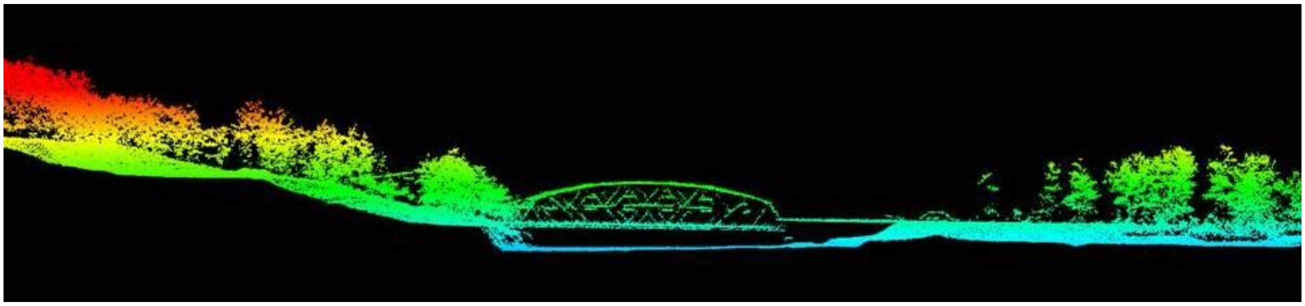
Bradfield to Behms Corridor Quality Assurance Report of DOT Roads to Resources Project

Technical Report · July 2015
DOI: 10.13140/RG.2.1.3415.5289

1 author:

Rick Guritz
University of Alaska Fairbanks



Alaska Department of Transportation
Roads to Resources Project
LiDAR & Imagery
Quality Assurance Report
Bradfield to Behms Corridor

Written by Rick Guritz

Alaska Satellite Facility

July 24, 2015

Contents

Study Area – Bradfield to Behms Corridor

The study area for this report is the Bradfield Canal to Behms Canal Corridor 31.8 kilometers long and 30.0 square kilometers in Misty Fjords National Monument, Southeast Alaska (see Figure 1).

Basis for Evaluation

The Software used for the evaluation includes:

- ESRI ArcMap and ArcCatalog 10.3
- Applied Imagery Quick Terrain Modeler v7.1.5 64-bit
- Blue Marble Geographic's, Global Mapper v13.1.2

Each block of LiDAR will be evaluated in the following ways:

- Check formatting and completeness of data delivery,
- Check completeness, clarity, and compliance of metadata,
- Assess the planimetric accuracy of the LiDAR data,
- Assess the vertical accuracy of the LiDAR data,
- Assess the LiDAR point cloud data of return density and classification accuracy,
- Access the LiDAR strips point cloud data for swath overlap,
- Assess the LiDAR bare-Earth and first-return surface data by mosaic and shaded relief analysis, identifying gaps, seams, anomalies, and hydro-flattening of data,
- Verify consistency of CAD derived products being provided by LiDAR contractor.

Itemized products to be evaluated include:

- Metadata
- Unclassified strips point cloud in LAS format
- Classified point cloud data in LAS format
- Bare-Earth surface (below canopy raster DEM)
- First-Return surface (top of canopy raster DSM)
- Intensity image composite
- Hydro-flattening break lines (single and double line) and lake polygons
- Contours in LandXML Format (elevation)
- Shaded relief mosaics
- Tile Index (full tile)

Format and Completeness of Data Delivery

Two separate data deliveries were required to correct all identified data anomalies. Each delivery provided data on an external computer disk organized by block and data type or smaller tile fixes were staged to FTP. The Caswell Lake block deliveries included:

- Metadata – A single metadata file for each major product type including both unclassified and classified point clouds, raster bare-Earth DEM, raster first-return DSM, raster LiDAR intensity image, elevation contours in CIV3D drawing files, elevation contours in LandXML format.

- LiDAR Data - LAS unclassified point cloud 51 swath strips and 110 tile based data, and classified point cloud data for 110 tiles of data.

- Ancillary Data – GPS aircraft position data for each flight mission, swath index files in both CIV3D Drawing format and as shapefiles, and sbet mission trajectory data.

- GIS Product Data – Raster gridded products including digital elevation models, first return digital surface models, and LiDAR intensity image products all in geotiff format. Vector products include boundary shapefile, and LiDAR tile index shapefile.

- CAD Data – Corridor wide boundary and breaklines file in CIV3D drawing format, and 2 foot contours in LandXML format, certified by a State of Alaska registered surveyor that the contours meet The RMS Error Standard for ASPRS Class 2 for vertical accuracy.

- Ortho Imagery – Color balanced ortho mosaic tiles in true color (RGB) and color infrared (CIR), a color balanced mosaic in true color (RGB) and color infrared (CIR) both in mrsid format, unbalanced tile based source ortho imagery, and vector data including agps, cutlines, and tiling scheme in both CIV3D and shapefile format.

The contents of the LiDAR point cloud files were verified to include the expected LiDAR classification layers. Then each layer was loaded into Quick Terrain Modeler (QTM) to verify coverage and extent of each classification layer. Each layer was captured to a computer graphic image in jpeg format for review. In some cases, the number of points within a classification layer needed to be separated into thirds to keep within available memory (16 GB) to optimize RAM use. In these cases up to three jpeg images were saved for the layer. Although this process was time consuming, it proved very useful in identifying omissions in coverage for particular classifications. By saving each layer in QTM, the number of points included in each classification layer was compiled to verify that the data was distributed appropriately between classification layers.

The contractor supplied GIS layers including corridor boundary and tile indexes were displayed and evaluated to insure consistency with the original project coverage feature class (see Figure 2). The gridded products were combined into large scale raster mosaics and stored as geotiff products for evaluation purposes. Shaded relief images were produced from each of the surface elevation products to evaluate completeness of coverage, gaps, seams, or other data anomalies. The LiDAR intensity image mosaic was also evaluated for completeness of coverage and quality of data. We observed variability in the LiDAR intensity data that is typical for this type of data. This is due to variations in sunlight conditions at time of acquisition and other factors that are beyond the contractor’s control.

Completeness, Clarity, and Compliance of Metadata

Each metadata file was examined for both content and clarity of the included metadata descriptions. Minor changes were requested when UAF spotted inaccuracies or omissions. In general, UAF found the metadata to be very good. To test for FGDC metadata format compliance, we used the USGS Metadata Parser (MP) program. No metadata errors were reported with testing of sample metadata files from each corridor.

Planimetric Accuracy of the LiDAR Data

Although there was no map identifiable features to survey and evaluate given the remoteness of these corridors, there was attempts to verify consistency of data between the LiDAR intensity mosaic and the ortho imagery products for good alignment of data. Several zoom windows were displayed and flickered between different image layers to verify alignment of data.

Vertical Accuracy of the LiDAR Data

There were nine checkpoints within the Bradfield to Behms Corridor. These included six barren ground, and three forested, for a combined nine points within the Bradfield to Behms Corridor. Using ArcMap, a Bare-Earth DEM mosaic was produced and elevation values for all checkpoints were extracted from the DEM. This was compiled into a spreadsheet and organized by land cover classifications on separate worksheets (see Figure 3). A vertical accuracy assessment was done for each land cover classification and compared to target accuracy specifications included in the LiDAR contract. UAF also looked at the combined class statistics which are included below. A root mean squared error at 95% confidence is used for barren ground, and a 95 percentile was used for all other land cover categories according to USGS methods. The contractual target vertical accuracies are listed to the right of each accuracy measurement. The accuracy measurement is colored green if it passed or red if it failed to meet the requirement. For Bradfield to Behms Corridor, all classes separately and combined were accurate enough to meet the target accuracy specification. **The Bradfield to Behms Corridor easily met the target accuracy specification for all land cover classes.**

Barren Ground Classes			Target
Statistical Summary:	Feet	Meters	Accuracy
Count = 6			
Min =	-0.28	-0.09	
Max =	0.29	0.09	
Mean =	0.01	0.00	
RMSE =	0.23	0.07	< 0.20 m
RMSE*1.96 (95%)	0.65	0.20	
Stddev =	0.23	0.07	

Table 1, Barren Ground Class (FVA) Accuracy Assessment Summary

Forested Classes			Target
Statistical Summary:	Feet	Meters	Accuracy
Count = 3			
Min =	-0.44	-0.14	
Max =	0.65	0.20	
Mean =	0.16	0.05	
RMSE =	0.34	0.10	< 0.20 m
95 Percentile =	0.48	0.15	
Stddev =	0.45	0.14	

Table 2, Forested Class (SVA) Accuracy Assessment Summary

Combined Classes			Target
Statistical Summary:	Feet	Meters	Accuracy
Count = 9			
Min =	-0.44	-0.14	
Max =	0.65	0.20	
Mean =	0.06	0.02	
RMSE =	0.33	0.10	< 0.20 m
95 Percentile =	0.65	0.20	
Stddev =	0.33	0.10	

Table 3, Combined (CVA) Accuracy Assessment Summary

LiDAR Point Cloud Data Density and Classification Accuracy

Point Density was determined using LAS tools provided by Aerometric. The application provides an ability to count point density creating ESRI ASCII GRID files for each tile. Global Mapper was used to read and display all of these grid files for the Bradfield to Behms Corridor (see Figure 4). We had to limit the point density per cell to a maximum of 30 points per meter clamping values greater to that value so that the color map would show sufficient color variation at the low end. Point density is displayed using a color map from blue (low) to red (high). The grid spacing used for the evaluation was 3 feet per pixel as specified in the contract. The First-Return of all valid classes (2-6, and 8-9), excluding withheld bit data classes (1 and 7). At least 90% of the cells should contain at least one LiDAR point. For the Bradfield to Behms Corridor, first-return density was confirmed to exceed 90% for all interior cells of the combined point cloud data.

Each unclassified point cloud was displayed and exported as a gridded surface at one meter posting and saved as a geotiff floating point surface. The data was then converted to a raster grid and each swath was evaluated for the number of swaths that overlap for the Wrangell corridor. The amount of swath overlap is also a very important requirement that needs verification. This maximizes the ability to penetrate dense forest canopy to get an adequate number of ground returns. The overlap for any given location in the corridor should be a minimum of two swaths. A swath overlap of one is colored red, then higher values are shades of green and blue (see Figure 5). For Bradfield to Behms corridor, we confirmed a minimum of two swaths overlap for the entire corridor.

Each classification layer in the LAS point cloud was loaded into Quick Terrain Modeler to verify extent and completeness of coverage, number of points per class, and accuracy of classification. Given the number of points included in some of the larger classifications, it was necessary to split the block into thirds (i.e. N, C, S or W, C, E depending on orientation). Some classes such as the low, medium, and high vegetation and water are texture mapped using a solid color such as light, medium, dark green, and blue respectively. Otherwise, the data is displayed with a color ramp for the elevation range of the corridor. The classifications included in the LAS point cloud data are listed below in order of class with point count totals per class:

Class # - Class Description	Point Count	Reference
Class 1 - Unclassified Data (Marked as Withheld)	6,498,724	Figure 6
Class 2 - Ground	41,481,306	Figure 7
Class 3 - Low Vegetation	193,446,794	Figure 8
Class 4 - Medium Vegetation	123,663,337	Figure 9
Class 5 - High Vegetation	777,667,126	Figure 10
Class 6 - Buildings	25,656	Figure 11
Class 7 - Error Points (Noise)	40,038	Figure 12
Class 8 - Ground Model Keypoints	9,209,385	Figure 13
Class 9 - Water	48,104,079	Figure 14

Table 7, LiDAR Point Cloud Classes Summary Table

Cross validation analysis was performed between each of the class layers and other source data. The individual layer product could be compared to the LiDAR point model stack, keeping in mind that in high vegetation, there may also be returns for medium vegetation, low vegetation. Other layers such as water can be compared to bare Earth elevations of the lakes and rivers by using the contractor supplied polygons of lakes and single and double break line polyline feature files. The elevation of point returns from each vegetation layer was evaluated to confirm that the height range matches the expected classification definition, low vegetation from 0-6 feet, medium vegetation from 6-15 feet, and high vegetation above 15 feet. Point below the minimum height are colored red and heights above the maximum are colored violet (see Figure 15-17). A composite coverage figure can be made by combining the model base color layers such as low, medium, and high vegetation and water can be assembled to show corridor wide coverage, creating a LiDAR point model stack for later comparison to the canopy height classification (see Figure 18).

LiDAR Gridded Products

There are three gridded products being delivered for each corridor. These include the bare-Earth DEM, the first-return DSM, and the LiDAR intensity image (see Figures 19-21). Each of these raster data sets are being delivered in geotiff floating point format. For each type of data, a mosaic product was produced from source tiles from each gridded raster product. Then for the two surface products, a shaded relief image is produced with consistent elevation and angle of the sun to provide consistency (see Figure 22 and 23). These derived images are then evaluated visually by zooming up to a quarter-tile per screen and panning through the mosaic, left to right, and top to bottom. Any height discontinuity between swaths or tiles would show up as a darker linear feature with a common orientation (to detect seams). Any regions of missing data will show up as dark edge to white interior of the missing data (to detect voids). Surface texture gives clues to intermittent vegetation (to detect corn rows) which may or may not be valid depending on the surface type. If anomalies are detected, then comparison to tile edges or swath edges can be made by bringing in other GIS layers for comparison.

Cross validation analysis was performed between each of the class layers and other source data. For example, a canopy height can be estimated by subtracting the bare-Earth surface DEM from the first-return DSM. This product could then be colorized based on a number of classifications to match each of the three vegetation layers (see Figure 24), i.e. 0-6 feet for low vegetation (light green), 6-15 feet for medium vegetation (medium green), and greater than 15 feet for high vegetation (dark green). Barren ground could have height differences from -1 to 1 foot. Other classes could draw attention to outlier heights such as values less than zero (red) and values greater than the tallest expected tree height (violet). The LiDAR point model stack product was compared to the canopy height classification product to verify consistency of data between LiDAR point cloud data and raster gridded data products. It should be pointed out that the graphic display of point data in Quick Terrain Modeler is different than displaying raster classification products in ArcMap, so some differences are expected, it gives a good idea of consistency of coverage in both products.

LiDAR Derived Products

Derived products include a variety of GIS layers including tile index, hydro break lines of lakes and streams, and topographic contours. Since, the primary customer of this data is DOT, the contour data is delivered as CAD CIV-3D Drawing files (.dwg) and Land XML contour files. Related source data was displayed and verified to be seamless and cover the full extent of the corridor. In some cases, disconnected islands, were not contoured if the task was too difficult for the contractor. Dan Ignotov of

DOT played an important supporting role in evaluation and acceptance of CAD data as each corridor was evaluated and approved. The CAD data for each corridor was reworked into road design layers that road design engineers would use to establish a proposed road corridor for future road development projects.

The elevation contours in Land XML format were loaded and displayed in CIV-3D to verify completeness of coverage to project boundary, and contractor supplied break lines within each corridor. Due to memory limitations, it was necessary to load the LandXML data into four adjacent areas to confirm coverage (see Figures 26-29). It was important to verify that drawing names used in each file would not conflict between tiles of a corridor. It was also important to check hydro break lines with adjacent contour data looking for inconsistencies, breaks or seams. **The CAD contour data in LandXML format for the Bradfield to Behms Corridor required a few minor corrections before being accepted.**

Ortho Imagery Products

The ortho-imagery deliverables were evaluated for completeness by Mitch Slife and Dayne Broderson of Geographic Network of Alaska (GINA). Radiometric quality was assessed for cloud/shadow, haze, blend, contrast, saturation, artifact, blurring, ghosting, color balance, and no-data masking. Geometric assessment of CIR and RGB sid and tiff products were made against available U.S. Forestry high resolution Ortho-imagery and USGS digital raster graphic topographic maps. Where this comparison was possible all RTR ortho-imagery was in firm spatial agreement with existing high resolution imagery. The tile based ortho-imagery data was compiled into true color (RGB) and near infrared (NIR) loss-less compression mosaics in both Alaska State Plane 1 and Alaskan Albers projection (see Figure 30-31).

