Dewberry
8401 Arlington Boulevard
Fairfax, Virginia 22031-4666
703 849 0396
703 849 0182 fax
www.dewberry.com

FINAL REPORT

QA/QC Assessment of Walton County, Florida LiDAR Dataset Produced by Sanborn Mapping

NOAA CSC Contract EA133C05CQ1048
Task Order No. T007

February 5, 2007

Table of Contents

References:

A — NOAA Coastal Services Center Task Order T007, 11 Mar 2006, Walton County Florida Lidar Validation

B — Part 3: *National Standard for Spatial Data Accuracy (NSSDA)*, "Geospatial Positioning Accuracy Standards," published by the Federal Geographic Data Committee (FGDC), 1998

C — Appendix A, *Guidance for Aerial Mapping and Surveying*, "Guidelines and Specifications for Flood Hazard Mapping Partners," published by the Federal Emergency Management Agency (FEMA), April 2003

D — *Guidelines for Digital Elevation Data*, Version 1.0, published by the National Digital Elevation Program (NDEP), May 10, 2004

E — *ASPRS Guidelines, Vertical Accuracy Reporting for Lidar Data*, published by the American Society for Photogrammetry and Remote Sensing (ASPRS), May 24, 2004

NOAA Guidance

Reference A tasked Dewberry to validate the bare-earth lidar dataset of Walton County, FL, both quantitatively (for accuracy) and qualitatively (for usability). This report primarily addresses the qualitative aspect as the quantitative assessment was addressed in the accompanying report. This Final Report incorporates changes recommended by NOAA is response to the Qualitative Assessment Report submitted by Dewberry on January 30, 2007.

Survey Procedures

To assess the quantitative accuracy of the LiDAR for Walton County, a verification survey was performed. The survey incorporated both GPS and conventional methods to ensure the desired accuracy was met. For logistical and spatial distribution reasons, the project area was further divided into a north and south zones. This allowed us to use NGS control as base stations so that each of our test areas was based on the same control points used by the LiDAR vendor, Sanborn. Since our goal was to establish checkpoints that were approximately three times more accurate than the LiDAR, our target $RMSE_h$ was approximately 5 cm and this dictated our survey approach. As with any survey of this nature, many factors play a role such as cost, accessibility, and limitations of the survey equipment as well as achievable accuracy. Therefore a balance was created in establishing checkpoints using a combination of GPS Fast Static, Real Time Kinematic (RTK) and conventional survey techniques.

The approach that was used is termed the "cluster" survey. This process establishes a series of secondary base stations dispersed throughout the project area which are then used as the basis in which RTK measurement (cluster points) and conventional measurements are derived. Since there is no adequate existing NGS control throughout the County it was imperative that these secondary base stations be established. The location of the secondary base stations are based on many factors such as; location within the county in relationship to other base station locations, land cover type, and proximity to existing NGS control. The process of establishing these secondary base stations were identified using existing aerial photographs from Google Earth. First, an ideal open area was identified that would support a GPS base station with no obstructions, secondly it needed to be within 15 miles of a NGS control monument to maintain the desired accuracy, and thirdly it had to be central within the range of RTK to be able to measure the different land cover types as required; therefore the area also had to have many land cover types near by. After these three criteria were met, fast static surveys were established from the existing NGS control to the secondary base stations. For the northern project area a total of three GPS units were used simultaneously, two on NGS control and the third on the secondary base station. This allowed for a minimum of two vectors into each secondary base. Due to logistical constraints the south was performed similarly except that only two GPS units were used at a time. However each secondary base had two vectors but one vector was measured one day from an NGS control point and a second vector was measured the next day from a second NGS control point.

By utilizing multiple station observations the accuracy was assessed with a least squares adjustment and the results indicated no outliers were apparent. For this survey no repeat observations were required. After each secondary base station was established it was used as the base station for all RTK measurements. The advantage of RTK is the ability to resolve phase ambiguities and achieve centimeter level accuracy with less than 1 minute of observations (at times only a few seconds is required). For the RTK measurements a few factors dictated where they were established. In choosing suitable locations, not only is the correct land cover type important but also the location in relationship to the flight lines. For example, if a line is flown in an east/west direction, two points along the same axis do not truly give a unique check on the data (other than land cover type). Our process involved ensuring we could cover as many different flight lines as possible. Again if the lines were flown east/west our checkpoints were collected in a north/south pattern. Again this allows us to cover more flight lines. For each suitable site, an RTK measurement was taken and a picture taken and documented. Additionally, inter-visible pairs were established to allow the use of conventional equipment to measure points within the forest where GPS would not be suitable. The two points allowed for the instrument setup as well as a back sight. Although there are two points typically in the land cover type grass, these are not considered two check points as they do not meet the required spacing between checkpoints. However they are used as an internal verification of the LiDAR. Again to ensure data integrity, some points were measured twice with the RTK, usually a few days later. If any measurement exceeded our 5 cm criteria it was rejected and a new observation was performed until a comparison of less than 5 cm existed.

The process of verifying our survey is done through a series of checks and balances utilizing multiple measurements, ties into existing control and through least squares adjustments. However the definitive check is how well the LiDAR fits to the checkpoints. It is understood that the survey is to verify the LiDAR but the LiDAR can also be used to verify the survey. If a good comparison is obtained, which was the case for this survey, we can conclude that both the survey and the LiDAR meet specification.

Quantitative Assessment

The findings of the quantitative assessment were favorable and met the desired acceptance criteria (see Table 1 – yellow portion). The major accuracy reporting statistics were as follows:

- **Based on NSSDA, FEMA, NDEP and ASPRS methodology: Tested 25.5 cm vertical accuracy at 95% confidence level in open terrain (29.4 cm criteria).**
- **Based on NSSDA and FEMA methodology: Tested 29.4 cm vertical accuracy at 95% confidence level in all land cover categories combined (36.3 cm criteria).**

The relevant criteria are summarized in Table 1. Criteria in yellow refer to NOAA-specific requirements in Reference A ($RMSE_h$ = 15-cm in open terrain only), whereas criteria in green refer to FEMA requirements in Reference C, but expressed in terminology used by the NDEP and ASPRS in references D and E.

Table 1 — DTM Acceptance Criteria for Walton County

Quantitative Criteria	Measure of Acceptability
$RMSE_h$ = NSSDA vertical accuracy statistic at 68% confidence level	15 cm in open terrain only
$Accuracy_z$ = NSSDA vertical accuracy statistic at the 95% confidence level = $RMSE_h$ x 1.9600	29.4 cm (15 cm $RMSE_h$ x 1.9600) in open terrain only
Fundamental Vertical Accuracy (FVA) in open terrain only = 95% confidence level	36.3 cm (18.5 cm $RMSE_h$ x 1.9600) for open terrain only
Supplemental Vertical Accuracy (SVA) in individual land cover categories = 95% confidence level	36.3 cm (based on 95^{th} percentile per category; this is a target value only, not mandatory)
Consolidated Vertical Accuracy (CVA) in all land cover categories combined = 95% confidence level	36.3 (based on combined 95^{th} percentile)

Figure 1 displays a map of the distribution of the errors found in the quantitative analysis. The larger errors are red and orange, and the smaller errors are yellow and green

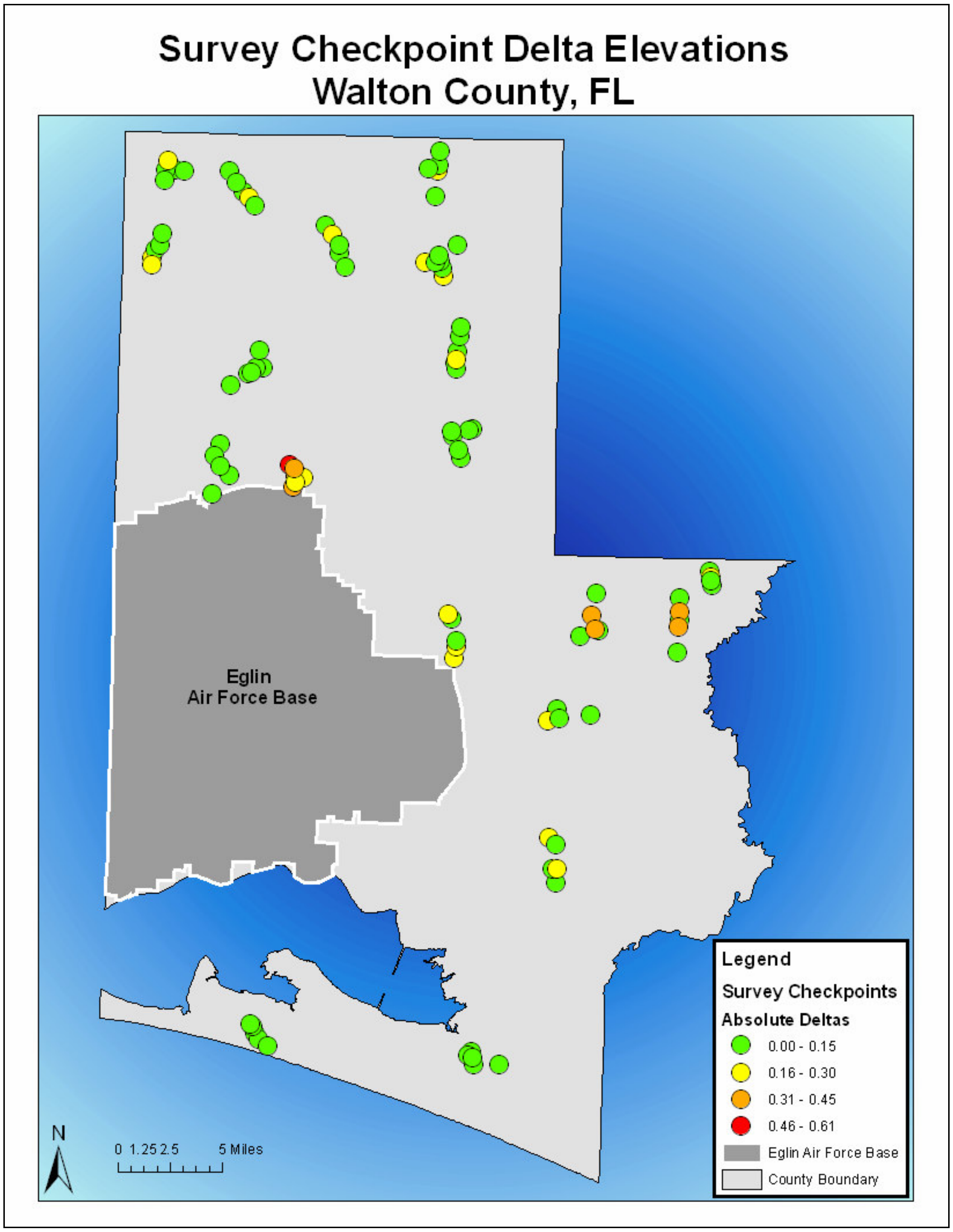


Figure 1: Walton County map with distribution of errors between the lidar data and the survey checkpoints.

