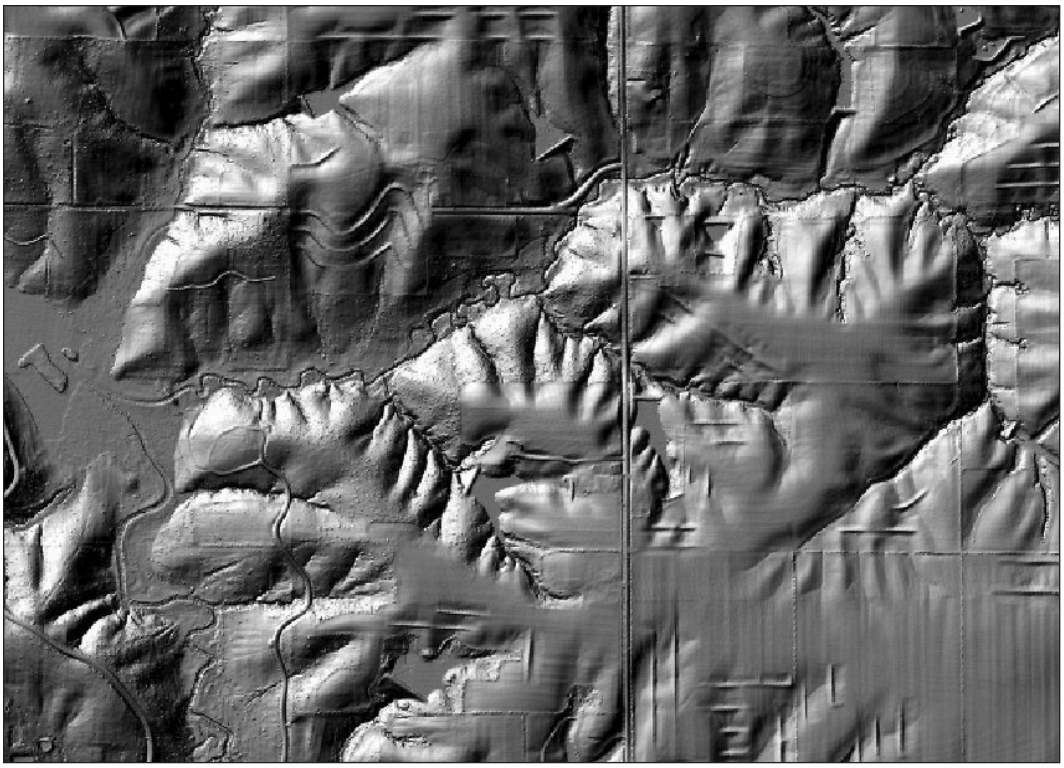

Figure 4. Portion of the bare earth shaded relief image of the Lake Darling watershed, showing natural and man-made features readily apparent in the lidar data.

