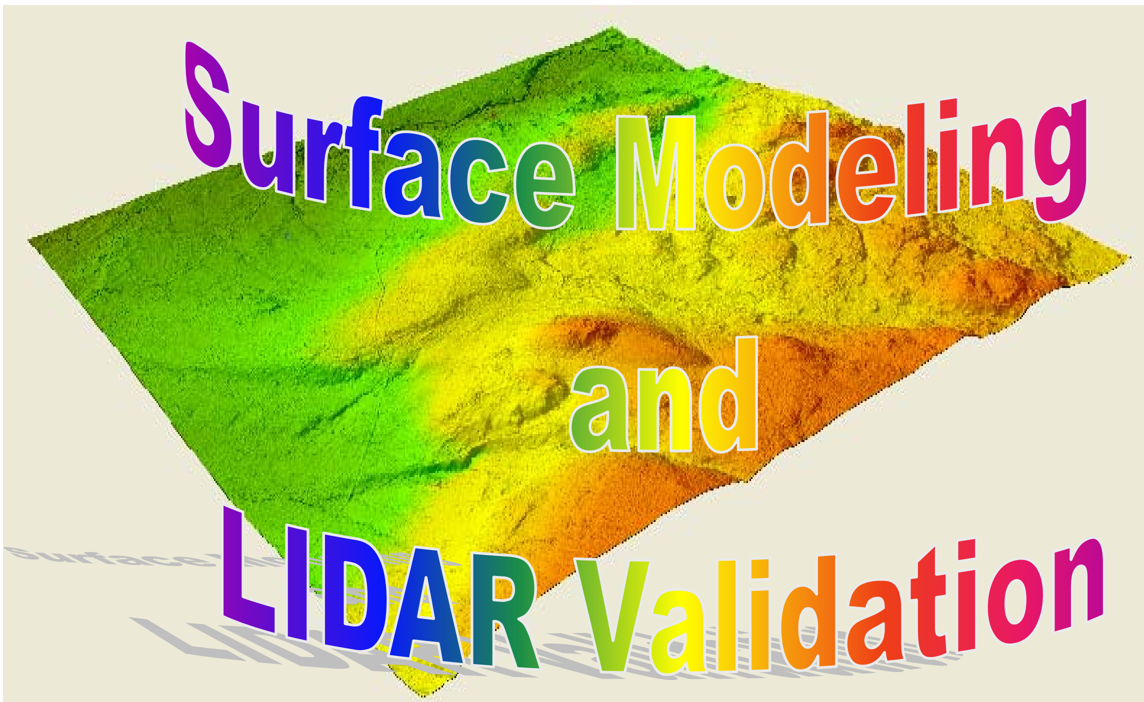


AGRG
Applied Geomatics Research Group

Ted MacKinnon
ADP Assignment 2
November 22, 2003

Contents

Introduction

LIDAR (Laser Imaging and Distance Ranging) is one of the newest Geomatics techniques to derive very precise topographic elevation data. The LIDAR sensor emits a series of rapid laser pulses of near infrared light in a swath towards the earth's surface and measures the time it takes for the reflected pulses to return the sensor. This measurement is accurate because the sensor knows the actual position of the aircraft from the DGPS (Differential Global Positioning System) and the orientation from the IMU (Inertial Measurement Unit).

The IMU measures the roll, pitch, and yaw and when combined with the GPS can correct any drift or movement from the aircraft. The system produces a series of point measurements that consists of geographic location and height of both natural and man-made features, and can be processed to produce several different interpolated products and be integrated into a Geographic Information System (GIS).

LIDAR has many different applications such as floodplain mapping, natural resource management, coastal zone mapping, and more. LIDAR data can be separated into ground only hits and non-ground hits. This allows for the removal of all non-ground features such as forest canopy and buildings. LIDAR data is ideal for representing an accurate model or a visual description of the earth's surface by interpolating an accurate Digital Elevation Model (DEM) from the ground only data. The derived DEM allows us to see features under forest canopy that normally would be missed since traditional DEMs are typically derived from photogrammetry, and do not have the vertical accuracy or resolution.

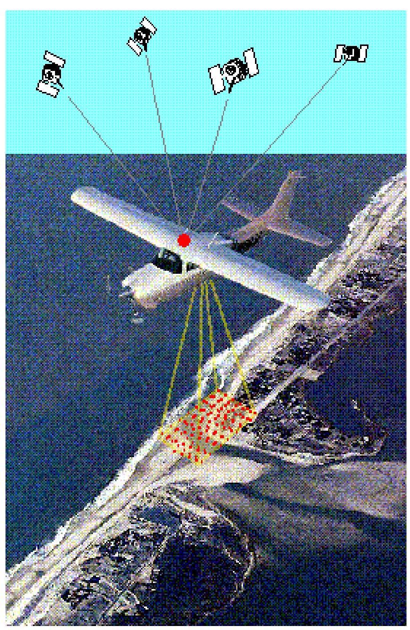


Figure 1 - The aircraft uses DGPS and an IMU to determine the location of where the aircraft is relative to the time it took the laser pulse to reflect back from the surface. The resultant data is a series of points representing the geographic location and height of both natural and man-made features.

Terra Remote Sensing collected the LIDAR point data for this project during the spring of 2003. They provided it in an ASCII file format classified into ground and non-ground files with a 2m spacing resolution. The files contained space delimited columns of GPS Week, GPS Time, Flight Line Number, Easting, Northing, Ortho Height, and Ellipsoid Height data. The Eastings and Northings were measured in UTM zone 20 WGS84. Each file represented a 4 km by 4 km tile, labeled by the coordinates of the lower left corner of the tile.

The study area or LIDAR tile used for this project was 333_4971, located in the community of Nictaux West, south of Middleton. Two forest plots surveys from the spring validation campaign were located within the tile. Both plots were hardwood dominated and had leaves starting to form, so it may be possible that the LIDAR points did not penetrate the Forest canopy, reaching the ground beneath the trees.

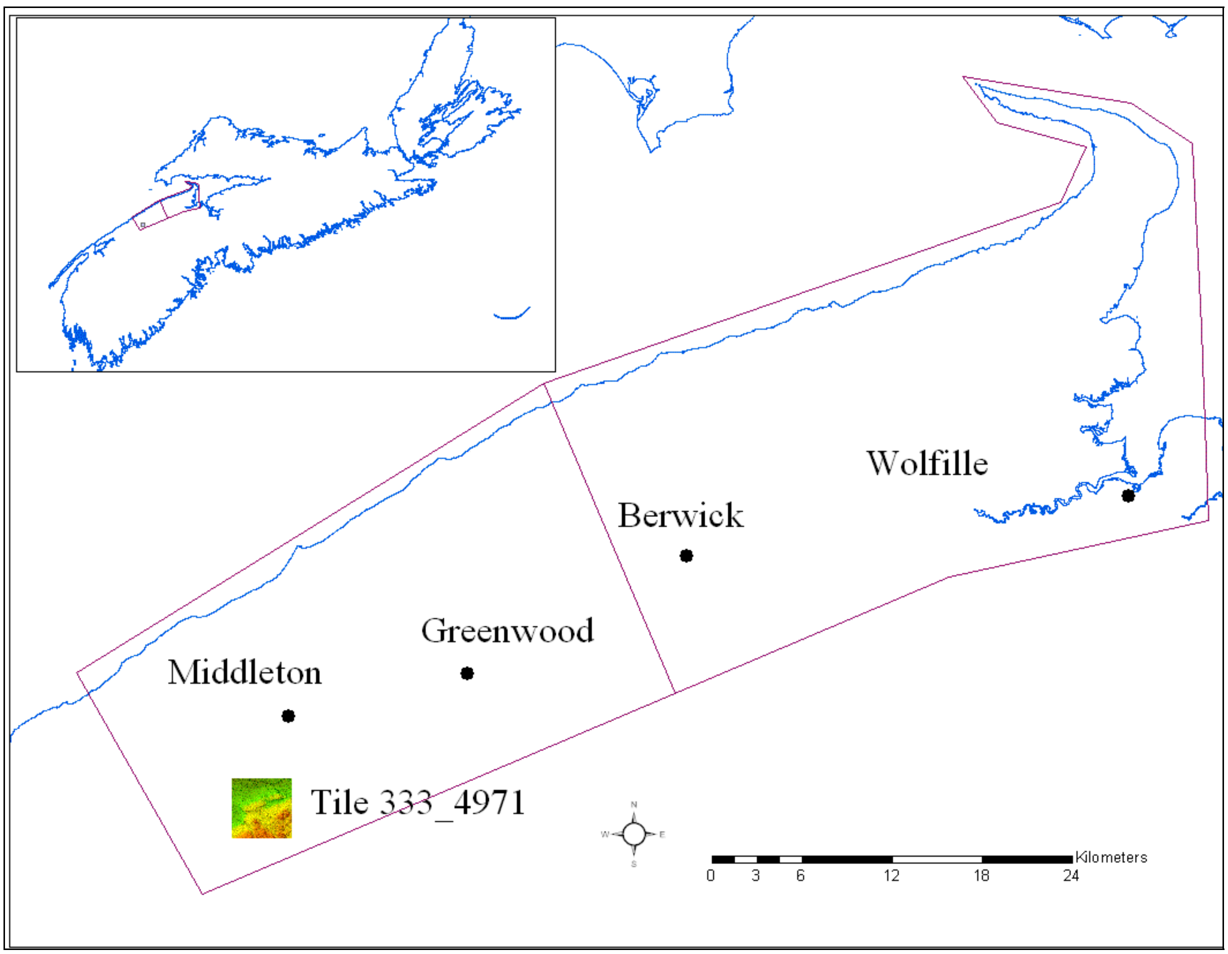


Figure 2 - The study area or LIDAR tile used for this project was a 4 km by 4km tile, located in the community of Nictaux West and in the lower left corner of the overall study area.

This project deals with the various techniques of generating a model of the earth's surface. Various ESRI algorithms were evaluated to determine which method is the best to construct a ground surface representation from LIDAR point data. TIN (triangulated irregular network), TOPOGRID, IDW (Inverse distance weighted), Kriging, Spline and some non-interpolation grid routines such as POINTGRID were examined.

A TIN is a vector based topological data model that is used to create a continuous surface representing terrain data. The resultant surface is produced from irregularly spaced elevation points. It represents the terrain surface as a set of triangle facets made up of triangles, nodes and edges. When viewed in three dimensions, each triangle forms a "facet" of the surface. Together, the triangle facets form a continuous surface model. For each of the vertices, the geographic location and the elevation values are encoded. Here, the orthometric heights for the points were applied.

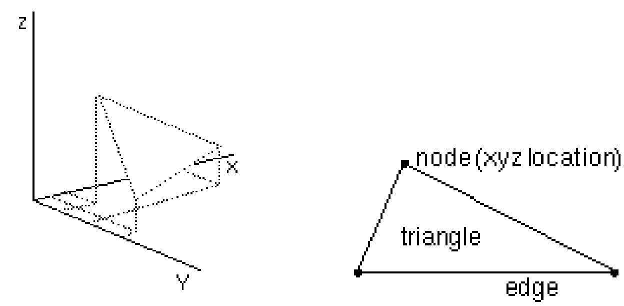


Figure 3 – A TIN is made up of triangular facets made up of triangles, nides, and edges defined by geographic location and elevation.

An advantage of a TIN is that the extra information is encoded for areas of complex relief without requiring large amounts of data to be collected from areas of simple relief such as break lines. Since the size of each facet is variable, smaller triangles and therefore a more detailed representation can be provided where there is a higher density of data points.

TOPOGRID is an interpolation method designed for the creation of hydrologically correct DEMs from comparatively small, but well selected elevation and stream data. Water is an erosive force that determines the general shape of most landscapes, thus most have many hills and few sinks, resulting in a connected drainage pattern. TOPOGRID applies this knowledge about surfaces and imposes constraints on the interpolation process that result in a connected drainage structure and a better representation of ridges and streams. This imposed drainage condition produces higher accuracy surfaces with less input data.

The program acts conservatively in removing sinks and will not impose the drainage conditions in locations that would contradict the input elevation data. Such locations normally appear in the diagnostic file as sinks. This information is used to correct data errors, particularly when processing large data sets[1].

The purpose of the drainage enforcement process is to remove all sink points in the output DEM that have not been identified as sinks in the input sink coverage. The

[1] ESRI ArcInfo Workstation Help File

program assumes that all unidentified sinks are errors, since sinks are generally rare in natural landscapes[2].

Kriging is a complex procedure based on the regionalized variable theory that generates an estimated surface from a scattered set of points with z values and involves an interactive investigation of the spatial behavior of the phenomenon represented by the values. The procedure requires knowledge about spatial statistics that can be conveyed in the command reference. ESRI recommends that before using the Kriging method, one should have a thorough understanding of the fundamentals of Kriging and have assessed the appropriateness of the data for modeling with this technique.

The Kriging method assumes that the spatial variation represented by the z values is statistically homogeneous throughout the surface and this spatial homogeneity is fundamental to the regionalized variable theory. Point sets that have pits, spikes, or abrupt changes are not appropriate for the Kriging technique.

The IDW interpolation method calculates cell values using a linearly weighted combination from a set of sample points. The weight is a function of inverse distance. The user can control the significance of known points upon the interpolated values, based upon their distance from the output point. Specifying the higher value will put more emphasis onto the nearest points, thus, nearby data will have the most influence, and the surface will have more detail. Specifying the lower value will provide a bit more influence to these of the surrounding points that are a little farther away.

SPLINE is a two-dimensional minimum curvature interpolation method that results in a smooth surface that goes through the input points. This basic minimum curvature technique is often referred to as thin plate spline method.

The POINTGRID command converts point data into a GRID cell format. Each cell of the grid is assigned an elevation according to the point or points that it overlays. If a cell has more than one possible point the code with the most occurrences in the cell is used. If no points occur within a cell, a NODATA value will be assigned.

---

[2] Goodchild, M. F. and Mark, D. M. 1987. The fractal nature of geographic phenomena. Annals of Association of American Geographers. 77, no 2: 265-278.

Processing the Raw LIDAR Data

Convert point data to an ArcInfo coverage.

The LIDAR files were imported from DVDROM to the AGRG Unix workstation where the majority of the processing took place. ESRI ArcInfo Workstation was used to process the data and build the DEM.

Coverages can be stored in either single or double coordinate precision. Single precision coverages can store up to seven significant digits for each coordinate while double-precision coverage can store up to fifteen significant digits. The coordinate systems that we use for our map projections use values that are larger then seven digits, so it is important that we use double precision coordinates to maintain accuracy.

The *PRECISION* command defines the coordinate precision of coverages. If the precision is set before the point coverages are generated then the coordinate precision of the output coverages will always be double-precision.

Usage: PRECISION <SINGLE | DOUBLE> {HIGHEST | LOWEST | SINGLE | DOUBLE}

Arc: **PRECISION DOUBLE**

The *GENERATE* command generated point coverages from the ASCII LIDAR data files. However, there are some limitations to this command and another procedure was implemented. The generate command can only handle importing the ID, x, and y columns, but there is more data columns then that. Thus the rest of the attribute data were brought into the coverage using a join item**.** An AWK script was used to copy the x and y data columns from the ASCII file and create a temporary file (comma delimited data that contained only ID, x and y data) to use with the *GENERATE* command. The ID is important for joining the other columns in the ASCII data win the coverages.

Usage: CREATEWORKSPACE <workspace>
Arc: **CREATEWORKSPACE groundonlyhits**

Orion% **awk'{print NR "," $4 "," $5}' 333_4971.gnd 333_4971.xy**

Usage: GENERATE <cover>

Arc: **GENERATE groundhits**
Generate: **Input 333_4971.xy**
Generate: **POINTS**
Generate: **QUIT**

The *BUILD* command created a Point Attribute Table **(**.pat file) for the point coverage and was used to join the other columns of the ASCII data.

Usage: BUILD <cover> {POLY | LINE | POINT | NODE | ANNO.<subclass>}

Arc: **Build groundhits POINT**

TABLES command changed the Arc prompt to the Tables environment where the remaining columns of data were appended. *T ABLES* allows for creation, query, simple analysis and display of the INFO database for the point coverage. The first step was to define the .dat file and then to populate the items with the space delimited records from the ASCII file. The .pat file then needed to be joined with the rest of the columns. A join item was added to both the files and populated with $RECNO. The INFO file and the COVER were generated from the same ASCII files thus should still have the same Ids and # of records.