Qualitative Assessment

Dewberry's qualitative assessment utilizes an interpretive and statistical based methodology to assess the quality of the data for a bare-earth digital terrain model (DTM). This process looks for anomalies in the data and also identifies areas where man-made structures or vegetation points may not have been classified properly to produce a bare-earth model. Overall the data are of good quality and should satisfy most users for an accurate bare-earth elevation data product.

Overview

Within this review of the lidar data, two fundamental questions were addressed:

- Did the lidar system perform to specifications?
- Did the vegetation removal process yield desirable results for the intended bare-earth terrain product?

Mapping standards today address the quality of data by quantitative methods. If the data are tested and found to be within the desired accuracy standard, then the data set is typically accepted. Now with the proliferation of lidar, new issues arise due to the vast amount of data. Unlike photogrammetry where point spacing can be eight meters or more, lidar point spacing for this project is two meters or less. The end result is that millions of elevation points are measured to a level of accuracy previously unseen for elevation technologies, and vegetated areas are measured that would be nearly impossible to survey by other means. The downside is that with millions of points, the data set is statistically bound to have some errors both in the measurement process and in the vegetation removal process.

As previously stated, the quantitative analysis addresses the quality of the data based on absolute accuracy. This accuracy is directly tied to the comparison of the discreet measurement of the survey checkpoints and that of the interpolated value within the three closest lidar points that constitute the vertices of a three-dimensional triangular face of the TIN. Therefore, the end result is that only a small sample of the lidar data is actually tested. However there is an increased level of confidence with lidar data due to the relative accuracy. This relative accuracy in turn is based on how well one lidar point "fits" in comparison to the next contiguous lidar measurement. Once the absolute and relative accuracy has been ascertained, the next stage is to address the cleanliness of the data for a bare-earth DTM.

By using survey checkpoints to compare the data, the absolute accuracy is verified, but this also allows us to understand if the vegetation removal process was performed correctly. To reiterate the quantitative approach, if the lidar operated correctly in open terrain areas, then it most likely operated correctly in the vegetated areas. This does not mean that the bare-earth was measured, but that the elevations surveyed are most likely accurate (including elevations of treetops, rooftops, etc.). In the event that the lidar pulse filtered through the vegetation and was able to measure the true surface (as well as measurements on the surrounding vegetation) then the level of accuracy of the vegetation removal process can be tested as a by-product.

To fully address the data for overall accuracy and quality, the level of cleanliness is paramount. Since there are currently no effective automated testing procedures to measure cleanliness, Dewberry employs a visualization process. This includes creating pseudo image products such as hillshades and 3-dimensional models. By creating multiple images and using overlay techniques, not only can potential errors be found, but we can also find where the data meets and exceeds expectations. This report will present representative examples where the lidar and post processing had issues as well as examples of where the lidar performed well.

Analysis

Process

Dewberry reviewed approximately 25% of the tiles at a detailed level. This entailed looking at each tile individually and examining it for any gross anomalies or blunders. If the tile had any questionable findings it was then reviewed at the micro level. Our effort concentrated on the coastlines to ensure homogenous data in these critical areas, but tiles were also selected based on a semi-random approach to ensure good distribution. An atomic waste dump within the extents of Eglin Air Force Base prevented lidar acquisition at this location and consequently no lidar data was delivered for this area. The resulting gap in the UTM tile index as well as the tiles that were reviewed can be seen in Figure 2.

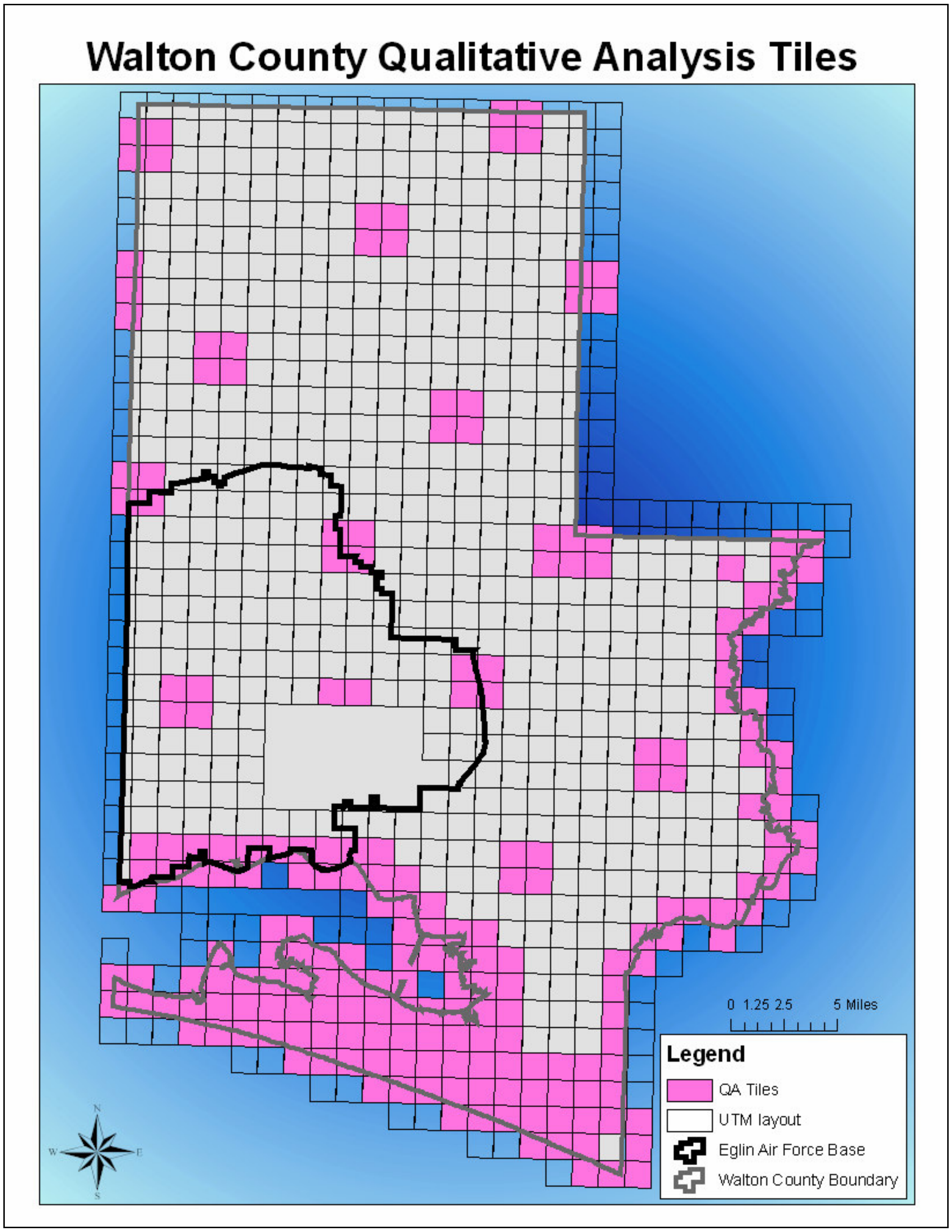


Figure 2: Location of QA tile review. Missing tiles within Eglin Air Force Base were not included in deliverable.

The process of identifying issues utilized two different software packages; ESRI and QT Modeler. Each package has a strength that the other does not possess. However the majority of analysis was performed with QT Modeler due to its robust speed and power for viewing large amounts of data. To begin, different iterations of 3-D models are created. Since the main objective is to review the bare-earth data, these models are built first. The LAS data is converted to both a surface model (similar to a raster lattice model) and a vector bare-earth point cloud. The raster model is unique in that it builds not only the DEM portion, but combines a hillshade with it. This is invaluable for reviewing the surface models as it is more robust at finding anomalies. The full point cloud data including intensity is also converted to both surface and point cloud models. The advantage of using intensity with full point cloud is that it provides a pseudo 3-D image similar to an aerial photograph that can be used as a reference. But the intensity also provides a surface roughness image that helps us classify the terrain type and to find anomalies with the sensor. Other models are also created that look at the density if required.

Our process always started with the bare-earth surface model and further reviewed by overlaying the bare-earth points. Each model is analyzed statistically by examining the minimum and maximum elevation values as well as point density. To reiterate, the data for the most part is exceptionally good. It exhibits excellent accuracy and vegetation classification, providing a good bare-earth data product.

Issues Found

The lidar data contained sporadic issues such as artifacts or small anomalies which is typical of any lidar dataset. However one issue that was repeatedly found throughout the dataset is a phenomenon we call "edge match" issues. Edge match issues are caused by defining a seam line between adjacent areas and removing points from overlapping scans. This in effect joins one scan to the next with no overlapping areas while also reducing noise. Some processing techniques include removing points from both scans along the defined seam line. The edge matching issues found in this dataset seem to have been a result of a different methodology. It appears that overlap was reduced between flight lines by removing points from one scan along a defined edge. For example instead of having an overlap of 30% it has been reduced to 10%. However because there is a slight offset between flight lines, this processing technique produces what appears to be a seam line shift where one dataset is higher than the other at the edge of the reduced overlap. In most cases the data is of excellent quality and the adjacent data matches quite well, but there are areas where it does not match to the same high level of accuracy. Figure 3 illustrates a DEM that includes an edge match issue where it is easy to see the offset between the center flight line and the neighboring flight lines to the north and south. Figure 4 illustrates the same tile but with the mass points visible revealing the edge where the two scans were merged. Figure 5 illustrates a cross section perpendicular to the seam line. In this example the offset is approximately 12 cm which is the average offset difference found in the data.