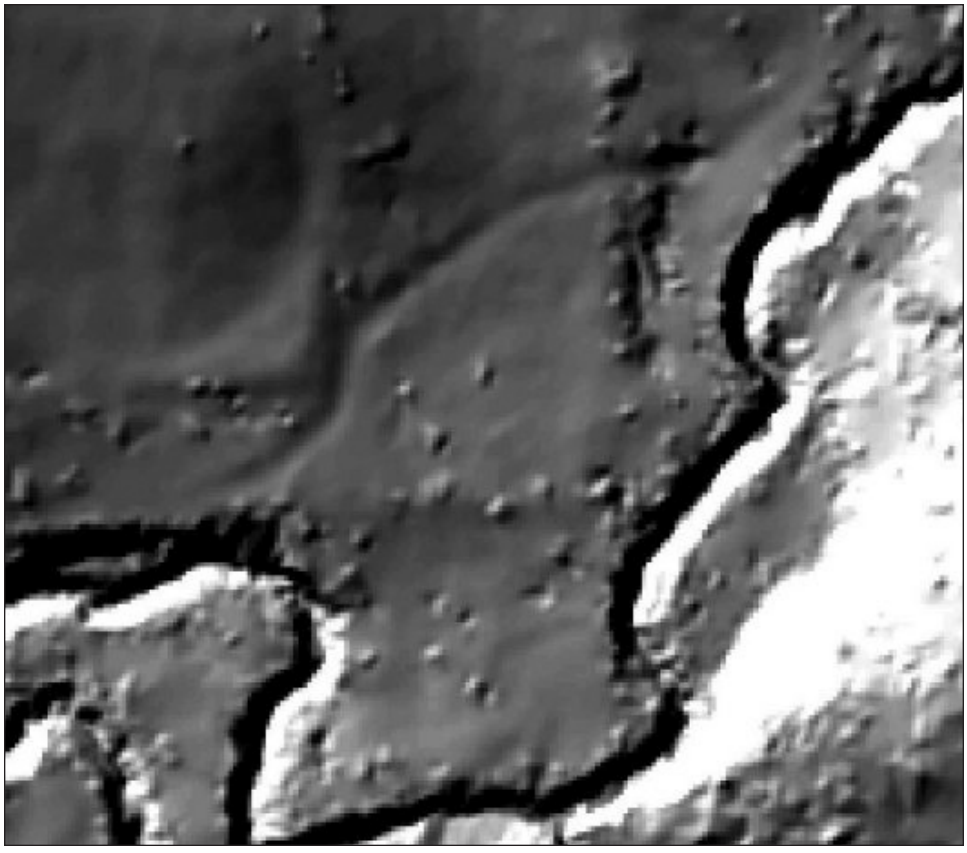

Figure 5. Portion of a bare earth shaded relief image showing artifacts (bumpy texture) in deciduous forest areas. These artifacts are lidar elevations classified as bare earth, but probably are from tree trunks, branches, or understory close to the ground and classified as bare earth by the vendor’s algorithm.

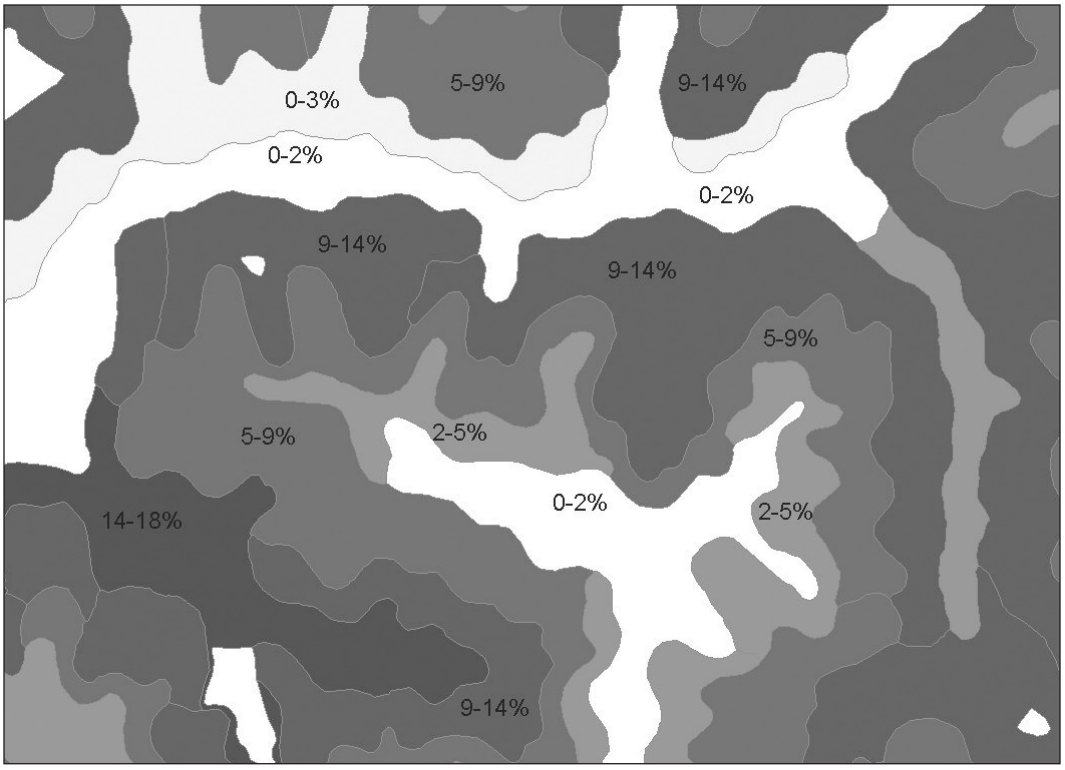

Figure 6. Soil survey soil polygons shaded by slope class range: light shades are lower slopes and darker shades indicate steeper slopes.