Usage: TABLES {info_directory} {user_name}

Arc: **TABLES**

Usage: DEFINE <info_file>

Tables: **DEFINE groundhits.dat**
Tables: **gps_week, 5, 5, I, 0**
Tables: **gps_time, 13, 13, n, 4**
Tables: **flt_line, 5, 5, I, 0**
Tables: **easting, 13, 13, n, 3**
Tables: **northing, 13, 13, n, 3**
Tables: **ortho_ht, 10, 10, n, 3**
Tables: **ellip_ht, 10, 10, n, 3**

Usage: SELECT {info_file} {RO}

Tables: **SELECT groundhits.dat**

Tables: **ADD FROM 333_4971.gnd**

Usage: ADDITEM <info_file> <item_name> <item_width> <output_width><item_type> {decimal_places} {start_item}

Tables: **ADDITEM groundhits.dat rec 10 10 I**

Usage: CALCULATE <target_item> = <arithmetic_expression>

Tables: **CALCULATE rec = $RECNO**
Tables: **SELECT groundhits.pat**
Tables: **ADDITEM groundhits.pat rec 10 10 I**
Tables: **CALCULATE rec = $RECNO**
Tables: **Quit**

Usage: JOINITEM <in_info_file> <join_info_file> <out_info_file> <relate_item>
{start_item} {LINEAR | ORDERED | LINK}

Arc: **JOINITEM groundhits.pat groundhits.dat groundhits.pat rec**

The projection of the coverage was defined with the *DEFINE* command.

Usage: PROJECTDEFINE <COVER : GRID : FILE : TIN > <target>

Arc: **PROJECTDEFINE COVER groundhits**
Arc: **PROJECTION UTM**
Arc: **ZONE 20 T**
Arc: **DATUM WGS84**
Arc: **PARAMETERS**

Now a spatially referenced point coverage of the ground LIDAR hits was generated. The above steps were repeated for the non-ground hits file.

DEM Construction

The *CREATETIN* command generated the TIN surface. Since the LIDAR data, contained only one elevation feature type, mass points, they were the main component of the tin. The points had an x, y, and a z value and all mass points have equal significance when building a tin.

CREATETIN is an interactive command with its own prompts. The name of the TIN to be created was specified. Weed tolerance was ignored because it is used to reduce the number of vertices on any linear features, and our data did not contain any line features. The proximal tolerance was set to the default, which was decided by the precision. No vertical exaggeration was applied to the TIN, so the z factor was ignored. The *COVER* command was used to indicate which coverages were used to create the TIN surface. A spot_item or spot_value (ortho_ht) was used to indicate which field the z values from the input coverages are located. The end statement was entered after all coverages were specified to begin the TIN construction process.

After the TIN was created it could be examined with the *DESCRIBETIN* command. This command displays information about the TIN and displays it on the screen for a visual examination. The min, max values, # of triangles and the boundary coordinates are all provided. To draw the TIN 9999, 9998 or 9997 was specified (9999 - draws triangle edges, 9998 - analytic hillshade, 9997 - hypsometric shading). Flat is used to highlight any flat triangles. Too many flat triangles indicate a poor triangulation of the surface. Pressing the following number keys will allow you to change your field of view.

1 zoom in
2 pan the TIN
3 zoom out
4 view the edges
5 hillshadde
6 hypsometric shading
9 quit.

(ArcMap can also be used to display and analyze the TIN surface.)

USAGE: CREATETIN <out_tin> {weed_tolerance} {proximal_tolerance} {z_factor} {bnd_cover | xmin ymin xmax ymax} {device}

Arc: **CREATETIN groundtin # 2 #**

USAGE: COVER <in_cover> {POINT | LINE | POLY} {spot_item | spot_value} {sftype_item | sftype} {densify_interval} {logical_expression | select_file} {weed_tolerance}

CREATETIN: **COVER groundhits point ortho_ht mass**
CREATETIN: **end**

USAGE: DESCRIBETIN <tin> {device } { flat}

Arc: **DESCRIBETIN groundtin**

```
Arc: describetin groundtin
                          Description of TIN groundtin

                                  TIN Topology

                Nodes = 1086934             Triangles = 2173809

                              Surface value range

                Min Z = 9.270               Max Z = 190.310

                                  TIN Boundary

                Min X = 333000.000          Min Y = 4971000.000
                Max X = 336999.990          Max Y = 4974999.990

                          COORDINATE SYSTEM DESCRIPTION

Projection                  UTM
Zone                         20
Datum                      WGS84
Units                     METERS            Spheroid                  WGS84
Parameters:
```

Figure 4 – Output from the *DESCRIBETIN* command for the groundtin.

Arc: **DESCRIBETIN groundtin 9999**

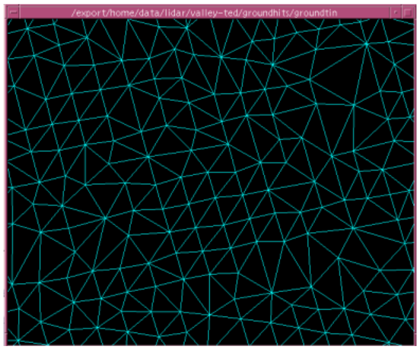


Figure 5 – Output from the *DESCRIBETIN* command for the groundtin with the display triangles option.

Arc: **DESCRIBETIN groundtin 9997**

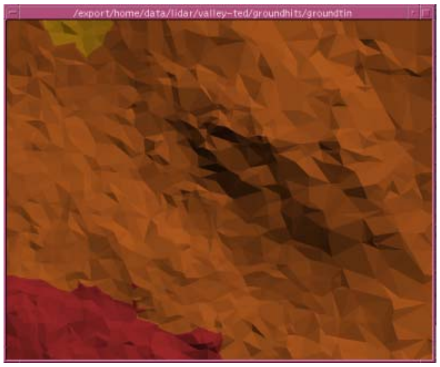


Figure 6 – Output from the *DESCRIBETIN* command for the groundtin with the display hypsometric shading option.

The TIN was then transformed into a grid or a lattice using both Quintic and Linear interpolations with a 2 m pixel resolution. A lattice is a surface interpolation of a grid, represented by equally spaced sample points referenced to a common and a constant sampling distance in the X and Y direction. Each mesh point contains the Z value of that location. Surface z values of locations between lattice mesh points are approximated by interpolation between adjacent mesh points. Quintic method provides a smoother surface then linear, and is best used to create ground surface DEMs. Linear is recommended to create the Digital Surface Models (e.g. all hits and non hits surface) with because it is known to interpolate sharp corners (such as buildings) better.

The command *TINLATTICE* converts the TIN to a lattice by interpolating the mesh points. The point z values are interpolated from the TIN using either the Linear or the Quintic method. The output lattice covers a rectangle area, so areas with NO DATA from outside the TINS zone of interpolation (convex hull) is represented by a null value of –9999.

After entering the *TINLATTICE* command, ArcInfo will display the extent of the grid and requires four more entries before the interpolation process begins. The first three lines can be left as the default values but the distance between the mesh points should be set to 2.

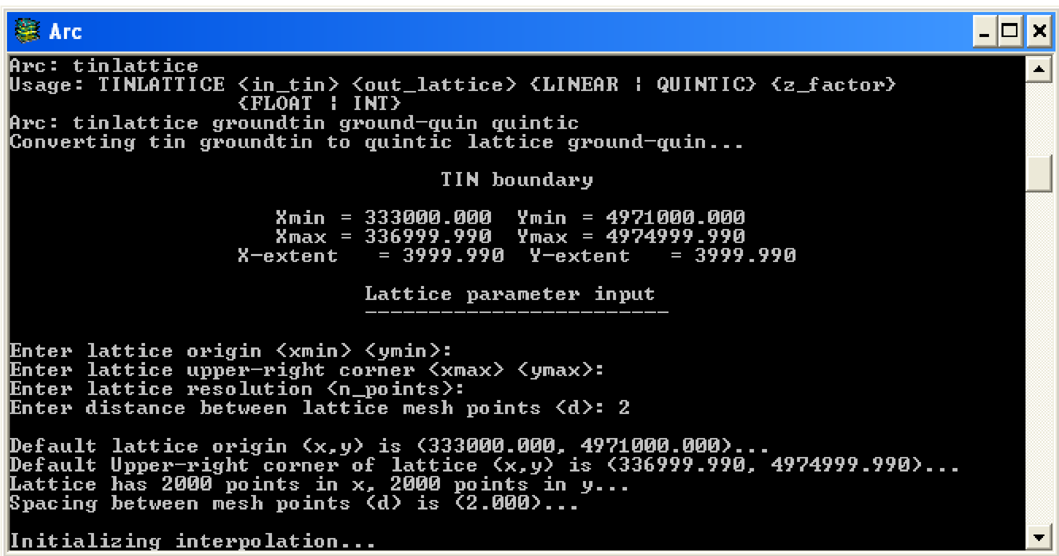


Figure 7 – The steps involved when using the *TINLATTICE* command.

Usage: TINLATTICE <in_tin> <out_lattice> {LINEAR | QUINTIC} {z_factor} {FLOAT | INT}

Arc: **TINLATICE groundtin ground-quin quintic**

The DESCRIBE command is useful to verify the details of the lattice. This will give you a quick overview with details such as cell size and coordinate system. ArcPlot or ArcMap can be used to display the resultant lattice for a visual examination.

Usage: DESCRIBE <geo_dataset>

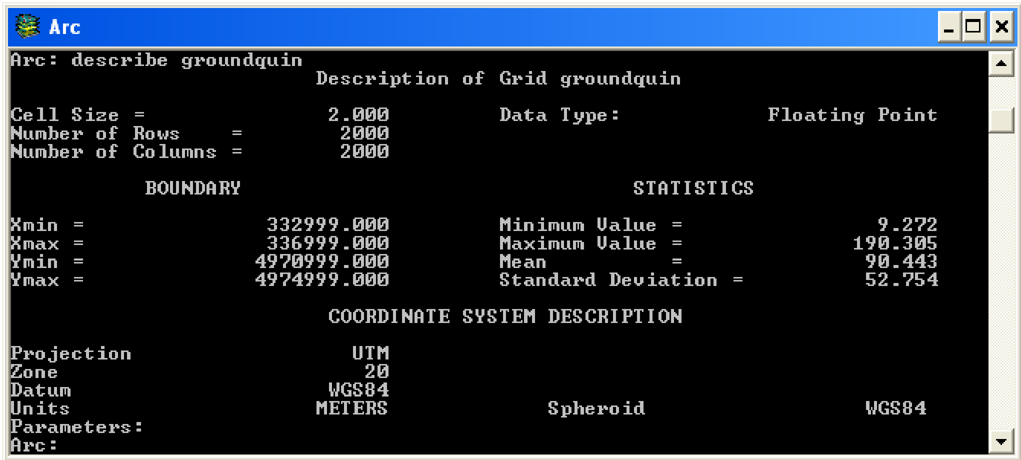


Figure 8 – Results of the *DESCRIBE* command will allow you to verify details of your file.

The above steps were repeated to create lattices of the non ground TIN. A lattice containing both the ground and the non ground hits together was created to provide an all-hits grid. The all-hits grid is an actual surface representation that includes all of the possible LIDAR points from the tile. Features such as fields and clear cuts are easily distinguished with the all-hits grid. This layer can be integrated with imagery data to create 3D models of the study area.

TOPOGRID is useful for generating a hydro-logically correct grid with elevation values from point, line, and polygon coverages. The LIDAR data contained only point data. The parameters for the TOPOGRID module were set with the following sub commands: *ENFORCE*, *DATATYPE*, and *POINT*. TOPOGRID will hold as much information as possible in the computer memory, so the GRIDALLOCSIZE variable was set.

The *DATATYPE* specified that the primary form of input data was a point coverage with elevations. *ENFORCE* was turned off to prevent the drainage enforcement routine from attempting to remove all sinks or depressions in the surface. *POINT* command specified which coverage and what attribute contained the elevation values. After all the options were specified *LIST* was used to verify that all the current parameters were correct, and *END* was the keyword command that indicated the conclusion of data input and initiated the computation process.

Usage: setenv GRIDALLOCSIZE <nrows*ncols*0.0000158>

Arc: **setenv gridallocsize 63.2**

Usage: TOPOGRID <out_grid> <cell_size>

Arc: **TOPOGRID groundtopo 2**
TopoGrid: **setenv gridallocsize 62**

Usage: ENFORCE <on |off>

TopoGrid: **enforce off**

Usage: DATATYPE <contour | spot>

TopoGrid: **datatype spot**

Usage: POINT <in_cover><elev_item>

TopoGrid: **point gnd333_4971 ortho_ht**
TopoGrid: **list**

```
ITERATIONS 30
ENFORCE OFF
DATATYPE SPOT
POINT gnd333_4971 ortho_ht
XYZLIMITS: xmin=333000.000 ymin=4971000.000 xmax=336999.990
ymax4974999.990 zmin=-26.938 zmax=226.518
Actual input elevation range  LowZ=9.270   HighZ=190.310
X limits are defaulted to the maximum X range of BNDs of all input coverages
Y limits are defaulted to the maximum Y range of BNDs of all input coverages
Z limits are defaulted to 20% of the Z range (of all input coverages with an
elev_item) below and above the lowest and highest elevations found
MARGIN: 0.000
TOLERANCES: tol1=2.500
          horizontal_std_err=1.000
            vertical_std_err=0.000
```

TopoGrid: **end**

The *IDW* is a GRID command that allows control over the significance of known points upon the interpolated values based on the distance from the output points. The interpolation determines the grid cell values with a linearly weighted combination of points.

Setting a larger *POWER* value will result in less influence from the surrounding points, thus the nearby data will have the most influence and the ground surface will have more detail. The default value was 2.

Usage: IDW(<point_cover | point file>, {spot_item}, {barriers}, {power}, {sample, {num_points}, {max_radius}}, {cellsize}, {xmin, ymin, xmax, ymax})

Usage: IDW(<point_cover | point file>, {spot_item}, {barriers}, {power}, {radius, {radius},{min_points}}, {cellsize}, {xmin, ymin, xmax, ymax})

GRID: **gndIDW = IDW(gnd333_4971, ortho_ht, #, 2, sample, 2, 2, 2)**

The *SPLINE* method is a GRID command that performs a two dimensional minimum curvature spline interpolation on the data and results in a smooth surface that goes through the input points. There are two methods of *SPLINE* that can be interpolated, *REGULARIZED* will give a smooth surface and smooth first derivatives, and *TENSION* fine tunes the surface according to the character of the modeled phenomenon. The greater the number of points used, the smoother the surface will be.

Usage: SPLINE(<point_cover | point file>, {spot_item}, {regularized | tension}, {weight}, {num_points}, {cellsize}, {xmin, ymin, xmax, ymax})

GRID: **gndspline = SPLINE(gnd333_4971, ortho_ht, #, #, #, 2)**

The *KRIGING* command was a GRID function that interpolates a grid from a point coverage using a kriging method based on a mathematical function to model the variation in z values. Kriging is a very computer-intensive process and speed of execution is dependent on the number of points in the input data set and the size of the search window. The *GRAPH* option is relatively quick, and should be executed before generating the final output grid to determine the most appropriate options.

BOTH, *GRAPH*, and *GRID* are keywords that specify whether a grid, or an INFO data file containing semi-variance values suitable for graphing, or both will be created. *METHOD* is a keyword that will specify the type of mathematical function used to model the semi-variance (*SPHERICAL* [default], *CIRCULAR*, *EXPONENTIAL*, *GAUSSIAN*, *LINEAR*, *UNIVERSAL1*, and *UNIVERSAL2*).

