Project report
TO0008 Rhode Island
8/31/2007

Submitted to:
Roald Haested inc.
Prepared by:
Dewberry
Fairfax, VA

EXECUTIVE SUMMARY

This Lidar project covered approximately 40 sq miles along the coastline of Washington County, Rhode Island and was acquired in December of 2006 providing a mass point dataset with an average point spacing of 3ft. The data is tiled, stored in LAS format and Lidar returns are classified in 2 classes: ground and non-ground.

Dewberry Fairfax performed a quality control of these data including a quantitative and a qualitative assessment.

First, the elevation exceeds the accuracy required for this project (accuracy equivalent to 2ft contours according to FEMA *Guidelines and specifications for Flood Hazard Mapping Partners*). These data were tested 0.20ft fundamental vertical accuracy at 95 percent confidence level in open terrain using RMSEx1.96 on 20 survey points.

Secondly, every tile was reviewed at macro level for data completeness: all tiles were delivered, no remote-sensing data voids were found and data are exempt of systematic errors. The cleanliness of the bare earth model was assessed on 20% of the tiles at micro level and exhibits an excellent quality. Minor errors were found (like cornrows and possible vegetation remains) but are not representative of the majority of the data.

In essence, this Lidar dataset is of outstanding quality and meets the needs of FEMA and FEMA contractors for flood mapping.

Breaklines were acquired over streams, lakes and coastline using existing orthophotos to establish hydrological features in the terrain model. These breaklines are 2D lines and were used with a 15ft gridded version of the ground Lidar points to generate 2ft engineering type contours.

TABLE OF CONTENT

PROJECT REPORT

1 Introduction

Lidar technology data gives access to precise elevation measurements at a very high resolution resulting in a detailed definition of the earth surface topology. As a consequence of this precision, millions of points with potential measurement and processing accuracy issues must be verified. This constitutes a challenge for the quality assessment aspect. Our expertise is to provide both a quantitative and qualitative evaluation of the Lidar mass points and it's usability for flood mapping.

Quantitative analysis addresses the quality of the data based on absolute accuracy of a limited collection of discrete checkpoint survey measurements. As the accuracy is tested in several land cover types (open terrain, vegetated areas) but always at ground level, the classification accuracy is indirectly evaluated, i.e. Lidar ground points will fit survey ground points in vegetated areas only if the vegetation is correctly removed by classification and if the Lidar penetrated the canopy to the ground. Although only a small amount of points are actually tested through the quantitative assessment, 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 as acquisition conditions remain similar from one point to the next.

To fully address the data for overall accuracy and quality, a qualitative review for anomalies and artifacts is conducted based on the expertise of Dewberry's analysts. As no automatic method exists yet, we perform a manual visualization process. This includes creating pseudo image products such as 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.

Within this Quality assurance-Quality control process, three fundamental questions were addressed:

- Was the data complete?
- Did the Lidar system perform to specifications?
- Did the ground classification process yield desirable results for the intended bare-earth terrain product?

The first part of this report presents the QaQc process and results. Then, the methodology used to generate the 2 ft contours from the Lidar data using the supplemental breaklines will be explained.

2 Quality Assurance

2.1 Completeness of Lidar deliverables

Once the data are acquired and processed, the first step in our review is to inventory the data delivered, to validate the format, projection, georeferencing and verify if elevations fall within an acceptable range.

2.1.1 Inventory and location of data

The project is separated in 3 areas along the coastline of Washington County, Rhode Island (RI) as illustrated in Figure 1:

- New Shorham (island, the planned project area only included the coast, 4.8 sq miles)
- North Kingston (11.7 sq miles)
- Westerley Charlestown (23.2 sq miles)

Data were acquired in fall 2006 during leaf-off and before snow conditions at low tide conditions. The average point spacing of the data is 3ft.

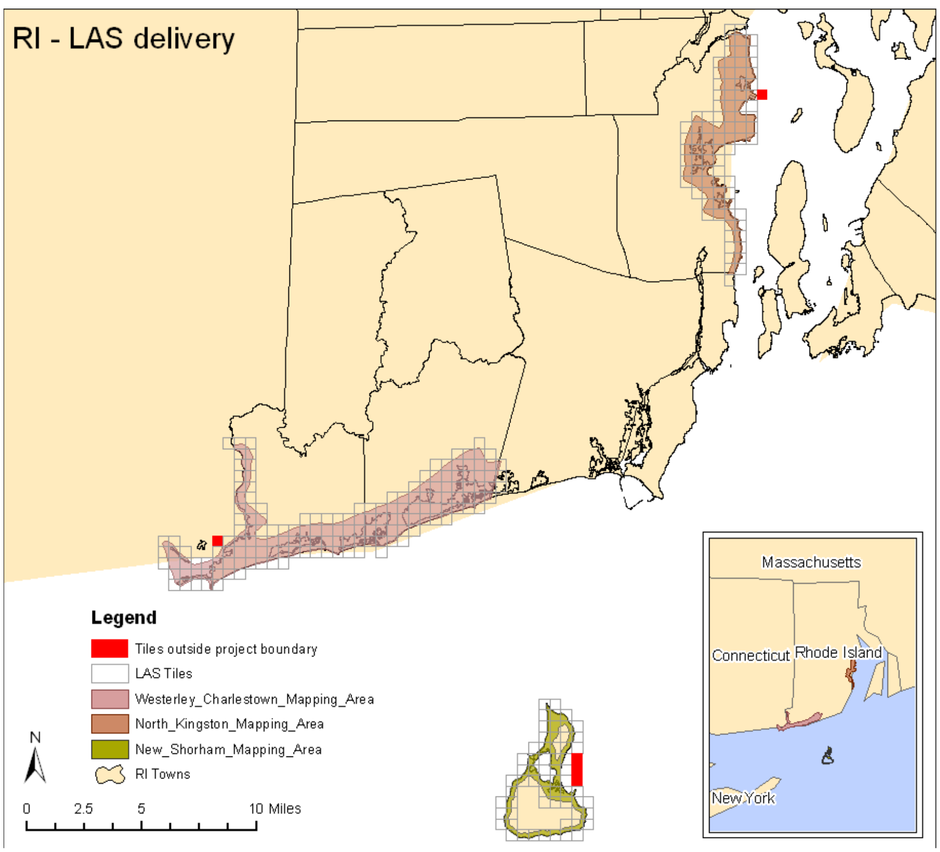


Figure 1 – inventory of the Las files

The tile scheme follows the State's existing orthophoto master tile scheme (10,000x10,000ft) divided in 16 subtiles of 2,500x2,500ft each, the tile number is based on the ortho number with a suffix from 1 to 16 (see illustration in Figure 2).