Bradfield to Behms Corridor Results and Recommendations

After significant effort testing, documenting data quality issues, consulting with Tetra Tech and the Alaska DOT staff, and testing two deliveries. UAF is confident that the Bradfield to Behms Corridor is of excellent quality. The spatial extent, coverage provided, vertical accuracy, completeness and consistency of products makes this corridor the first of five corridors we recommend acceptance. Tetra Tech has worked hard to address all of the identified quality issues to date. UAF has thoroughly documented the results of its assessments, including a complete record of all quality issues to date and what the solution was for each case.

Upon completion of writing this report and reviewing the results of our assessments, UAF recommends that the Bradfield to Behms Corridor be accepted. We are very pleased with the quality of data for this corridor. Tetra Tech is applying lessons learned from data quality issues found in each corridor to subsequent corridor deliveries.

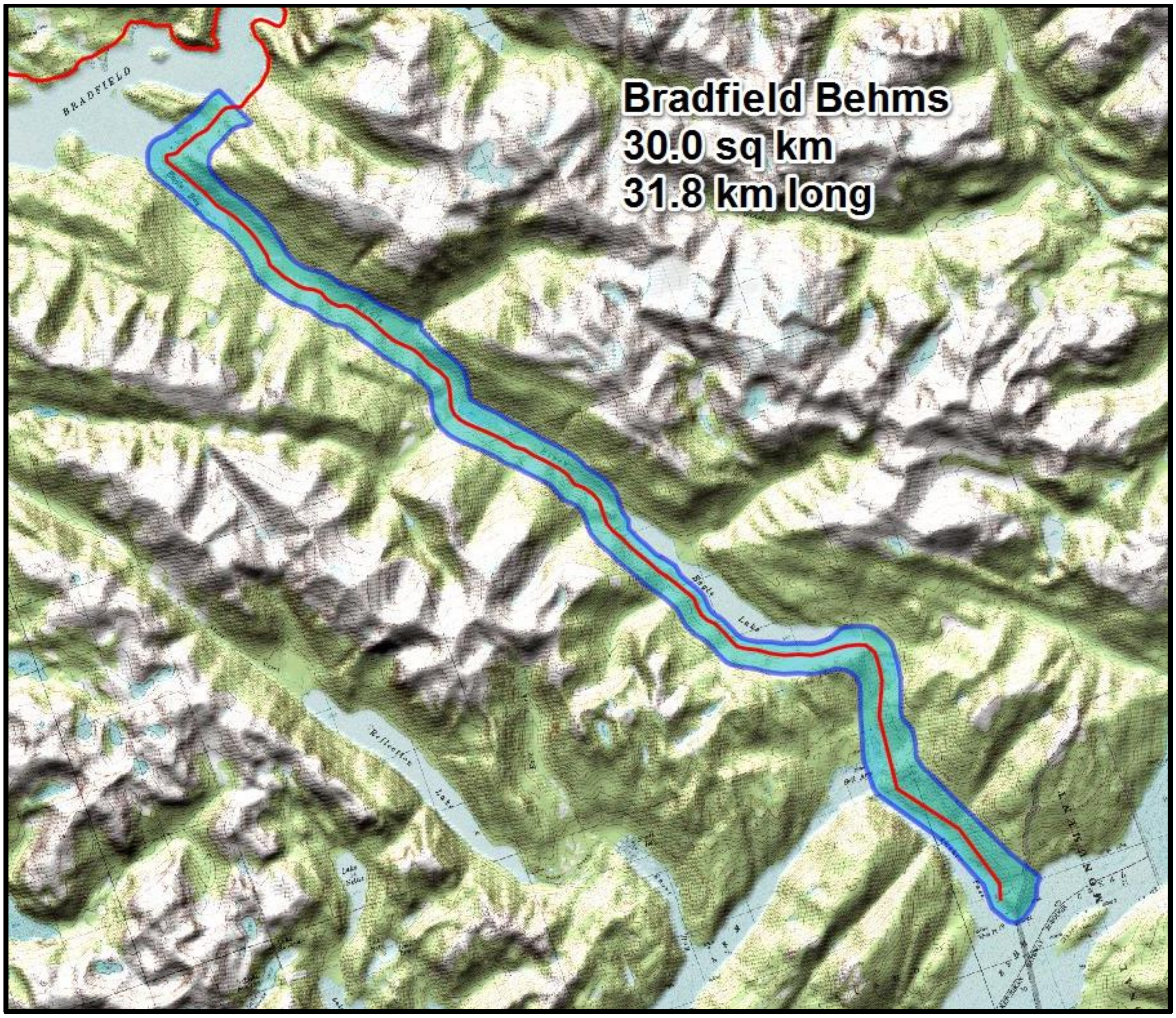


Figure 1, Tetra Tech Bradfield to Behms Corridor Region of Interest (ROI)

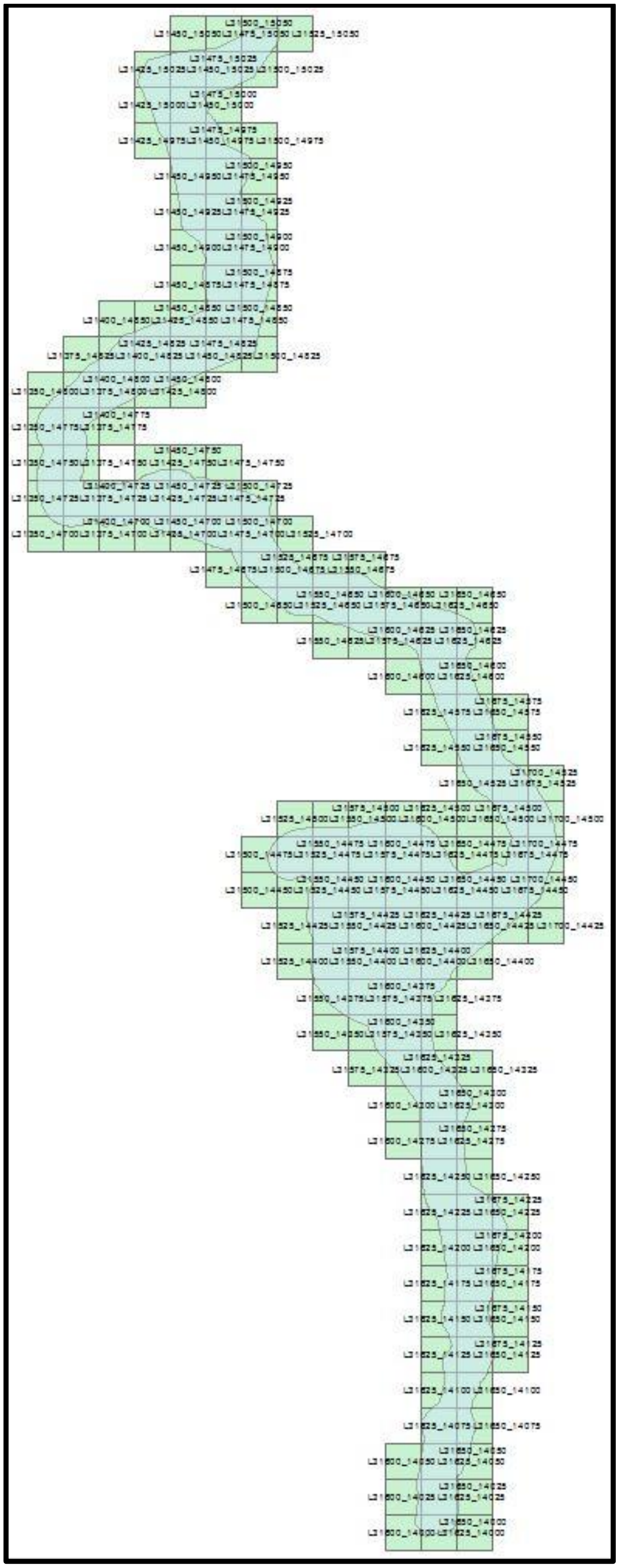
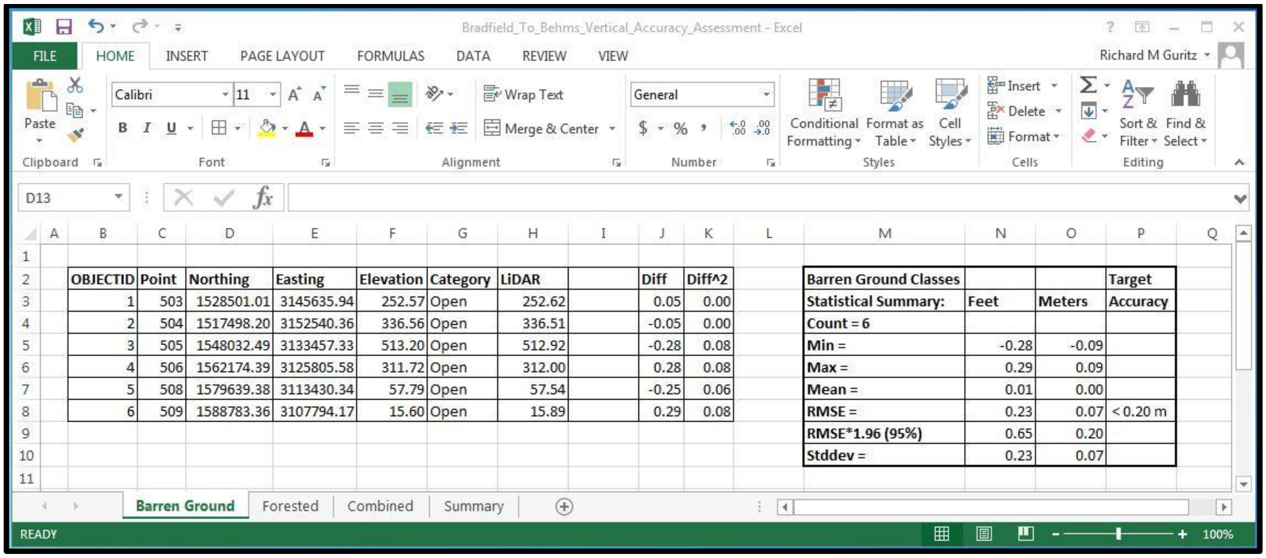


Figure 2, Contractor Supplied GIS Layers. Full tile (green), and project boundary (blue region).



OBJECTID	Point	Northing	Easting	Elevation	Category	LiDAR	Diff	Diff^2
1	503	1528501.01	3145635.94	252.57	Open	252.62	0.05	0.00
2	504	1517498.20	3152540.36	336.56	Open	336.51	-0.05	0.00
3	505	1548032.49	3133457.33	513.20	Open	512.92	-0.28	0.08
4	506	1562174.39	3125805.58	311.72	Open	312.00	0.28	0.08
5	508	1579639.38	3113430.34	57.79	Open	57.54	-0.25	0.06
6	509	1588783.36	3107794.17	15.60	Open	15.89	0.29	0.08

Barren Ground Classes			Target
Statistical Summary:	Feet	Meters	Accuracy
Count = 6			
Min =	-0.28	-0.09	
Max =	0.29	0.09	
Mean =	0.01	0.00	
RMSE =	0.23	0.07	< 0.20 m
RMSE*1.96 (95%)	0.65	0.20	
Stddev =	0.23	0.07	

Figure 3, Vertical Accuracy Assessment using Bill McClintock Checkpoint Survey.