Usage: KRIGING(<point_cover | point_file>, {spot_item}, {barrier_cover | barrier_file}, {BOTH | GRAPH | GRID}, {output_variance}, {method}, {SAMPLE, {num_points}, {max_radius}}, {cellsize}, {xmin, ymin, xmax, ymax})
Usage: KRIGING(<point_cover | point_file>, {spot_item}, {barrier_cover | barrier_file}, {BOTH | GRAPH | GRID}, {output_variance}, {method}, {RADIUS, {radius}, {min_points}}, {cellsize}, {xmin, ymin, xmax, ymax})

GRID: **gndkrig = kriging(gnd333_4971, ortho_ht, #, grid, #, #, #, #, #, 2)**

The *POINTGRID* command will create a grid from a point coverage. Each cell in the resultant grid will be assigned a value according to the point that it overlays. Once the initial parameters are set, ArcInfo will prompt for the cell size and what portion of the coverage to be converted.

Usage: POINTGRID <in_cover><out_grid>{value_item}{lookup_table}{weight_table}

Arc: **POINTGRID gnd333_4971 groundpntgrd ortho_ht|**
Converting points from gnd333_4971 to grid groundpntgrd
Cell Size (square cell): **2**
Convert the Entire Coverage(Y/N)?: **y**
Enter background value (NODATA | ZERO): **nodata**
Number of Rows = 2001
Number of Columns = 2001

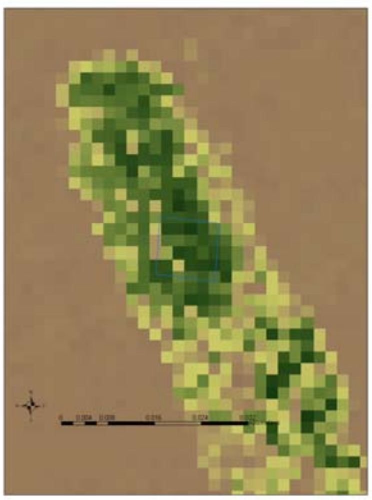
Vegetation Height Model

A vegetation heights grid is a useful tool to derive the heights of non ground features such as forest stands and man made features. This Grid is a simple byproduct from the all-hits and the ground-hits grids. The GRID program provides a full suite of operators to perform analysis between multiple grids. This method allows us to derive a new vegetation heights grid by subtracting the ground only grid from the all hits grid.

USAGE: OOUTGRID = INGRID1 – INGRID2

GRID: **vegheight = allhits - groundhits**

The resultant vegheight grid will have cell values that represent the height of the feature it represents. If you query a cell, you will get the height of the contents that the 4 m^2 cell represents (forest, building etc). ArcMap can be used to find the cell values of features from the vegheight layer. This layer is important for analyzing whether or not the LIDAR actually penetrated through the forest canopy.



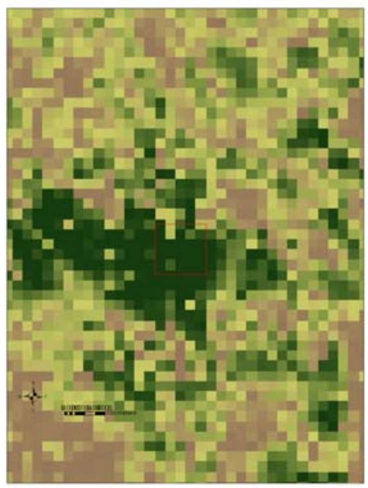


Figure 9 – Forest study plots (plot3 left, plot2 right) overlaid on top of the vegheight grid in ArcMap. Brown colors represent low heights, and green colors represent higher heights. From the vegheight layer, it was determined that the average height for plot 2 was 14.87 m and plot 3 was 10.97 m.

RGB COMPOSITE

PCI image software was used to visualize the DEM, create a RGB composite image, a color shaded relief image and then export elevation data for the GPS error anaysis. In order to import the lattice into PCI, it must be converted to an ASCII file.

Usage: GRIDASCII <in_grid> <out_ascii_file> {item}

Arc: **GRIDASCII groundquin groundquin.,grd**

The ASCII file (.grd) containing the elevation data from the exported grid was imported into a new PCIDSK file (.pix) with the *FIMPORT* command. This command will import the georeferencing data in addition to image data. Only file formats supported by the GDB (GeoGateway Data Base) library may be imported with *FIMPORT*, so it is important that the imported file have a .grd file extension.

Nearest neighbor was chosen as the resampling method and Band was chosen for the layout of the image. Band stores all the data together and produces good results when not all of the bands are being accessed at all times. The choice of layout is primarily based on performance.

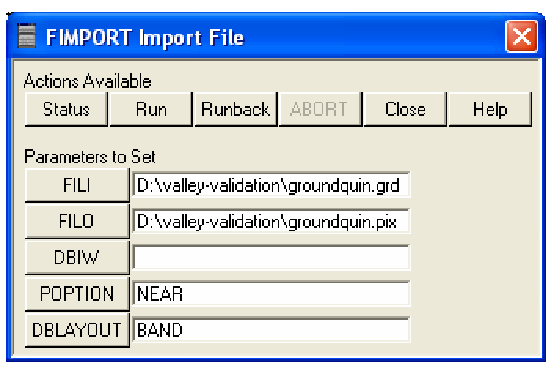


Figure 10 – FIMPORT command parameters XPACE window.

Eight 8-bit channels were added to the PCIDSK file (.pix) using the PCIMOD command. The new channels were needed to hold the results of the following procedures.

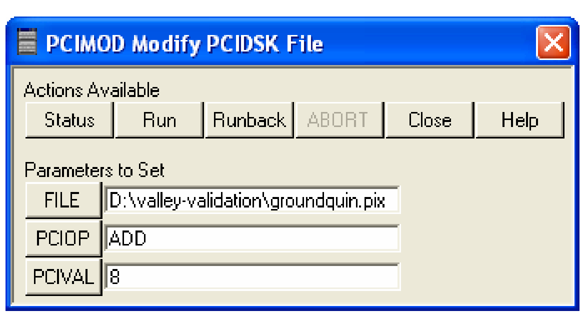


Figure 11 – PCIMOD command parameters XPACE window.

The DEM was scaled from a 32-bit real high-resolution image to an 8-bit lower-resolution image in order to better represent a correct pseudo color encoding. PCI has a default input value range of 0 to 255, thus we want our data with the highest elevation represented by the 255 value and the lowest by the 0 value, with all the other values between that.

The DEM contained a NODATA value of -9999, set by ArcInfo. When the data was imported to PCI, the software interprets the range of values from -9999 to your highest value (191m). This will result in all your data at one end of the possible values. So a histogram of your values would show all of your data near the 255 value.

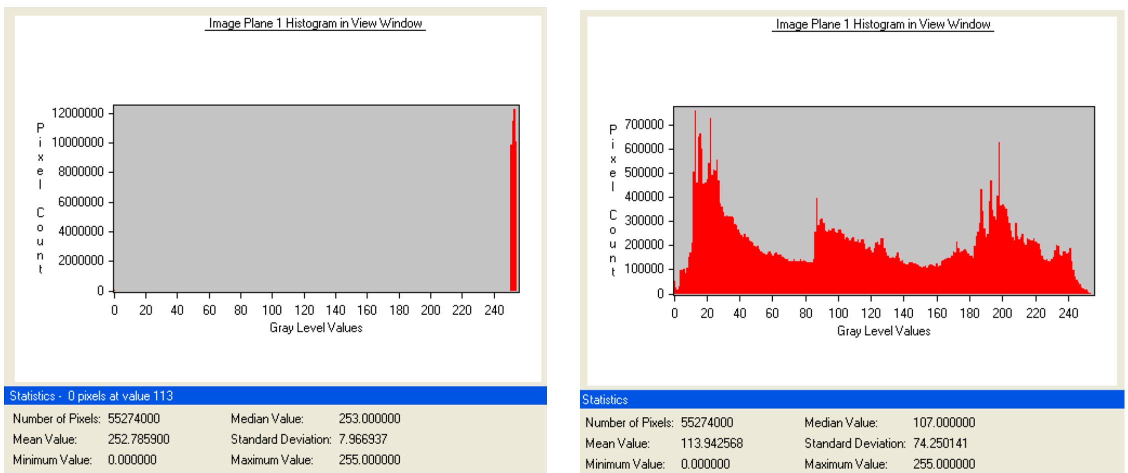


Figure 12 – PCI histograms, with the -9999 NODATA (left) and after using the SCALE command (right).

Loading the lattice into ArcMap will allow you to get the values of the highest and lowest points on your grid. The highest elevation for the ground hits only DEM was 190.30 m and the lowest elevation of this area was 9.27 m. The *SCALE* command will scale the data portion of your file to fit into the data range so you can use all the possible data.

Starting the input range from 9 to 191 means that all of the DEM values less then 9 will be omitted, thus our NODATA value of –9999 will not be taken in effect and only the existing data will be scaled into the usable range of 0-255. The Linear option was chosen so that the data would be scaled equally or linearly among the elevations between 9 and 191. After the scaling process is complete, the output histogram of your values should fully occupy the full range.

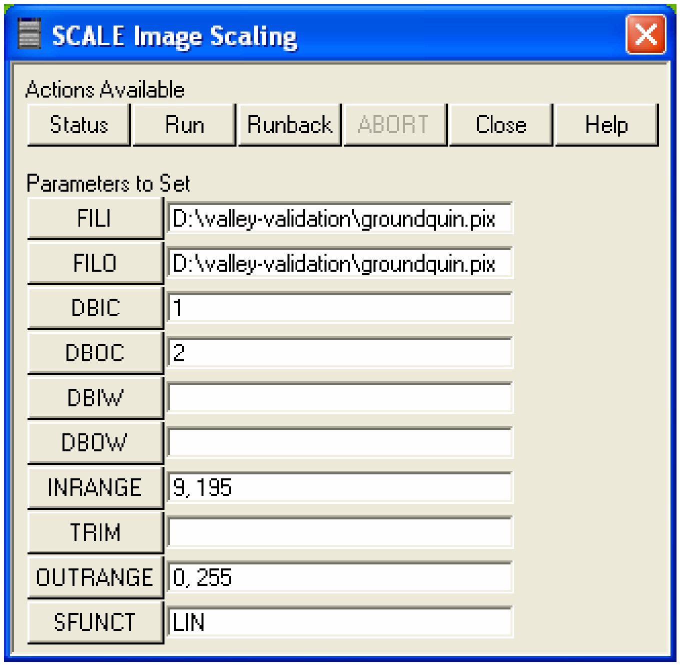


Figure 13 – SCALE command parameters XPACE window.

The DEM image is a 32-bit (real) image, thus when viewed using PCI IMAGEWORKS Software it had to be loaded into a 32-bit image plane. Three 8-bit image planes and one 32-bit image was allocated prior to loading the image. The image was then loaded, and the image plane appeared black because the default setting is to load the first three image planes that have no data, thus appearing black. The DEM will appear in the view by setting the view to Black and White and selecting the fourth image plane. A linear stretch was applied to enhance the image. The DEM should appear in gray scale where the dark areas represent low elevation and the light areas represent high elevation.

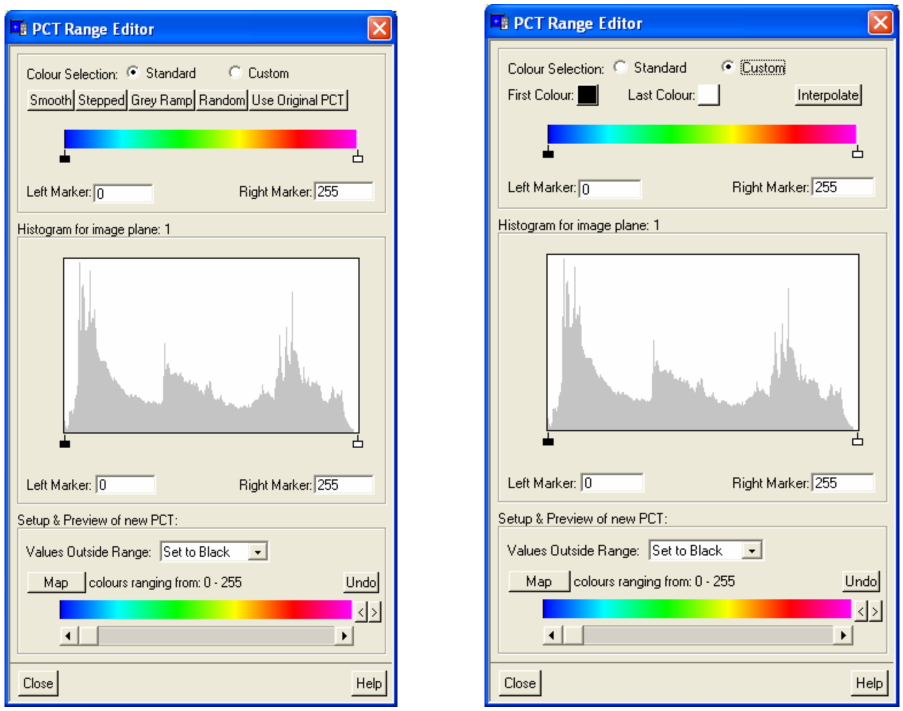


Figure 14 – The PCT Range Editor window, default setting (left) and the custom option (right).

To look at the image in pseudo color, the color option was set to PC (Pseudo Color). The PCT Range Editor window was opened by selecting PCT Range from the EDIT menu. The smooth button was selected and then the custom option was chosen to give the image a more suitable pseudo color range. The colors ranged from blue (representing low elevations) to green to yellow and to red (representing high elevations). The custom pseudo color table was then saved to a segment to the PCIDSK file by selecting SAVE PCT from the FILE menu. A name and description was added and then the new segment was created.

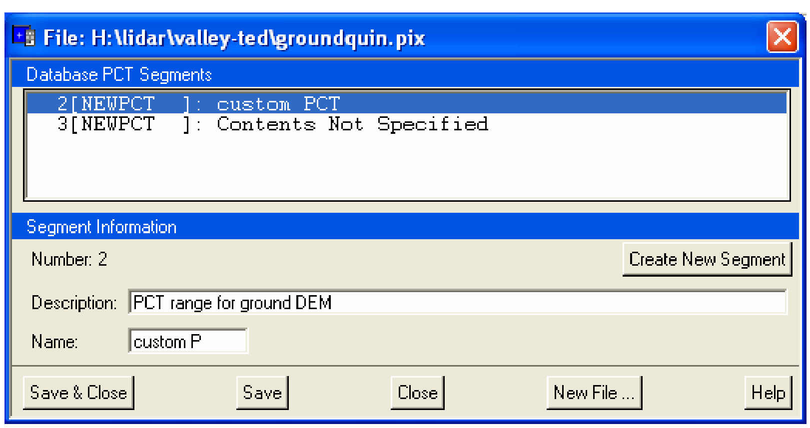


Figure 15 – Save the PCT segment as part of the PCIDSK file.

The *PCE* command was used to encode the scaled DEM channel using the PCT segment, to create three output channels representing red, green, and blue components. *NORM* was selected as the encoding method because it completely encodes the input channels into the three output channels where as the other option only encodes the non-zero input pixels. Now the image was loaded into FOCUS and was set to RGB, and appeared the same as it did with the pseudo color range.

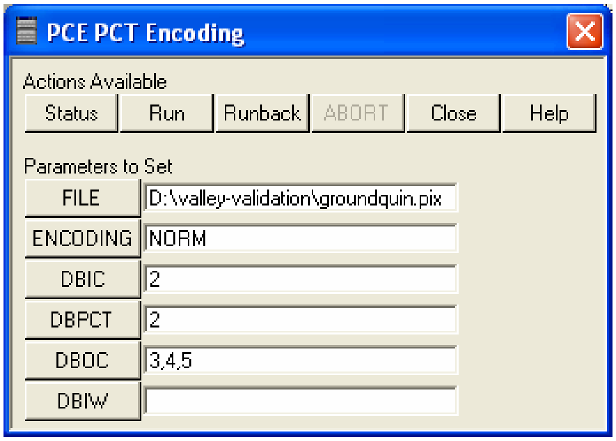


Figure 16 – PCE command parameters XPACE window.

Color Shaded Relief

A shaded relief image was produced with the *REL* command using the 32-bit DEM. This method gives the image a texture look, by making the slopes facing a user-specified light source (315,45) appear bright and those facing away appear dark. The shaded grey level at a point is calculated from the cosine of the angle between the normal vector to the surface (slope and aspect) and the direction of illumination. All surfaces not illuminated by the light source are set to zero.