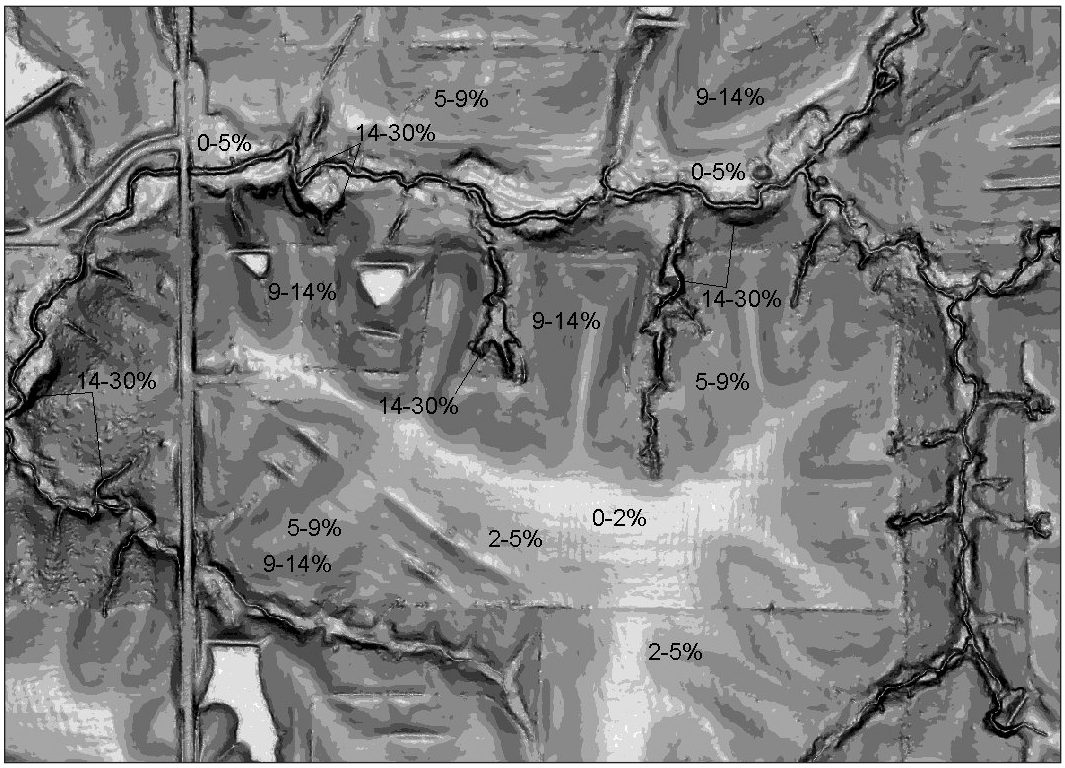

Figure 7. Slope class ranges derived from Lake Darling lidar data. While low slope areas on lidar look similar to the soil polygons, lidar shows more detail on steep slopes such as gullies and stream channels.

the near-infrared portion of the spectrum (lidar used for Lake Darling had a wavelength of 1064 nm), the intensity of lidar return is directly related to the near-infrared reflectance of the target material. An image constructed from the intensity component of the returns (Figure 8) looks very much like a black and white near-infrared aerial photograph. An intensity image has one interesting peculiarity: tree shadows point away from the flight lines, so one can see shadows pointing in opposite directions close together at the edge of two flight lines. Because intensity is recorded from each lidar return, it is possible to construct first return intensity images as well as last return intensity images, and have them look quite different. This may especially occur in forested areas where the first return might mainly represent the treetops, while the last return intensity represents many features, including the forest floor.

VERTICAL ACCURACY TEST AND INFLUENCE OF LAND COVER

Usually, one of the first questions new users of lidar have is about the vertical accuracy of the elevation data. In the Lake Darling project, the stated accuracy was 15 cm (.5’) RMSE (root mean square error) in bare earth areas and 37 cm (1’) in vegetated areas. Because there are no high accuracy geodetic monuments in the watershed and access to survey grade GPS equipment was unavailable, another way to test the vertical accuracy needed to be found. Fortunately, a digital terrain model and associated 2’ contours produced by aerial photography and photogrammetric techniques for a road project were available from the Washington County engineer’s office. This digital terrain model and contours were created by a local aerial photography firm and had a stated vertical accuracy of 6.1 cm (.2’). The area covered by the model was over 2 miles long and a quarter of a mile wide. The digital terrain model consisted of elevation points and break lines (Figure 9) in CAD format. Using the 3D_ANALYST extension in ArcGIS, the photogrammetrically derived terrain model was converted into a triangulated irregular network or TIN, and interpolated into a 1 meter elevation grid. The lidar elevation grid was subtracted from the grid made by photogrammetry to produce a simple difference grid. The overall average difference between the two grids was only 3.3 cm (.11’). To compare the two grids to their stated accuracies, the RMSE had to be calculated. First, the simple difference grid was multiplied by itself to create the squared difference grid. Using a polygon coverage of land cover from 2005, the mean squared difference was calculated for each land cover class using the zonal statistics command in ArcToolBox. By using the spatial calculator function in the SPATIAL_ANALYST extension, the square root of the values in the “mean” field