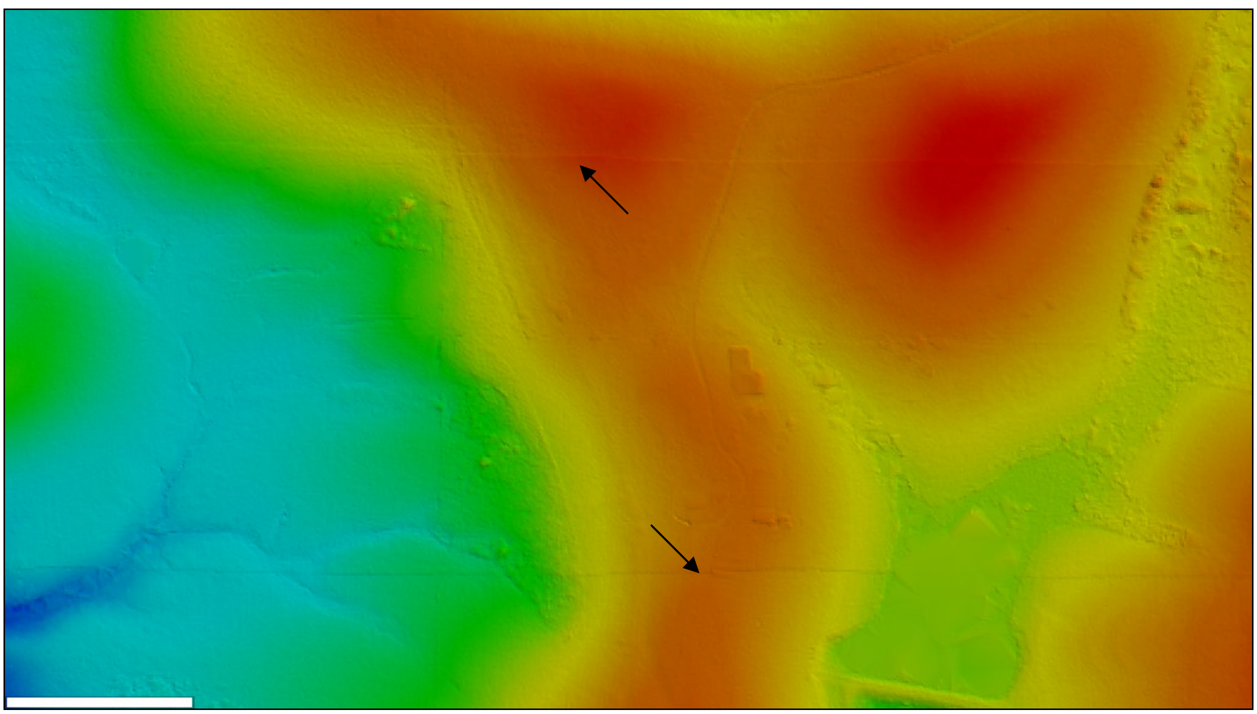


Figure 3: DEM (Tile 509) color coded by elevation illustrating the height offset between adjacent flight lines.

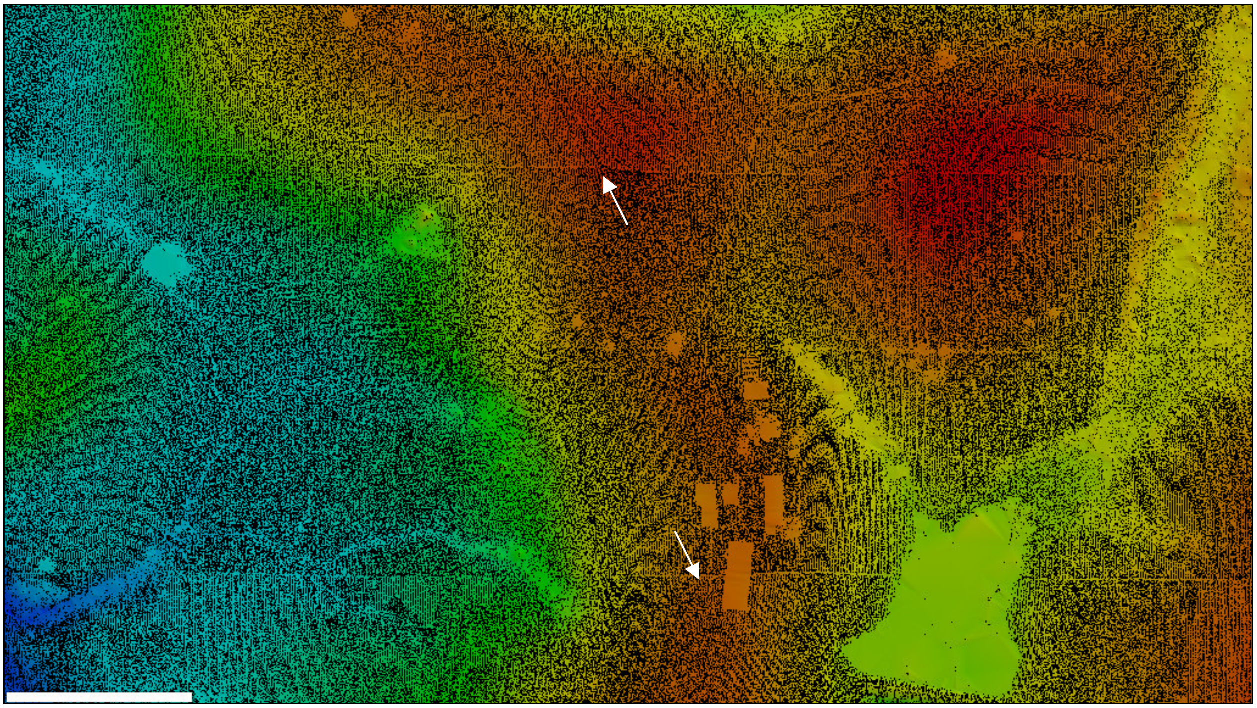


Figure 4: DEM and mass points (Tile 509) illustrating edge of flight lines.

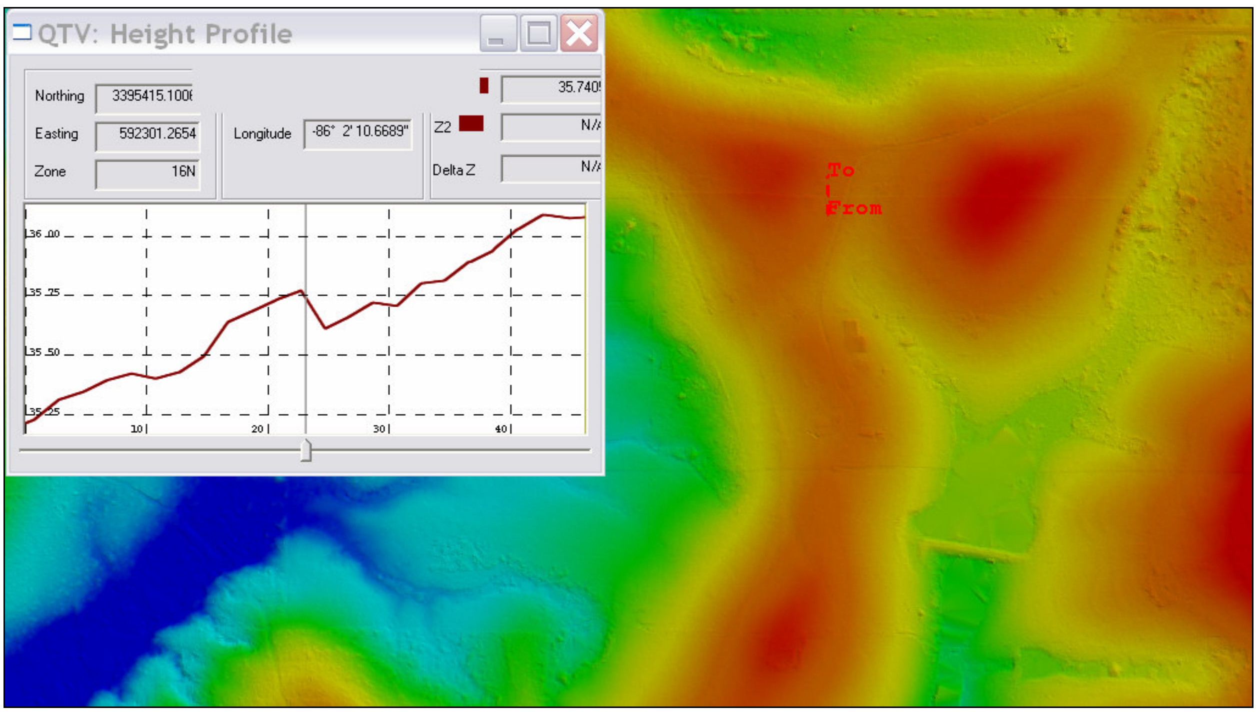


Figure 5: DEM (Tile 509) and cross-section displaying elevation offset.

Although similar edge matching issues were found sporadically throughout the dataset and in some cases across multiple tiles (Figure 8), this does not indicate that the data are of poor quality. An offset of 12 cm is small enough to retain good relative accuracy between flight lines and the checkpoint survey validated the vertical accuracy. There were a few instances however where the offset was as high as 50 cm (Figures 6 and 7) and further investigation is recommended.

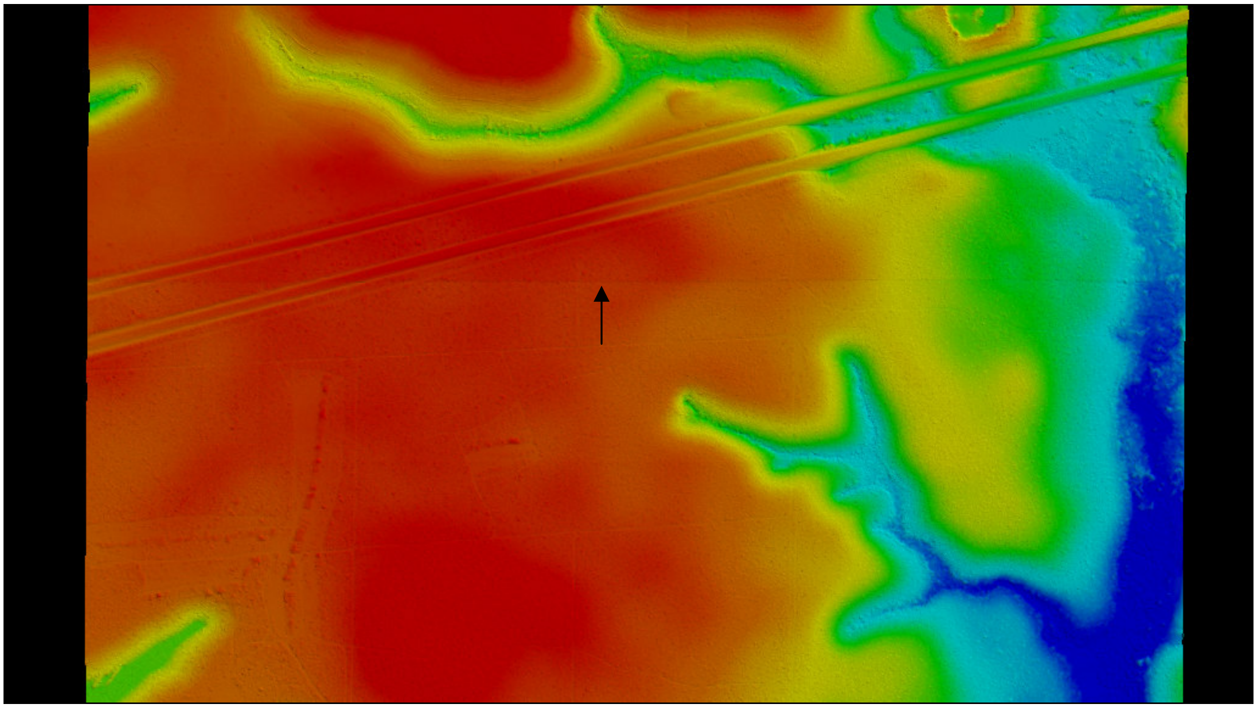


Figure 6: DEM (Tile 547) displaying edge match issue.