The 32-bit DEM was used as the input source, to avoid creating any artifacts. An elevation exaggeration of 5 was applied to enhance the 3D of the surface. 315 was set as the azimuth angle and 45 for the elevation angle. The output pixel size of 2m was chosen.

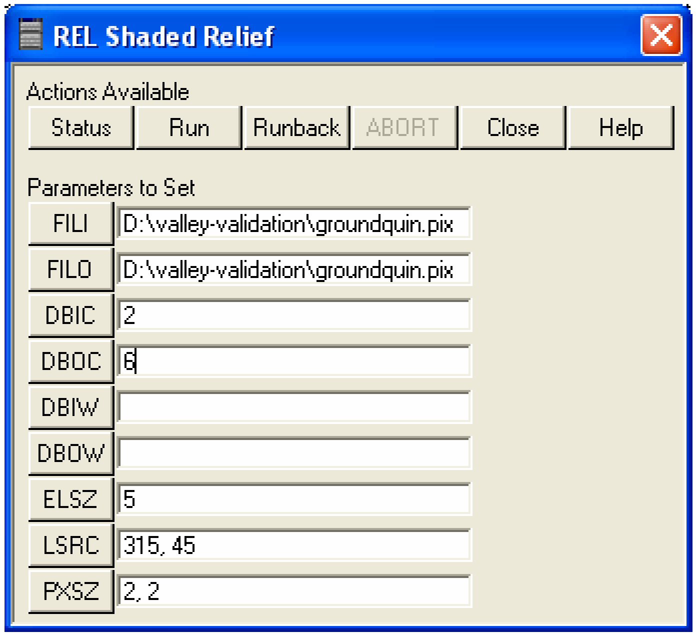


Figure 17 – REL command parameters XPACE window.

The *Model* command which implements a high level modelling language that can be used for raster GIS and imagery applications, was used to integrate channels created by PCE with the shaded relief to create a Color Shaded Relief image. This gives the surface a more pleasing visual appearance, and makes it easier to depict the elevations of the image.

By using a model (a special programming language), we can depict how channels of the imagery data and attribute data should be combined. The results is three new channels consisting of fifty percent of the Shaded relief and fifty percent of the pseudo color. The percentage of each can be changed to emphasize either channel, but they must add up to one hundred. In the model, the % sign represents a channel in the PCIDSK file. The following is the model that creates the CSR:

%7 = %3 * 0.5 + %6 * 0.5;
%8 = %4 * 0.5 + %6 * 0.5;
%9 = %5 * 0.5 + %6 * 0.5

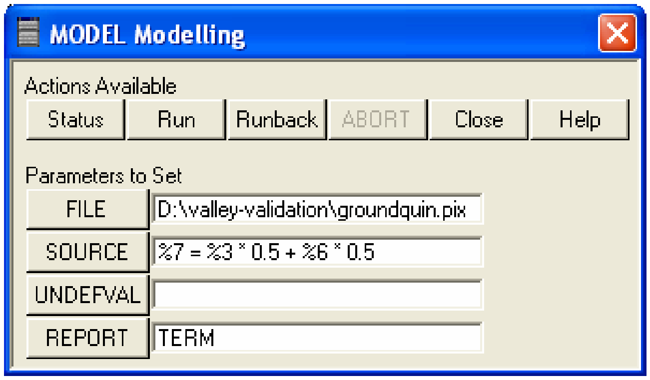


Figure 18 – MODEL command parameters XPACE window.

The GPS ESRI shape files were imported into PCI IMAGEWORKS, by selecting LOAD VECTORS from the FILE menu. The GPS data was then saved as segments of the PCIDSK file.

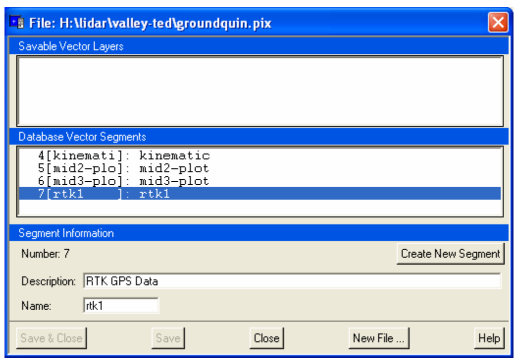


Figure 19 – Save the PCT segment as part of the PCIDSK file.

Export Elevation Data

The *VSAMPLE* command was used to extract the pixel values from the image that corresponded underneath the GPS vector segment and the results were exported to a text file. The resultant text file will contain the extracted elevation data from under the GPS points to be used for error analysis.

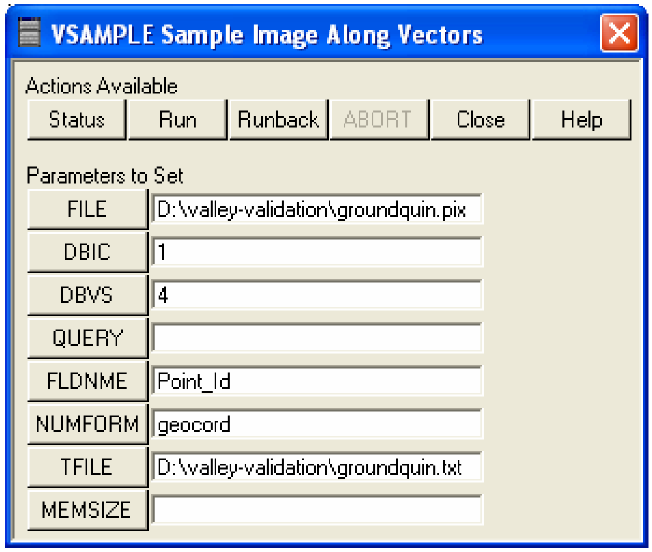


Figure 20 – VSAMPLE command parameters XPACE window.

The final PCIDSK Image files were then exported as geo-tiffs using the *FEXPORT* command. This command will transfer image and auxiliary information (optional) from the source file to an output file.

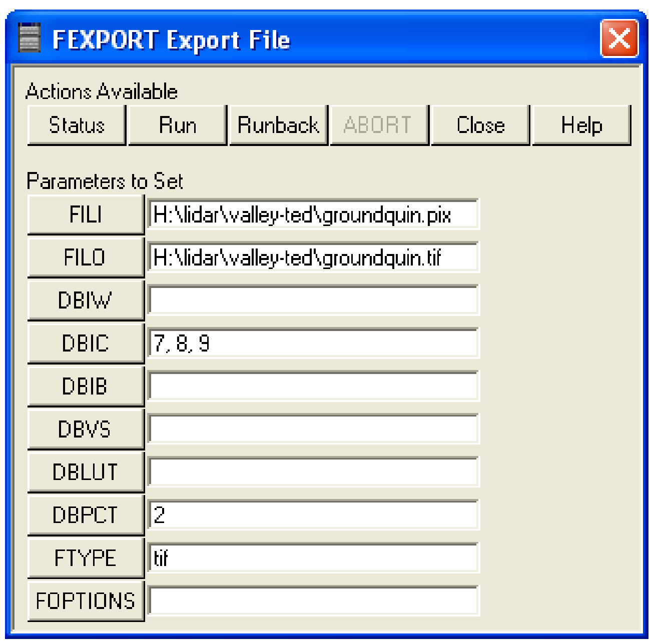


Figure 21 – FEXPORT command parameters XPACE window.

The exported geo-tiff files were compressed using MrSID (Multi-resolution Seamless Image Database) Software by Lizard Tech that utilizes a lossy compression technique based on wavelet technology. The MrSID image file format is designed specifically for transporting, managing and storing images. MrSID reduce the size of high-resolution images (compression ratios of 30:1 to 50:1) while minimizing the quality and integrity of the original. More information on Lizard Tech and MrSID is available from www.lizardtech.com. This type of compressed files are suitable for most applications such as ESRI ArcMap.

USAGE: mrsid_encode [Image Type] [-mos] –I Input [-o Output] [-n Nlev] [-b Blocksize] [-c Compression Ratio] [-t Temp Dir] [-v Version] [-bgc color] [-tpc color] [-h help]

mrsid_encode –tiffg –i groundquin.tif –o groundquin.sid

LIDAR Validation

GPS Collection / Field Work

In May 2003, The AGRG was involved in an extensive GPS survey to validate the LIDAR data collected by Terra. The specifications for the LIDAR data was to be within 30 cm, and 95% of the data must meet the specifications, so it was really important that the GPS data have better accuracy then the LIDAR data.

The GPS measurements were collected with a high precision Leica real time kinematic (RTK) differential Global Positioning System that was accurate up to two centimeters. The GPS was collected during the LIDAR campaign to ensure that our GPS measurements represented the actual conditions that the LIDAR sensor collected. Most of the RTK data was collected utilizing a moving vehicle. This was mainly because it was the most efficient way to collect an abundant amount of data and Terra will only guarantee to meet the specifications on hard flat surfaces.

The RTK unit on a pole and the Total Station system were used to collect GPS in areas that were not accessible to a vehicle such as fields and in the forest. Phase code GPS cannot penetrate dense forest canopy so it was important to use the Total Station unit to obtain our orthometric height in the forest. The RTK unit would provide the Total Station with its initial GPS coordinates and then it would use high a precion laser and prism to measure both distance and angles and then compute the coordinates of the new position.

Figure 22 – Setting up the Total Station unit outside of a forest stand.

A concern with LIDAR data is how well the laser actually penetrates through the forest canopy, so several forest plots were investigated and studied. Two of the various forest plots examined by the AGRG are located within this LIDAR tile.

Each forest plot represented a 100 square meter area. Detailed descriptions of the vegetation and the topography within each plot were recorded. The height of the trees were measured with a SUUNTO height meter and a spherical densitometer were used to estimate the percent of crown closure. The total station provided geographic location of corners of the plots and provided a profile of the surface beneath the canopy. All of the plot details were recorded and entered into a spreadsheet.

Figure 23 – Each forest plot represented a ten meter by a ten meter grid. Plot 2 (red box) in the left image was located west of a recent clear cut in an Aspen stand. Plot 3 (green box) in the right image was located within a hedgerow and was also an Aspen stand.

Note: The air photos above were photographed 2 months after the actual forest plots were sampled, and was intended for a geographical representation and not the actual forest conditions at the time of the plots.

LIDAR surface compared with GPS

This validation approach dealt with comparing the orthometric heights of the GPS points with the DEMs created from the LIDAR interpolations. The GPS data was imported into a Microsoft Excel spreadsheet using the DBF file associated with the shapefile. The LIDAR data was also imported into the same spreadsheet using the text file created with the *VSAMPLE* command in PCI. The VSAMPLE data contained only elevation information for the associated GPS points.

The Orthometric heights for the LIDAR and the GPS measurements were then used to calculate a column containing the difference in elevation for each point. The difference column was then used to create a column containing the absolute difference. Statistics were then calculated using these two columns of data. The mean **[AVERAGE (difference)]** was calculated from the average of all the difference values. The magnitude of deviation **[AVERAGE (absolute difference)]** was calculated from the average of all the absolute difference values. The Standard Deviation **[STDEV (difference)]** was calculated from the difference values. The average magnitude **[AVEDEV (difference)]** was calculated from the difference values. The root mean square **[SQRT (AVERAGE (Difference^2))]** was calculated from the square root of the average of all the difference values.

The statistics were graphed once all the calculations were complete. A graph representing a comparison between the orthometric height of the LIDAR plots and the measured orthometric height from the GPS. A second graph was plotted to demonstrate the variance of each measured point with the derived LIDAR points.

LIDAR points compared with GPS

Another approach of LIDAR validation was to compare the measured GPS points to the LIDAR points within a specified radius of the GPS points. A search radius of 2 m was specified to compare only the LIDAR points that were close to the GPS points. Most of the GPS data was collected using RTK with a vehicle, typically on a paved road, so using a radius of 2 m would ensure that the LIDAR points you are comparing are on the same level surface (e.g. shoulder of road vs. center of road). The Orthometric heights of the LIDAR points within the search radius were then statistically compared to those of the GPS orthometric heights.

This section was done utilizing an AML (TOOL1.AML) program written by a previous AGRG student. The AML was a GUI (Graphical User Interface) style setup that leads the user through the procedure. The actual computation was done using ArcPlot and involved several other AML programs. The operation was repeated several times for the different GPS data files (RTK – vehicle, RTK – Pole, and Total Station).

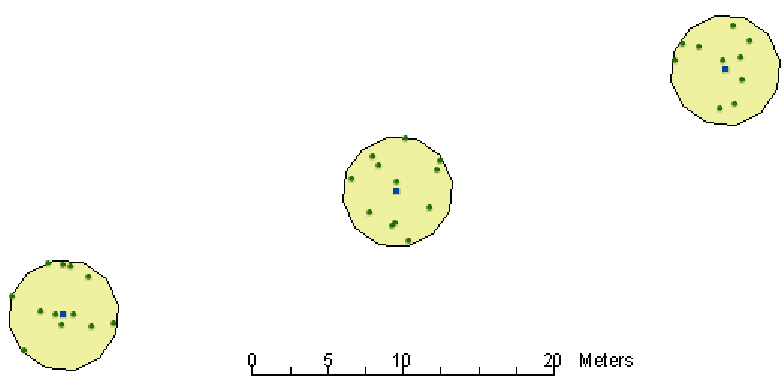


Figure 24 – GPS points (blue square) buffered with a 2 m radius (yellow circle) and the LIDAR points (green circles) that are within the buffered zone.

The following is a brief overview of the TOOL1.AML program. The GPS shapefile had to be converted into a coverage prior to running this program.

USAGE: &RUN <AML program>

Arc: **&RUN tool1.aml**

- Select the GPS point coverage
  plot2
- Select the elevation field
  ORTHO_HT
- Specify a search radius in meters
  2
- Choose the workspace where the coverage to be processed is located
  Groundhits
- Create or choose an existing directory to store all of the new coverages and info files (NOTE: All the coverages that are going to be processed must have exactly the same fields or the AML will not work properly)
- Select the elevation field
  ORTHO_HT
- Specify the desired percentage of Standard Deviation
  25
- A screen will appear with all the details on all your new coverage and info files that you created
- A histogram will then be displayed

The AML will produce:

- A polygon coverage of the buffered GPS points
  clipcov2
- A point coverage of the LIDAR points within the buffered area
  mrg_pnts2
- A point coverage of the GPS points
  results2
- A table (info .dat file) with elevation differences between GPS point and the LIDAR points within the buffer
  pntdist2.dat
- Summary Statistics for the above table
  pntstats2.dat
- Detailed Statistics for each point
  pntstats2gr.dat

The statistics provided from the AML are: Frequency of LIDAR points within the buffer, minimum elevation difference, maximum elevation difference, mean elevation difference, and the standard deviation. A scatter plot comparing the Orthometric height difference between the LIDAR and the GPS with the Distance from the GPS point for each file was plotted. The expected result is that there should be a larger variance with the results, the further you go away from the GPS point.

Visual Comparison

All the tiff files, grids, point coverages, shape files and info tables were loaded into ArcMap. Each generated surface was examined and evaluated. The LIDAR points were examined and compared to the surfaces that were interpolated from them to ensure that the data had been properly classified by Terra and that no erroneous points were classified wrong.

Classifying the GPS points with different colors allowed any erroneous points to be identified. The LIDAR points from within the buffered radius was colorized to demonstrate which points had a greater difference of 30 cm.

The DEM and generated surfaces were also loaded into ArcScene to examine and evaluate them with a three dimensional perspective view. A LandSat 7 image, and an IKONOS image were used to drape over top of the surface models. AGRG aerial photographs were mosaiced together with PCI ORTHOENGINE to provide a higher resolution image to use with the surface models when analyzing the forest plots.

Vegetation Heights

ArcMap, ArcScene and an Excel spread sheet were used to evaluate whether or not the LIDAR penetrated the forest canopy. The Vegetation Heights surface model, the LIDAR points, the Total station points, the forest plot polygons and the aerial photo mosaic were all loaded into both ArcMap and ArcScene.