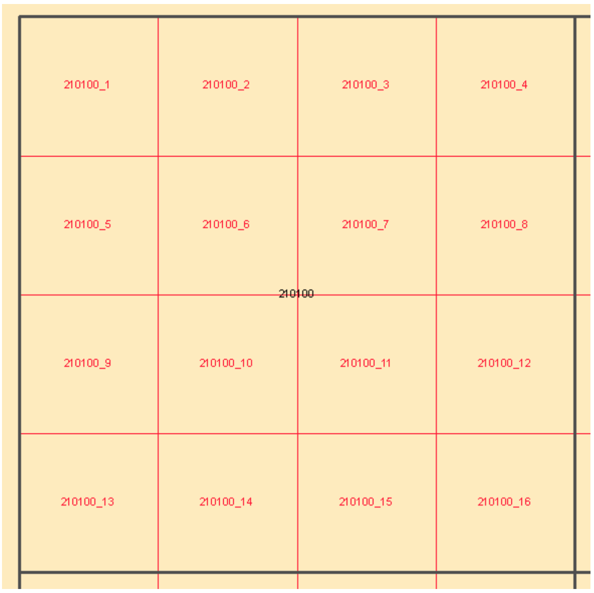


Figure 2 – Tile scheme (Lidar in red, ortho in black)

312 tiles are delivered and used for mapping purpose; the inland tiles of New Shorham area were acquired during the data collection phase by our subcontractor Terrapoint but are not part of the project area. Moreover, all the expected data has been received.
Data were delivered in LAS format version 1.0, each record includes the required class code (code1 for non-ground and code 2 for ground) along with additional information like: flightline number, intensity, return number. Although the initial Scope of Work stated that the data should be in LAS 1.1 this was not possible as the Lidar system is only capable of collecting data and storing it as LAS1.0. It should be noted that the 1.0 format basically provides the same information and no data integrity is lost.

Although the LAS file header does not include a projection definition, it was verified that the spatial reference for the data is:

- Projection: State Plane - Rhode Island
- Horizontal Datum: NAD83
- Vertical Datum: NAVD88
- Units: US Survey Feet

2.1.2 Statistical analysis of tile content

To verify the contents of the data and to validate the data integrity, a statistical analysis is performed on all the data. This process allows us to statistically review 100% of the data to identify any gross outliers. This statistical analysis consists of:

1. Extract the header information (number of points, minimum, maximum, elevation)
2. Compare the Lidar file extent with the tile extent

Each tile was queried to extract the number of Lidar points and all tiles are within the anticipated size range, except for where fewer points are expected (because of water features or near the project boundary). Figure 3 presents a map of the number of records in each LAS files (full point cloud), highlighting tiles less than 100 000 points.

Additionally the minimum and maximum elevation values in each file were also computed and mapped. Figure 4 shows the z min value for each tile; it can be noticed that a majority of the files have z minimum values between -5ft and -1feet, for the most part over water bodies. Besides, users should be aware that as the contract for acquisition and processing identified only two classes: Class “1” for unclassified and Class “2” for ground, there is no distinction for water and therefore some water points may be classified as ground if the water elevations are equal to that of the surrounding terrain.

One tile was identified as having a Z minimum elevation of -23 ft and although it is an erroneous point over water, it is correctly classified as class 1. No noticeable anomaly has been found concerning the z max value.

The geographic extent of the LAS files were compared to the extent of the tiles. This process ensures that the data within the LAS file is spatially contained with the extents of each corresponding tile boundary. Figure 5 shows that files were truncated along the project boundary. Moreover, all the files are located within the tile extents.

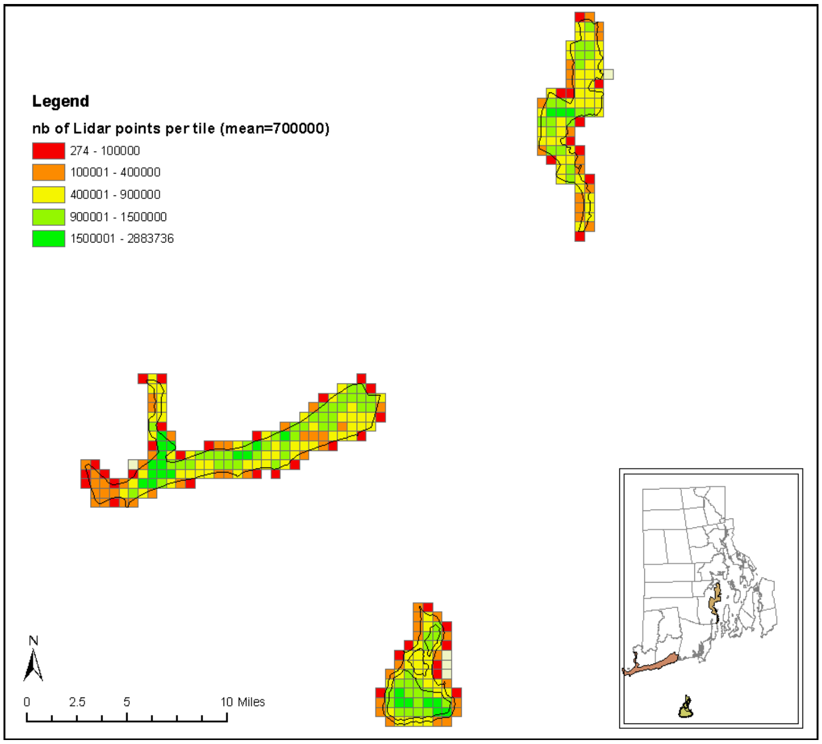


Figure 3 - Number of records in LAS files located by tiles.

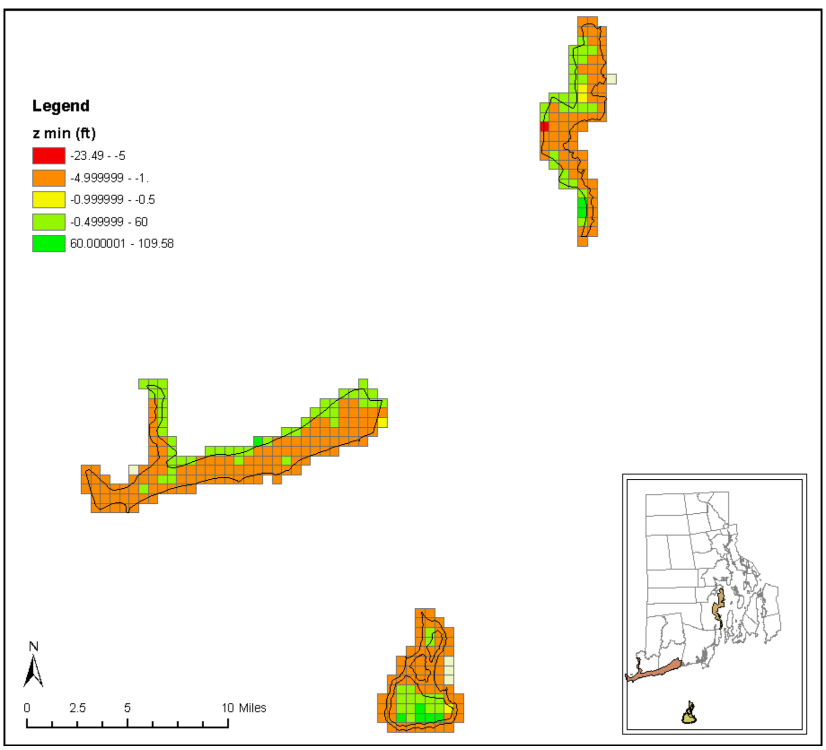


Figure 4 – Minimum elevation in each LAS file (all classes)