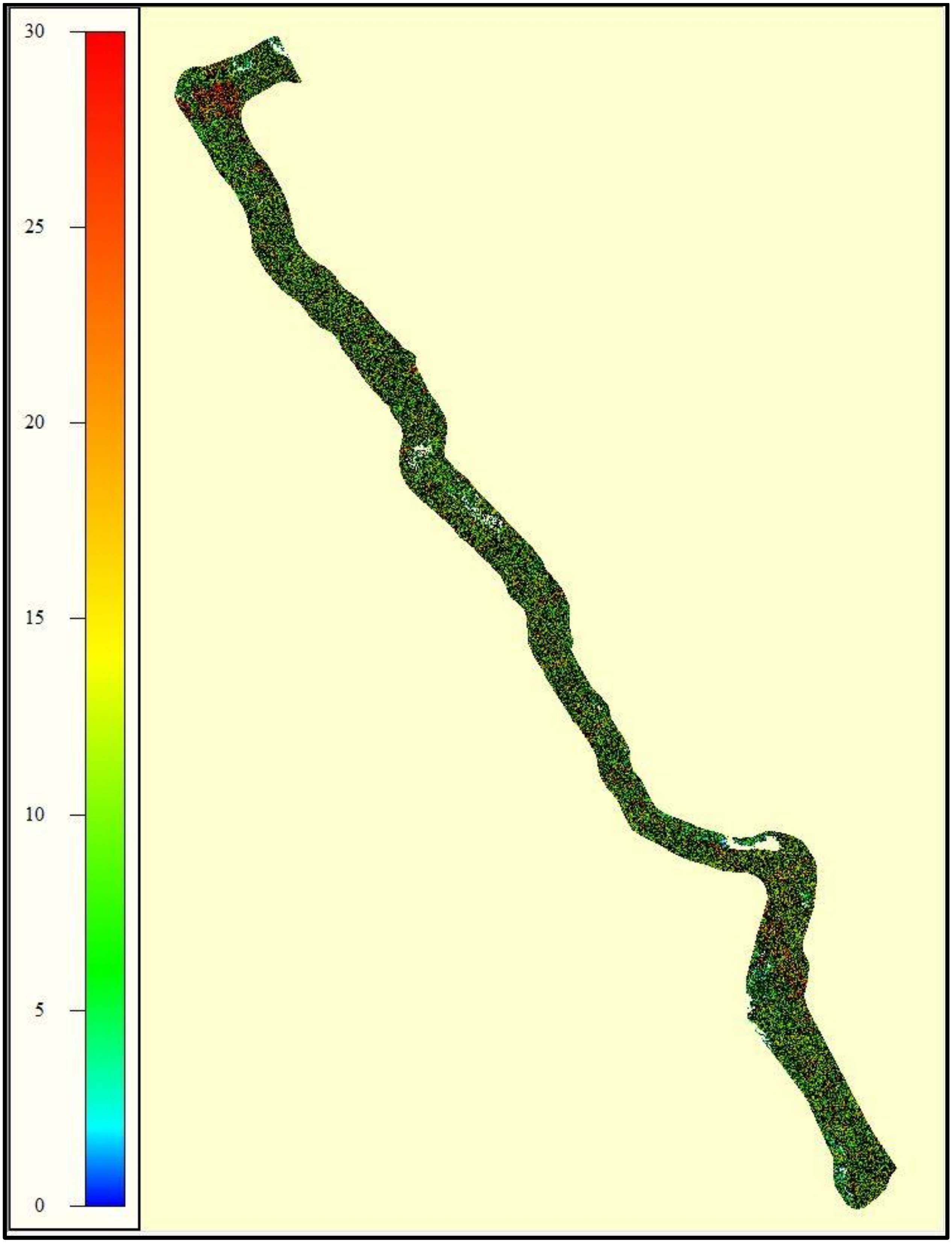


Figure 4, LiDAR Density - Point Count

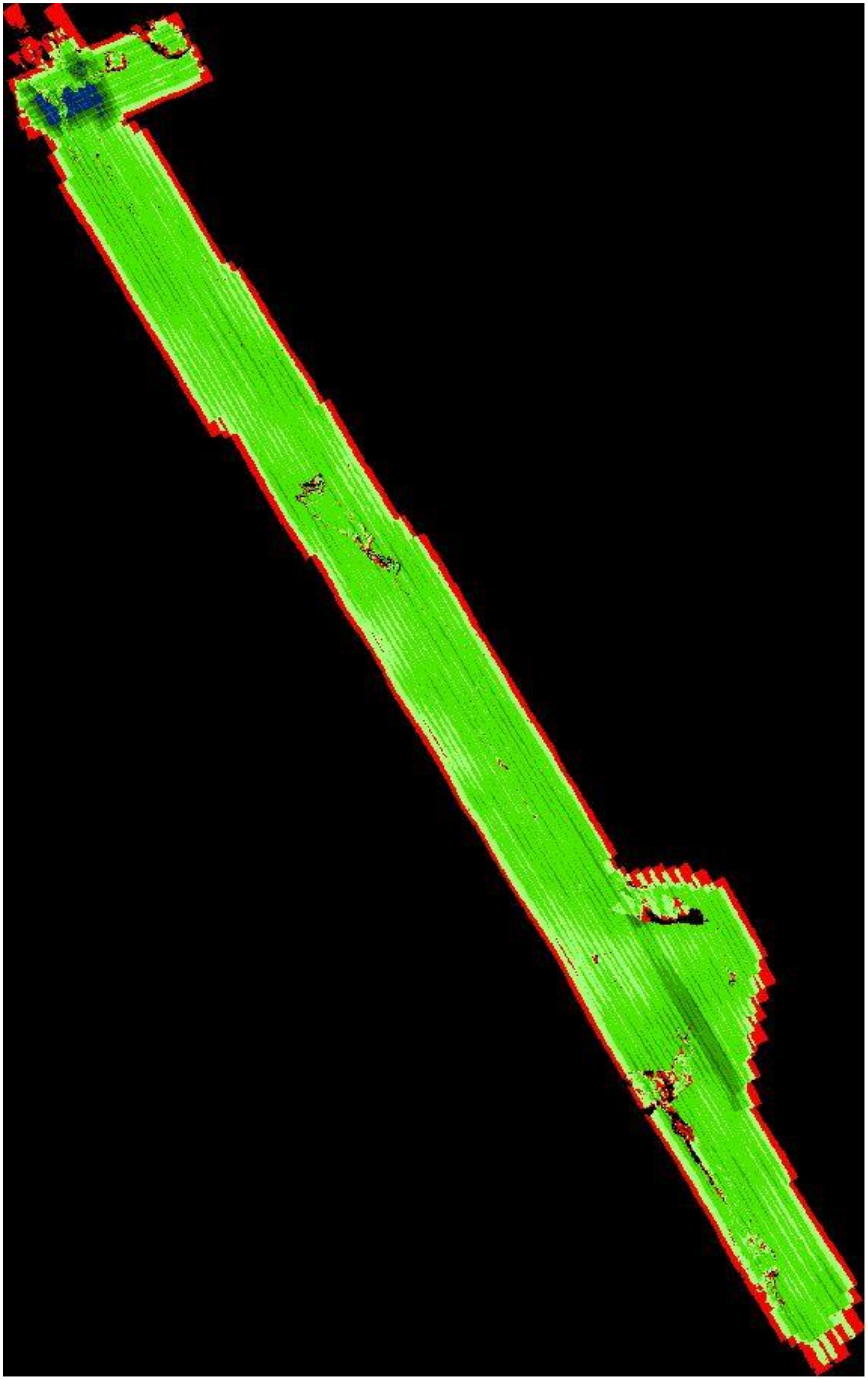


Figure 5, Swath Overlap Analysis

Figure 6, Class 1 – Unclassified.

Figure 7, Class 2 – Ground.

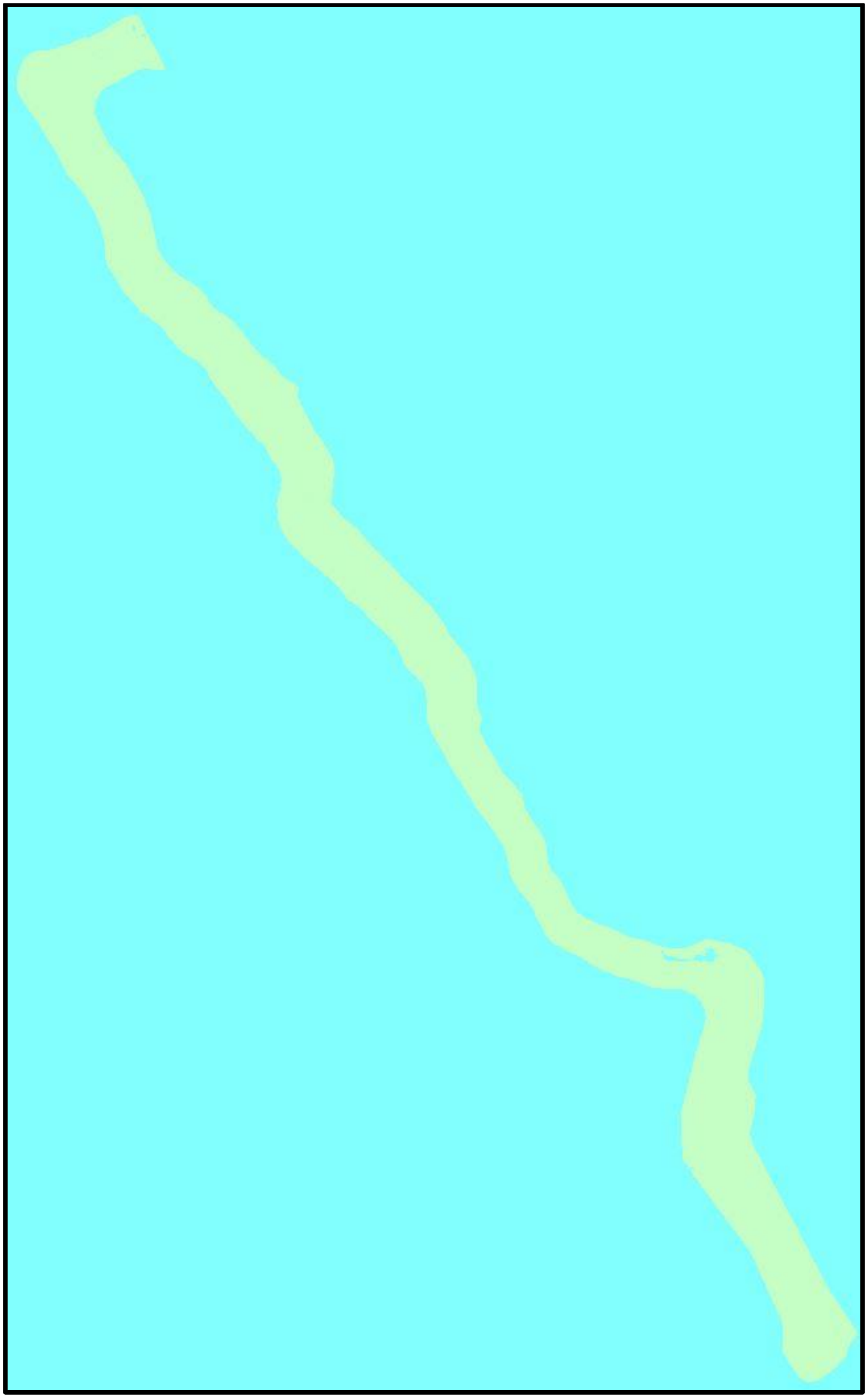


Figure 8, Class 3 – Low Vegetation.

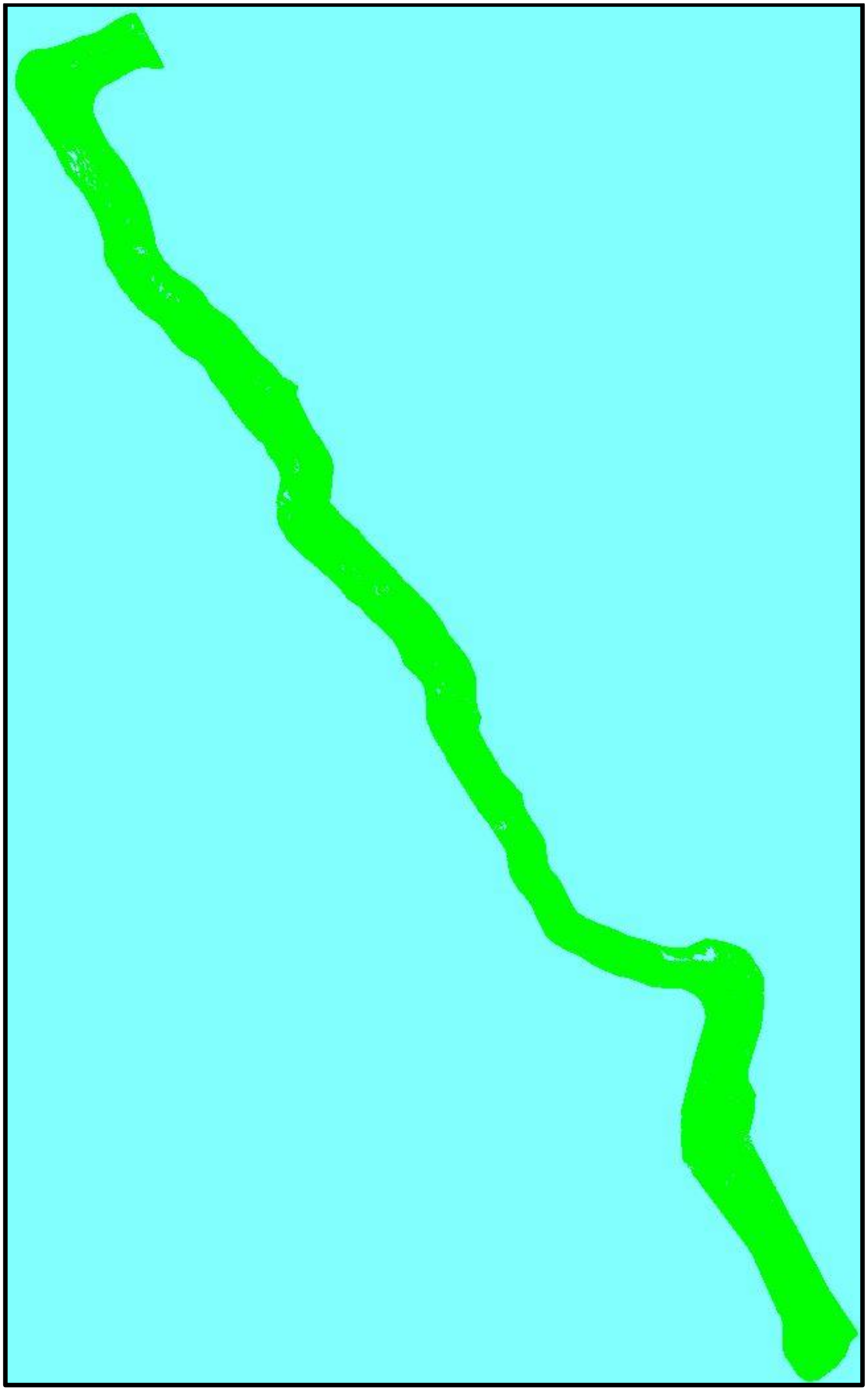


Figure 9, Class 4 – Medium Vegetation

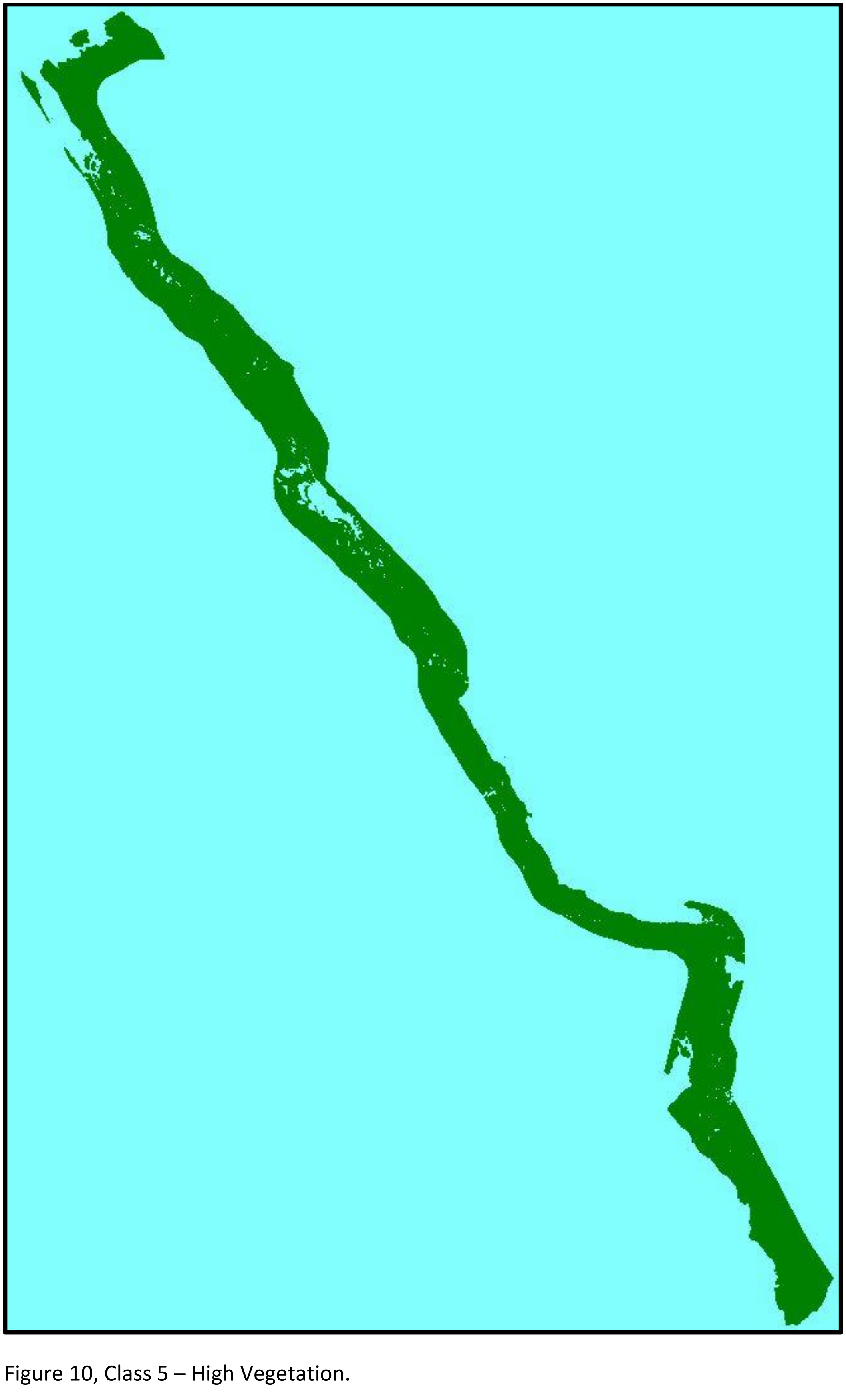


Figure 10, Class 5 – High Vegetation.