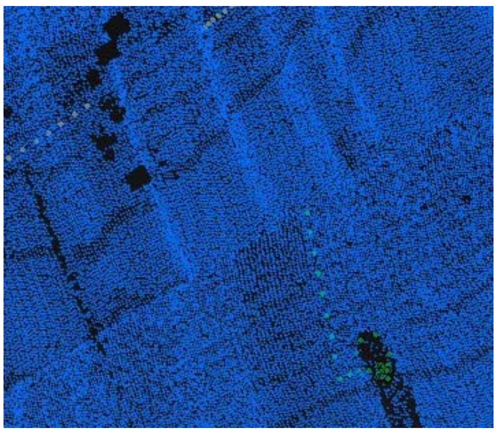
Ideally a ten meter by ten meter plot would incorporate 25 two meter grid cells (5 x 5) but neither of the two plots landed exactly on a 5 x 5 grid. One forest plot landed within a 5 x 6 area (30 grid cells) and the other plot landed within a 6 x 6 area (36 grid cells). This should not make much difference to the results but is important to note. Each cell value in the grid contained a vegetation height, so the height for each cell that landed within the plot area was entered into the forest polygon. The spread sheet then contained two columns of heights that could be compared; the measured heights and the interpolated heights.

The ground only and the non-ground LIDAR points were evaluated with the forest plot polygons to see how much LIDAR was actually penetrating to the ground. Simply symbolizing the two point coverages with different colors allowed the simply calculation of the ratio of ground hits and non ground hits actually land within the polygon.

Validation Results

Visual Comparison

After visually analyzing the ground only and non-ground LIDAR points it was determined that Terra did not properly separate the data. It seems that their algorithm has placed ground only hits into the non-ground hits file. This mismatch of points often results in the surface model having a "wood grain" appearance. This wood grain effect was portrayed in all of the interpolated surfaces.



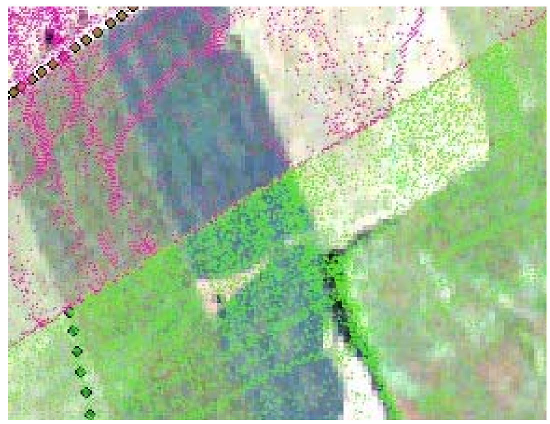


Figure 25 – After a visual analysis of the LIDAR data, it was determined that Terra has not properly separated the ground and non-ground points. Left image (about 1.5km) shows ground only hits (blue) and the right image shows the non-ground hits (green and pink).

The surface model interpolated with the Quintic TIN method best represents a true ground surface. The LIDAR points are so close and so abundant that the usual limitations from this method are not experienced. If the LIDAR contained fewer points, the triangle facets that created the TIN would be much longer and give the surface a stretched look.

The Linear TIN method was similar to that of the Quintic except for it was slightly rougher. The Linear method process took less time then the Quintic method. The Linear TIN grid of all the LIDAR points represents the buildings better then the Quintic method.

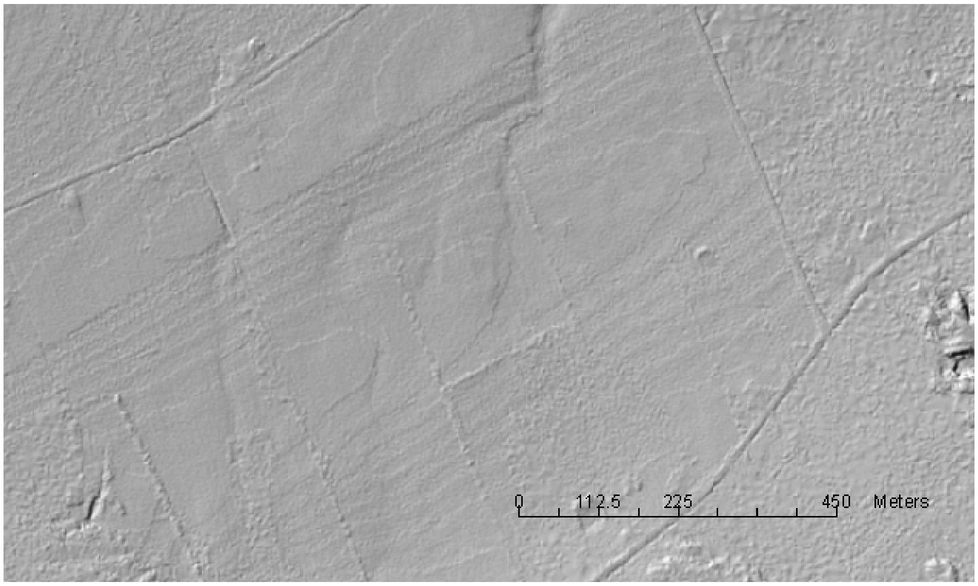


Figure 26 – Hillshade created using the grid interpolated with the Quintic TIN method.

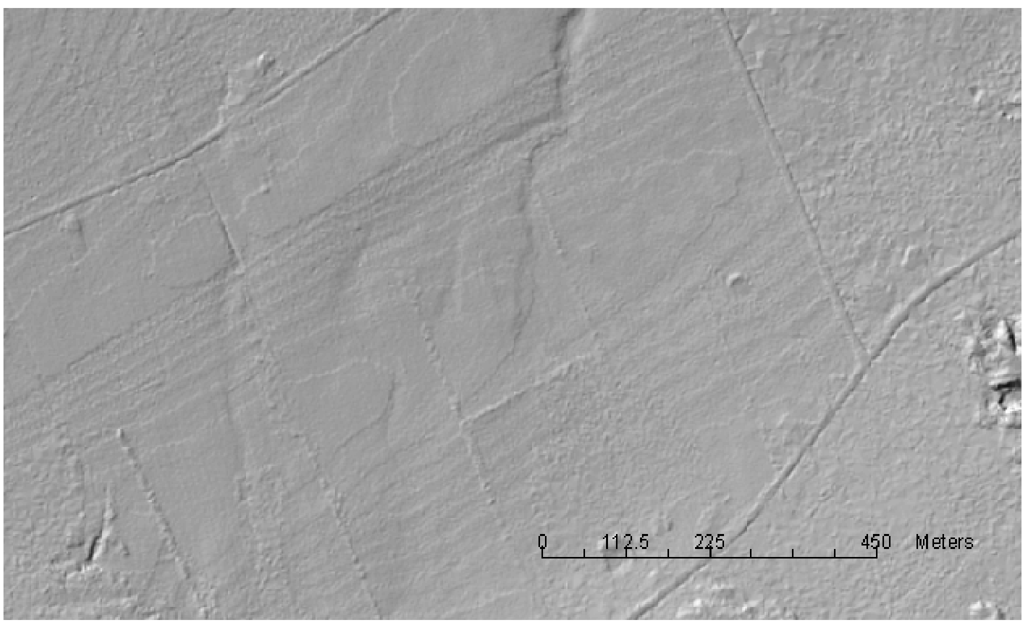


Figure 27 – Hillshade created using the grid interpolated with the Linear TIN method.

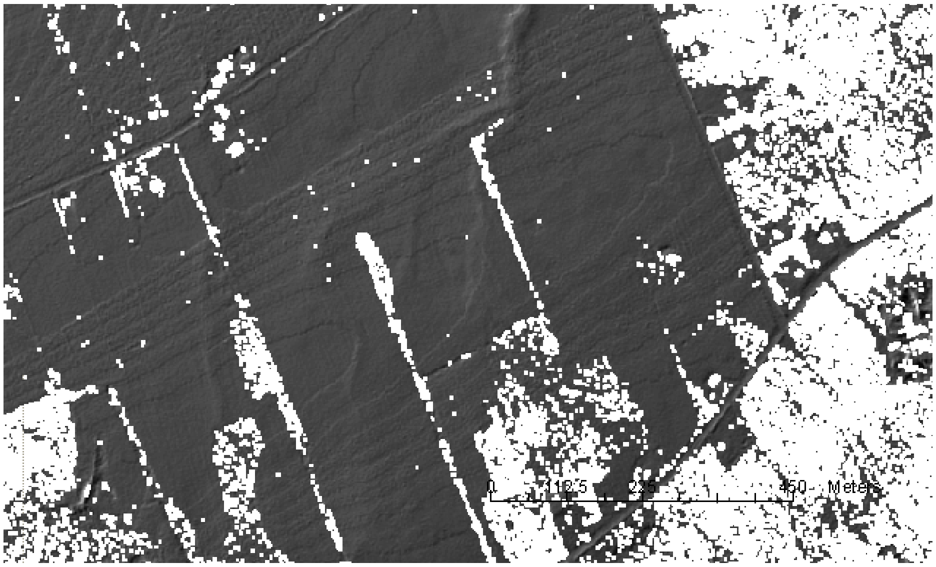


Figure 28 – A hillshade created from the IDW interpolation method.

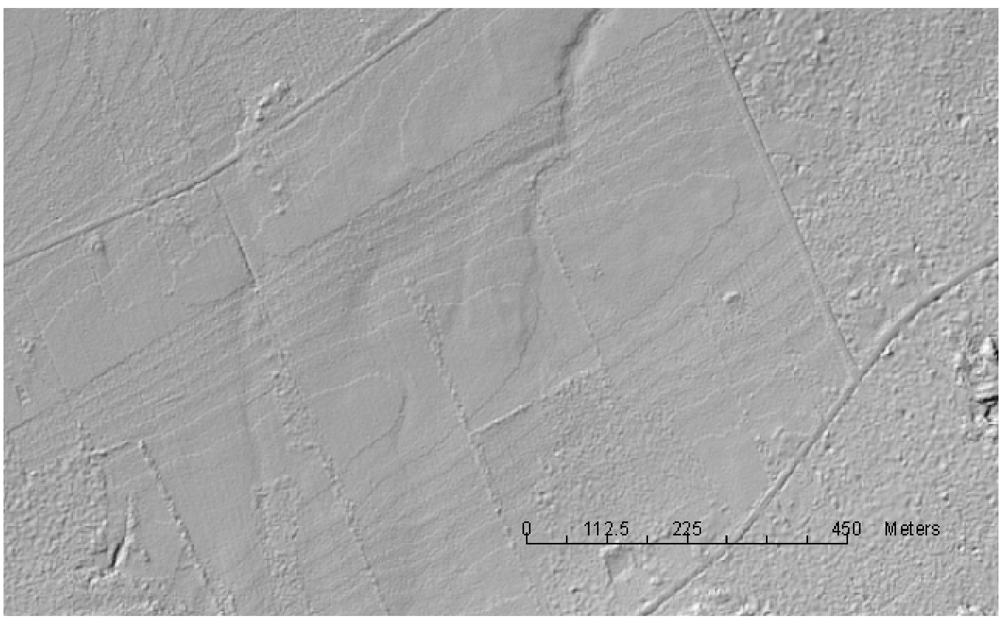


Figure 29 – A hillshade created from the Spline interpolation method.

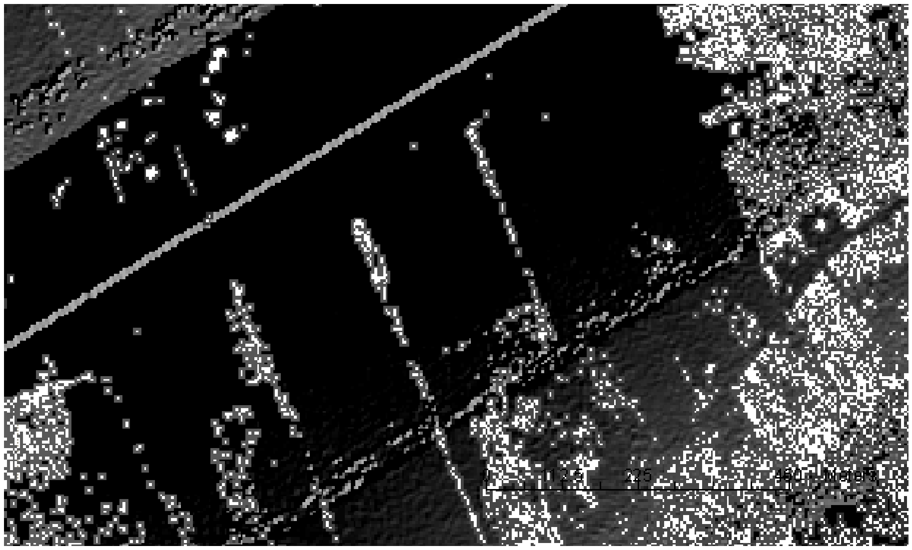


Figure 30 – Hillshade created using the grid interpolated with the POINTGRID method.

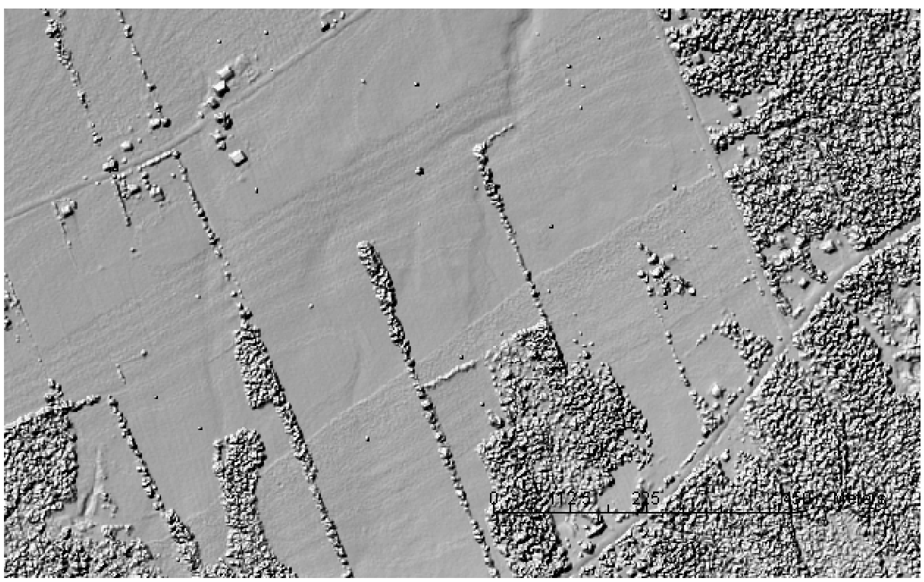


Figure 31 – A Hillshade surface model created using all of the LIDAR points. The blocky features in the top left corner are buildings and the rough features on the left hand side is forest.

The other surface interpolation techniques had various different results. IDW and POINTGRID grids contained empty grid cells when they experienced areas with no points, so the results were very blocky rough looking surface. The Spline method produced a grid that was similar to the Linear TIN method but was more textured. Problems with the TOPOGRID and Kriging commands prevented any surfaces from being generated with these method, thus they could not be visually compared with the others grids.

LIDAR surface compared with GPS

Several statistical calculations were applied to the data. The LIDAR error was usually better then the accepted 30 centimeter specification but the percentage ranged from eighty to ninety five percent. The following tables are results of the Orthometric height comparison between the LIDAR surface and the GPS points.

RTK(POLE) GPS vs LIDAR GRID (ground only from TIN using Quintic Interpolation)

Number of GPS Points	24
Number of GPS Points > 30 cm	1
Percent of Points Less then 30 cm	4.2%
Mean	-0.06
magnitude of deviation	0.10
Std. Dev.	0.11
Avg. Mag	0.09
RMS	0.12

RTK (Vehicle) GPS vs LIDAR GRID (ground only from TIN using Quintic Interpolation)

Number of GPS Points	580
Number of GPS Points > 30 cm	61
Percent of Points Less then 30 cm	10.5%
Mean	0.08
magnitude of deviation	0.15
Std. Dev.	0.17
Avg. Mag	0.13
RMS	0.19

TS vs LIDAR GRID (ground only from TIN using Quintic Interpolation)

Number of GPS Points	58
Number of GPS Points > 30 cm	13
Percent of Points Less then 30 cm	22.4%
Mean	0.04
magnitude of deviation	0.21
Std. Dev.	0.32
Avg. Mag	0.20
RMS	0.32

LIDAR points compared with GPS

Several statistical calculations were applied to the data with the TOOL1.AML. The LIDAR error was usually better then the accepted thirty centimeter specification but the percentage ranged from eighty to ninety four percent. The forest plot data did not meet the specs with this method, but they were not actually expected to. The histograms of the error were good typical bell curve shaped centered around 0 m, so there was no obvious shift in the data. The histogram of the RTK (vehicle) data does not appear to have a good curve shape, at first, due to a few erroneous points. Removing these bad points or scaling the data will produce a similar shape to the RTK (POLE) data. The following tables are results of the Orthometric height comparison between the LIDAR surface and the GPS points.