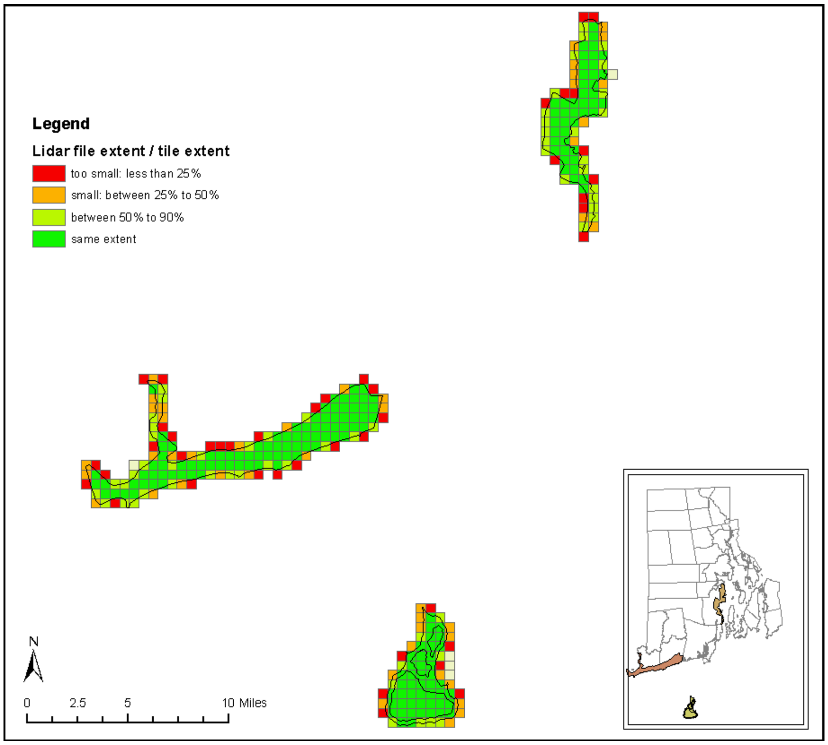


Figure 5 – Comparison between the Lidar file extent and the tile extent

2.2 Quantitative assessment

2.2.1 Inventory of survey points

A quantitative analysis of the accuracy has been conducted on the data using 20 check points distributed in 4 land cover types (open terrain, weeds/crops, forest, urban). These survey checkpoints were provided to Dewberry from Roald Haestad Inc. collected under the guidance from Dewberry. A list can be found in Appendix A and the complete field survey report from Roald Haestad Inc. is provided with the data. Although the FEMA guidelines state that a minimum of 20 checkpoints per land cover type should be utilized, this would be financially impractical where the survey checkpoints would cost more than the Lidar data collection. Therefore 20 points were selected which satisfies the National Standard for Spatial Data Accuracy (NSSDA) accuracy requirement from which the FEMA guidelines are based upon. The NSSDA states; "A minimum pf 20 checkpoints shall be tested, distributed to reflect the geographic area of interest and the distribution of error in the dataset. When 20 points are tested, the 95 percent confidence level allows one point to fail the threshold given in product specifications." It must be noted that the NSSDA does not define the corresponding project size for 20 checkpoints as this could be for 1 square mile or 1000 square miles.

2.2.2 Vertical Accuracy: elevation comparison

Using the ground truth checkpoint survey as the reference, the elevation at the same x and y positions are interpolated from the Lidar data. The method used to extract the elevation from the Lidar mass points at a given location is to create a triangular irregular network from the ground classified points and to interpolate the elevation at the given x and y coordinates using the 3 nearest Lidar neighbors. To compare the two types of measured elevations, statistics are then being computed following two different guidelines further explained in the following sections.

2.2.3 Vertical Accuracy Assessment Using the RMSE Methodology

The first method of testing vertical accuracy will use the FEMA specifications which essentially follows the NSSDA procedures. The accuracy is reported at 95% confidence level using the Root Mean Square Error (RMSE) which is valid when errors follow a normal distribution. This methodology measures the square root of the average of the set of squared differences between dataset coordinate values and coordinate values from an independent source of higher accuracy for identical points. The vertical accuracy assessment compares the measured survey checkpoint elevations with those of the TIN as generated from the bare-earth Lidar. The survey checkpoint's X/Y location is overlaid on the TIN and the interpolated Z value is recorded. This interpolated Z value is then compared to the survey checkpoint Z value and this difference represents the amount of error between the measurements.

The following tables and graphs outline the vertical accuracy and the statistics of the associated errors. Figure 6 illustrates the distribution of the elevation differences between the Lidar data and the surveyed points by land cover type and sorted from lowest to highest. Weeds and crops points tend to have the highest errors and both vegetated categories have a positive bias (all errors are positives). This means that the elevation interpolated from the Lidar dataset is higher than the surveyed point. This could be explained by a non-penetration of the Lidar beam al the way trough the ground or by a misclassification of some points actually belonging to the vegetation class. However, all the differences largely remain within acceptable ranges, and do not constitute an issue.