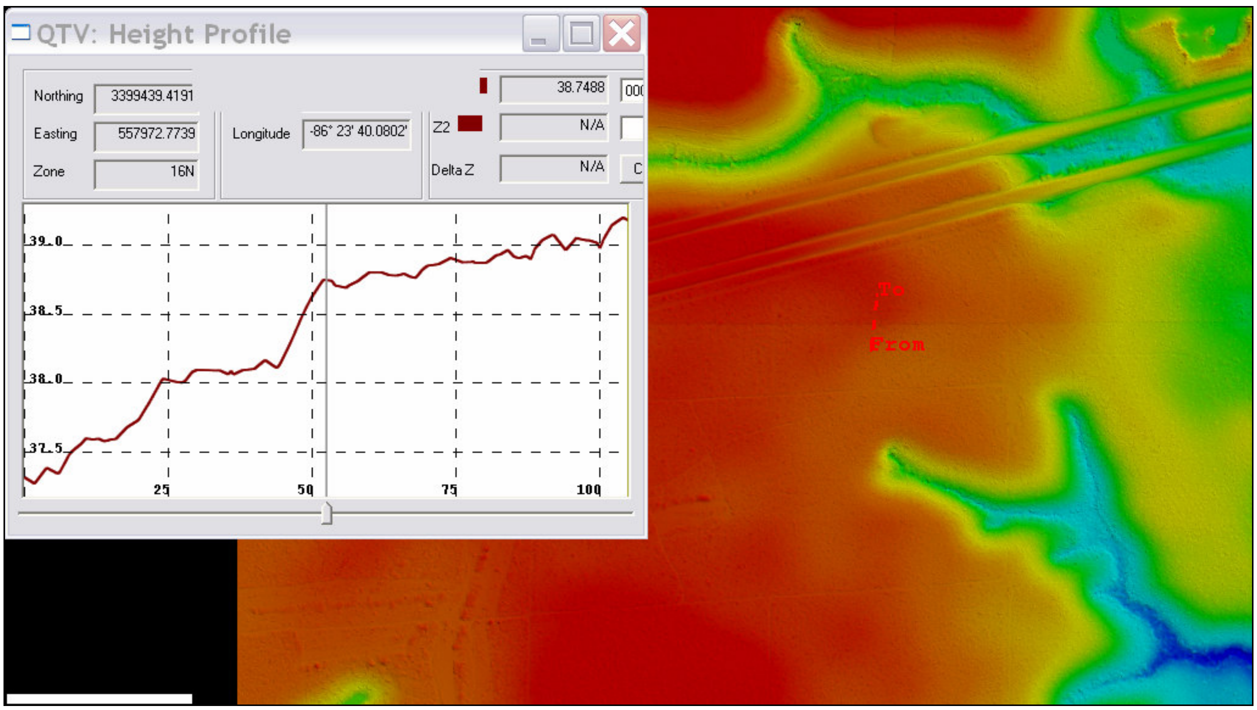


Figure 7: DEM (Tile 547) and cross section revealing offset distance of greater than 50 cm.

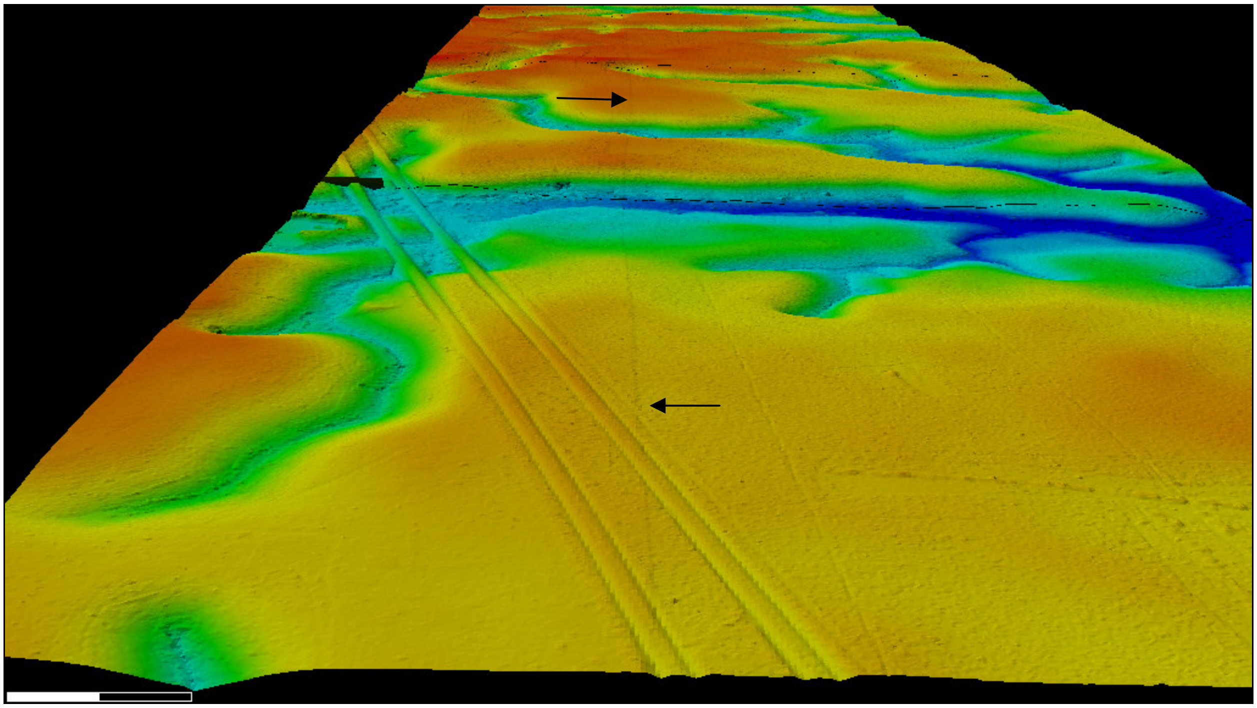


Figure 8: Visible offset across multiple tiles.

Another issue that was found intermittently throughout the project area is a problem called the "cornrow effect". There are a few different schools of thought on what causes this phenomenon but one of the most common is a problem with the sensor itself. It is believed that the sensor loses sync during turbulent flight conditions causing the forward and backward rotations of the mirror to be at two different elevations, producing a high-to-low furrow effect resembling crops in a field. This affect can also be found in areas of overlap between two adjacent flight lines that have slightly different vertical elevations typically caused by a weak GPS solution. Although this phenomenon is found frequently throughout the dataset, the differences in elevation are not large enough to have a significant effect on any anticipated modeling. Figures 9, 10, and 11 display examples of cornrows found within the data.

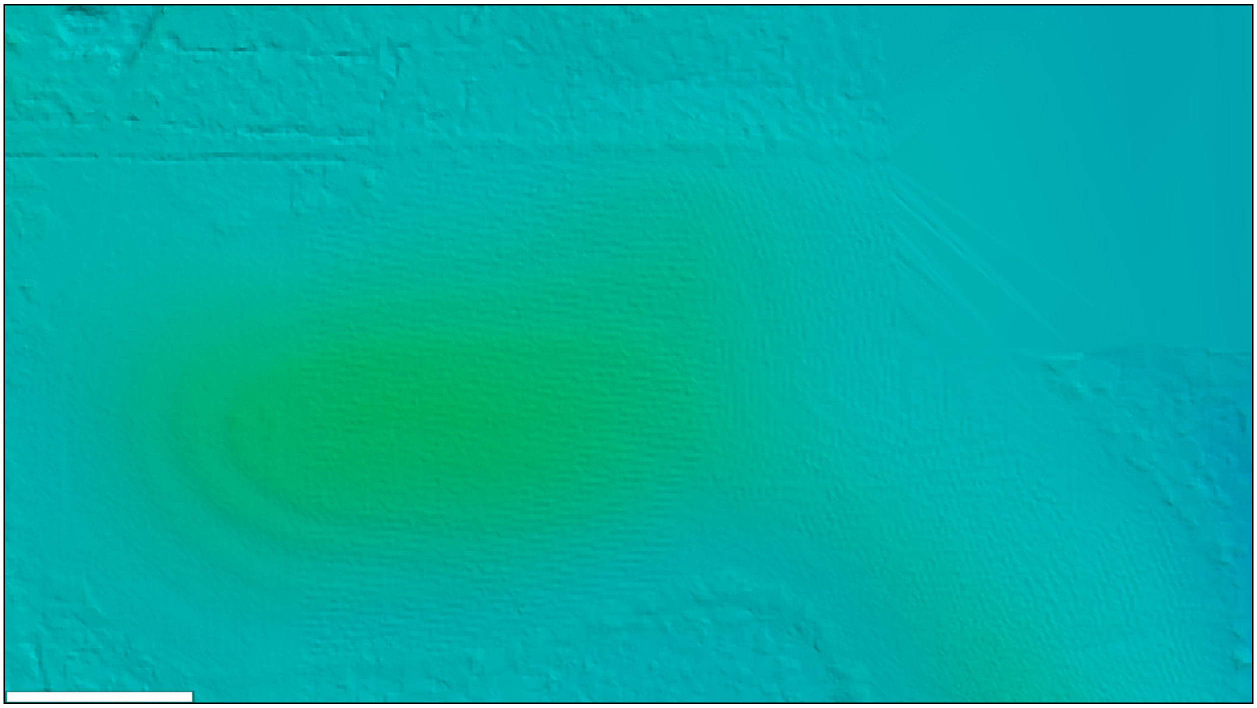


Figure 9: DEM (Tile 726) illustrating "Cornrow Effect"