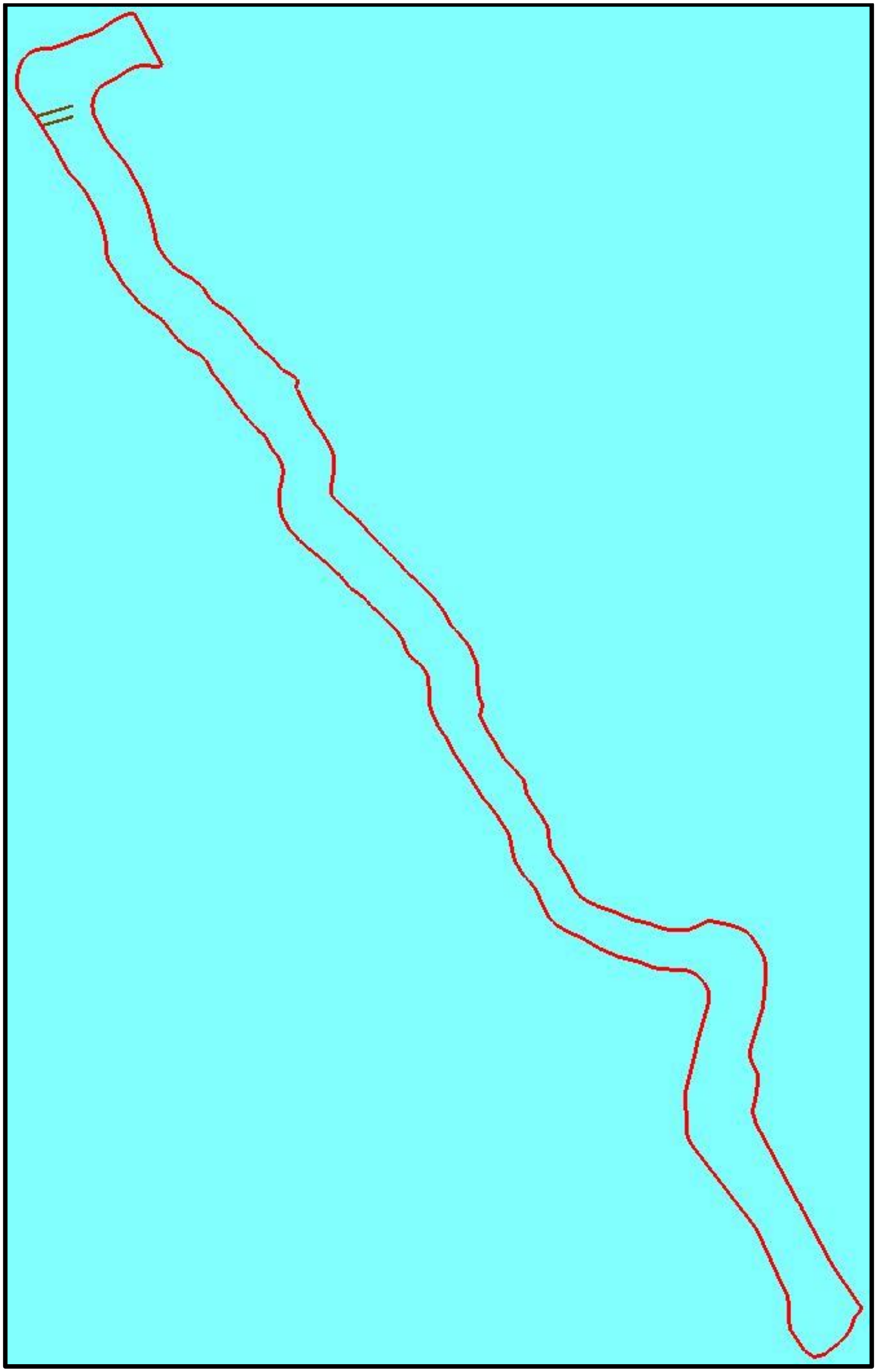


Figure 11, Class 6 – Buildings.

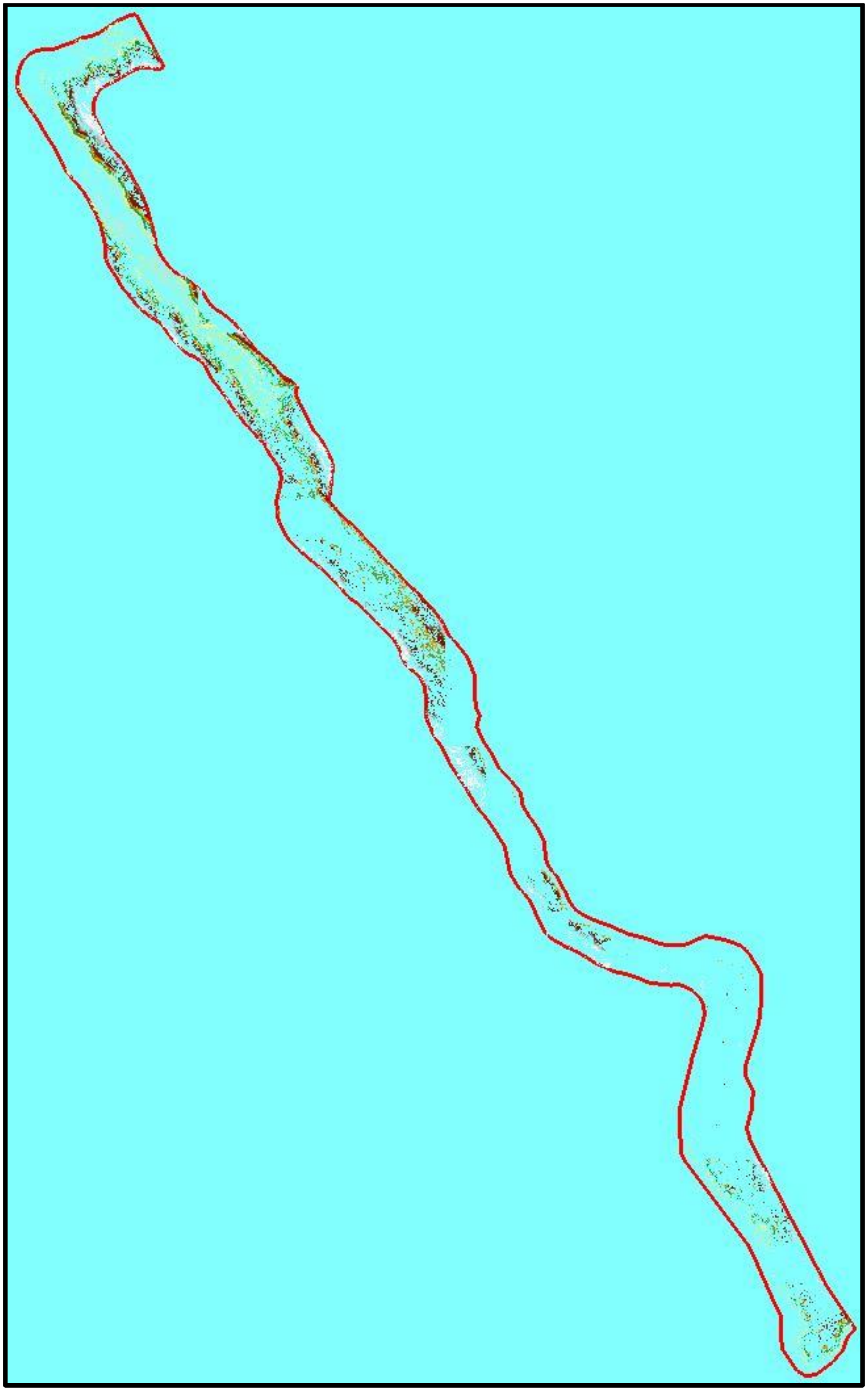


Figure 12, Class 7 – Noise.

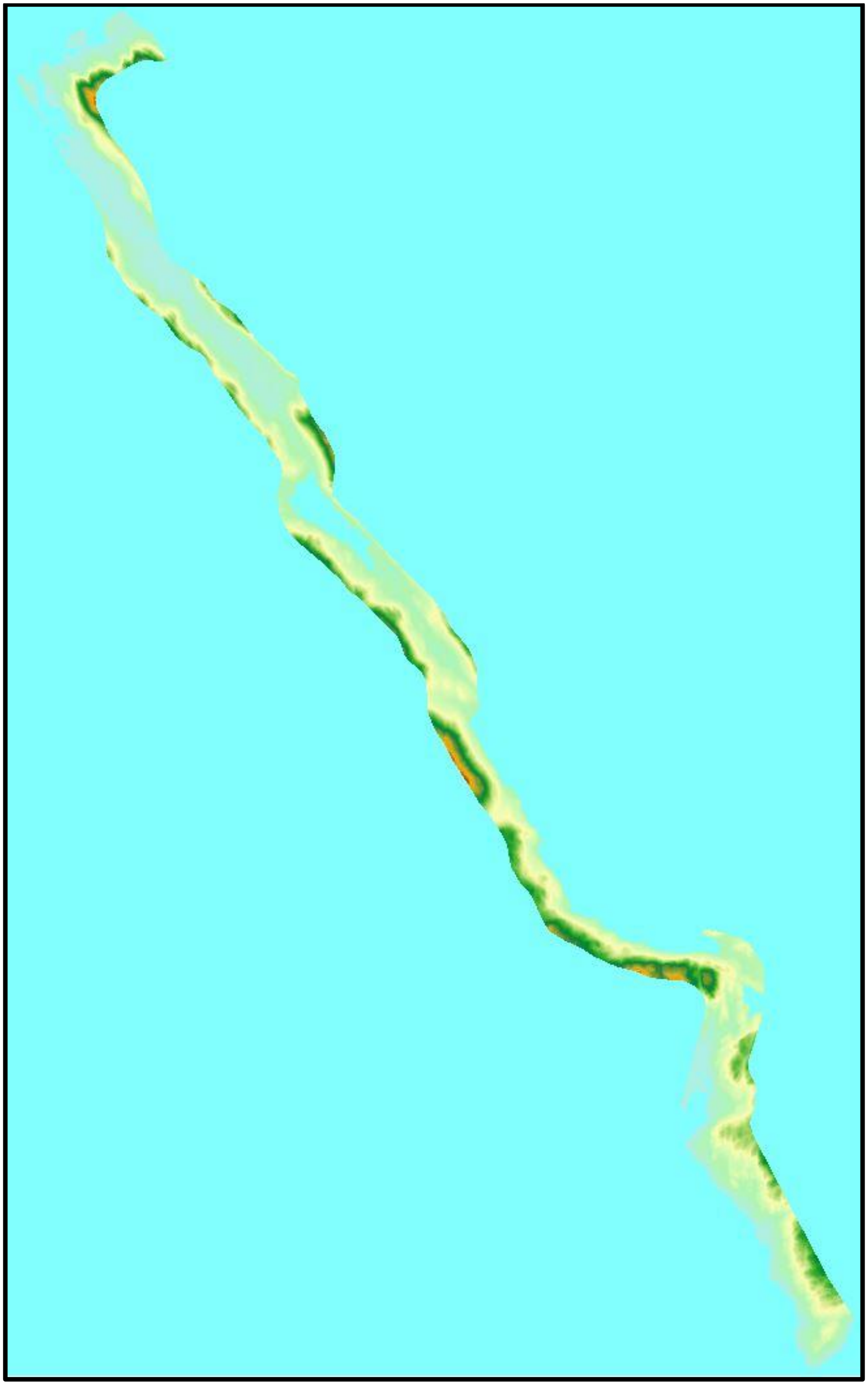


Figure 13, Class 8 – Ground Model Key Points.

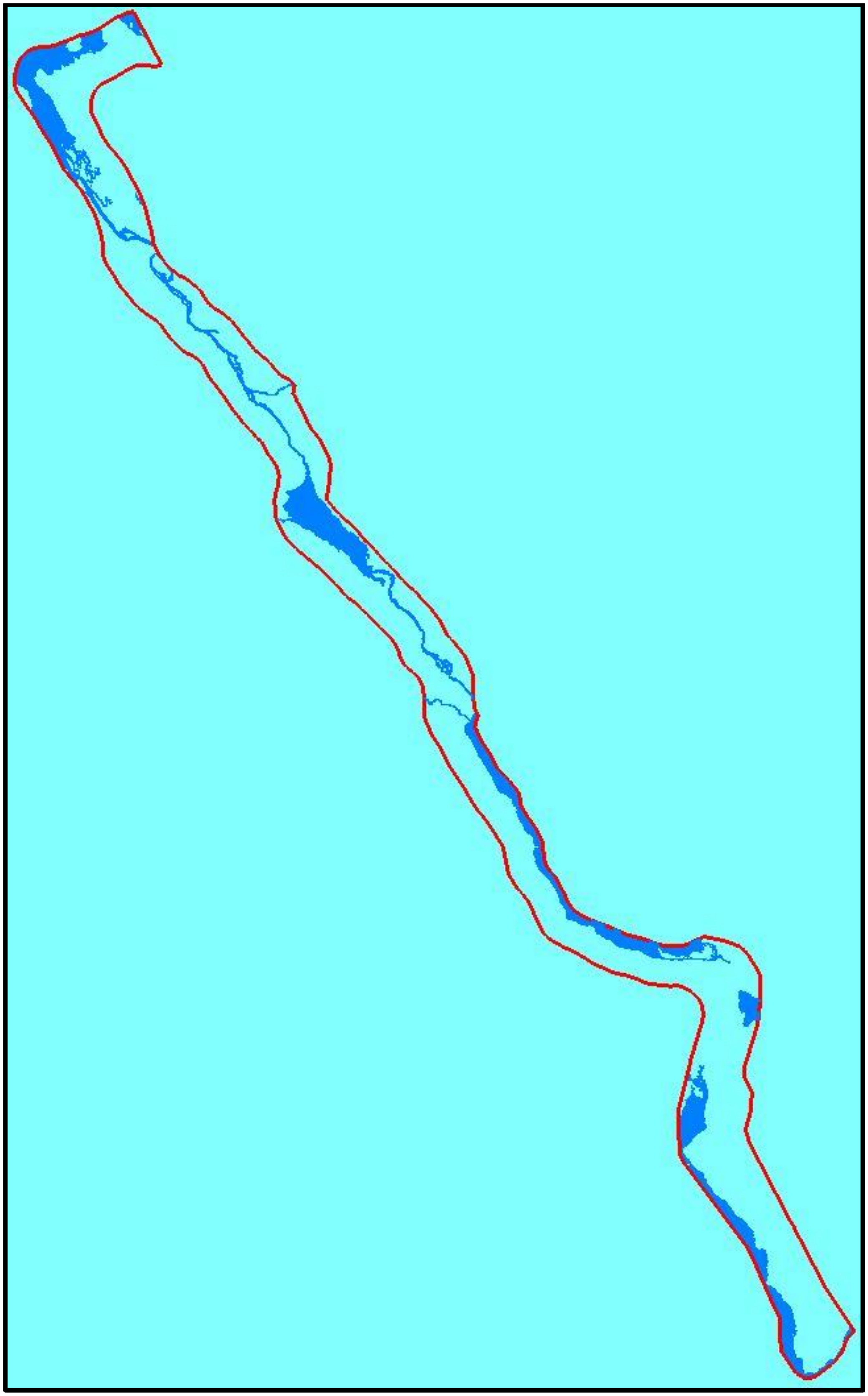


Figure 14, Class 9 – Water.

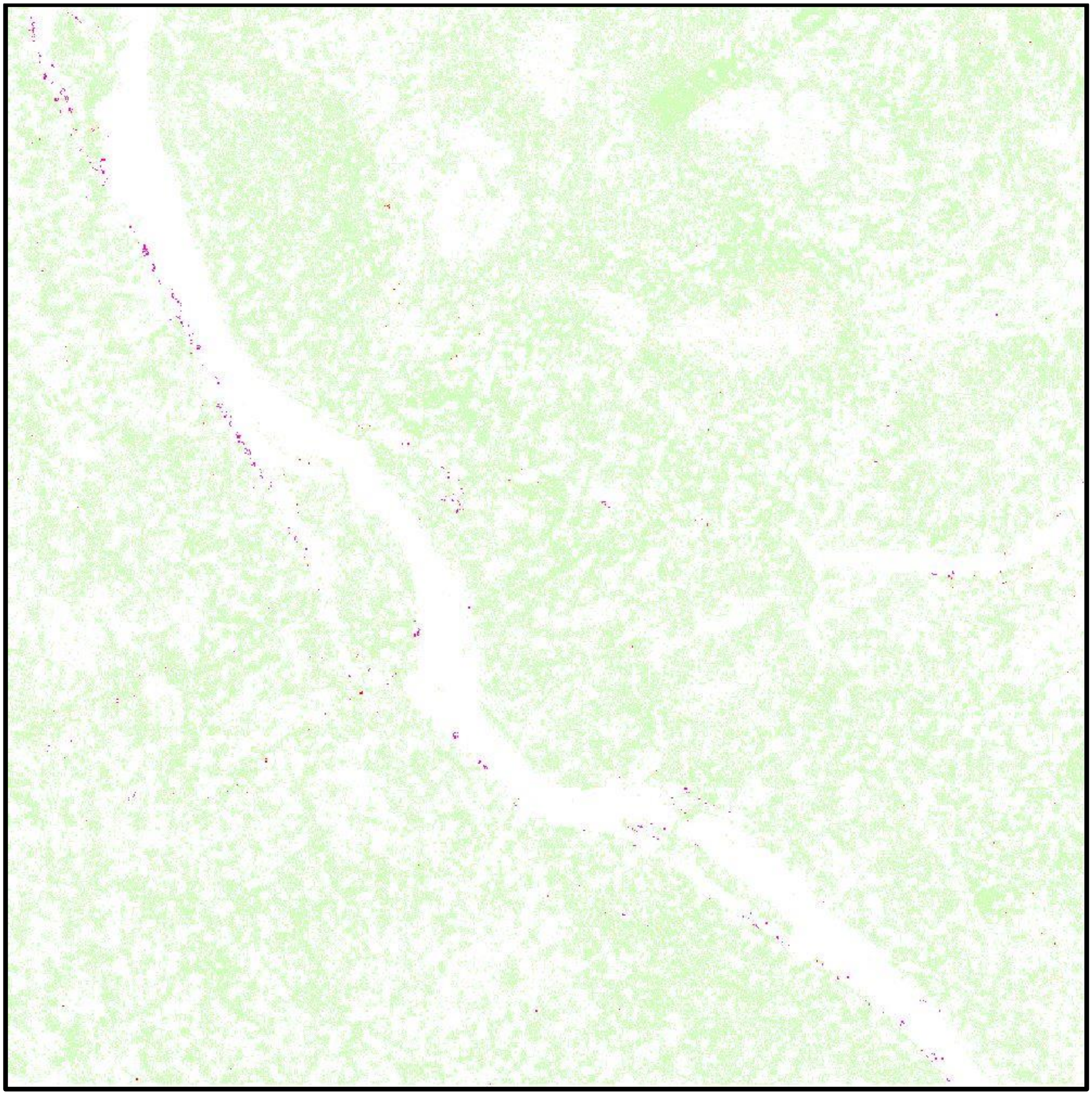


Figure 15, Low Vegetation LiDAR Return Height Check.