Table 1 – Descriptive statistics (FEMA guidelines) by land cover category

100 % of Totals	RMSE (ft) Spec=0.61ft	Mean (ft)	Median (ft)	Skew	Std Dev (ft)	# of Points	Min (ft)	Max (ft)
Consolidated	0.18	0.05	0.04	0.14	0.18	20	-0.31	0.45
Open Terrain	0.10	-0.02	-0.06	0.99	0.11	5	-0.13	0.16
Weeds/Crop	0.28	0.24	0.23	-0.11	0.18	5	0.00	0.45
Forest	0.10	0.09	0.10	-0.07	0.05	5	0.02	0.16
Urban	0.19	-0.11	-0.01	-0.49	0.17	5	-0.31	0.06

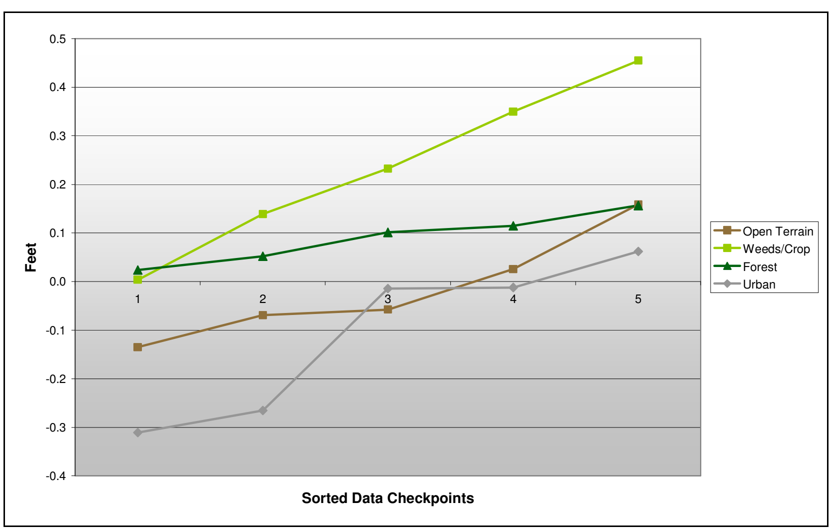


Figure 6 - Elevation differences between the interpolated Lidar and the surveyed QAQC checkpoints

2.2.4 Vertical Accuracy Assessment Using the NDEP Methodology

The RMSE method assumes that the errors follow a normal distribution and experience has shown that this is not always the case as vegetation and manmade structures can limit the ground detection causing errors greater than in unobstructed terrain. The NDEP methodology therefore assumes that the data does not follow a normal distribution and tests the open terrain (bare-earth ground) separately from other ground cover types.

The *Fundamental Vertical Accuracy* (FVA) at the 95% confidence level equals 1.96 times the RMSE in open terrain only (as previously explained: the RMSE methodology is appropriate in open terrain). *Supplemental Vertical Accuracy* (SVA) at the 95% confidence level utilizes the 95th percentile error individually for each of the other land cover categories, which may have valid reasons (e.g. problems with vegetation classification) why errors do not follow a normal distribution. Similarly the *Consolidated Vertical Accuracy* (CVA) at the 95% confidence level utilizes the 95th percentile error for all land cover categories combined. This NDEP methodology is used on all 100% of the checkpoints.

Table 2 - Accuracy using NDEP methodology

Land Cover Category	# of Points	FVA — Fundamental Vertical Accuracy ($RMSE_z$ x 1.9600) Spec=1.19 ft	CVA — Consolidated Vertical Accuracy (95th Percentile) Spec=1.19ft	SVA — Supplemental Vertical Accuracy (95th Percentile) Target=1.19 ft
Consolidated	20		0.355	
Open Terrain	5	0.200		0.154
Weeds/Crop	5			0.434
Forest	5			0.148
Urban	5			0.302

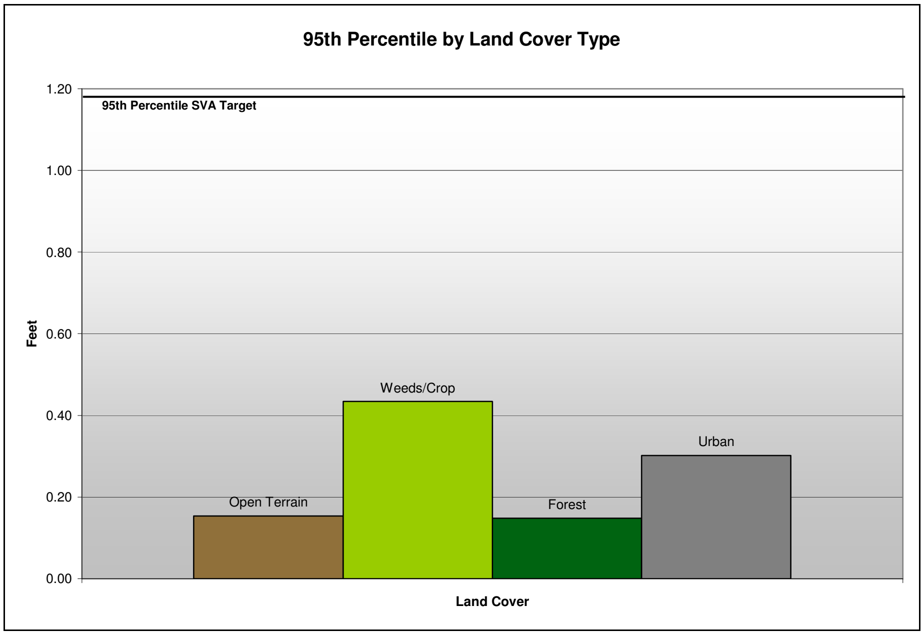


Figure 7 - 95th Percentile by Land Cover Type

The target objective for this project was to achieve bare-earth elevation data with an accuracy equivalent to 2 ft contours, which equates to an RMSE of 0.61 ft when errors follow a normal distribution. With these criteria, the Fundamental Vertical Accuracy of 1.19 ft must be met. Furthermore, it is desired that the consolidated Vertical Accuracy and each of the Supplemental Vertical Accuracy also meet the 1.19 ft criteria to ensure that elevations are also accurate in vegetated areas. As summarized in Table 2, this data:

- Does satisfy the NDEP's mandatory Fundamental Vertical Accuracy criteria for 2 ft contours.
- Does satisfy the NDEP's mandatory Supplemental Vertical Accuracy criteria for 2 ft contours.
- Does satisfy the NDEP's mandatory Consolidated Vertical Accuracy criteria for 2 ft contours.

2.2.5 Vertical Accuracy Conclusion

Utilizing both methods of vertical accuracy testing, this data meets and exceeds all specifications. The consolidated RMSE of 0.18ft is less than the FEMA requirement of 0.61ft. This data is of excellent quality and should satisfy most users for high accuracy digital terrain models.

2.3 Qualitative assessment

2.3.1 Protocol

The goal of this qualitative review is to assess the continuity and the level of cleanliness of the data. The acceptance criteria we have reviewed are the following:

- If the density of points is homogeneous and sufficient to meet the user needs,
- If the ground points have been correctly classified (no manmade structures and vegetation remains, no gap except over water bodies),

- If the ground surface model exhibits a correct definition (no aggressive removal, no over smoothing, no inconsistency in the post-processing), in a context of flood modeling, special attention is given to stream channels,
- If no obvious anomalies due to sensor malfunction or systematic processing artifact are present (data holidays, spikes, divots, ridges between tiles, cornrows…).