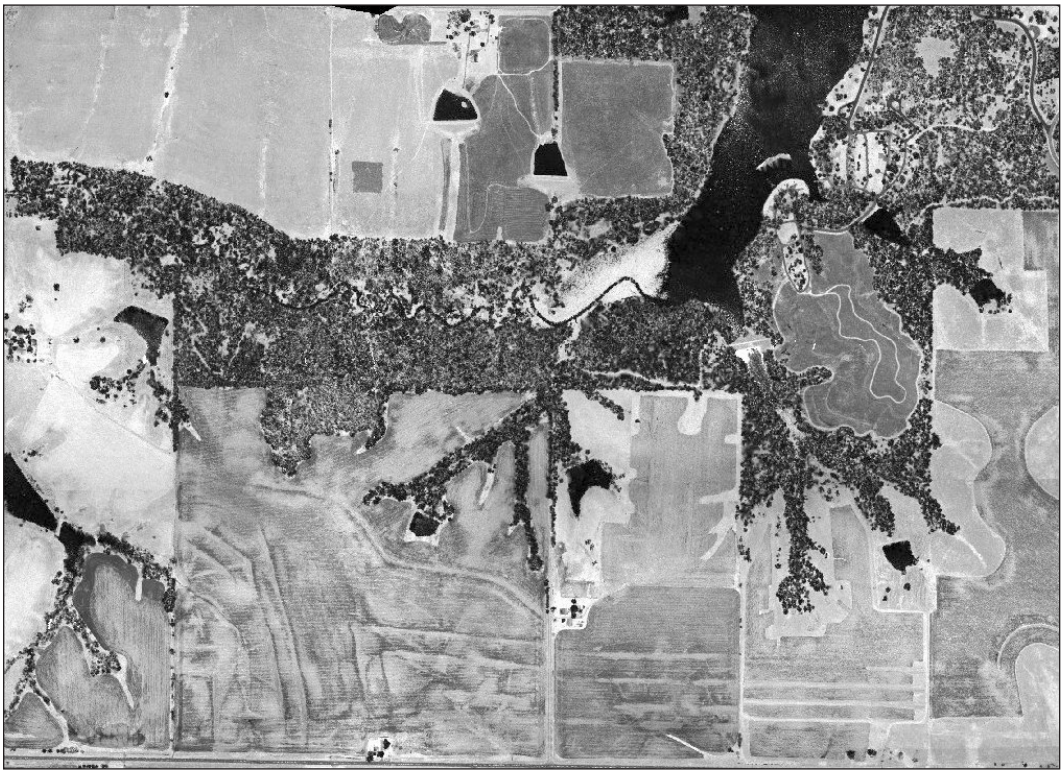

Figure 8. Portion of a lidar intensity image of the Lake Darling watershed, constructed from bare earth return intensity values.

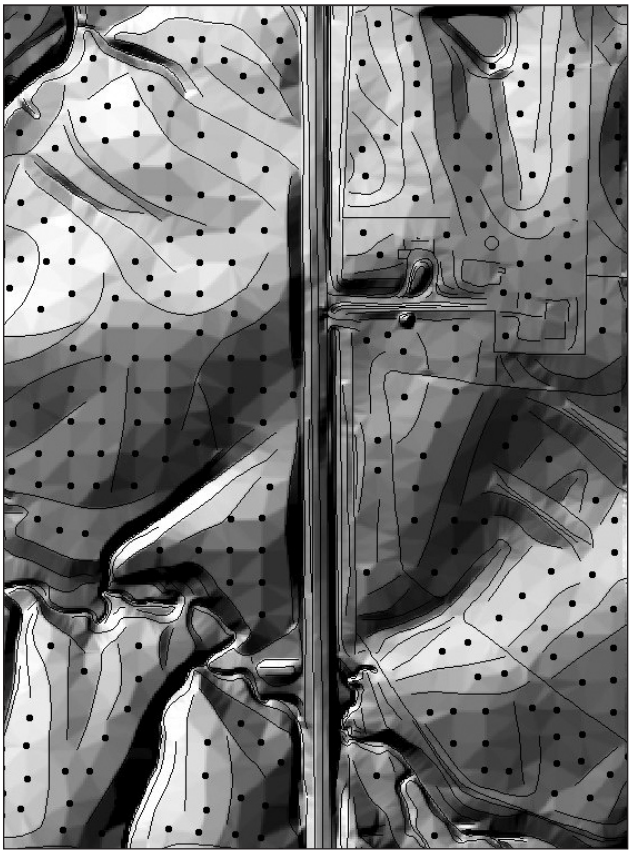

Figure 9. Portion of shaded relief digital terrain model derived from low altitude aerial photos. The black dots are elevation mass points and the black lines are break lines.