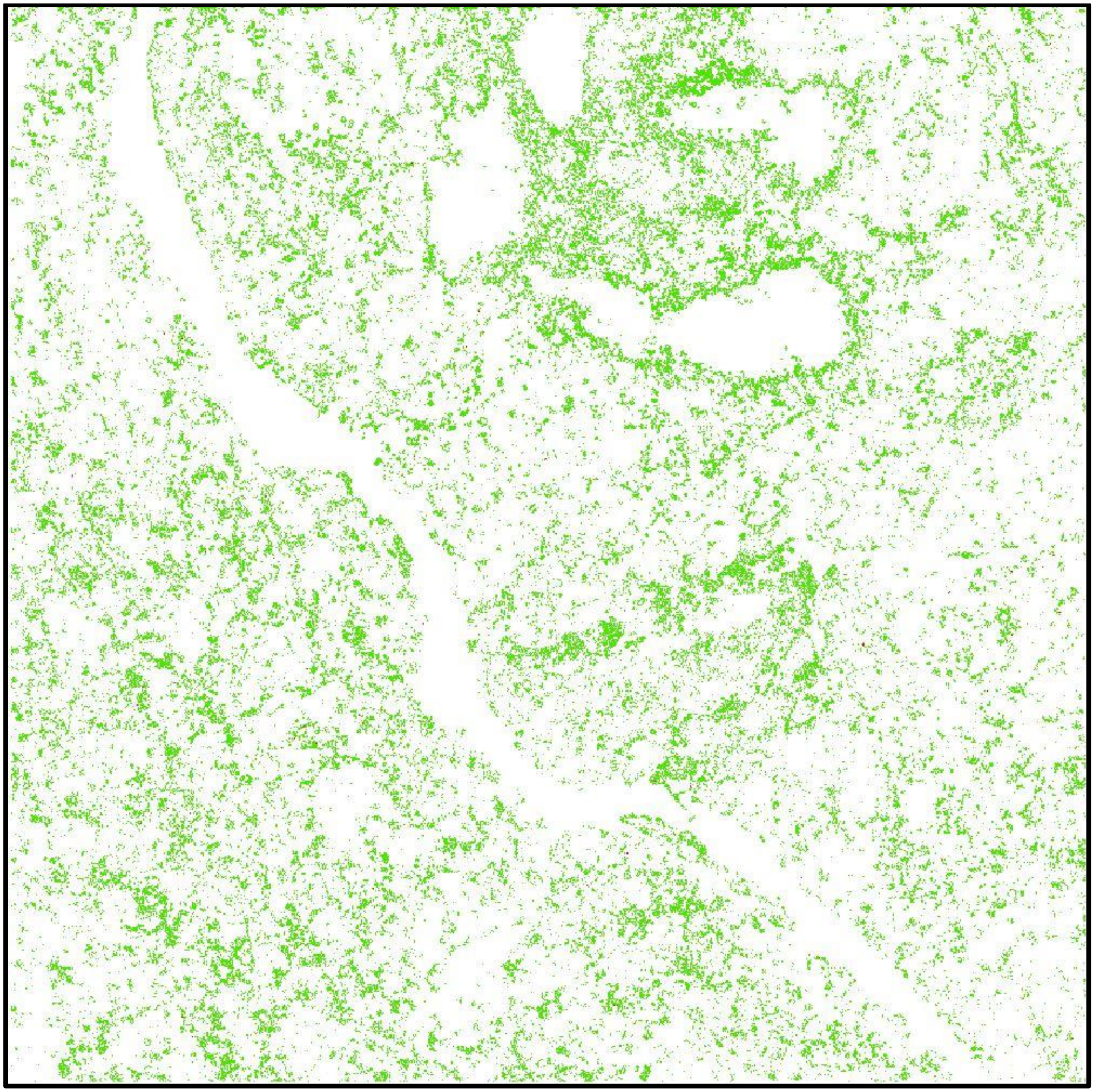


Figure 16, Medium Vegetation LiDAR Return Height Check.

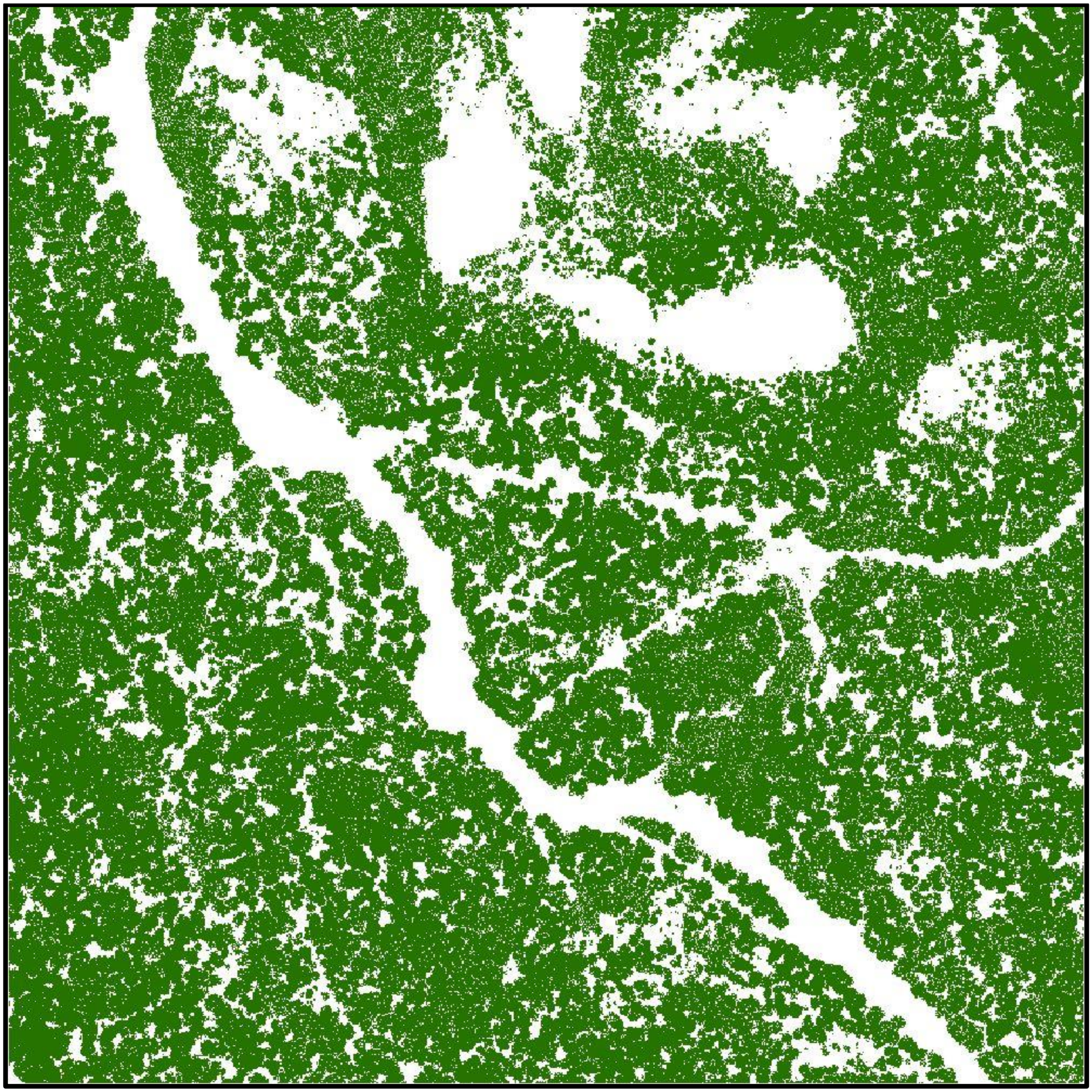


Figure 17, High Vegetation LiDAR Return Height Check.

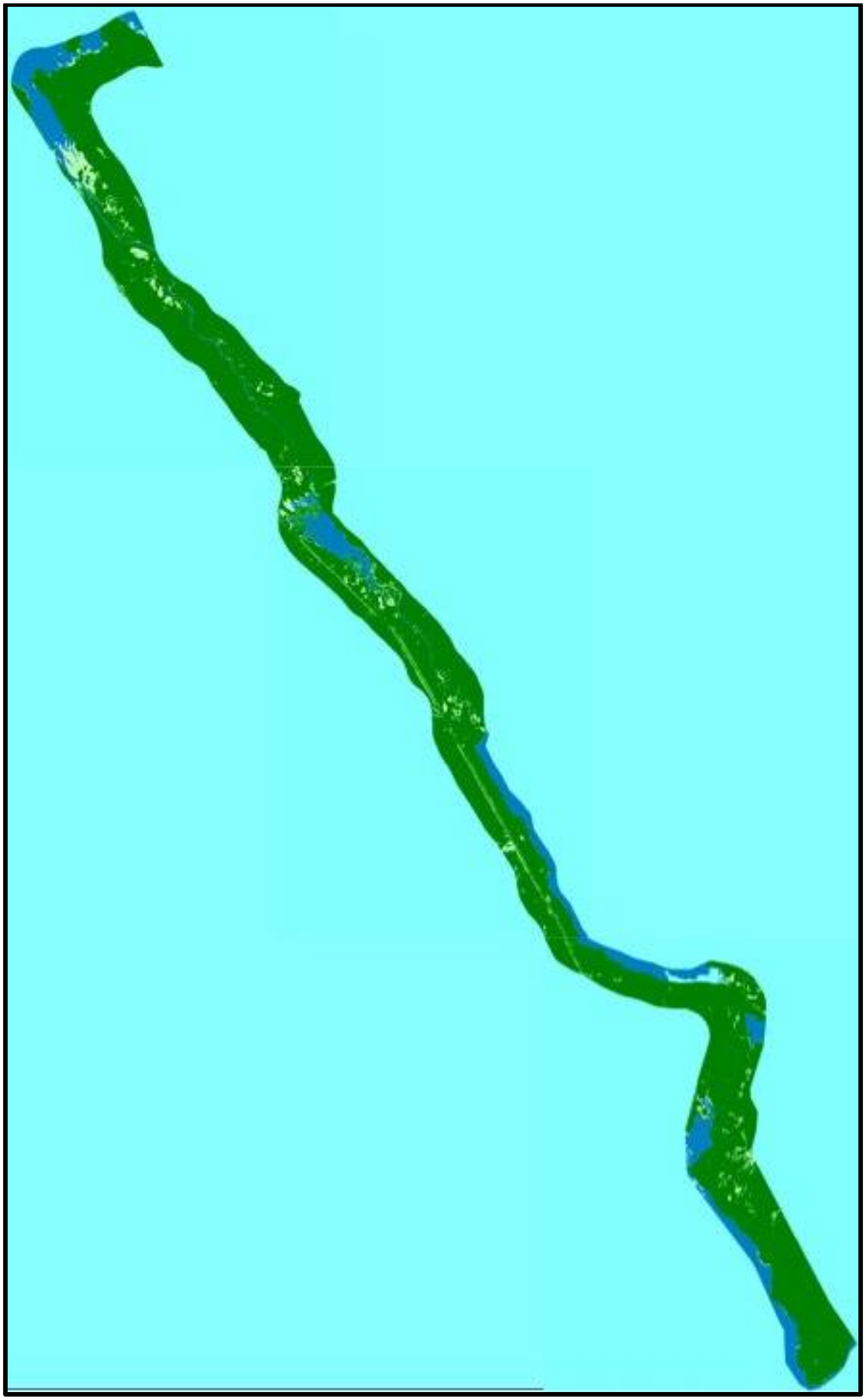


Figure 18, Stack of models from classified point cloud (water, low veg, medium veg, and high veg).

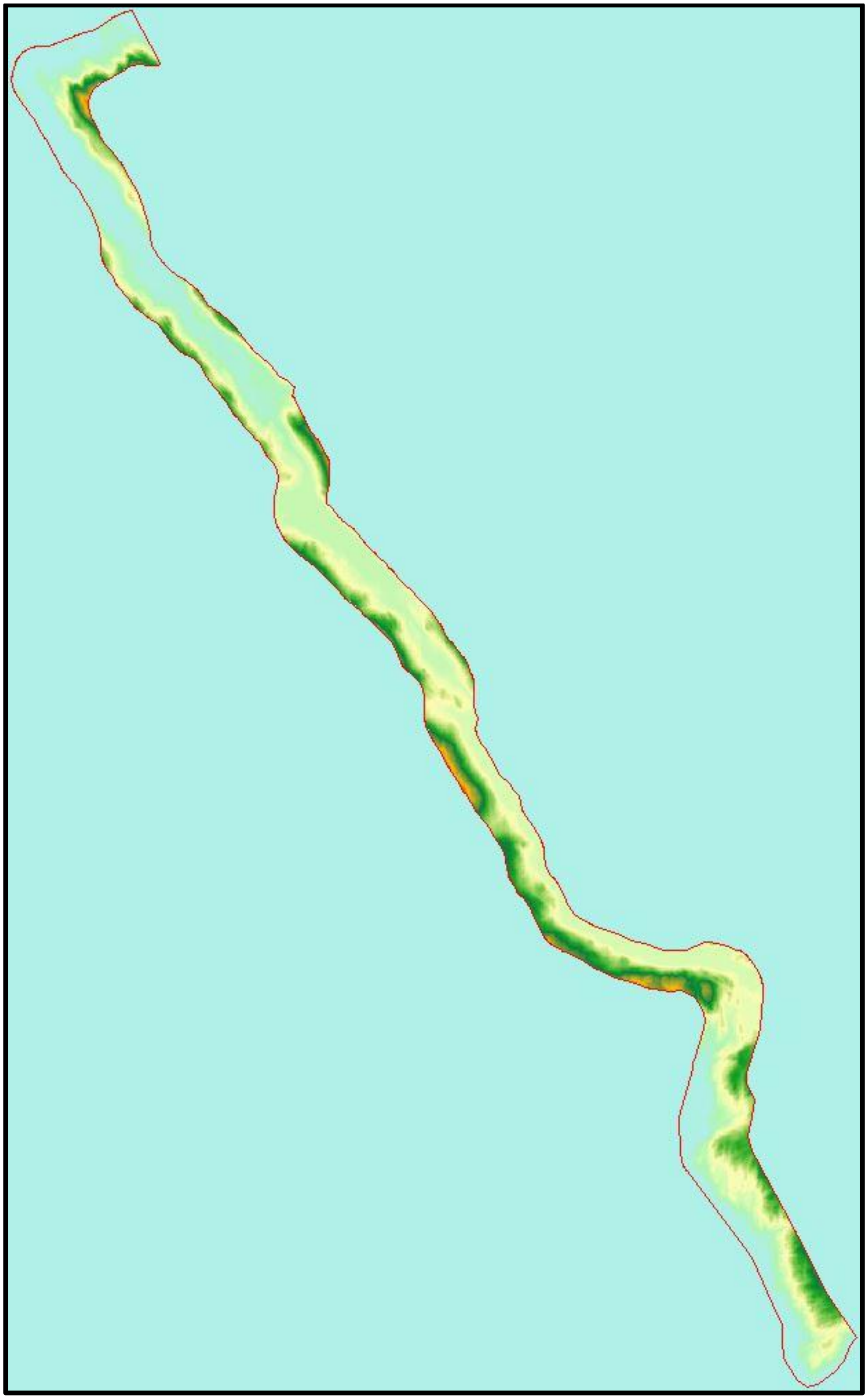


Figure 19, Bare-Earth Gridded DEM Mosaic.