GPS (RTK-POLE) vs LIDAR POINTS

Number of GPS Points	73
Number of GPS Points > 30 cm	4
Percent of Points Less then 30 cm	5.5%
Mean	0.06
magnitude of deviation	0.12
Std. Dev.	0.14
Avg. Mag	0.11
RMS	0.15

GPS (RTK-Vehicle) vs LIDAR POINTS

Number of GPS Points	2783
Number of GPS Points > 30 cm	590
Percent of Points Less then 30 cm	21.2%
Mean	-0.16
magnitude of deviation	0.22
Std. Dev.	0.66
Avg. Mag	0.19
RMS	0.68

Vegetation Heights

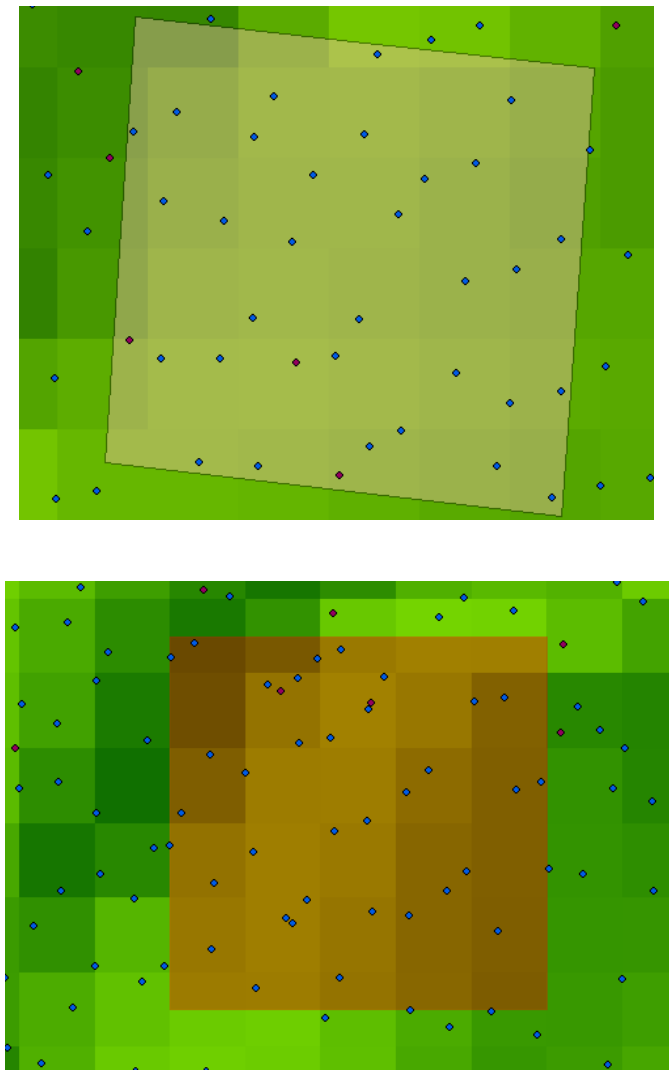


Figure 32 – Ground only (red) points and non ground points (blue) compared with both of the 10 x 10 meter forest plots (plot2 on the top and plot 3 on the bottom) and the vegetation heights surface model.

The results from comparing the ratio of ground verses non ground points within each forest proved that not very much of the LIDAR pulses are actually penetrating through the canopy to the surface. Eight percent of the ground LIDAR penetrated through the forest in plot 2 and five percent of the LIDAR penetrated through the forest in plot 3.

The average height for each plot derived from the vegetation heights layer were both different then the actual heights measured in the field. The average height of the

LIDAR in plot 2 was ninety two centimeters higher then the actual average height of the trees. The average height of the LIDAR in plot 3 was five and a half meters lower then the actual average height of the trees.

The following tables are summary results representing both plot 2 and plot 3.

Forest Plot 2 Details

		Plot Size	10 m x 10 m plot
Date	2003/05/18		5 x 6 2m pixels
Number of Trees	18		
Gnd_Veg	grass, 8 spruce saplings (2m), 20 HW saplings (1-2m leaf on)		
Gnd_Veg_Ht	.02 to .05 m		
Gnd_Moisture	Dry		
Slope_Position	Flat	Per_Rock	1%
Slope	Flat/minor	Per_Soil	0
Aspect	South	Per_Veg	99%
Terrain_Roughness	Flat (some slight Undulating)	Per_Water	0
Terrain_Ht_Var	0.3 m	Per_Snow	0
Crown_Closure	69%	Per_Other	0
Comments	~ mostly Trembling Aspen, some spruce ~ stand is near edge of clear cut, irregular landscape around the plot ~ lots of Granite Outcrop around the plot		

AVERAGE Height (measured)	13.95 m
AVERAGE Height (LIDAR)	14.87 m
AVERAGE Height (DNR)	13.00 m

Forest Plot 3 Details

		Plot Size	10 m x 10 m plot
Date	2003/05/18		6 x 6 2m pixels
Number of Trees	33		
Gnd_Veg	Ferns, grass, HW (about 1m)		
Gnd_Veg_Ht	0.06 m		
Gnd_Moisture	Moist	Per_Rock	0
Slope_Position	-	Per_Soil	0
Slope	Flat	Per_Veg	100%
Aspect	-	Per_Water	0
Terrain_Roughness	Flat (some slight Undulating)	Per_Snow	0
Terrain_Ht_Var	0.3 m	Per_Other	0
Crown_Closure	69.80%		
Comments	~ stand is within hedgerow between 2 fields ~ Leaf Conditions: All had leaves starting ~ some dead logs		

AVERAGE Height (measured)	16.51 m
AVERAGE Height (LIDAR)	10.96 m
AVERAGE Height (DNR)	N/A

Conclusions

After several visual inspections of the surface models derived from the LIDAR points, and reading the ERSI help section about each module it was determined that the best representation of a true ground DEM surface was from the Quintic TIN method. The Linear TIN method was processed in less time and was best suited for the al-hits surface. Adding Color to the shaded DEM definitely helps to optimize the relief and make it more appealing to the human eye.

The statistical validation on this tile with both GPS comparison methods determined that the LIDAR did meet the 30 cm specification with between 80 and 95% meeting the specifications. It is important to note that it is important to calculate your statistics and validate the data to ensure that your data actually meets the specifications. A Visual comparison of the points and surface models found that there were classification errors of the points.

Although the LIDAR is capable of penetrating through the vegetation canopy, allowing for detailed measurements of the ground topography, much more research should be applied to this topic. Even though the occurrence of ground hits under the canopy was sparse, a high-resolution ground surface model can still be produced.

After a thorough quantitative and visual analysis of the LIDAR points and the products derived from it, it was determined that the data meet the specs. The resultant LIDAR DEMs are a great high-resolution alternative to the 20m provincial DEM conventionally used to represent the topography of the valley.

Overview of GPS vs LIDAR GRID (ground only from TIN using Quintic Interpolation)

	Mean	Magnitude of Deviation	Standard Deviation	Average Magnitude	Root Mean Square	Percent that meets Specifications
RTK(CAR) GPS vs LIDAR GRID	0.08	0.15	0.17	0.13	0.19	89.50%
RTK GPS vs LIDAR GRID	-0.06	0.1	0.11	0.09	0.12	95.80%
TS vs LIDAR GRID	0.04	0.21	0.32	0.2	0.32	77.60%
GPS (RTK-POLE) vs LIDAR POINTS	0.06	0.12	0.14	0.11	0.15	94.50%
GPS (RTK-Vehicle) vs LIDAR POINTS	-0.16	0.22	0.66	0.19	0.68	78.80%

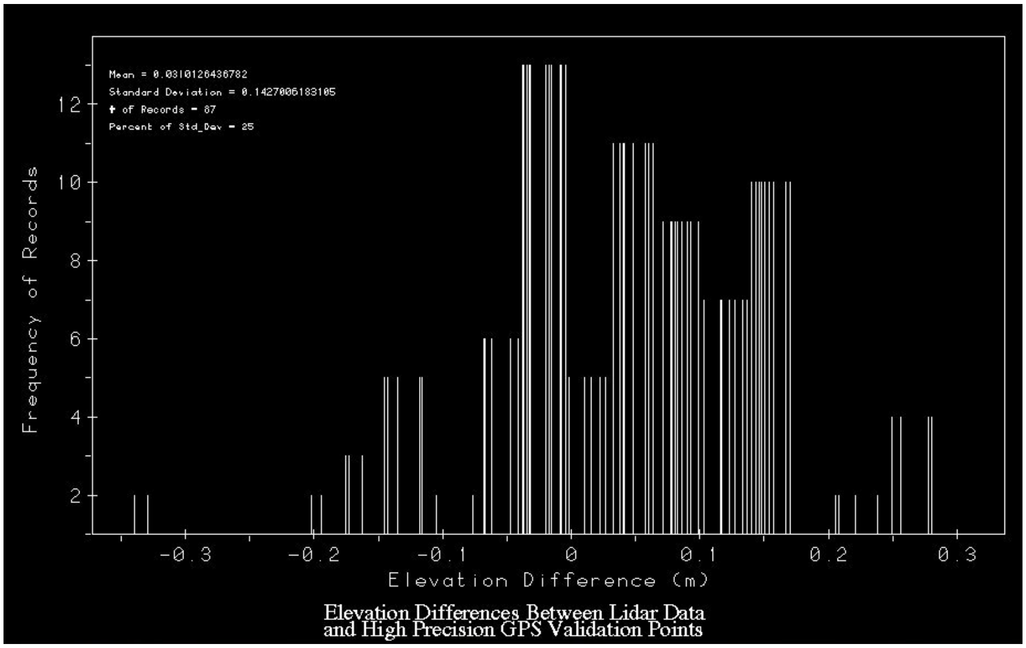


Figure 33 – RTK (POLE) Histogram created with ArcPlot.

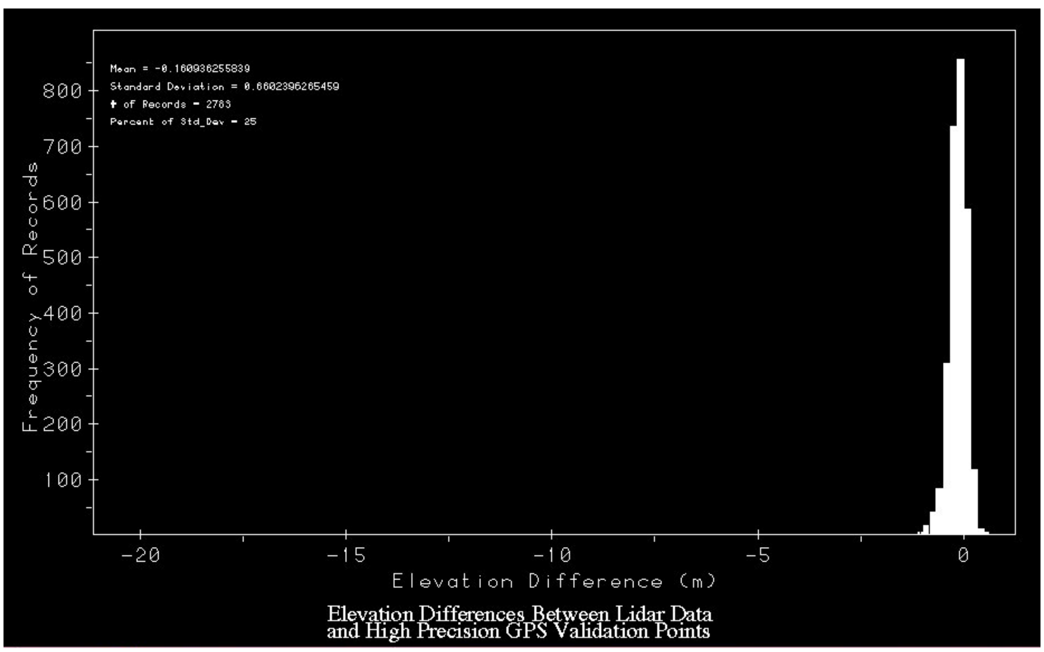


Figure 34 – RTK (Vehicle) Histogram created with ArcPlot.

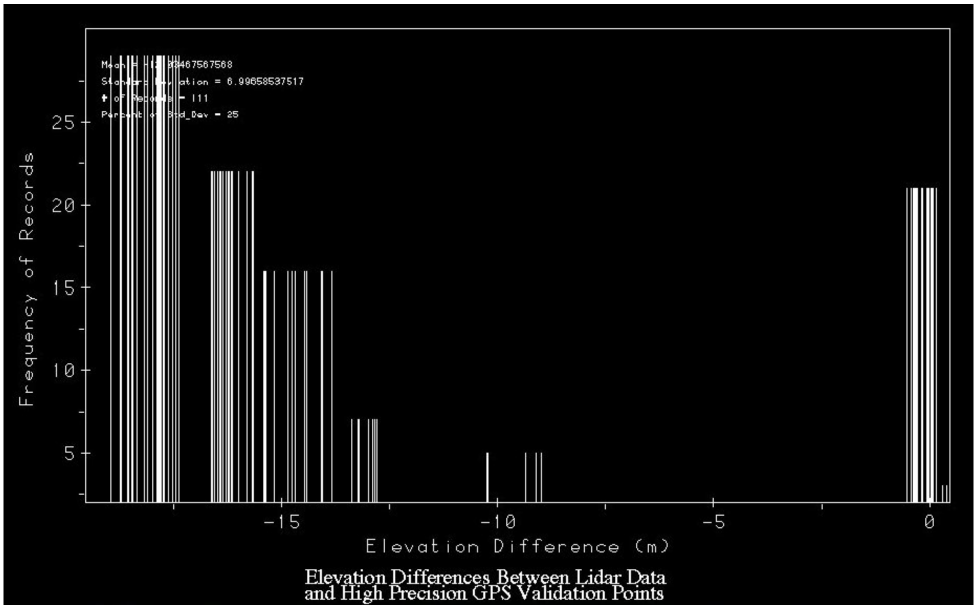


Figure 35 – Total Station Histogram for Plot 2 created with ArcPlot.

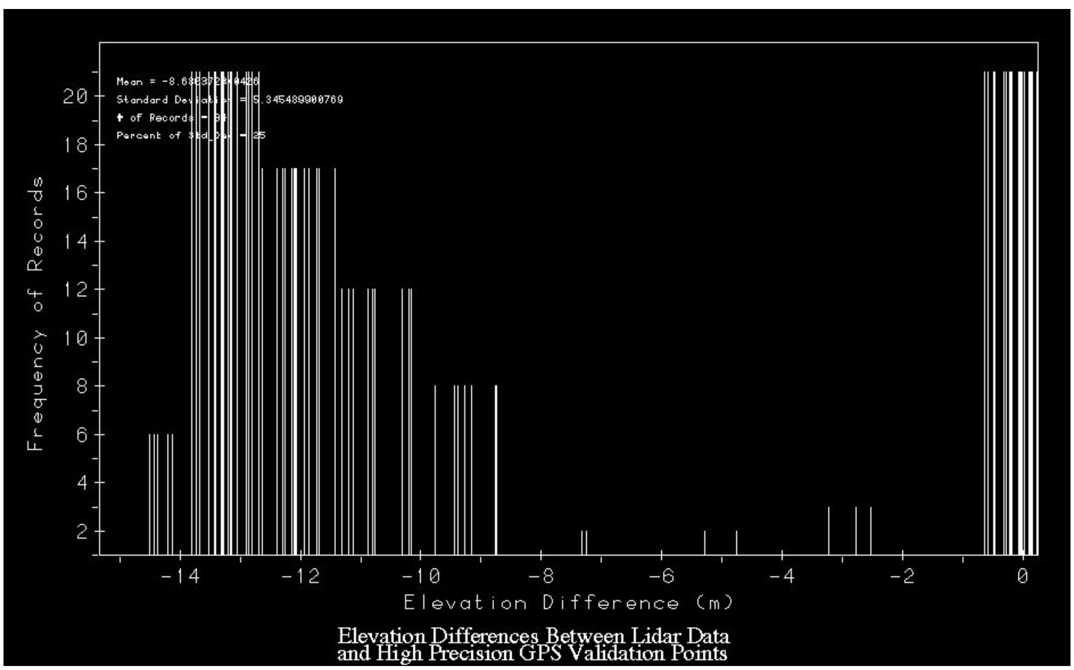


Figure 36 – Total Station Histogram for Plot3 created with ArcPlot.