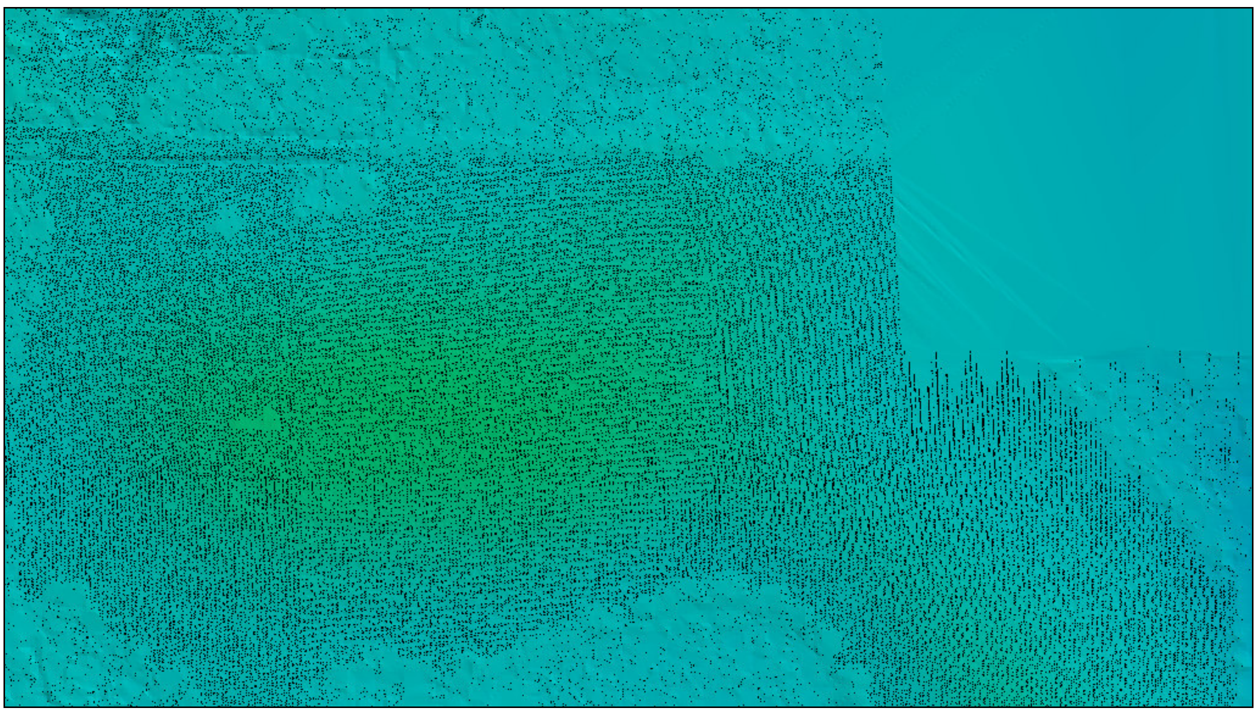


Figure 10: Cornrows with corresponding points. By overlaying the points, it is easy to see the scanning pattern of the mirror. This example shows an area where multiple flight lines converge hence the multiple scan directions.

Dewberry

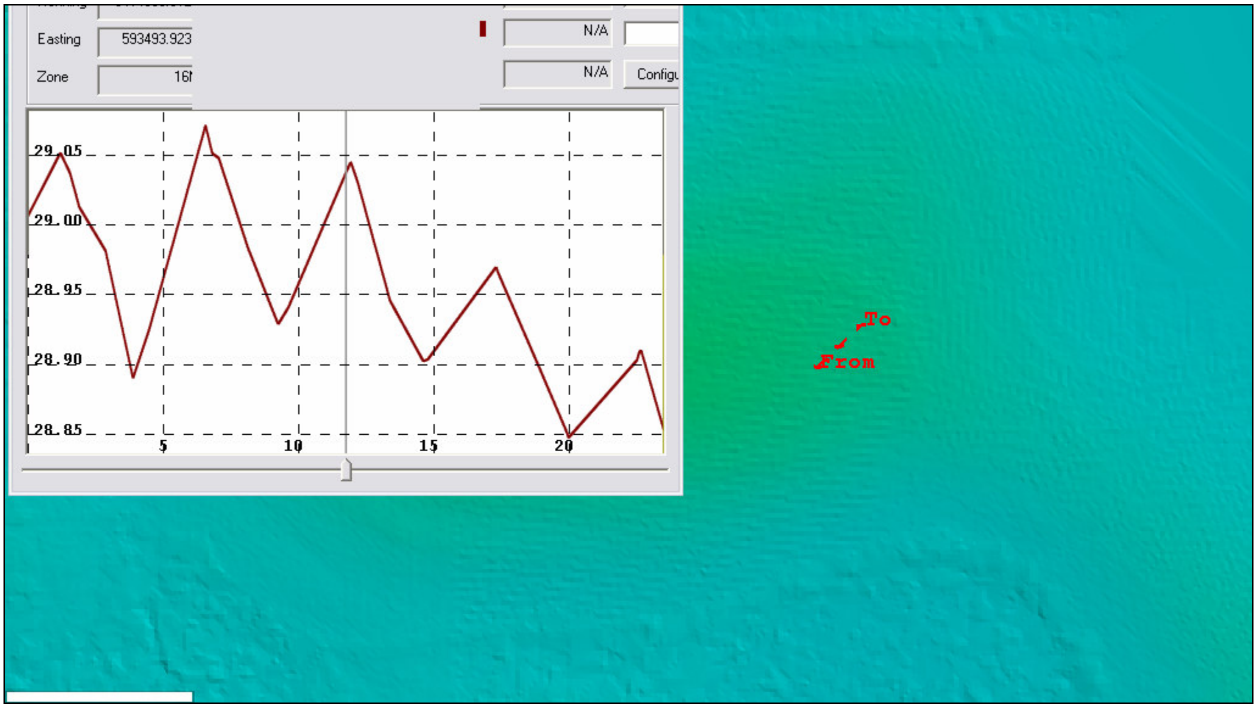


Figure 11: DEM (Tile 726) with cross section drawn across corn rows. The elevation differences vary between 6 and 17 cm.

Overall the level of cleanliness for this product is excellent and meets the guideline for lidar to be generally clean of artifacts. At times though the classification process can be in error and minor issues are presented. There were a few sporadic instances where buildings were not completely removed during the classification process and sharp edges could be seen in the data (see Appendix A). One issue that was found throughout the dataset was the presence of divots. These are caused by one or more "low" points that were left in during the classification process, producing what looks like a hole in the terrain (Figures 12, 13, and 14). There were quite a few divots found in the data most of which were less than 1.5 meters deep. Although considered common for certain types of sensors, these anomalies should be re-examined based upon the scale of the analysis performed on the area in question. While these divots will not have a significant effect on the overall quality of the data, they could affect small scale analysis.

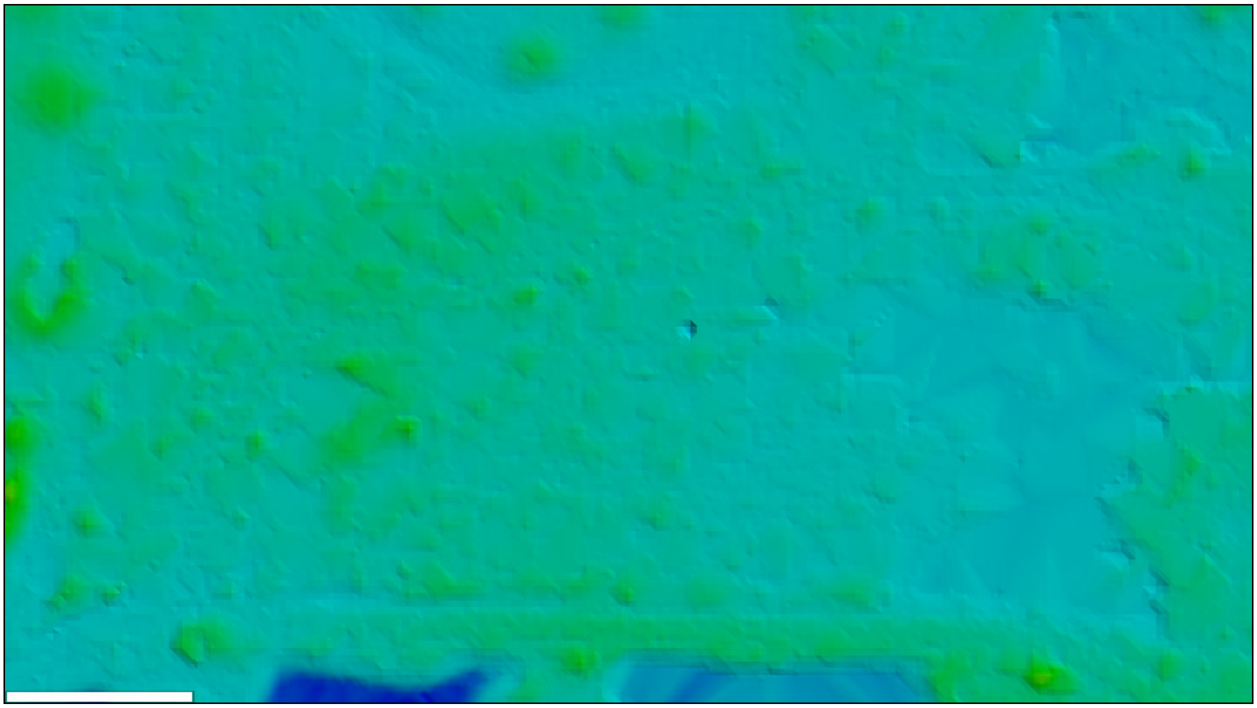


Figure 12: DEM (Tile 75) illustrating divot.

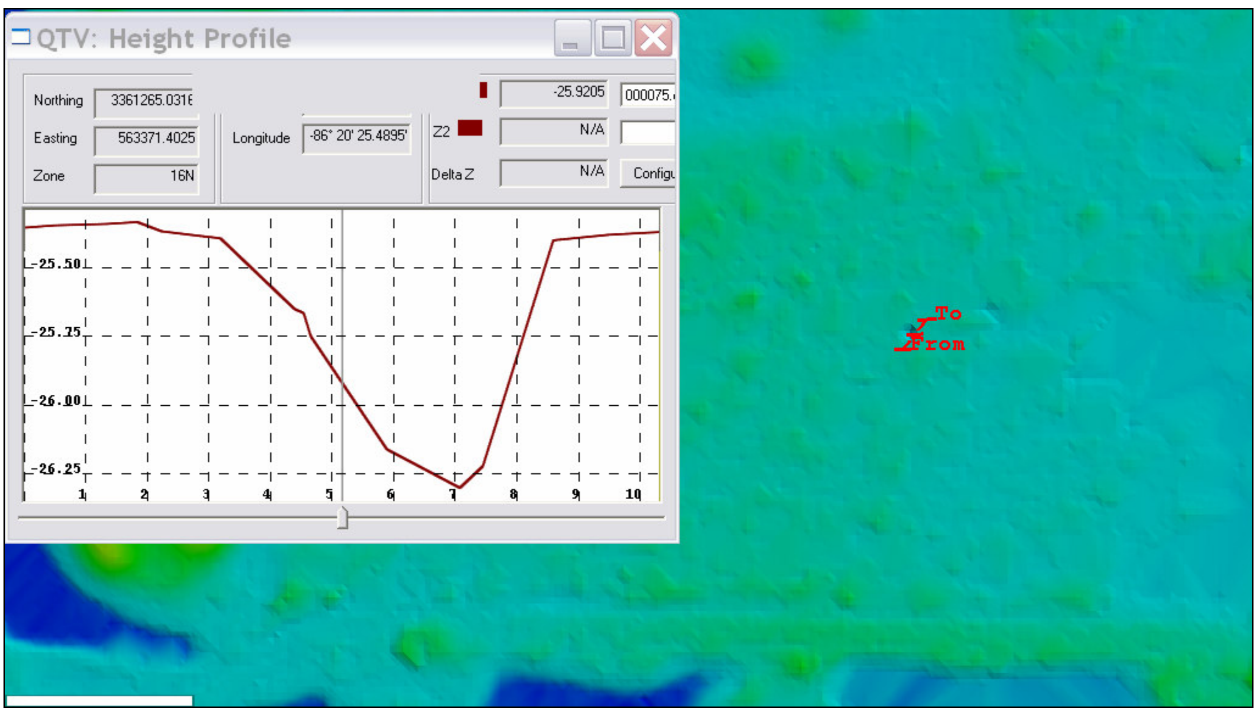


Figure 13: DEM (Tile 75) illustrating cross section of divot. The graph shows that this divot is approximately 75 cm deep.