Dewberry analysts, experienced in evaluating Lidar data, performed a visual inspection of a bare-earth digital elevation model (bare-earth DEM). Lidar mass points are first gridded with a grid distance of two times the full point cloud resolution. Then, a triangulated network is built based on this gridded DEM and is displayed as a 3D surface. A shaded relief effect is applied, which enhances 3D rendering. The software used for visualization allows the user to navigate, zoom and rotate models and to display elevation information with an adaptive color coding in order to better identify anomalies.
One of the variables established when creating the models is the threshold for missing data. For each individual triangle, the point density information is stored, if it meets the threshold, the corresponding surface will be displayed in green, if not it will be displayed in red (see Figure 8).

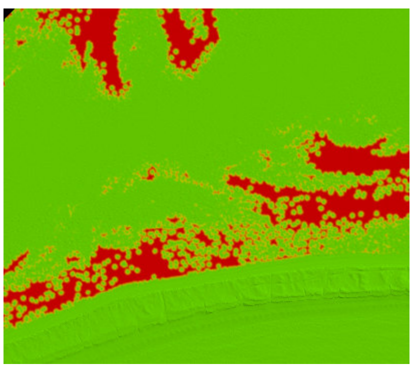


Figure 8 – Ground model with density information (red means no data)

The first step of our qualitative workflow is therefore to verify data completeness and continuity using the bare-earth DEM with density information, displayed at a macro level. If, during this macro review of the ground models, we find potential artifacts or large voids, we use the digital surface model (DSM) based on the full point cloud including vegetation and buildings to help us better pinpoint the extent and the cause of the issue. Moreover, the intensity information stored in Lidar data can be visualized over this surface model, helping in interpretation of the terrain.
Finally, in case the analyst suspects a systematic errors relating to data collection, a visualization of the 3D raw mass points is performed, rather than visualizing as a surface. This particular type of display helps us visualize and better understand the scan pattern and the flight line orientation.
The process of importing, comparing and analyzing these two later types of models (DSM with intensity and raw mass point), along with cross section extraction, surface measurements, density evaluation, constitutes our micro level of review.

2.3.2 Quality report

We reviewed all the tiles at a macro level, and 20% at a micro level.

Overall, the data are consistent and of excellent quality, no data void, no major anomaly was found. The bare earth product exhibits a precise definition of all hydro features (see Figure 9 and Figure 10).

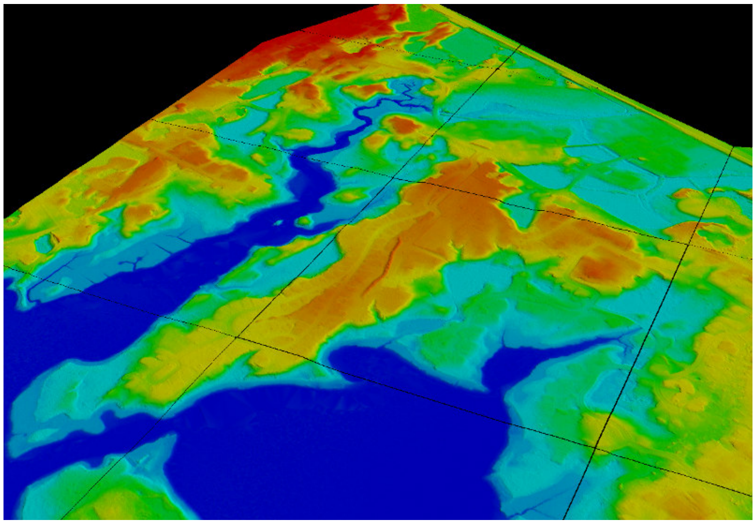


Figure 9 – Around tile 340190_9: Good example of the topographic definition

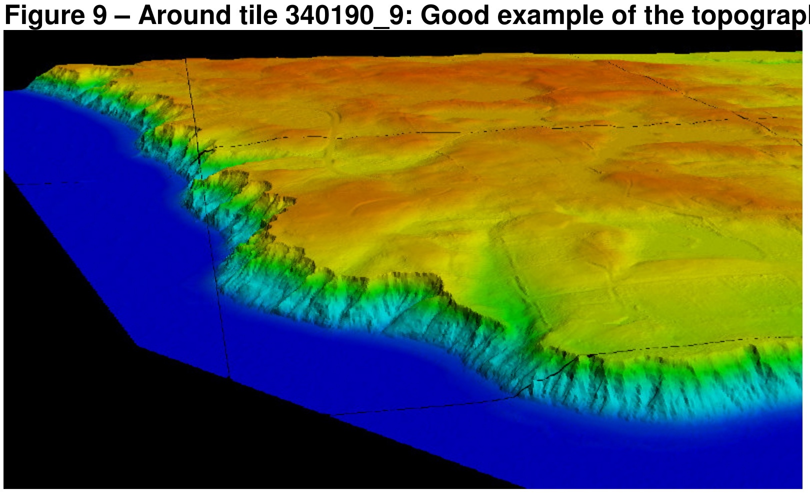


Figure 10 – Around tile 310030_6: Coastal view

The project contained very few errors in the data and these issues are somewhat common with Lidar data and are within acceptable limits.
The three types of issues are:

1. “Cornrow”
2. Cleanliness of artifacts
3. Hydro enforcement of stream channel

Cornrows were typically seen throughout the project. There are multiple reasons as to why this happens but the end result is that adjacent scan lines are slightly offset from each other. This will give the effect that there are alternating rows of higher, and then lower elevations. Although this is common with Lidar data, as long as the elevation differences are a less than 20 cm and that the occurrences are minimized, it is acceptable since it is within the noise and accuracy levels. However this also can be an indication that the sensor is mis-calibrated, or offsets exist between adjacent flight lines so each area identified is analyzed. Our review found only one significant instance of the cornrow effect, but the remainder of this effect was within acceptable limits. Figure 11 illustrates the highest cornrow error.