of the table was calculated, and the RMSE was found for each land cover class. The zonal statistic tool also computes a "count" of cells for each class and a "sum" of the elevations within that class. By calculating the sum of all the "count" field values and "sum" field values for all the classes, and dividing the total sum by the total count, the average squared difference for the entire dataset was found. By taking the square root of this value, the RMSE was found for the whole area. Initially, RMSE between the lidar DEM and the photogrammetry DEM was found to be .79 feet or 24.1 cm.

Upon examination of the squared difference image, it was apparent that the terrain in several areas had changed significantly between the time of the airphoto flight in 2000 and the lidar flight in 2005. These mainly included areas where the installation of sediment retention structures and dams, and road grading had occurred. When these areas were digitized and excluded from the squared difference calculation, the overall RMSE was found to be .57 feet or 17.4 cm (Figure 10). The RMSE of the row crop area was .46' (14.3 cm), grass areas .62' (18.9 cm), and forested areas .85' (25.8 cm). If the DEM derived by photogrammetric means is accepted as the higher accuracy source, then the lidar meets its stated accuracy of 15 cm in the bare ground areas, and under 37 cm in the vegetated areas. This appears to be a good test of lidar accuracy because it includes many types of land cover conditions, not just a few high accuracy locations at benchmarks on roads or nearby ditches.

COMPARING OLD AND NEW DATA

One of the first tests of any new lidar data set is to compare it with the existing DEM derived from the 10' contours from the USGS topographic quadrangle mapping projects of the latter half of the last century. Displayed at smaller scales, it is difficult to see much difference between the shaded relief images derived from the 30 meter resolution National Elevation Dataset or NED (http://ned.usgs.gov/) and lidar shaded relief. Only when the display is zoomed into larger scales is it possible to see the marked differences between the 30 meter NED (Figure 11) and lidar DEM (Figure 12). Visible on the lidar image (but not on the 30 meter NED shaded relief image), are man-made features such as roadways, ditches, fence lines, terraces and dams. Natural features such as stream channels, gullies, and floodplains are also visible.

Lidar excels at mapping topographically challenged areas: areas with little relief. Figure 13 is a shaded relief image, derived from the 30 meter NED, that shows typical glaciated terrain in north-central Iowa, east of Spirit Lake in Dickinson County. Figure 14 shows the same area using 1 meter resolution lidar, which focuses the indistinct mounds seen on the NED shaded relief into sharply defined, circular, and elongated features. These are interpreted to be remnants of ice walled lakes, which formed on the surface of the glacier. These lakes had varying amounts of sediment deposited in them, and after the ice melted, these sediments formed indistinct, low mounds (Quade et al., 2004).

Figure 15 shows the Missouri River floodplain north of Council Bluffs, Iowa, in a view, which again, uses the 30 meter resolution NED to create a shaded relief image. It reveals numerous defects in the original conversion of widely spaced contours on a very flat surface. With a 10' contour interval, there is not enough information to interpolate features on the floodplain adequately. The shaded relief image reveals cross-shaped artifacts within the DEM, which were created by the interpolation software's attempt to connect widely spaced data. Figure 16 shows the great improvement afforded by interpolating a surface from closely spaced lidar points (about 2 meter lidar postings). Missouri River meander scars, levees along drainage ditches, fence lines, interstate lanes, railroad right-of-ways, borrow pits, and sewage lagoons are all visible on the lidar shaded relief image.

When using shaded relief images for on-screen digitizing, geological mappers will need to become accustomed to recognizing and separating man-made as well