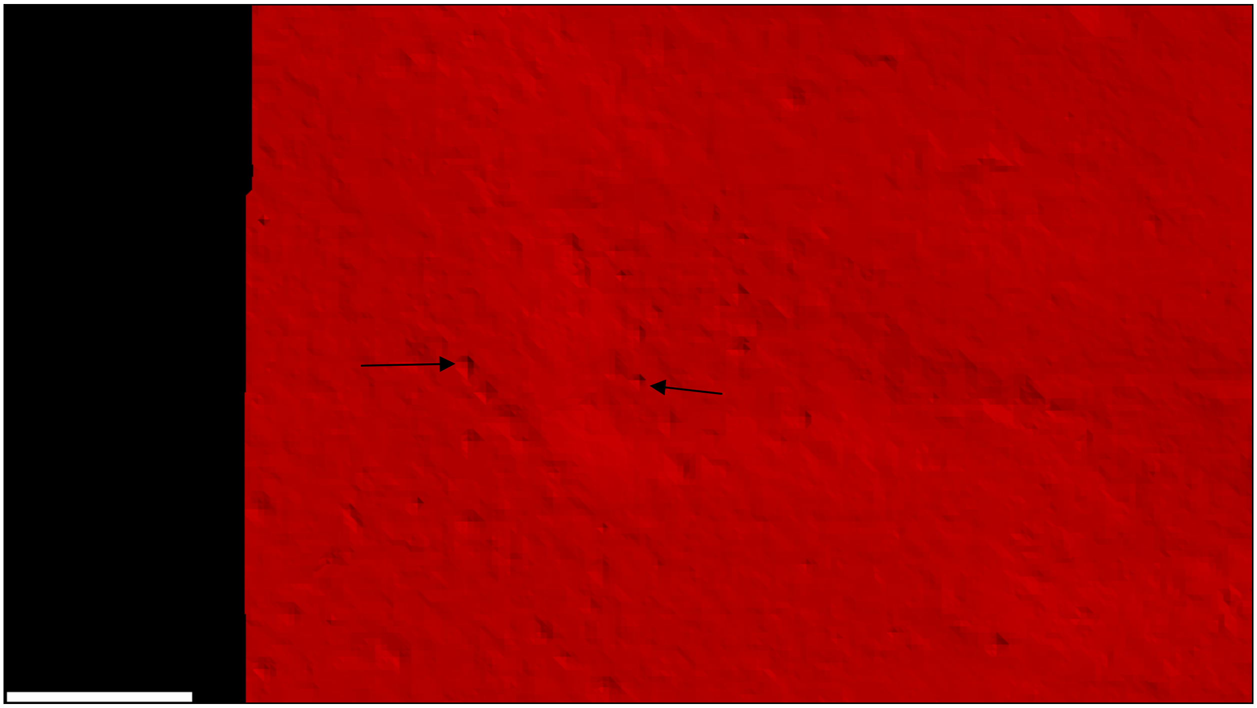
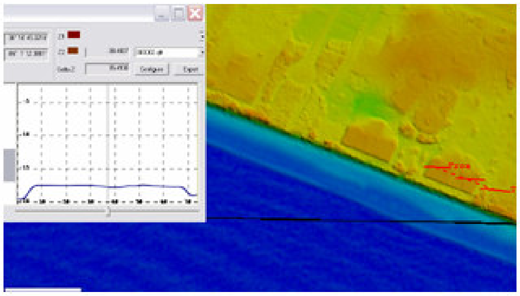


Figure 14: DEM (Tile 151) depicting multiple divots.

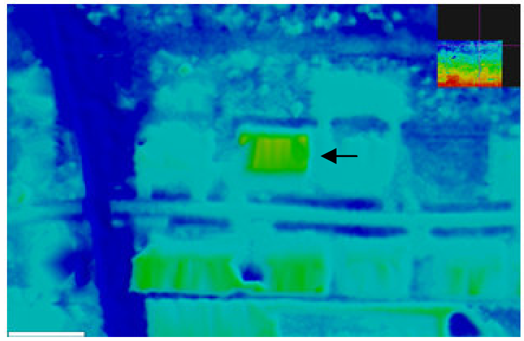
Conclusions

Overall the data are of excellent quality. The processing performed exceptionally well given the low relief and swampy terrain. No major issues were found with this data that makes it unusable and most issues have minimal impact. For example the cornrow effect is within accuracy specifications, so although it ideally would be best to minimize this phenomenon it is also a by-product of lidar data. As for the divot issue, some may be legitimate while others are not. Regardless, they have minimal effect on the overall quality of the dataset. Our review found that one point in a tile of 700,000 points can cause a divot which in the grand scheme has minimal impact considering the ratio of good to bad points. Therefore having twenty tiles with divots is not a major issue as users can easily edit the data to fit their needs if required. The edge match issue could potentially be reviewed due to the fact that it is so widespread although the offset distances are not significant enough to render the data unsuitable for 2 foot contours. Additional examples of these findings are illustrated in Appendix A.

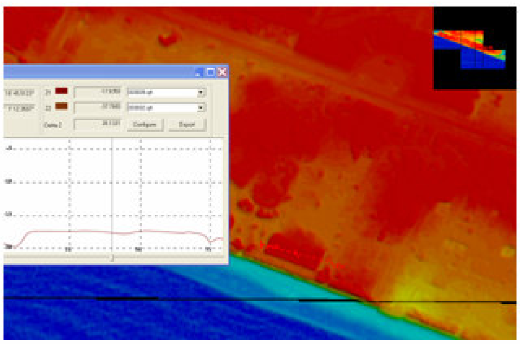
Appendix A – Potential issues identified during the QA review.