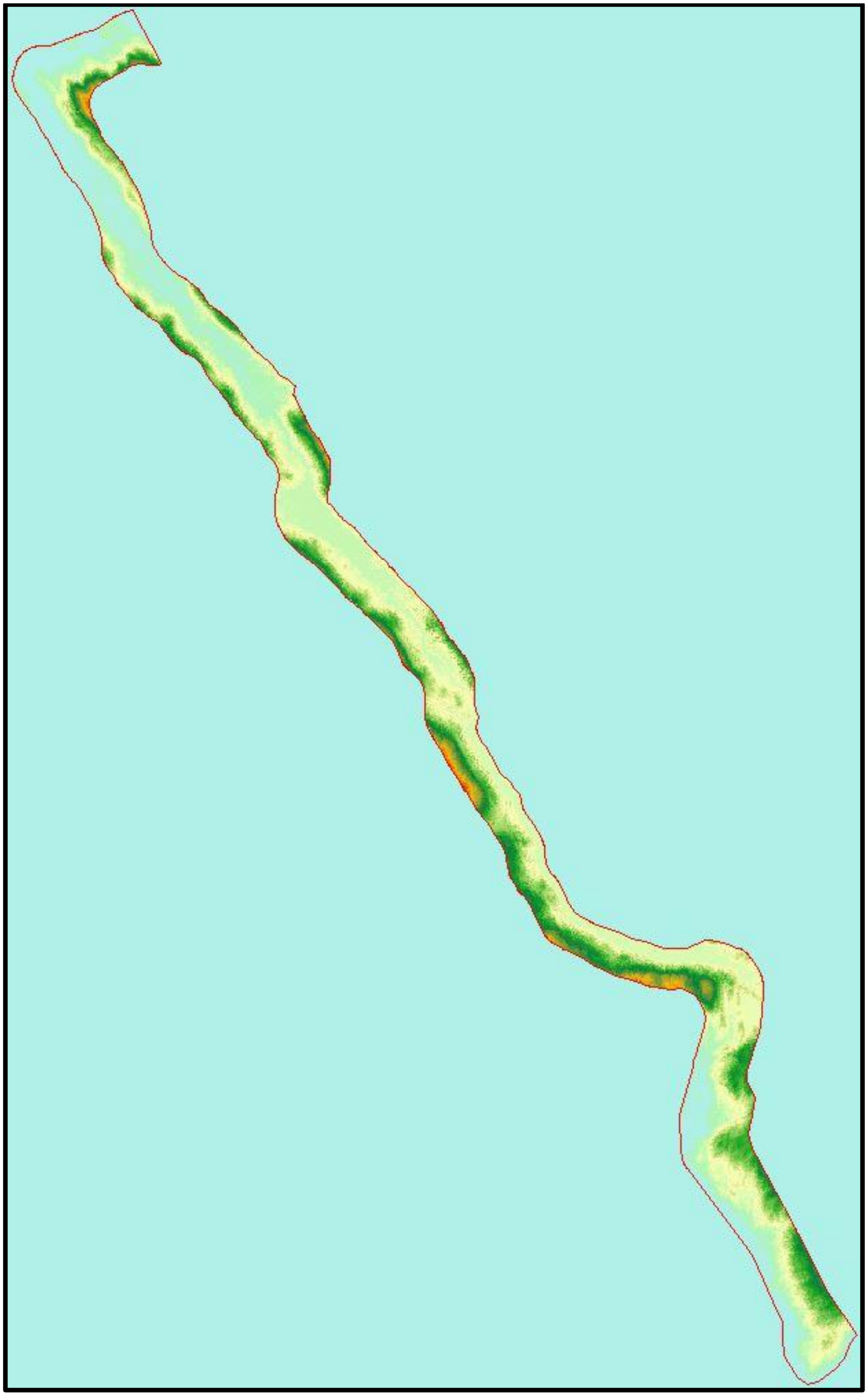


Figure 20, First-Return Gridded DSM.

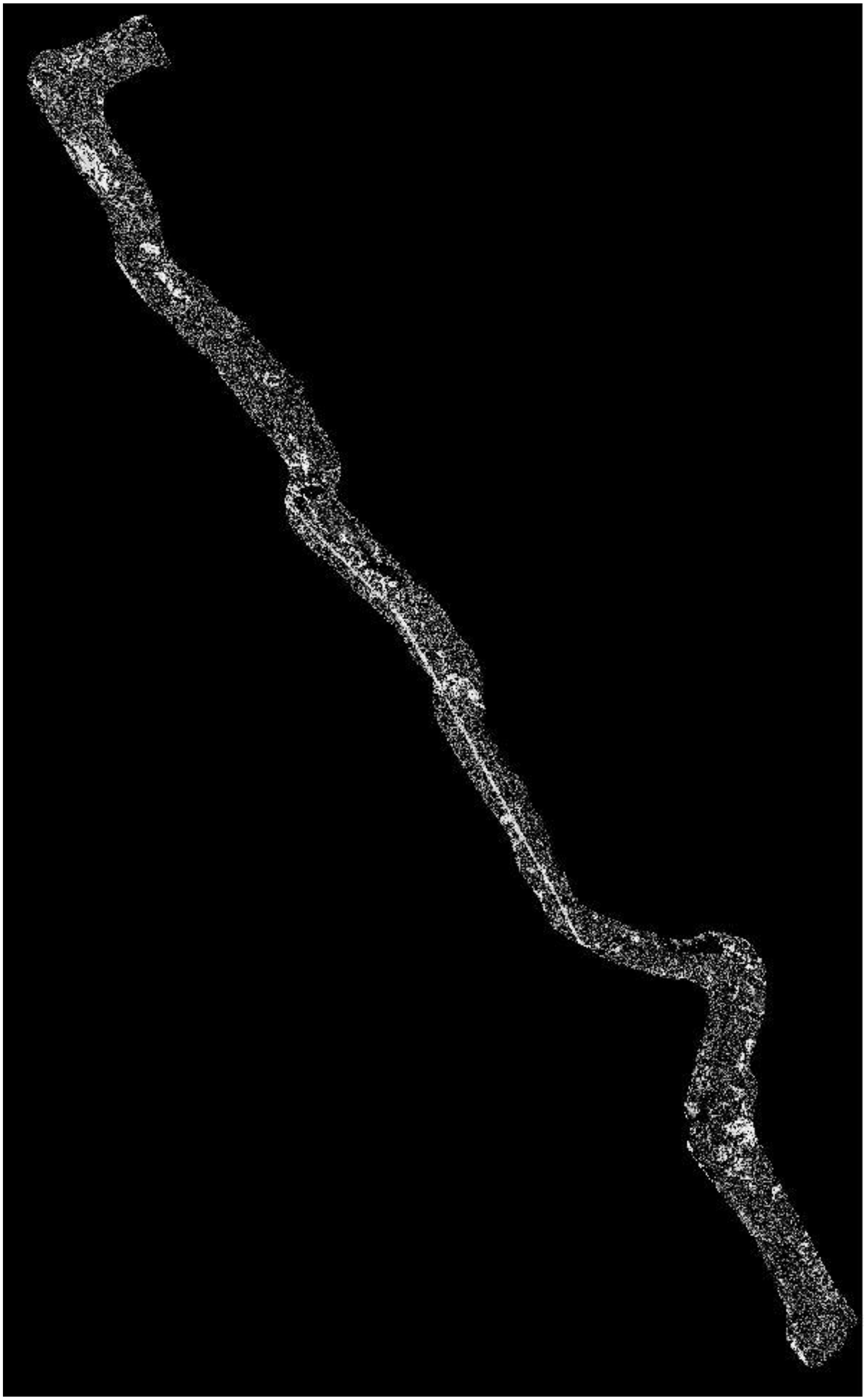


Figure 21, LiDAR Intensity Gridded Mosaic.

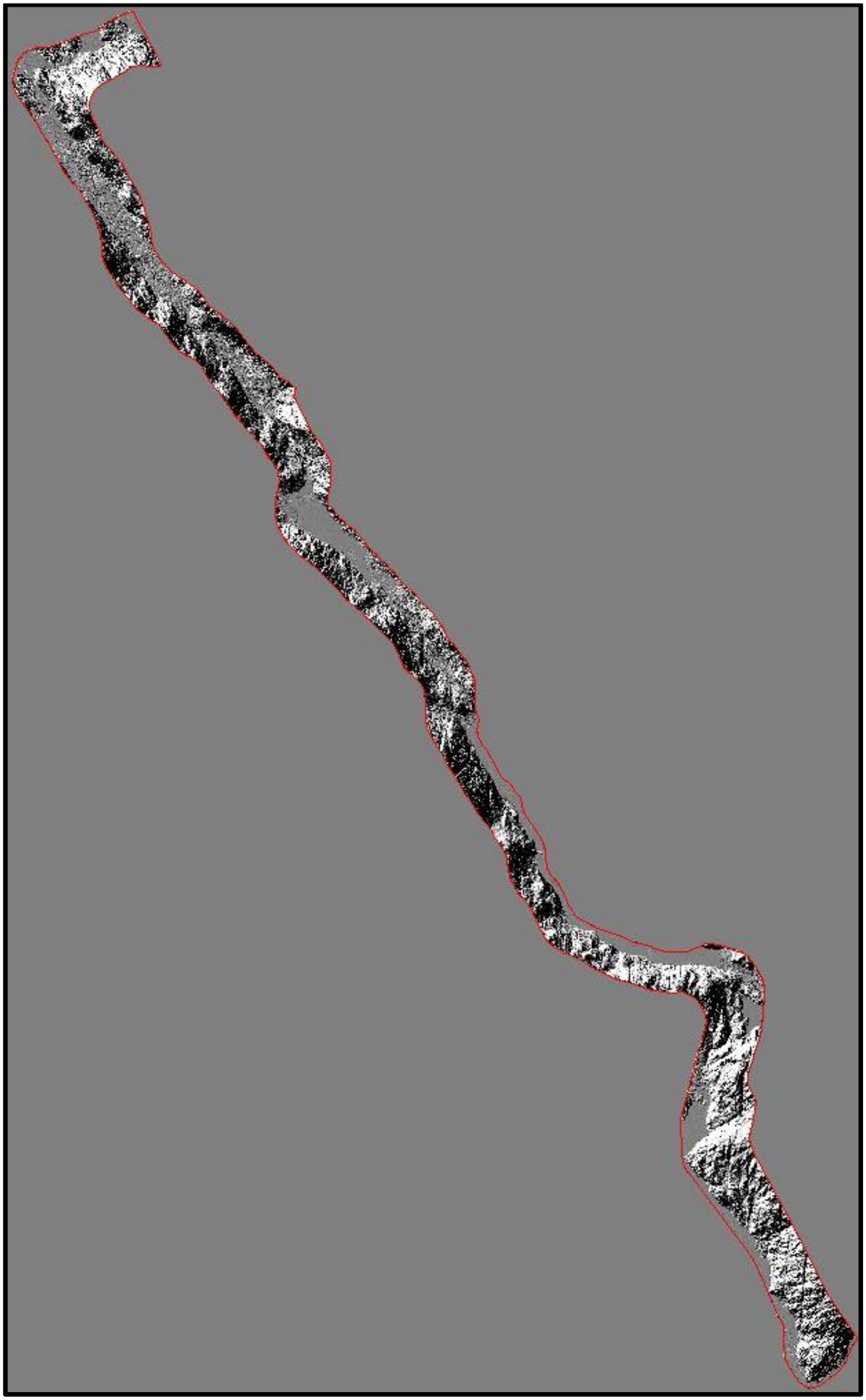


Figure 22, Shaded Relief Bare-Earth Gridded DEM.

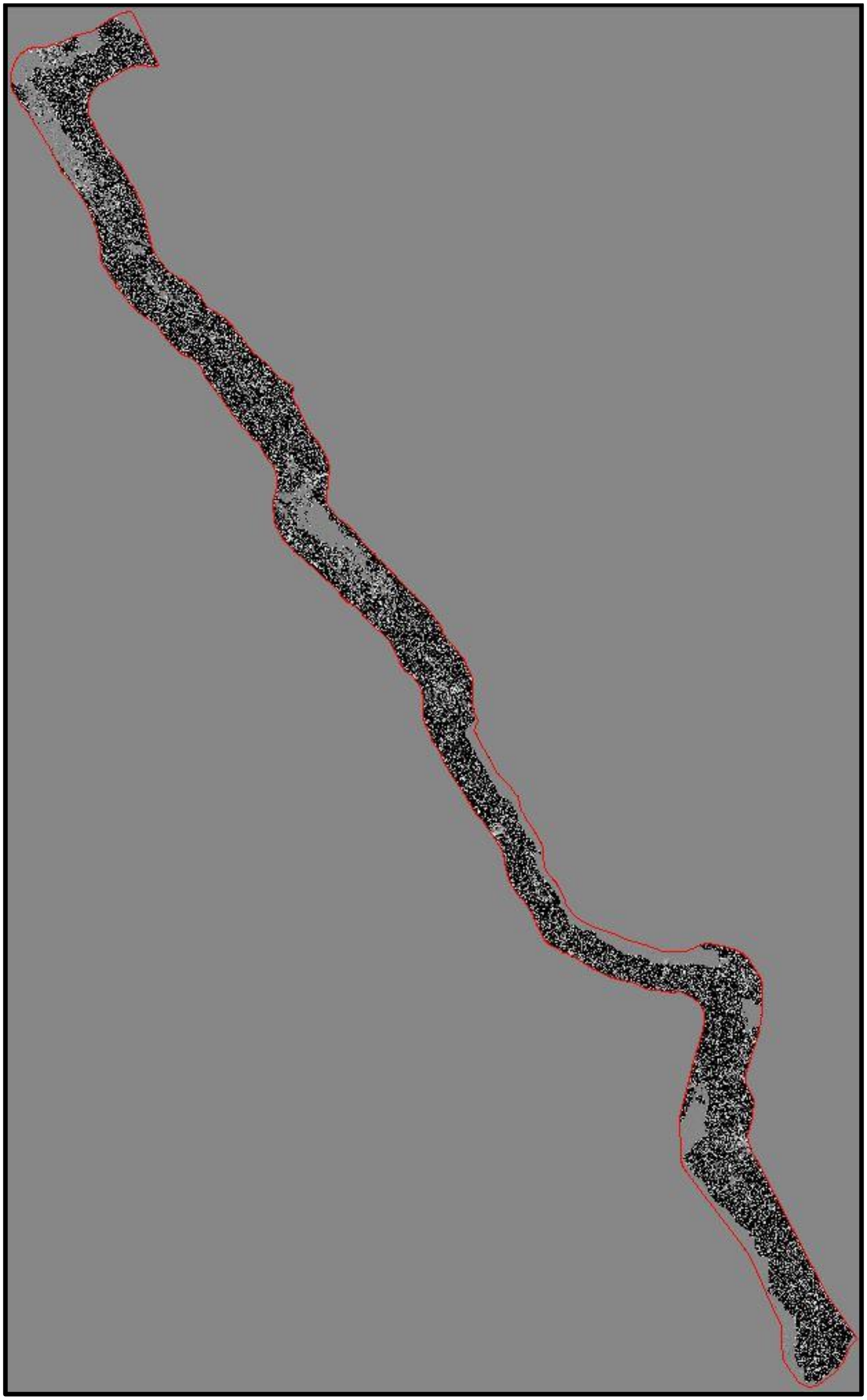


Figure 23, Shaded Relief First-Return Gridded DSM.