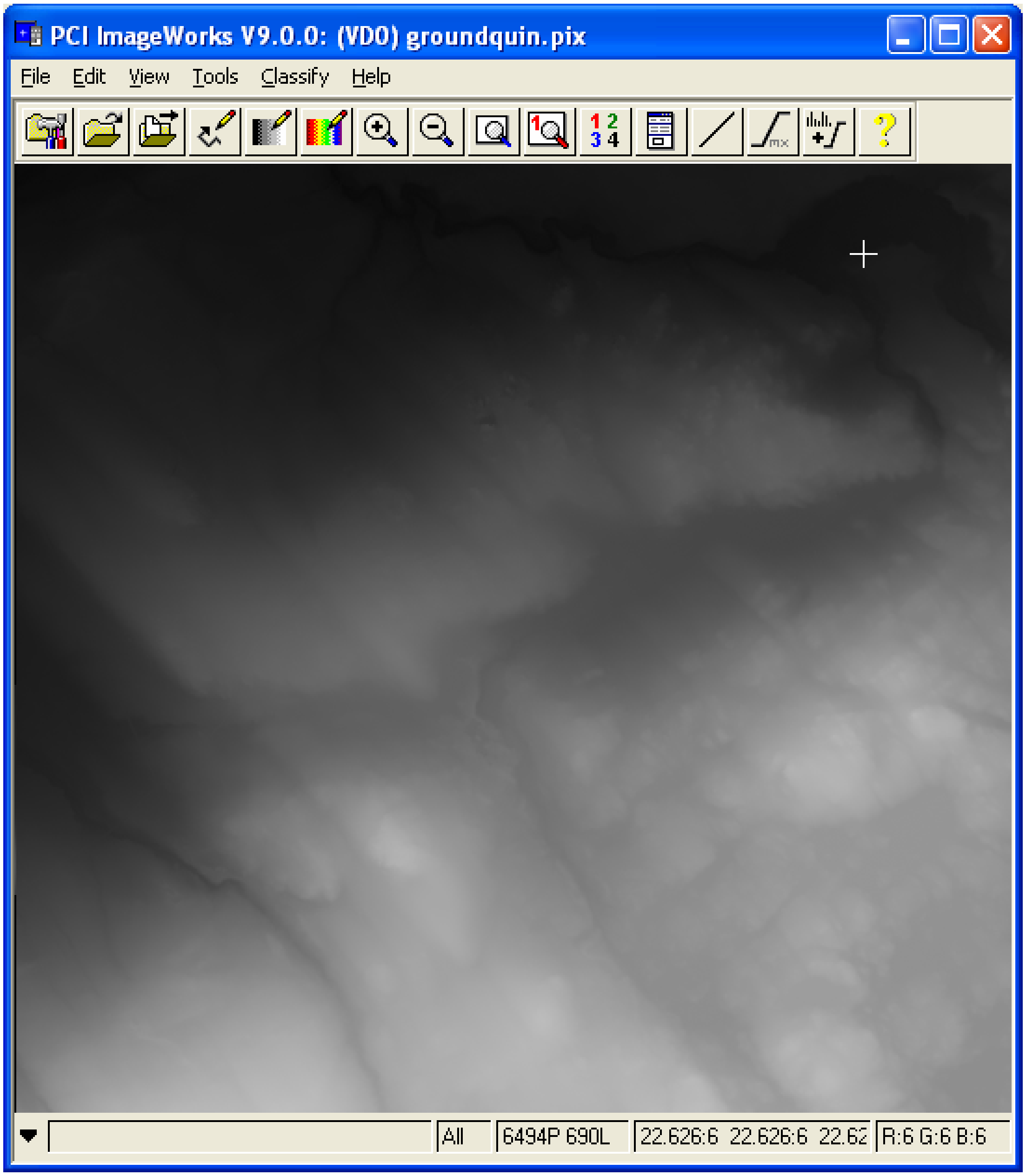


Figure 37 – The 32-bit Digital Elevation Model of the ground surface using the Quintic interpolation method. (Viewed with PCI ImageWorks)

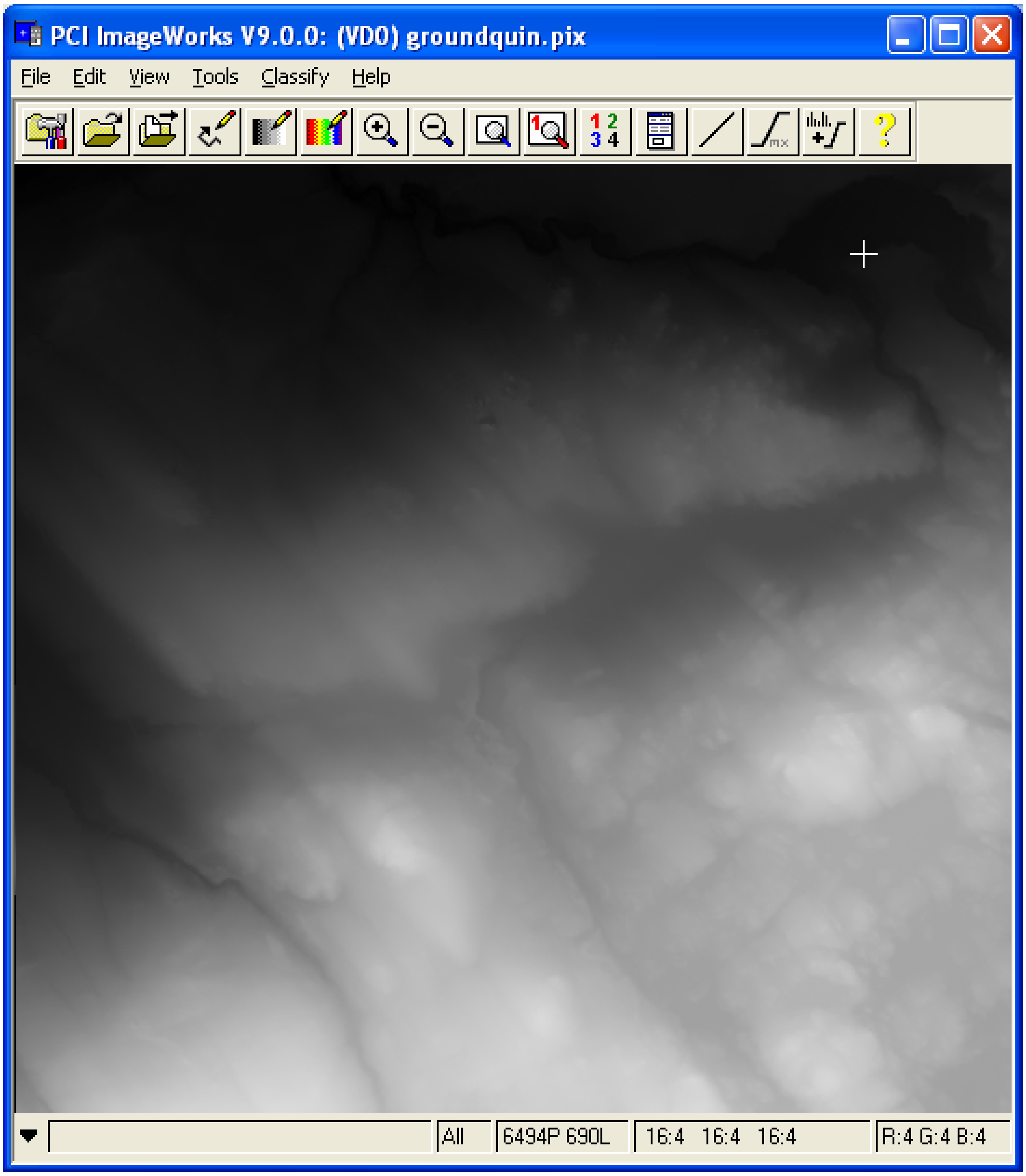


Figure 38 – The 8-bit scaled Digital Elevation Model of the ground surface using the Quintic interpolation method. (Viewed with PCI ImageWorks)

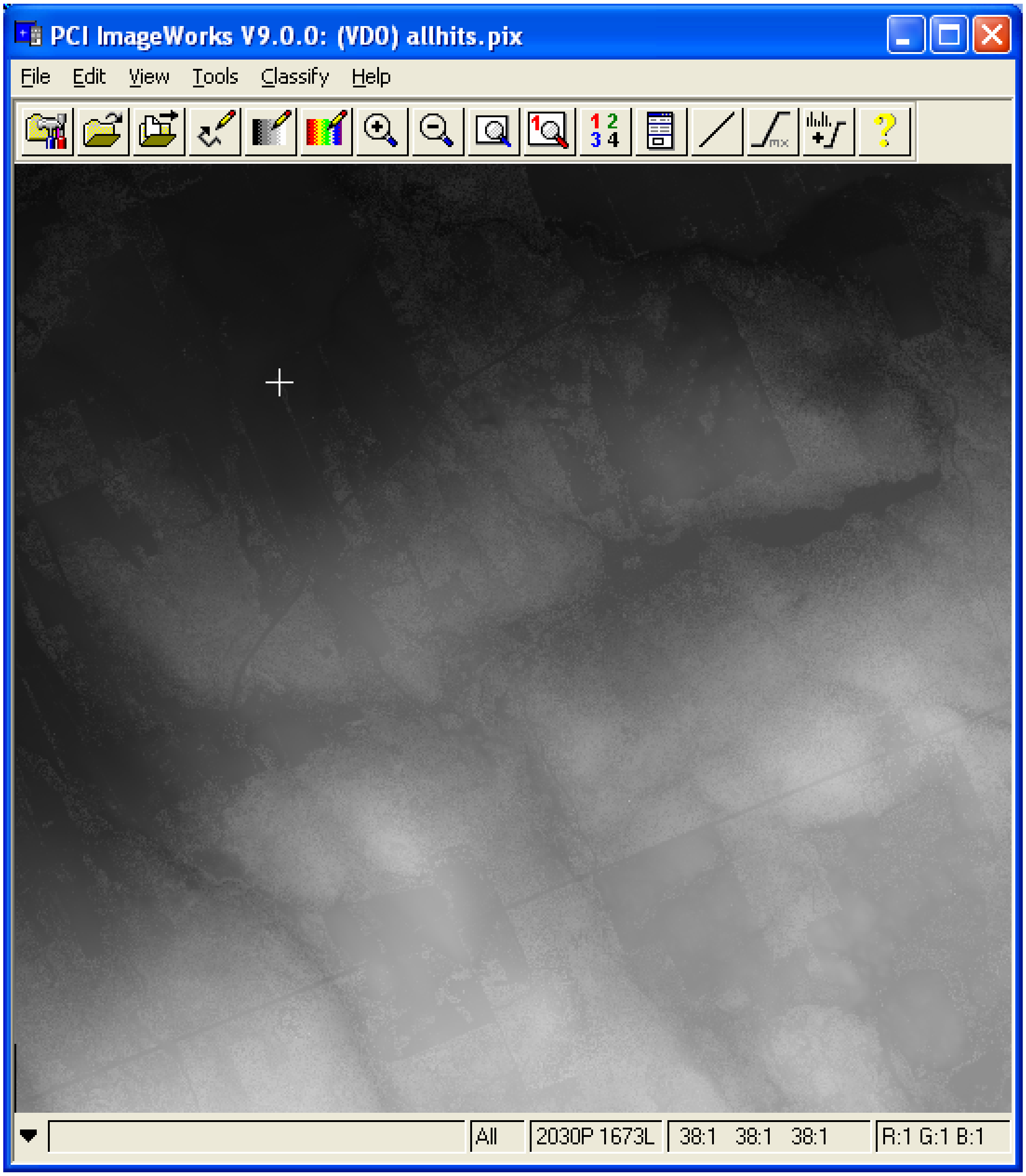


Figure 39 – The all hits Digital Surface Model. (Viewed with PCI ImageWorks)

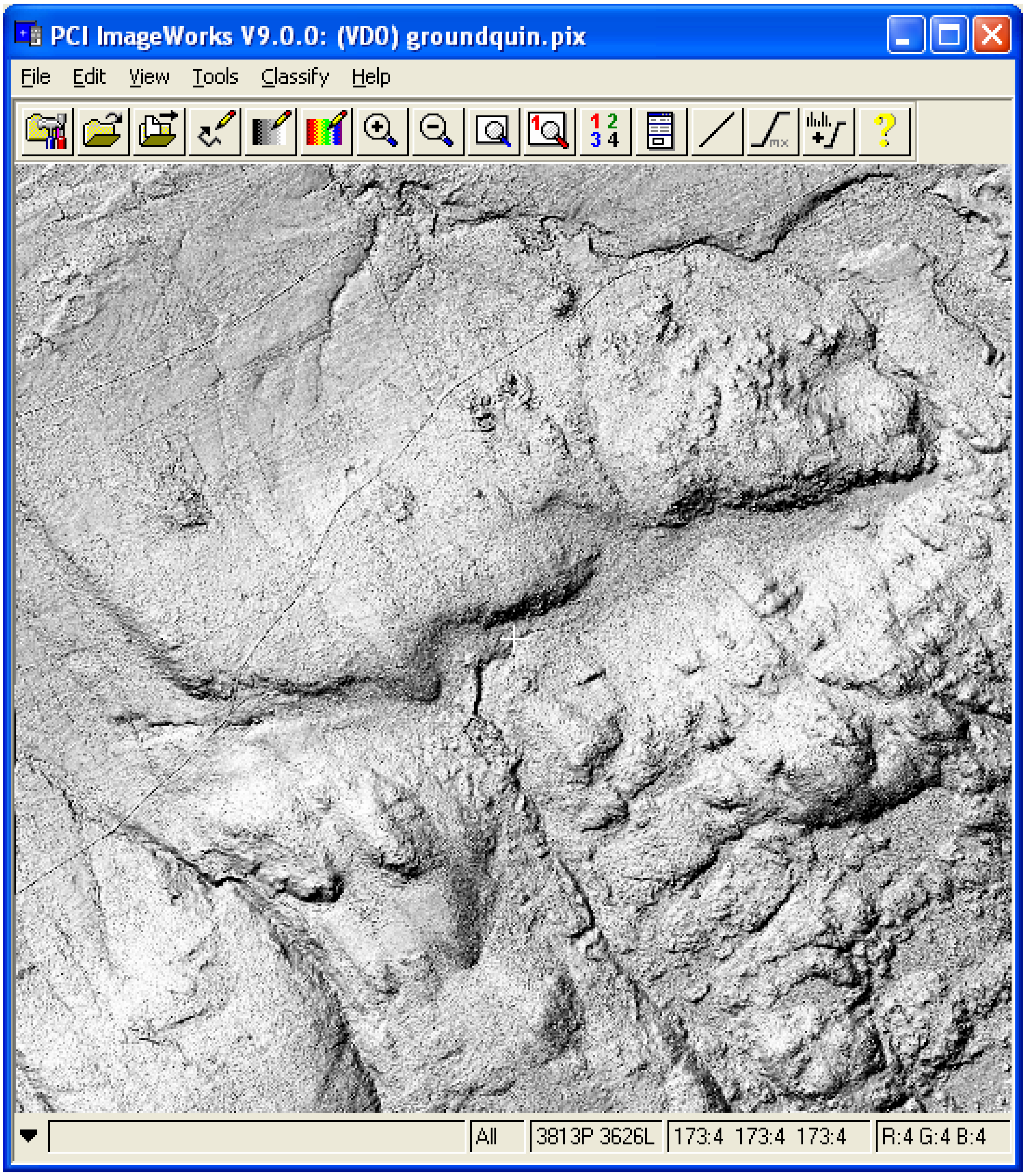


Figure 40 – The Shaded Relief model of the ground surface using the Quintic interpolation method. (Viewed with PCI ImageWorks)