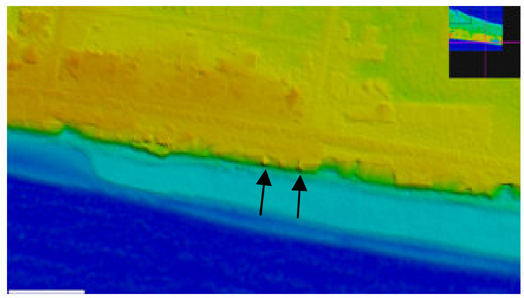
building_09_xs.bmp



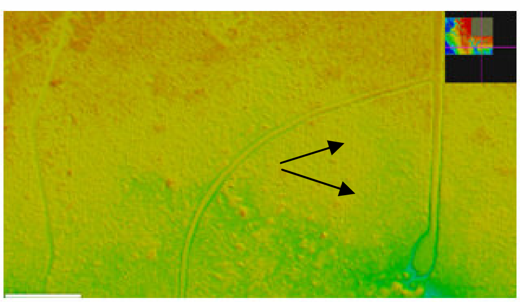
building_80.bmp



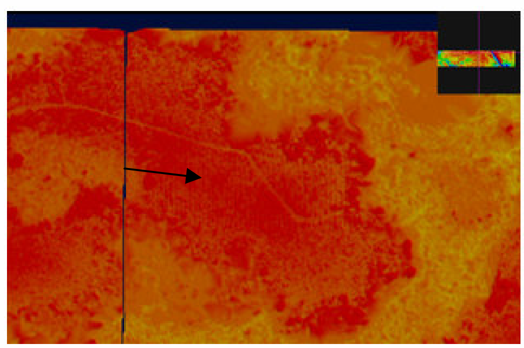
buildings_left_9.bmp



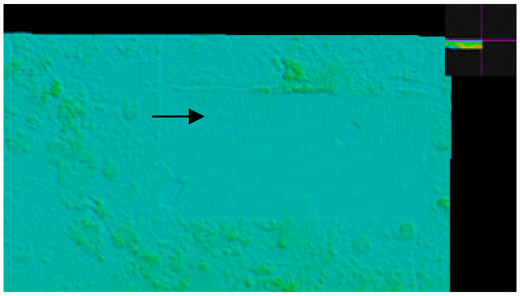
corners_of_houses_left_in_73.bmp



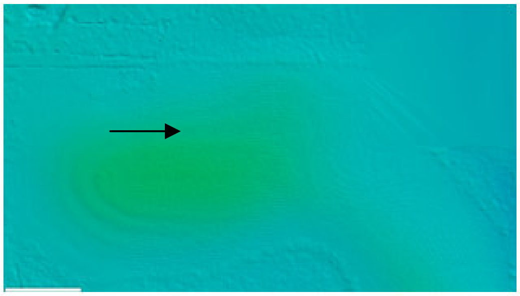
cornrows_187.bmp



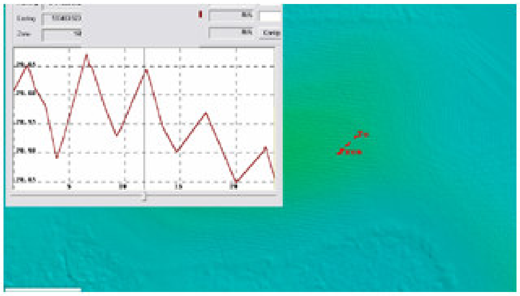
cornrows_32.bmp



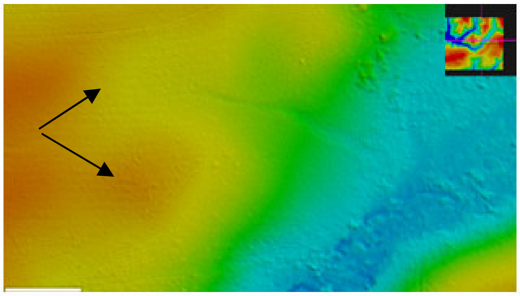
cornrows_60.bmp



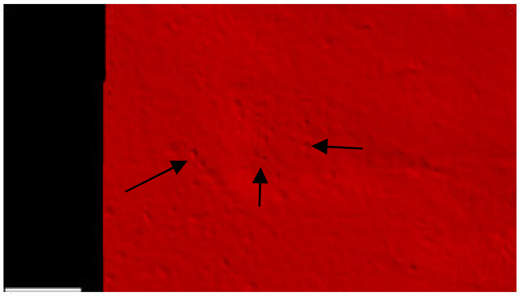
cornrows_726.bmp



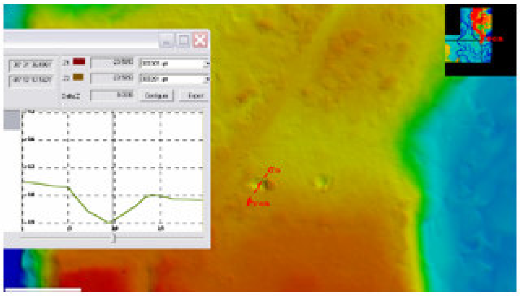
cornrows_726_xs.bmp



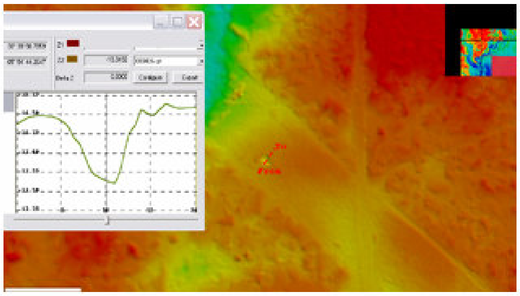
cornrows_scanlines_674.bmp



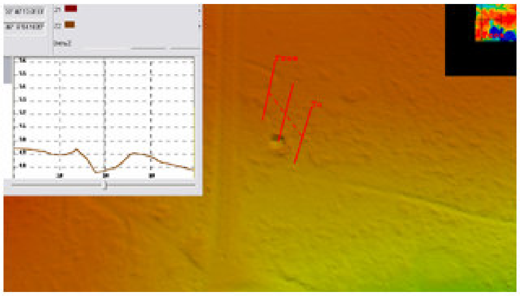
divot_151.bmp



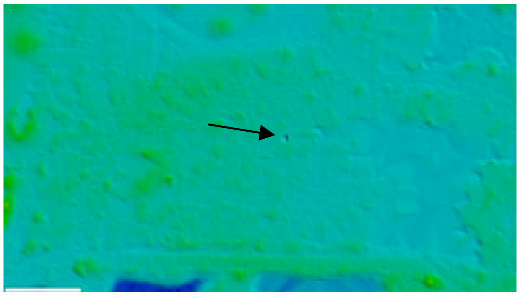
divot_282_xs.bmp



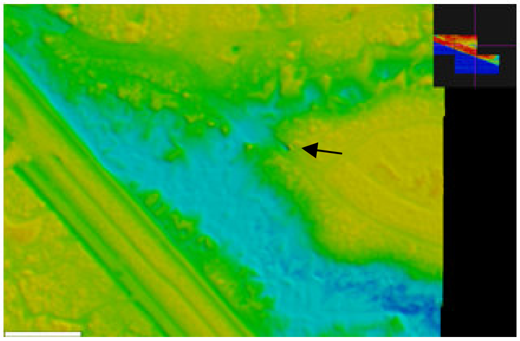
divot_461_xs_2.bmp



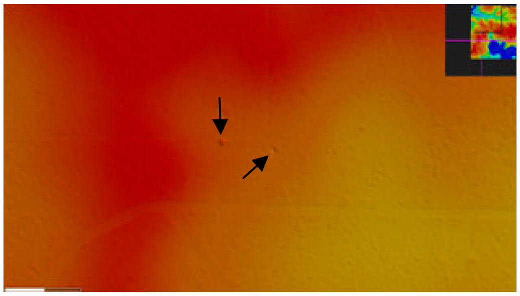
divot_625.bmp



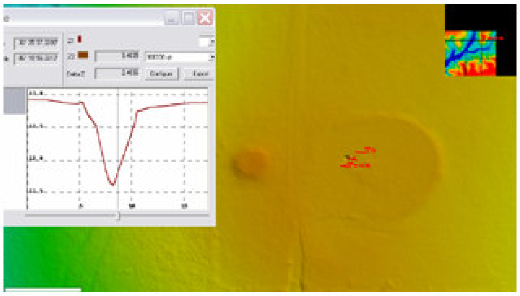
divot_75.bmp



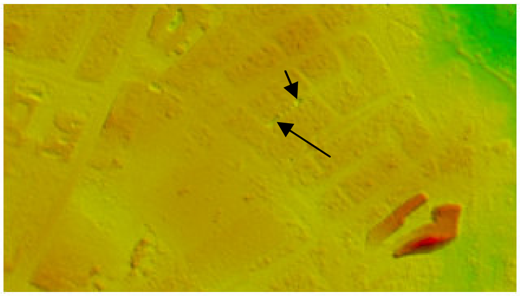
divot_in_road_9.bmp



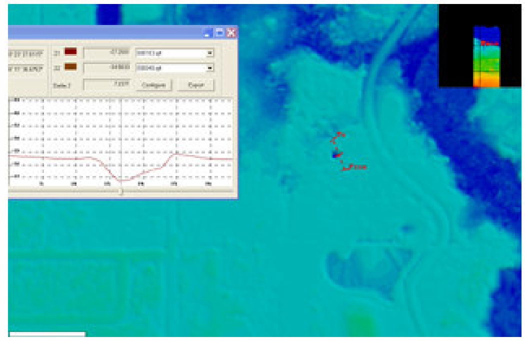
divot_mound_625.bmp



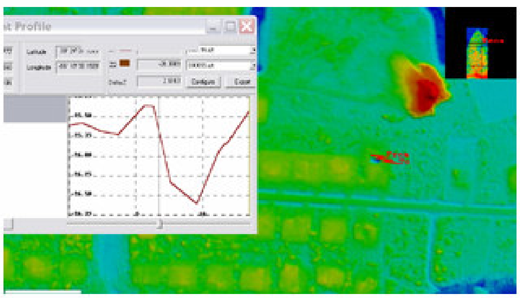
divot_xs_365.bmp



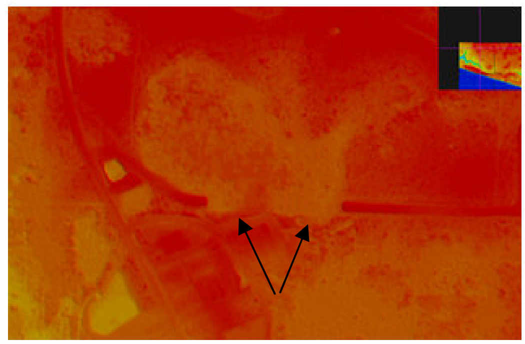
divots_09.bmp



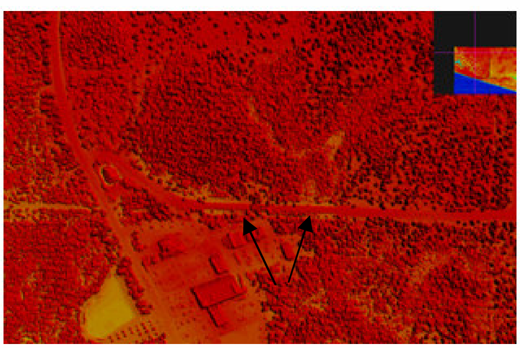
divots_103_xs.bmp



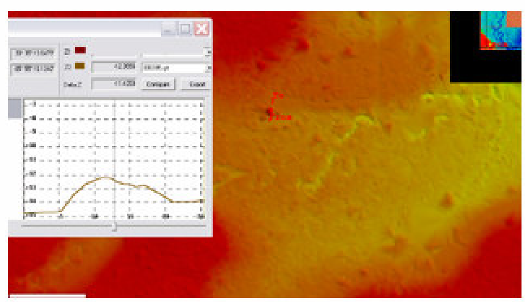
divots_116.bmp



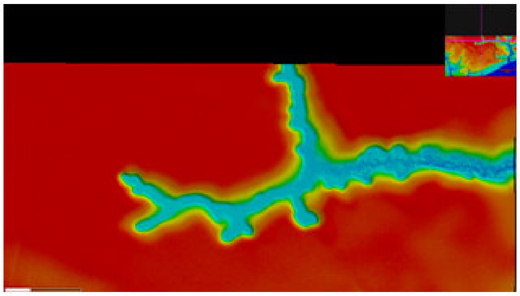
loss_of_road_27.bmp

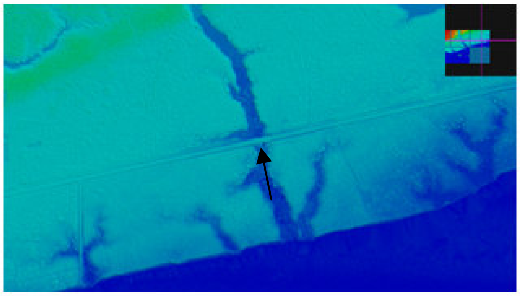


loss_of_road_27_fullpointcloud.bmp