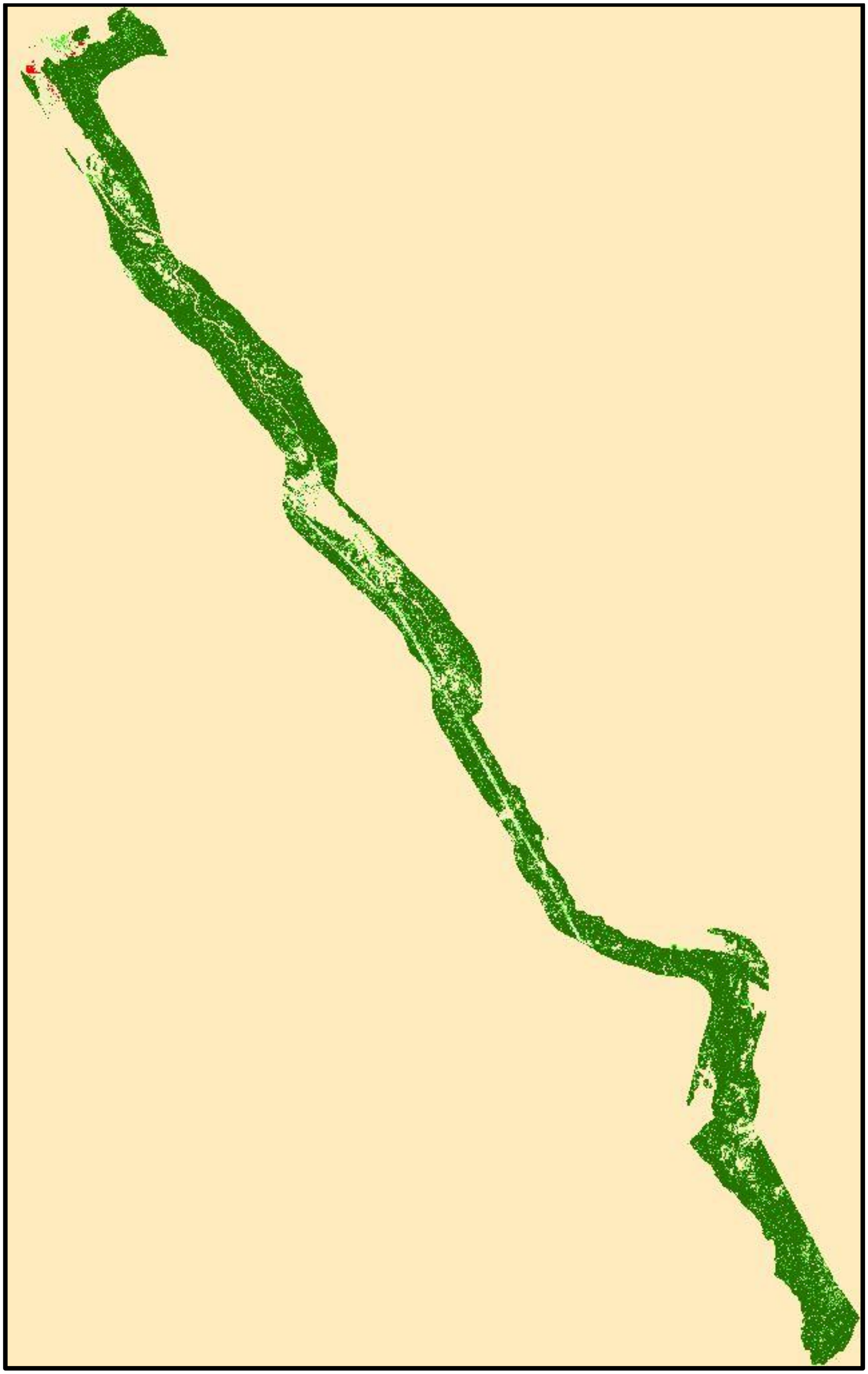


Figure 24, Canopy Height Classification.

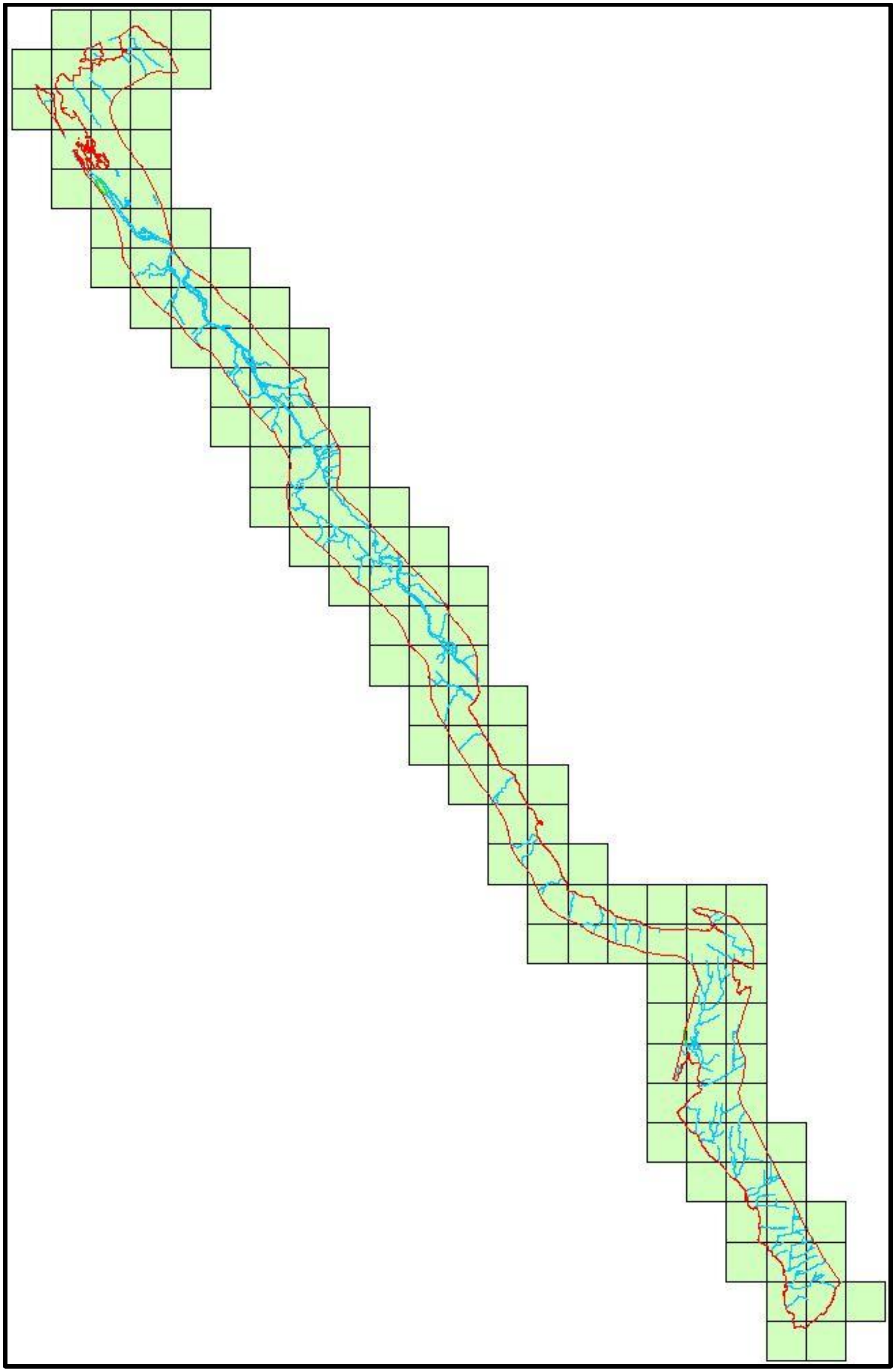


Figure 25, Contractor Supplied Hydro and Boundary CAD Layers.

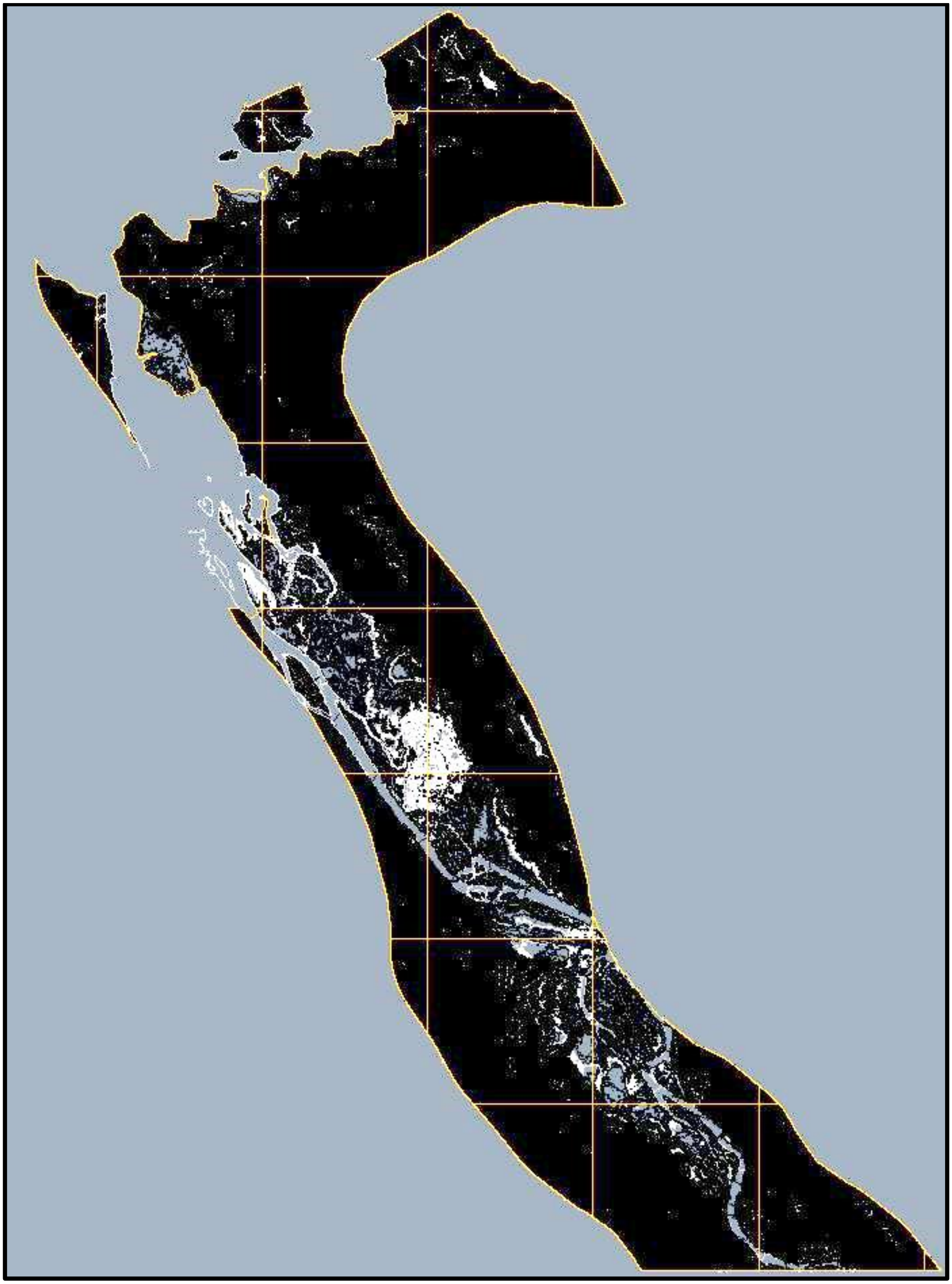


Figure 26, Bradfield to Behms LandXML Contours 15775 to 15950.

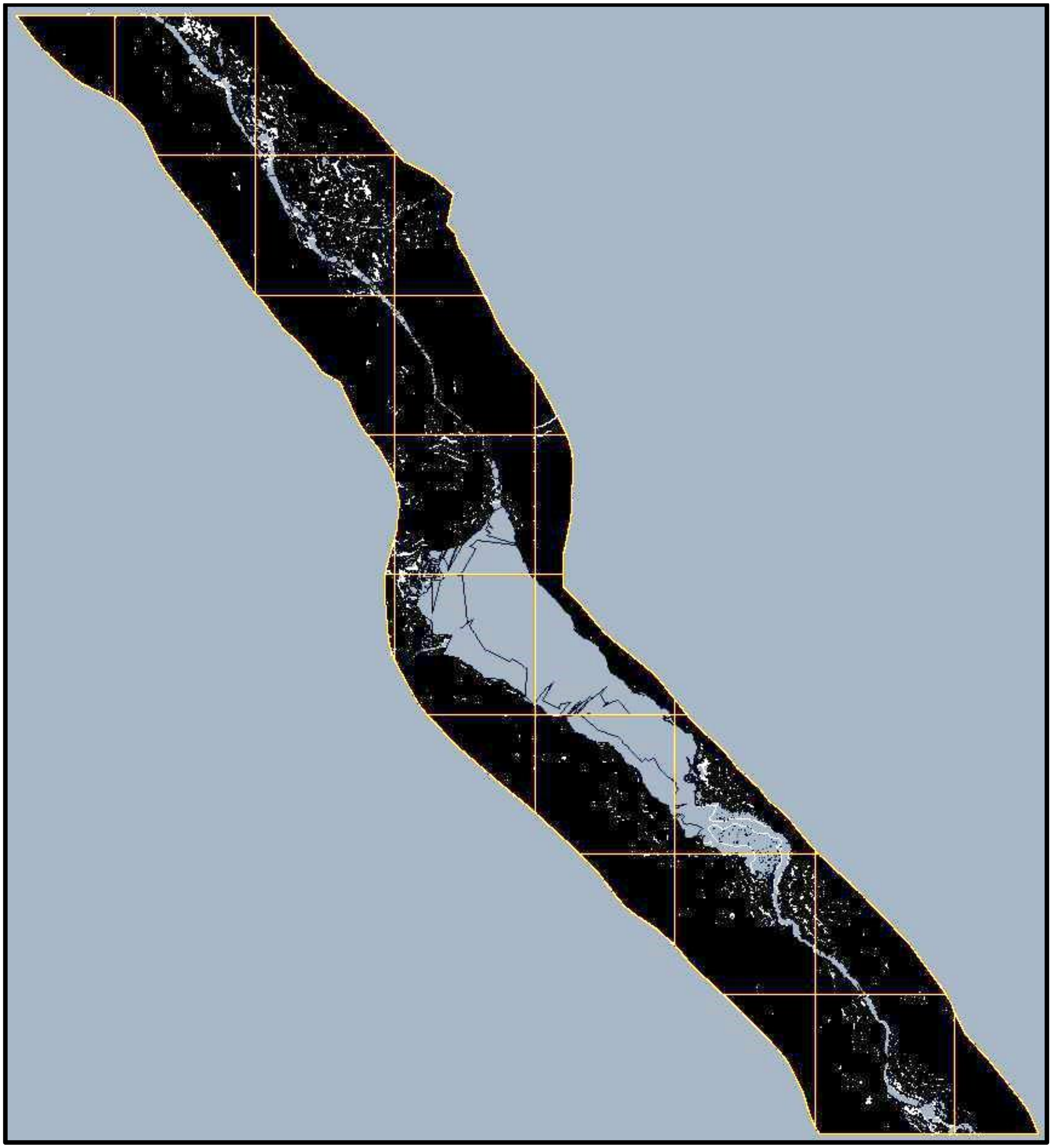


Figure 27, Bradfield to Behms LandXML Contours 15575 to 15750.