OID	LAND_USE	ZONE_C	COUNT	AREA	MIN	MAX	RANGE	MEAN	STD	SUM	RMSE_ft	RMSE_CM
0	Residential	1	36665	36665	0.00000	18.04	18.04	0.262028	0.823366	9607.27	0.511887	15.6023
1	Water	2	26069	26069	0.00000	10.053	10.053	0.234739	0.848544	6119.4	0.484499	14.7675
2	Pasture	3	152359	152359	0	41.5079	41.5079	0.383737	1.24089	58465.8	0.619465	18.8813
3	CRP	4	160257	160257	0	38.4666	38.4666	0.252299	0.855599	40432.7	0.502294	15.3099
4	Timber	5	121869	121869	0	106.069	106.069	0.71515	2.79821	87154.6	0.845665	25.7759
5	Wildlife/Wooded	6	1825	1825	0.00001	7.3607	7.36069	0.965743	1.03896	1762.48	0.982722	29.9534
6	Road	7	44609	44609	0	107.915	107.915	0.316614	0.942917	14123.9	0.562685	17.1506
7	Row Crop	8	353980	353980	0	328.666	328.666	0.214536	0.903116	75941.3	0.463180	14.1177
8	Alfalfa	9	9979	9979	0	8.07532	8.07532	0.179604	0.476496	1792.27	0.423797	12.9173
	TOTALS		907612							295399		

Mean Squ. Difference = 295399/907612 = .3255

Square root of MSD = .5705

RMSE = .57' or 17.4 cm

Figure 10. Root mean square error (RMSE) calculation of photogrammetrically derived DEM and lidar DEM, after 2000/2005 landscape-change areas removed from calculation.

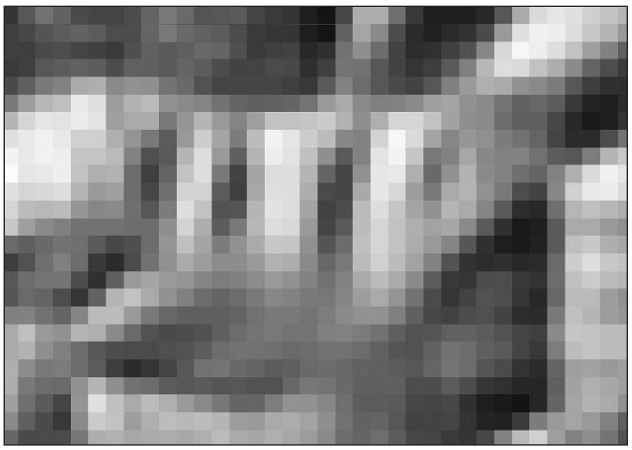

Figure 11. Portion of a shaded relief image made from a National Elevation Dataset (NED) 30 meter resolution DEM. Area is from Lake Darling watershed in Washington County, Iowa.

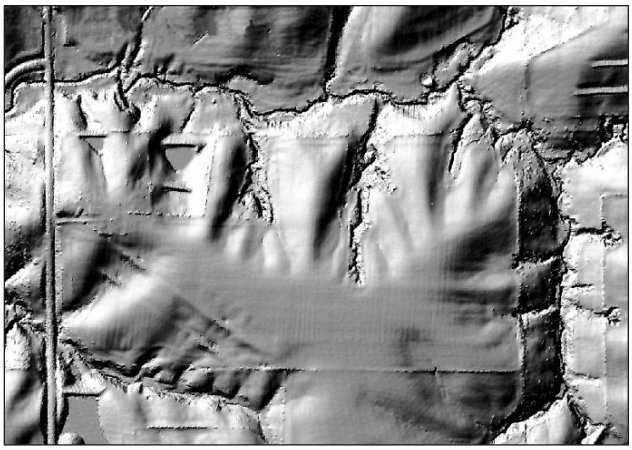

Figure 12. Portion of shaded relief image made from a 1 meter resolution lidar DEM for the same area in Washington County, Iowa.