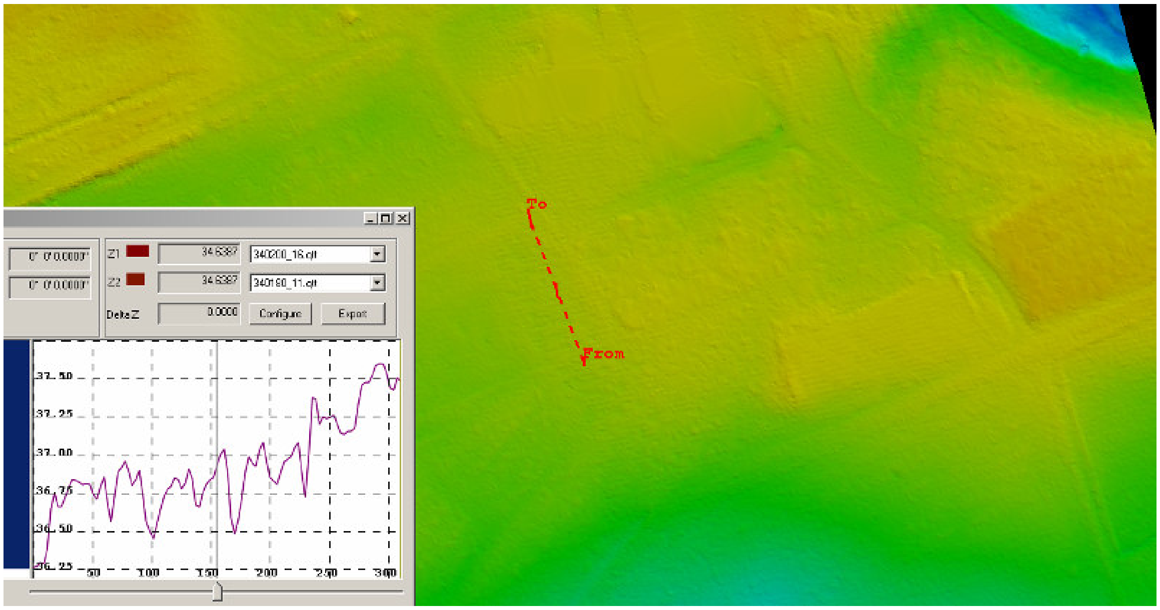


Figure 11 – Tile 340200_8: Possible corn row artifact

The cleanliness of the bare earth data is of excellent quality and meets the requirement of this project. In a few isolated tiles we have found potential vegetation artifacts most likely due to a misclassification (see Figure 12). Due to the large spectrum of geographic patterns, there are instances where the algorithms erroneously classify the data. However it is evident that these potential areas are relatively small and easily within the specification of being 95% cleaned of artifacts.

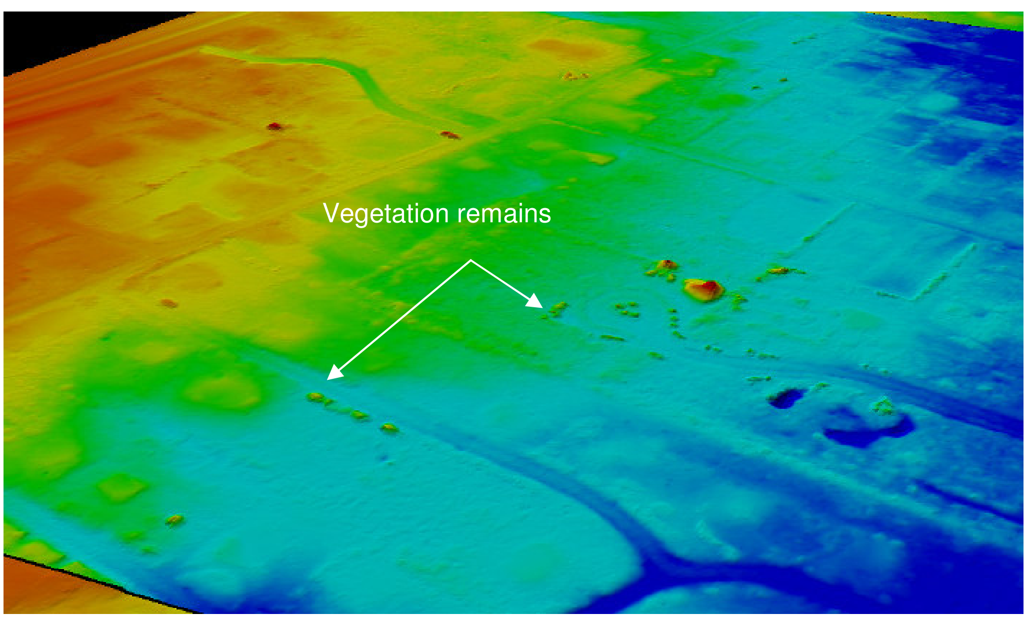


Figure 12 – Tile 290110_2: Possible vegetation artifacts in bare-earth model

One minor issue for the bare-earth terrain is the classification of bridges. Some users may require bridges to be removed (classified to non-ground) while others may require them classified as ground. For the user community if this is an issue this is easily remedied since it is clearly identifiable and the data can be reclassified. Figure 13 illustrates various scenarios of bridges partially removed or retained; generally speaking large bridges have been removed even though this was not a specific requirement in the scope of work.

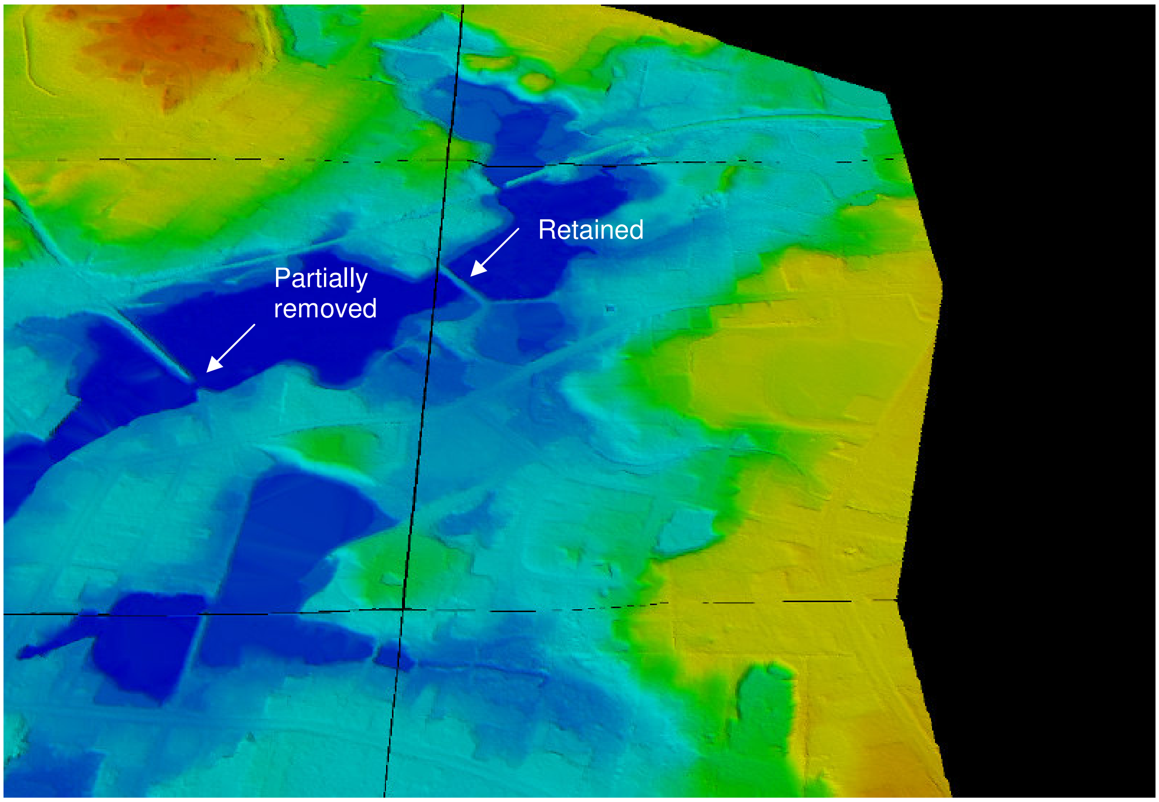


Figure 13 – Around Tile 34180_9: inconsistent Bridge editing. Bridges are sometimes totally retained or sometimes portions of the bride remain in the bare earth terrain.