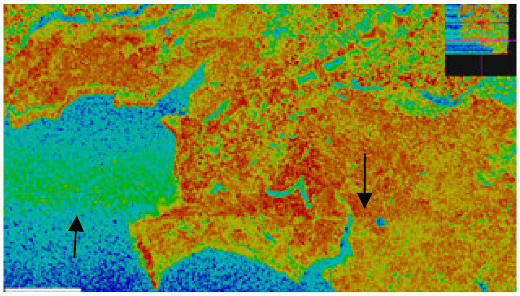
mound_xs_365.bmp

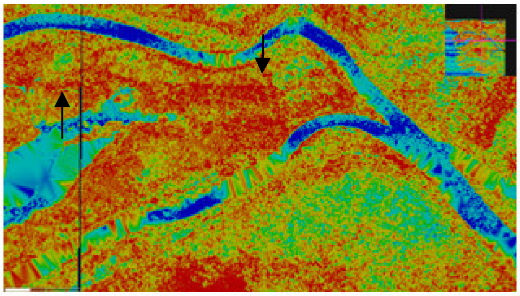




potential_oversmoothing_205.bmp

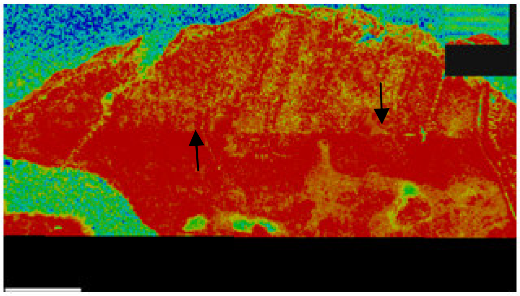
road_left_in_209.bmp

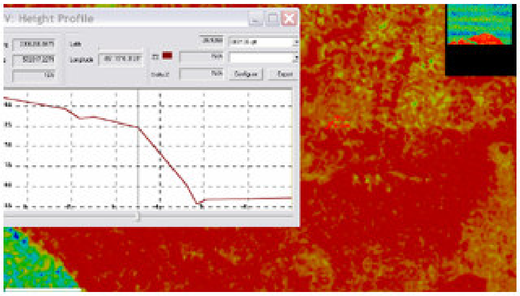




scanline_107.bmp

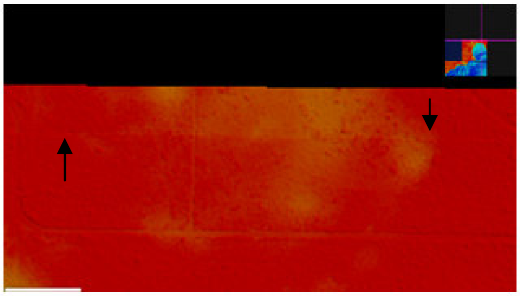
scanline_126.bmp



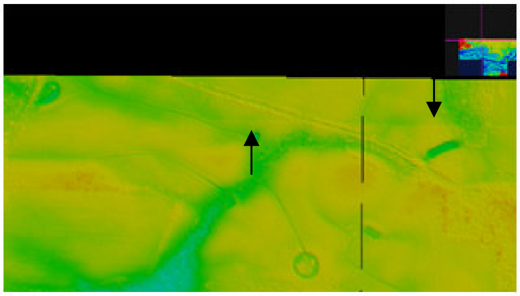


scanline_138.bmp

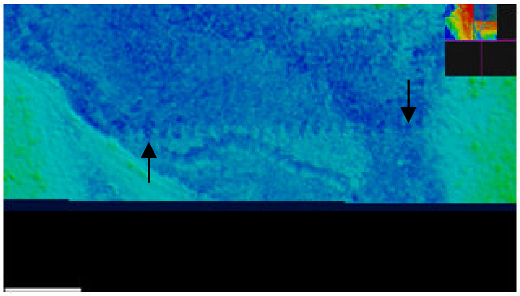
scanline_138_xs.bmp



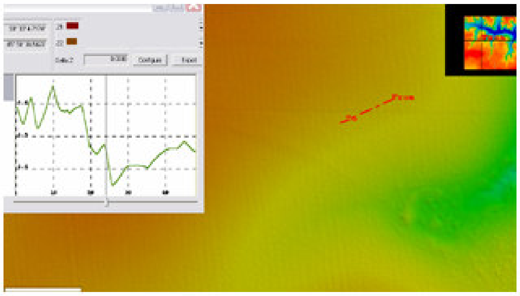
scanline_171.bmp



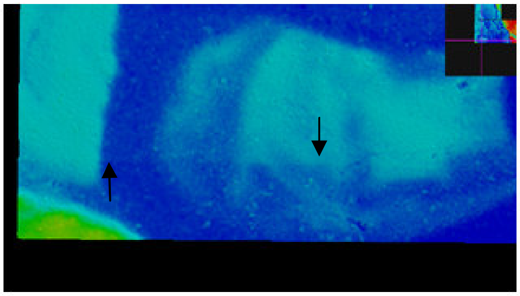
scanline_172.bmp



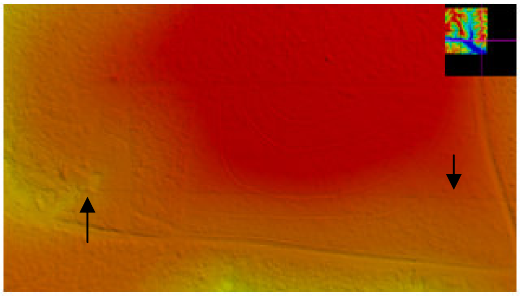
scanline_187.bmp



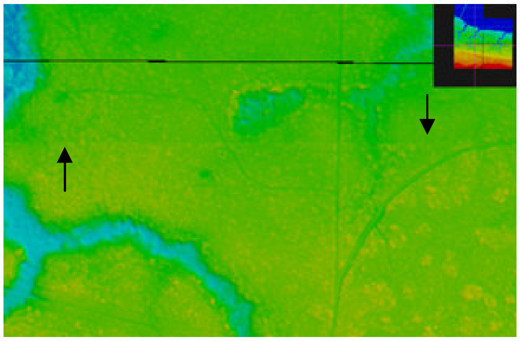
scanline_330_xs.bmp



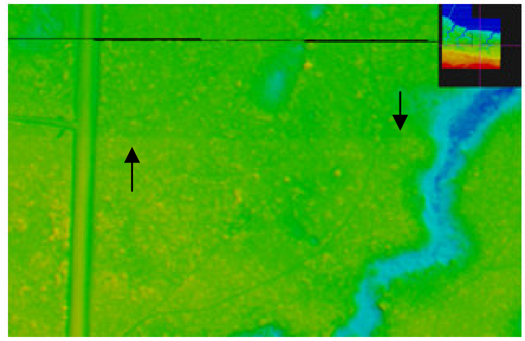
scanline_334.bmp



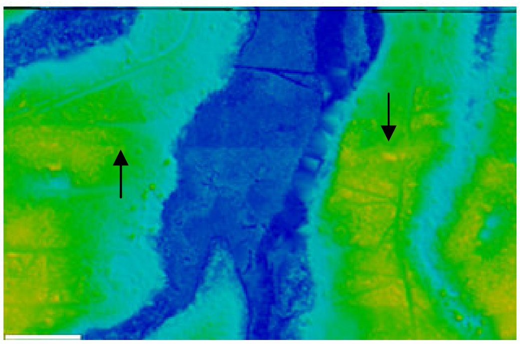
scanline_501.bmp



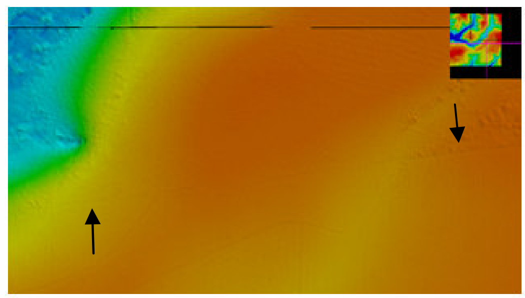
scanline_62.bmp



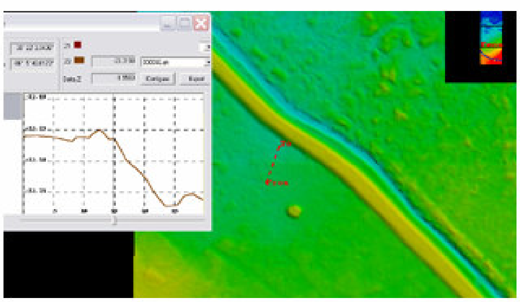
scanline_63.bmp



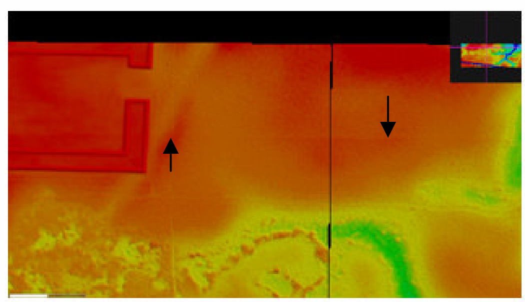
scanline_65.bmp



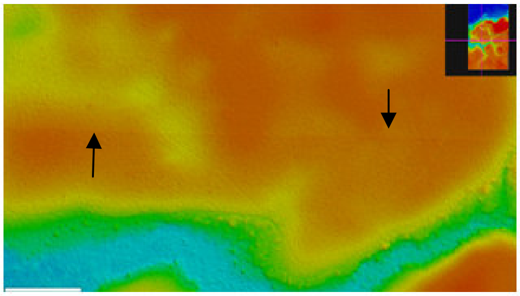
scanline_655.bmp



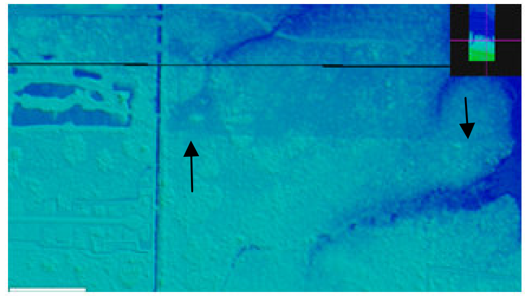
scanline_66_xs.bmp



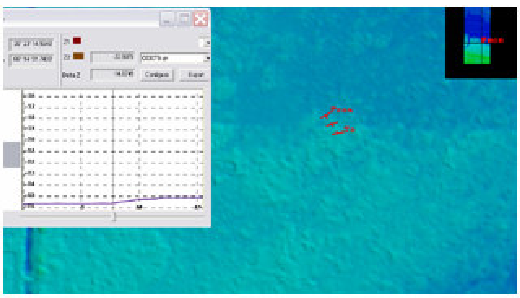
scanline_67_68.bmp



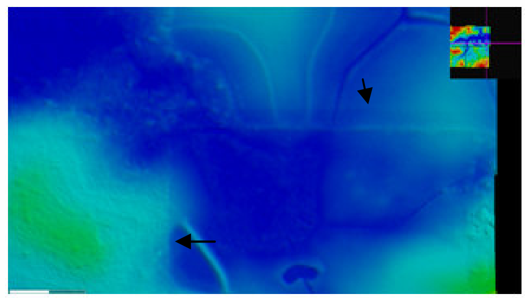
scanline_70.bmp



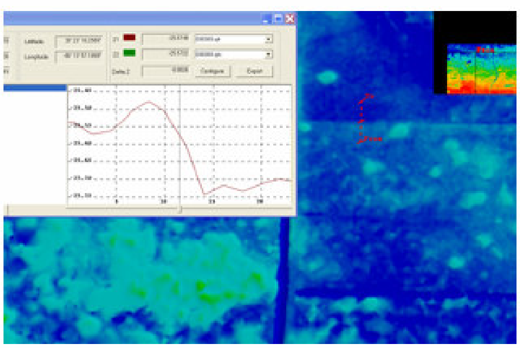
scanline_79_a.bmp



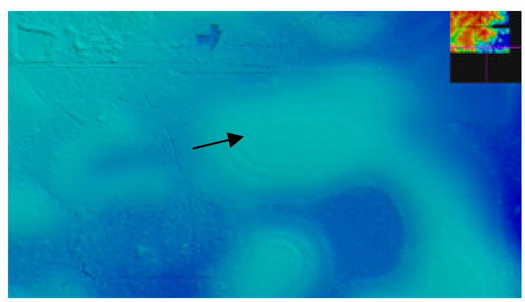
scanline_79_xs.bmp



scanline_856.bmp



scanlines_80_xs.bmp



scans_merge_cornrows_726.bmp