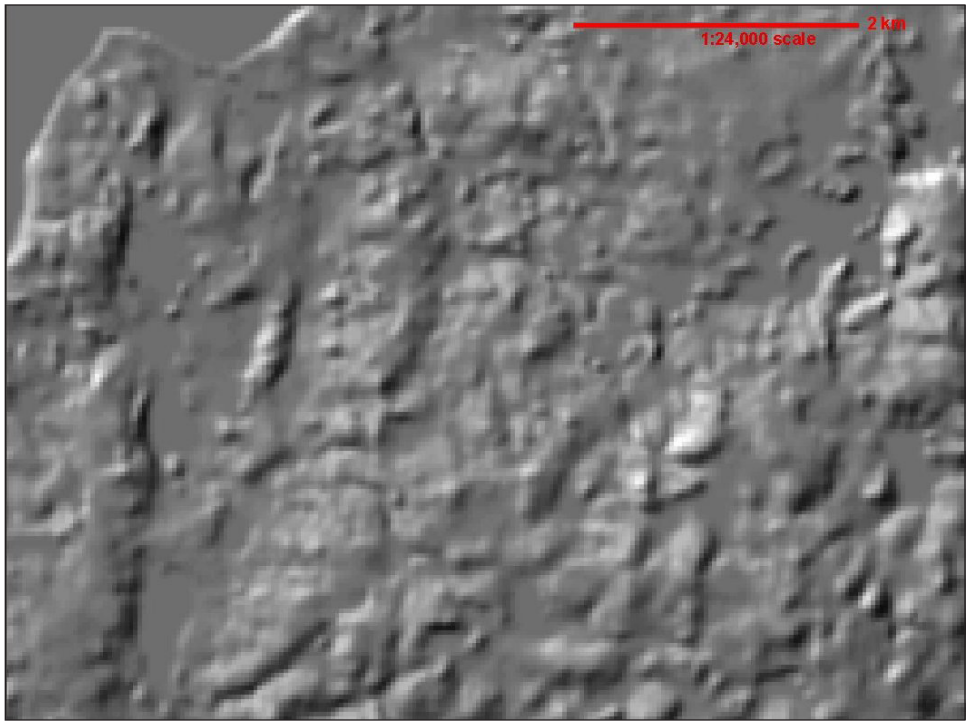

Figure 13. Portion of a shaded relief image showing recently glaciated terrain near Spirit Lake in Dickinson County, Iowa. The shaded relief was created from a 30 meter resolution DEM from the National Elevation Dataset (NED).

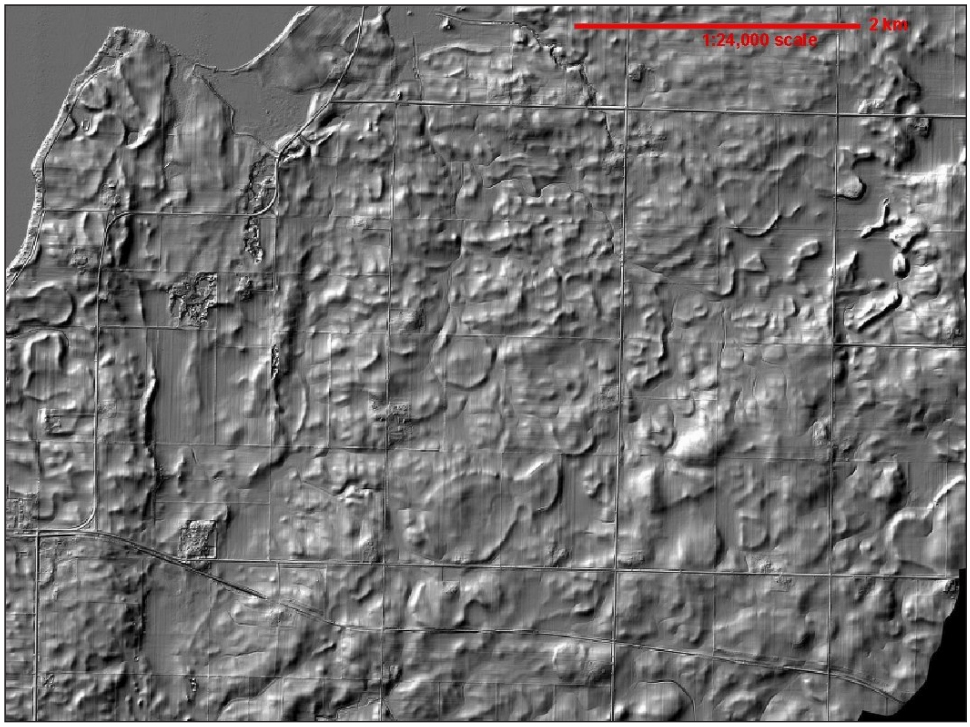

Figure 14. Portion of a lidar-derived shaded relief image of the same area of glacial terrain near Spirit Lake in Dickinson County, Iowa. Notice how the shapes of subtle, low relief glacial features are now readily apparent.