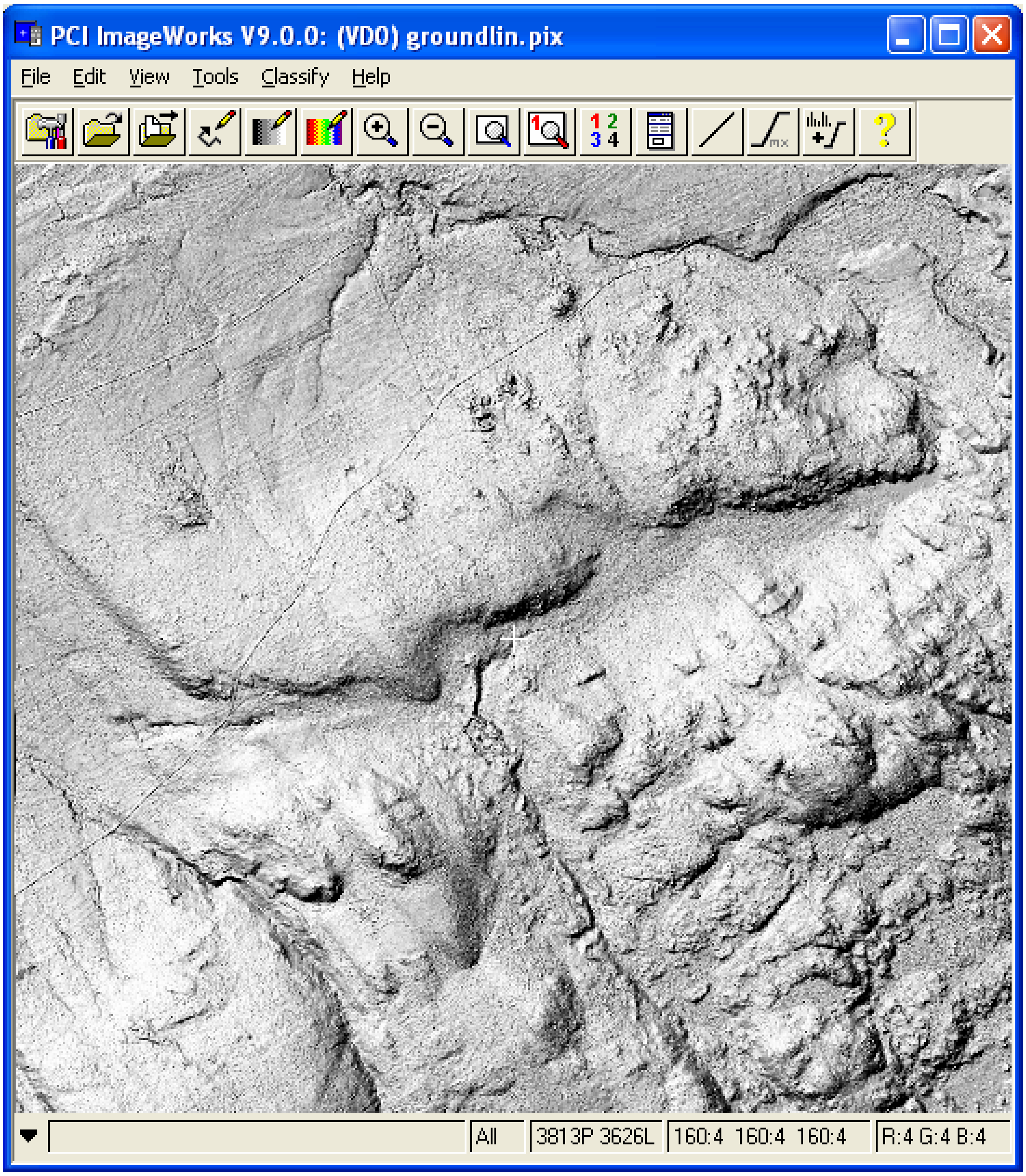


Figure 41 – The Shaded Relief model of the ground surface using the Linear interpolation method. (Viewed with PCI ImageWorks)

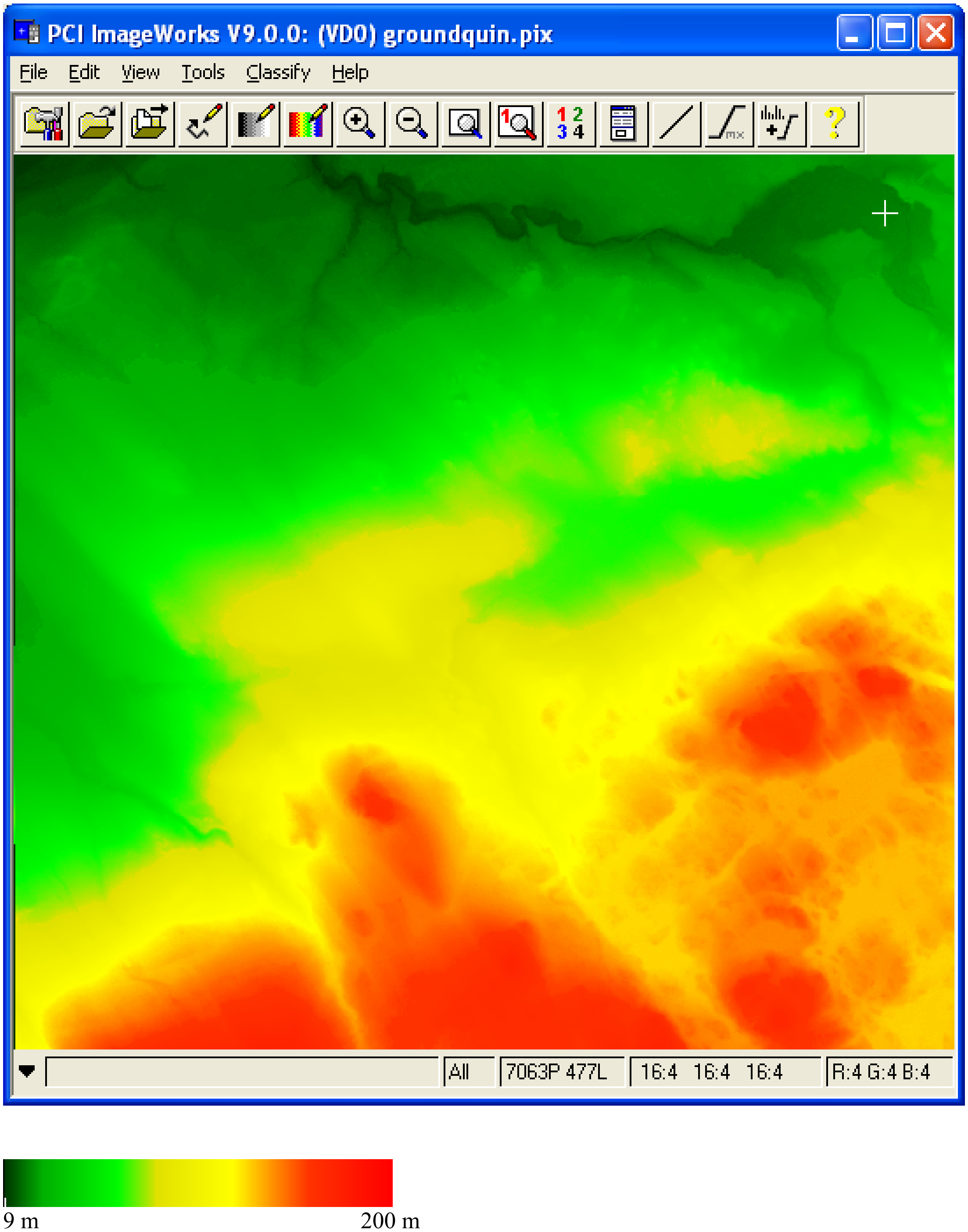


Figure 42 – The RGB composite of the ground surface using the Quintic interpolation method. (Viewed with PCI ImageWorks)

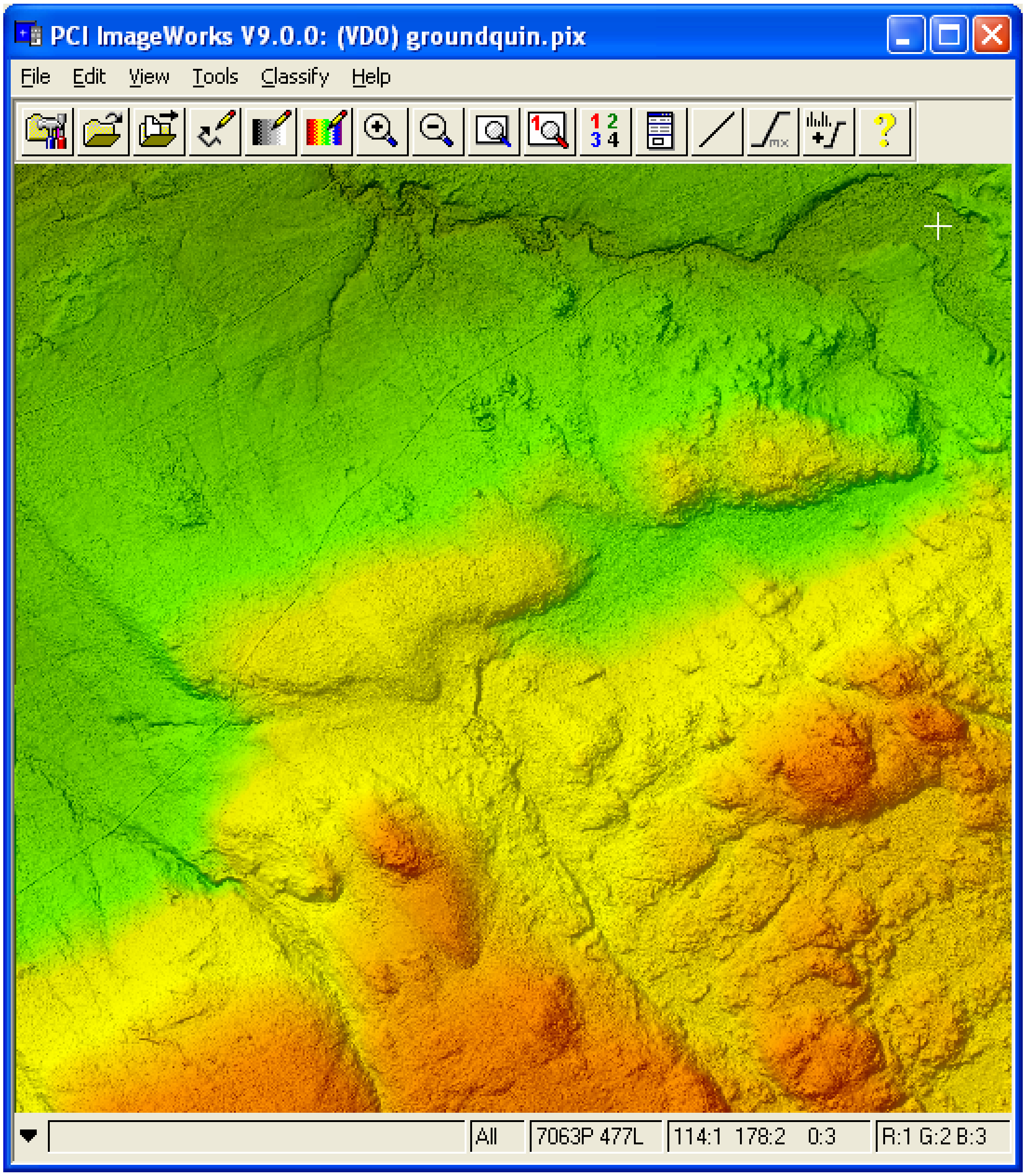


Figure 43 – The Color Shaded Relief model of the ground surface using the Quintic interpolation method. (Viewed with PCI ImageWorks)

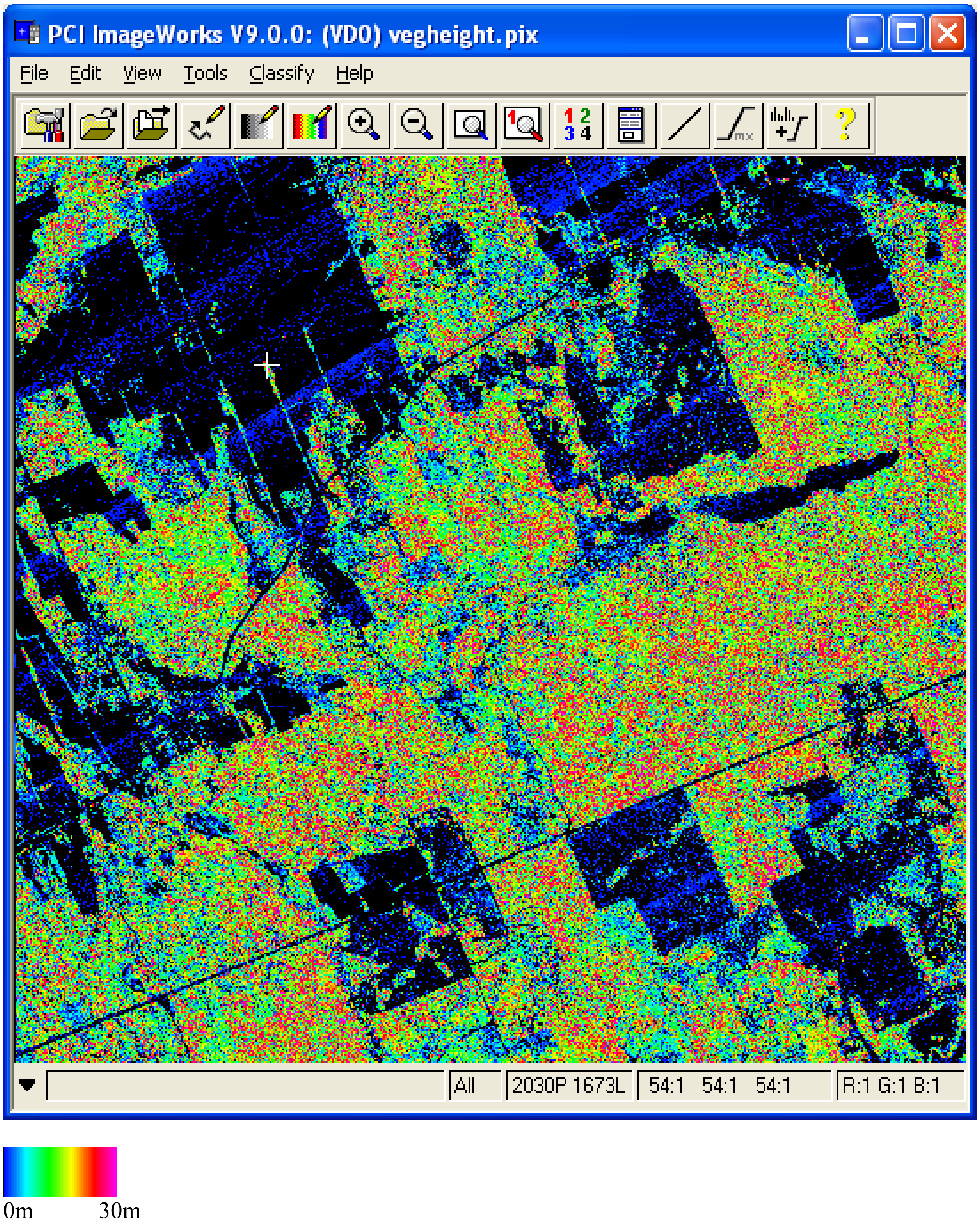


Figure 44 – The Vegetation heights surface model Coded to show Heights of Trees. (Viewed with PCI ImageWorks)