3 Topographic data production

3.1 Breaklines

Supplemental 2D breaklines were acquired by Dewberry over this entire mapping area using existing 2004 orthophotos from the Rhode Island Department of Transportation (RIGIS03/04 Ortho). Only hydrographic features needed to support the terrain model for hydrology have been considered. Table 3 lists the items available in the two delivered shapefiles and Figure 14 gives an illustration.

Table 3 – Breaklines acquisition specifications

File	Item represented	Feature Code
Line feature class	Stream <20ft	1
	Connector or bridges	3
	Center line of stream <20ft	2
	Coastline (interrupted at river mouth and ports)	4
Polygon feature class	Stream >20ft or big lakes with sea opening	1
	Lakes, ponds > 1acre	(1), 2, 4
	Artificial pond	3
	Islands	5

Figure 14 – Breaklines as colored polygons and polylines over the orthophoto

The following comments describe in detail what is provided:

- Features are only acquired in mapping area.
- Small ponds are not acquired.
- When features are over project boundaries, it has been entirely acquired.
- Breaklines (lines and areas) has an attribute feat_types (from 1 to 4 for lines, from 1 to 5 for polygons).
- Code 5 is for island, overlapping or not with associated water polygon,
- Polygons of river have a centerline.
- Polygons have no hidden part, but associate center line may have one.
- Coastline is one line except:
  - when it crosses a river mouth, it is interrupted by a polygon,
  - When interrupted by a pier.

Several steps of preprocessing have been done to render these features class usable as breaklines in a triangulated irregular network (TIN). The 3 major types of features we used are:

- Small streams lines and adjacent lines for large streams as hard breaklines, without elevation;
- Large river polygons and water bodies to remove the points inside water;
- Waterbodies (lakes, bays, sea) with a constant elevation as hard replace polygon, this means that the polygon outline is used as a hard breakline and the TIN elevation inside the polygon is replaced by the elevation selected in the attribute table. This elevation was computed as the minimum value of the removed points inside the polygon minus 0.5ft.

3.2 Contours

Our goal was to produce engineering grade contour lines at a 2ft interval. As stated in the scope of work, the emphasis was made on the accuracy of contours as opposed to aesthetically pleasing.

ArcGis 9.2, 3D analyst was used for the contouring. First the hydro features were edited to obtain the breaklines needed to constraint the TIN.
LAS files were then converted to a 3D point vector format usable in ArcGis (multipoints) with a selection of the ground class only, the points were used to create a raster at a coarser resolution than the source data, i.e. whereas the data have a 3ft point spacing we used a 15ft cell size for the raster. This would introduce a smoothing for the final contours while keeping the georeferencement of the lines within an acceptable range. The raster was then exported to a point format (storing the elevation information as an attribute) usable for the TIN construction. Water points were removed from this point file using the waterbody polygons. These mass points and the breaklines were then used to build a TIN from which 2ft isolines were computed. Short lines (<150ft, except at the project boundary) and closed contours inside rivers were erased and a final line simplification was applied to clean the contours (summarized in Figure 15). A final quality control was applied to assess the topology consistency of the contour lines (no intersections, no self-intersection, and no dangles except at the project boundary).
The result of this process can be seen in Figure 16.

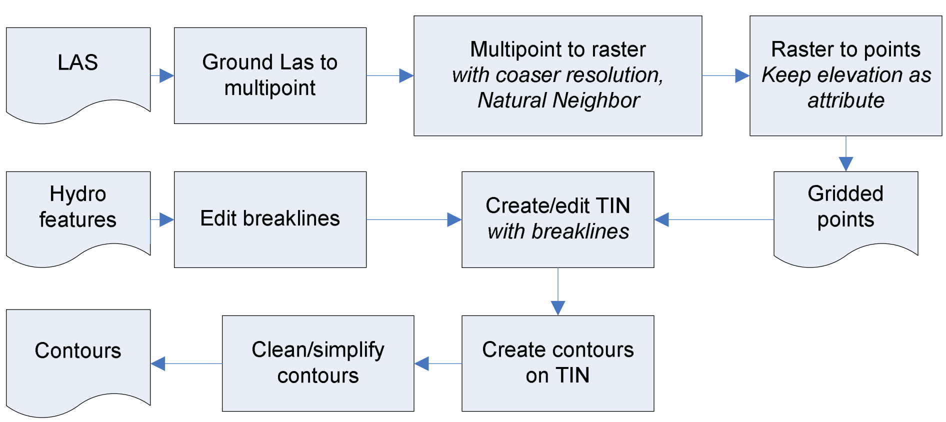


Figure 15 – Simplified process flow for contouring

Figure 16 – Example of the generated contours with breaklines over orthophoto

4 Conclusion

Overall the data exhibited excellent detail in both the absolute and relative accuracy. The level of cleanliness for a bare-earth terrain is of the highest quality and no major anomalies were found. The figures highlighted above are a sample of the minor issues that were encountered and are not representative of the vast majority of the data which is of excellent quality.