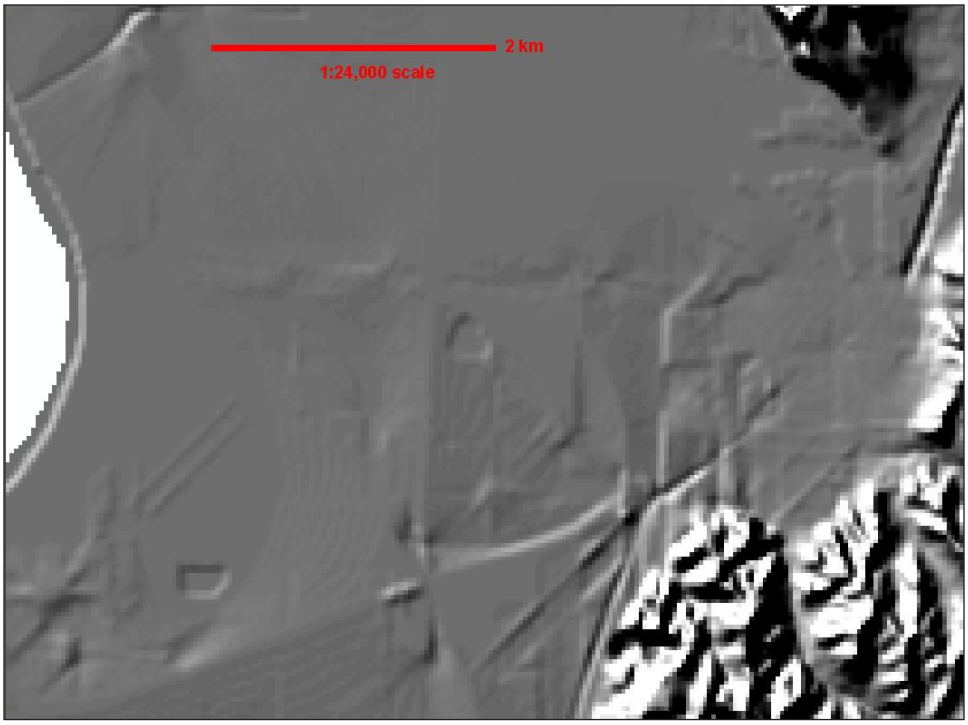

Figure 15. Portion of a shaded relief image showing the Missouri River floodplain north of Council Bluffs, Iowa. The shaded relief image was created from a 30 meter resolution DEM from the National Elevation Dataset (NED). Notice the cross-shaped features that are artifacts of the interpolation of the original 10’ contours from USGS topographic maps.

Figure 16. Portion of a lidar-derived shaded relief image of the same area on the Missouri River floodplain (see Figure 15). Notice the much finer detail showing the interstate cloverleaf, river meander scars, borrow pits, and a ditch and levee system. Lidar DEM was obtained from the Pottawattamie County GIS Department.

as geomorphic features. Because shaded relief images can represent the encoding of relatively small changes in slopes, mappers will need to build up criteria for recognition of everyday features using the clues in contrast, shading, shape, texture, pattern, and context contained in these images. In the past, geological mappers learned how to interpret aerial photos by poring over example after example of natural and man-made features. They also learned how to interpret geological features by looking at geomorphic signatures on topographic maps. Lidar will require relearning and reinventing both techniques by moving the geomorphic scale down to the realm of the airphoto, roughly at resolutions from 1 to 5 meters. While qualitative information on slopes was available by way of stereo viewers and aerial photos, there has never been as much quantitative slope information available until now with the advent of lidar data. With digital elevation data derived from lidar, new computer assisted classification strategies for geomorphic feature interpretation can be developed, as can new types of imagery to support manual interpretations.

SUMMARY

Large-scale lidar acquisitions will provide mapping professionals with an increase of new, high quality elevation data to use as base maps for their projects. To take full advantage of this new data source, those who are mapping need to be aware of how lidar data are collected and what data reduction processes commercial vendors use to make deliverable products for their clients. In many cases, mappers will want to manipulate the raw lidar returns into their own TINs, DEMs, and derived products, but sometimes they will only have access to vendor-supplied, finished products that have undergone unknown procedures to make the visual appearance more appealing. Mappers can use shaded relief images derived from lidar DEMs or TINs for on-screen digitizing, as well as new derivative products such as terrain slope and lidar intensity to identify geologic features and other features. Anyone using lidar data will be interested in the absolute vertical accuracy of elevations and will need to know how land cover type affects that accuracy.

REFERENCES

Baltsavias, E.P., 1999, Airborne laser scanning: basic relations and formulas: ISPRS Journal of Photogrammetry and Remote Sensing, no. 54, p. 199-214.

Quade, D.J., Giglierano, J.D. and Bettis, III, E.A., 2004, Surficial geologic materials of the Des Moines Lobe of Iowa, Phase 6: Dickinson and Emmett Counties: IDNR/IGS, OFM-0402, 1:100,000 scale, accessed at http://www.igsb.uiowa.edu/gsbpubs/pdf/ofm-2004-2.pdf.

Wehr, A. and Lohr, U., 1999, Airborne laser scanning – an introduction and overview: ISPRS Journal of Photogrammetry and Remote Sensing, no. 54, p. 68-82.