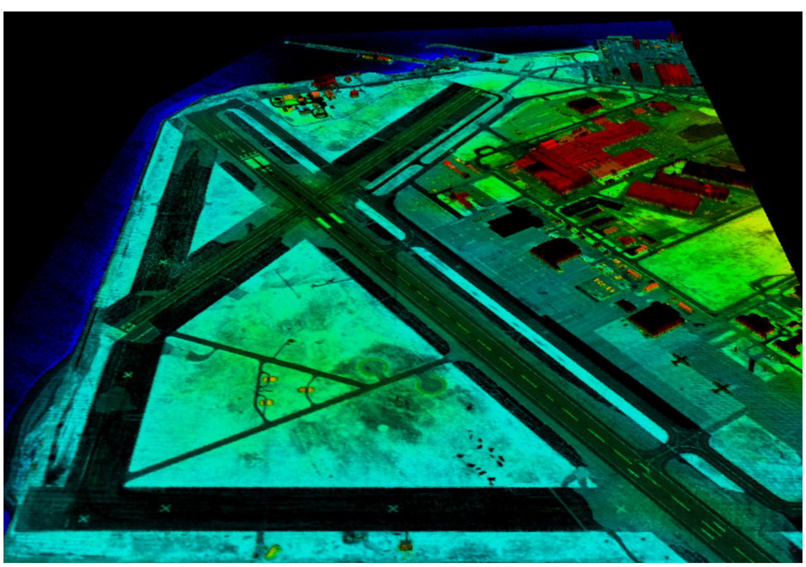


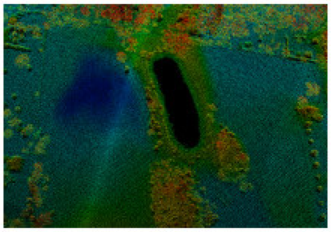
Figure 17 – Tile 350190_10: airport, surface model colored by elevation and shaded with Lidar intensity. This is an excellent example of the level of details given by the data.

Appendix A Control points

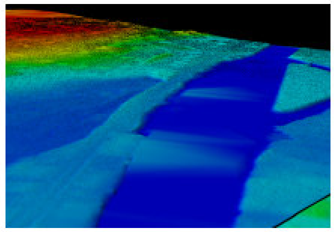
pointNo	e	n	elevation	zLidar	LandCoverType	DeltaZ
2406	313590.50	25470.41	159.76	159.63	Open Terrain	-0.13
2201	282309.01	104477.94	24.60	24.53	Open Terrain	-0.07
2106	255270.13	90029.32	3.46	3.40	Open Terrain	-0.06
2501	347091.09	185828.28	17.84	17.87	Open Terrain	0.03
2303	339245.11	179067.27	18.88	19.04	Open Terrain	0.16
2204	280816.08	103843.06	18.77	18.77	Weeds/Crop	0.00
2101	245104.28	88027.33	4.31	4.45	Weeds/Crop	0.14
2305	344175.21	170448.92	18.62	18.85	Weeds/Crop	0.23
2403	309073.53	38913.97	8.56	8.91	Weeds/Crop	0.35
2504	352627.49	184080.36	7.03	7.48	Weeds/Crop	0.45
2216	284240.72	103079.16	15.45	15.48	Forest	0.02
2516	348075.60	188408.93	34.39	34.44	Forest	0.05
2307	340755.11	179974.14	6.50	6.60	Forest	0.10
2111	235911.59	101924.22	21.81	21.93	Forest	0.11
2407	305725.46	35779.72	34.03	34.18	Forest	0.16
2509	350409.00	184332.37	11.49	11.18	Urban	-0.31
2105	241015.06	95255.28	36.72	36.46	Urban	-0.27
2205	292048.17	109882.38	32.15	32.14	Urban	-0.01
2404	312510.77	32624.90	6.20	6.19	Urban	-0.01
2304	341240.69	177512.96	4.41	4.47	Urban	0.06

Appendix B Contact sheets of potential qualitative issues

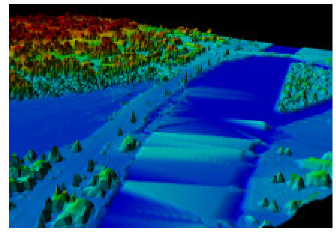
RI - screen shots of potential issues



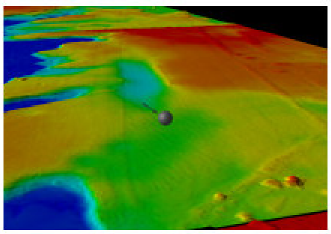
230090_3_voidInData_maybeWater_qtcFcp.png



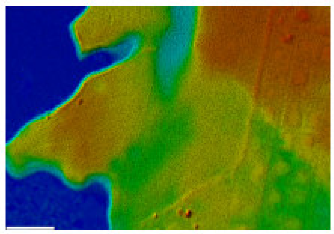
250100_15_poorDefinitionOfstreamChannel_ground.png



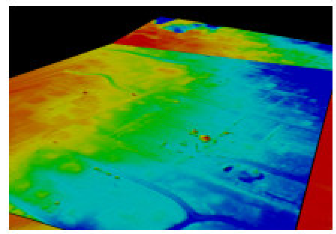
250100_15_poorDefinitionOfstreamChannel_qttFpc.png



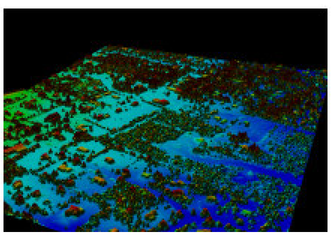
280110_4_cornRow_possibleAnomaly.png



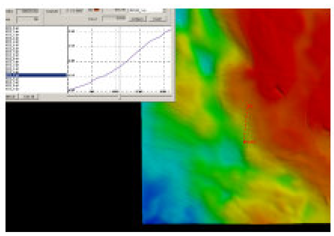
280110_4_CornRow_possibleAnomaly_fcp.png



290110_2_possibleVegetationArtifacts.png

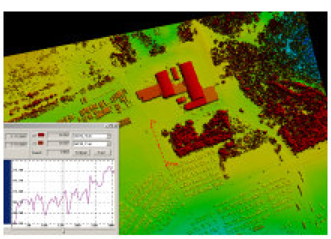


290110_2_possibleVegetationArtifacts_qttFcp.png

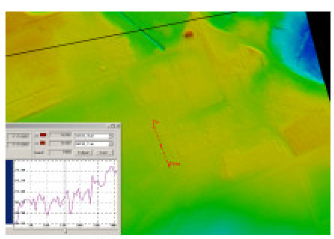


300030_5_minorCornRowEffect_qttGround.png

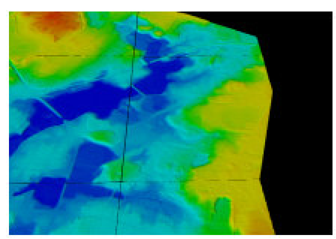
300040_4_possibleVegetationArtifact_qttGroundOrtho.png



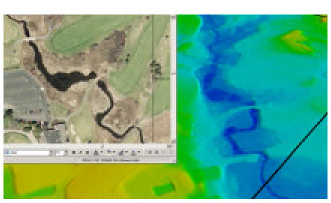
340200_12_cornRowArtifact_qttFpc.png



340200_8_cornRowArtifact_qttGround.png



34180_9_bridgesPartialHydroEnforcment_qttGround.png



350210_9_bridgeWithoutHydroEnforcment_qttGroundOrtho.png