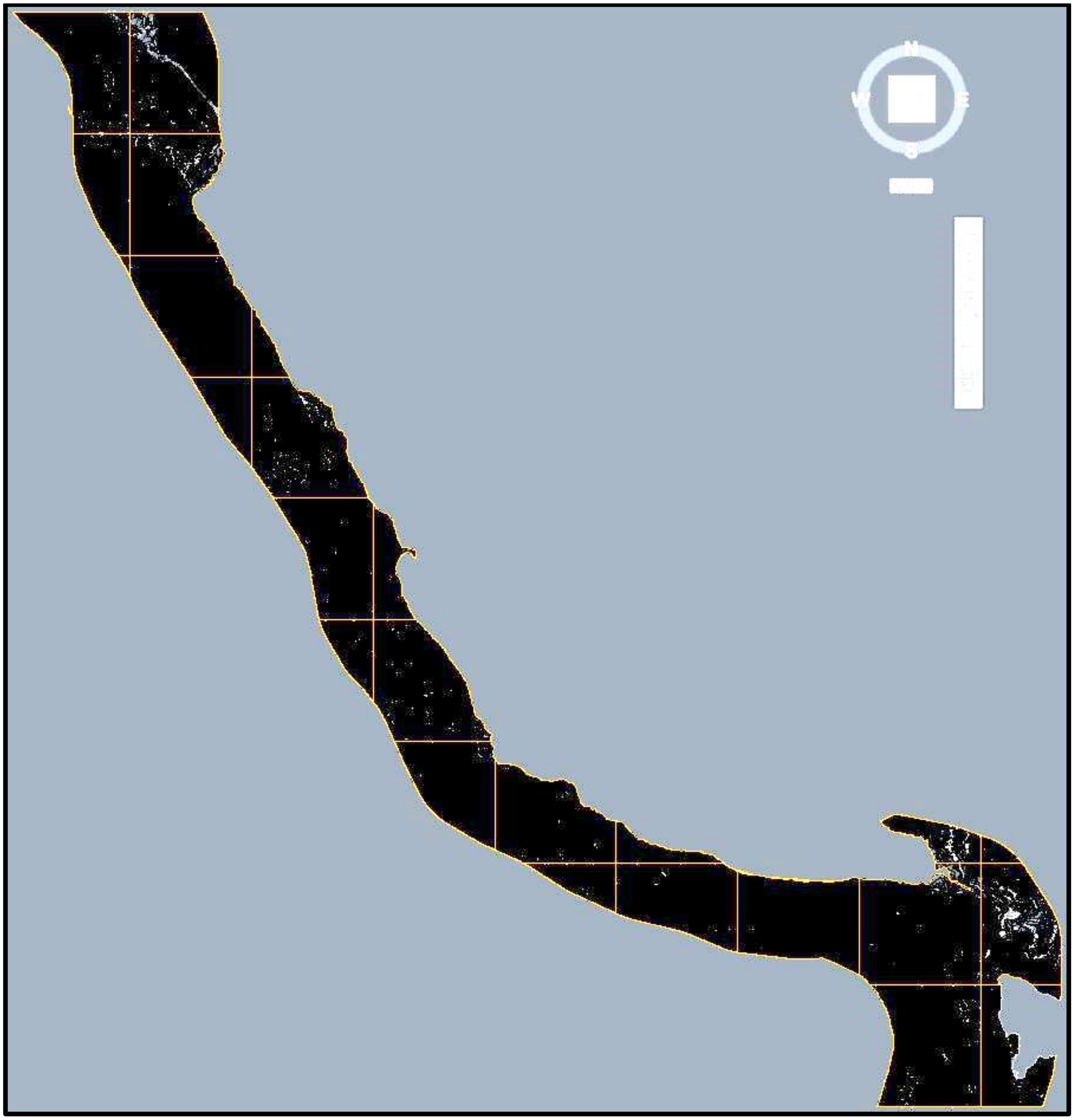


Figure 28, Bradfield to Behms LandXML Contours 15350 to 15550.

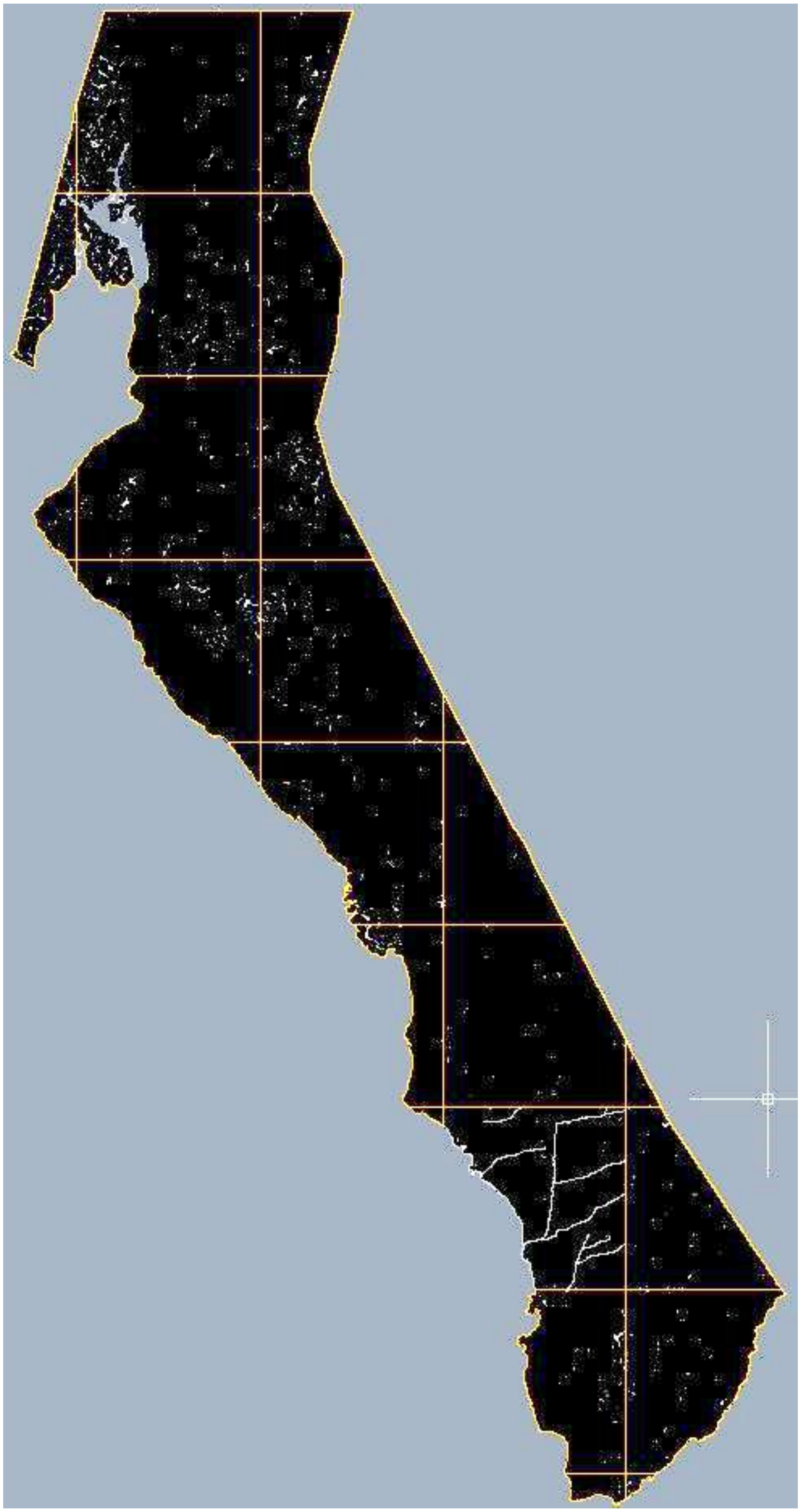


Figure 29, Bradfield to Behms LandXML Contours 15125 to 15325.

Figure 30, Ortho Rectifies Image Mosaic (RGB).

Figure 31, Ortho Rectified Image Mosaic (CIR).

Delivery History and Reported Data Quality Issues

Delivery 1 – Tetra Tech delivered to UAF the first delivery disk with Bradfield to Behms Corridor on 1/16/2015. The following provides a brief description of each quality issue found in the data provided in the first delivery and when they were reported.

UAF completed its initial review of Bradfield to Behms Corridor – On 5/8/2015, UAF completed quality assurance review of Bradfield to Behms Corridor. Provided summary of data quality issues of concern: 1) First Return DSM: Tile Seams and Discontinuity issues in lower elevation areas of tiles: L31025_15925 and L31050_15925. The forested areas look good, not sure why this looks the way it does in the shaded relief image. Please explain. 2) Land XML: A) 15625 to 15675 - Not sure why the lake contours are so confused. Contour lines jump around and look bad. Is there anything that can be done to clean this up? B) L31050_15875 - Tile includes a number of disconnected islands. Only four of these load any contour data. It looks like processing of this tile was interrupted and not finished processing. These islands should include contours to match up with adjacent tiles North and South (see Figure 32). C) L31050_15925B - Island segment drawing name is incorrectly named L31050_15925_B which may cause confusion. This drawing name needs to be corrected (see Figure 33). D) L31425_15375 - Bad tile edge, multiple drawing lines extending away from shore. Not sure what is going on here. Please clean up edge of data. E) L31475_15275 - Incorrect drawing name of L31465_15275. This causes a loading conflict with tiles from adjacent data. Please correct drawing name (see Figure 34).

Tetra Tech responded – On 2/7.2014, Renee provided initial responses to each data quality issue in a spreadsheet. Explained that seam in shaded relief is due to tidal changes that can't be fixed. Offered fixes for drawing name issues in LandXML tiles to be staged via ftp delivery.

UAF completed the ortho imagery assessment of Bradfield to Behms Corridor – On 5/14/2015, UAF completed review of the ortho imagery deliverables for Bradfield to Behms Corridor. Haze bleaching in small pockets were found in the imagery (see Figure 35). Also, cut line contrasts were observed in the imagery (see Figure 36). In consultation with Dan Ignotov of DOT, these image quality concerns are not a concern.

UAF requests fixes for selected issues – On 6/2/2015, provided spreadsheet responding to Tetra Tech feedback on data quality issues. UAF accepts without change the seams in L31025_15925 and L31050_15925 due to tide influences, and lake contours that look bad in LandXML from 15625 to 15675 that would require too much time to fix.

Delivery 2 – Tetra Tech staged fixes for Bradfield to Behms Corridor on 6/4/2015. Redelivery for Bradfield to Behms as requested. Only those files that were updated are included. They were placed in the same directory structure and should replace any previously delivered file.

UAF confirms latest CAD data looks good – On 6/19/2014, UAF confirmed that the CAD data that the fixes have been confirmed. UAF recommends acceptance of the Bradfield to Behms Corridor, the fourth of five corridors.

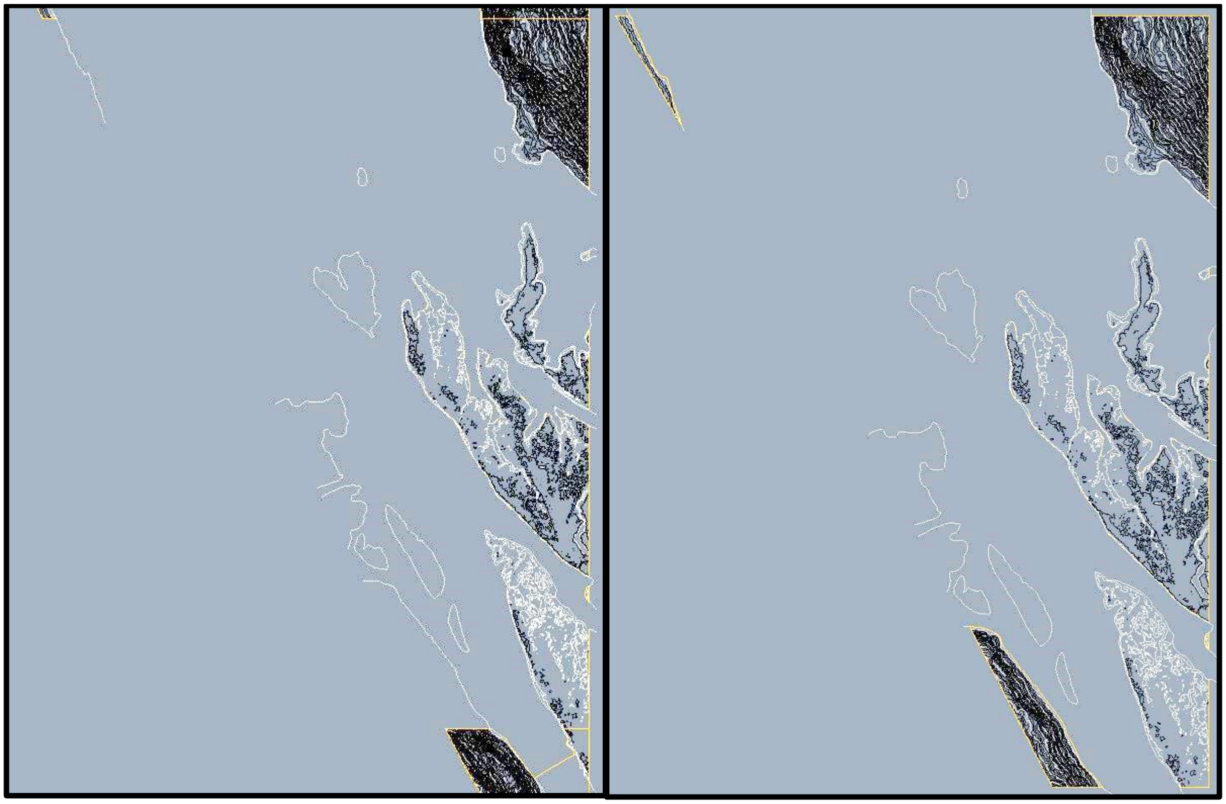


Figure 32, Bradfield to Behms LandXML L31050_15875 - Tile includes a number of disconnected islands. Only four of these load any contour data. Need land segments done at minimum to provide continuity with adjacent tiles.

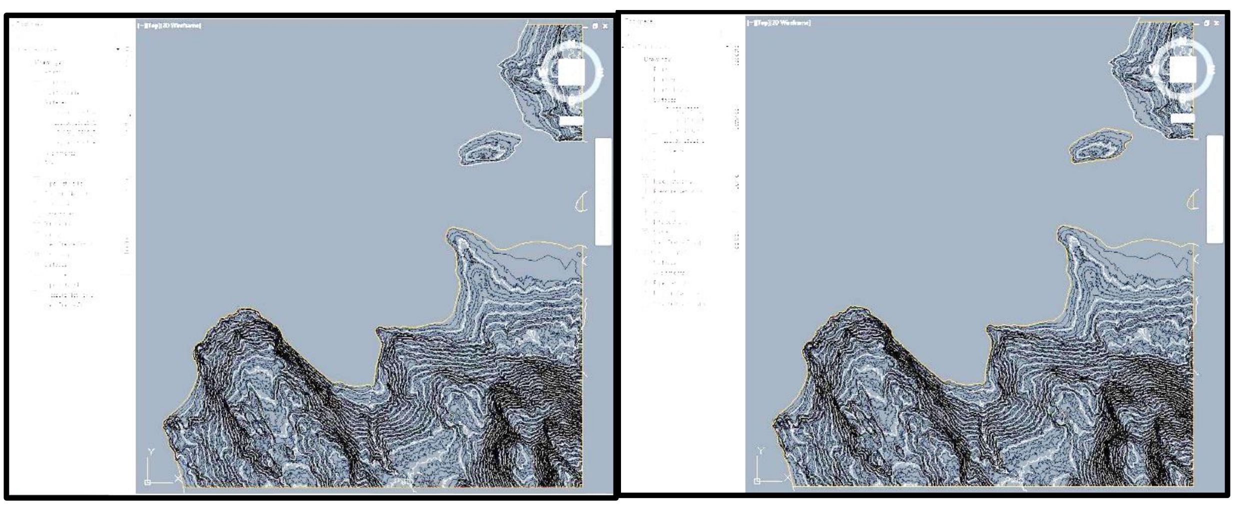


Figure 33, Bradfield to Behms LandXML L31050_15925B - Island segment drawing name is incorrectly named L31050_19925_B which may cause confusion. This drawing name needs to be corrected.

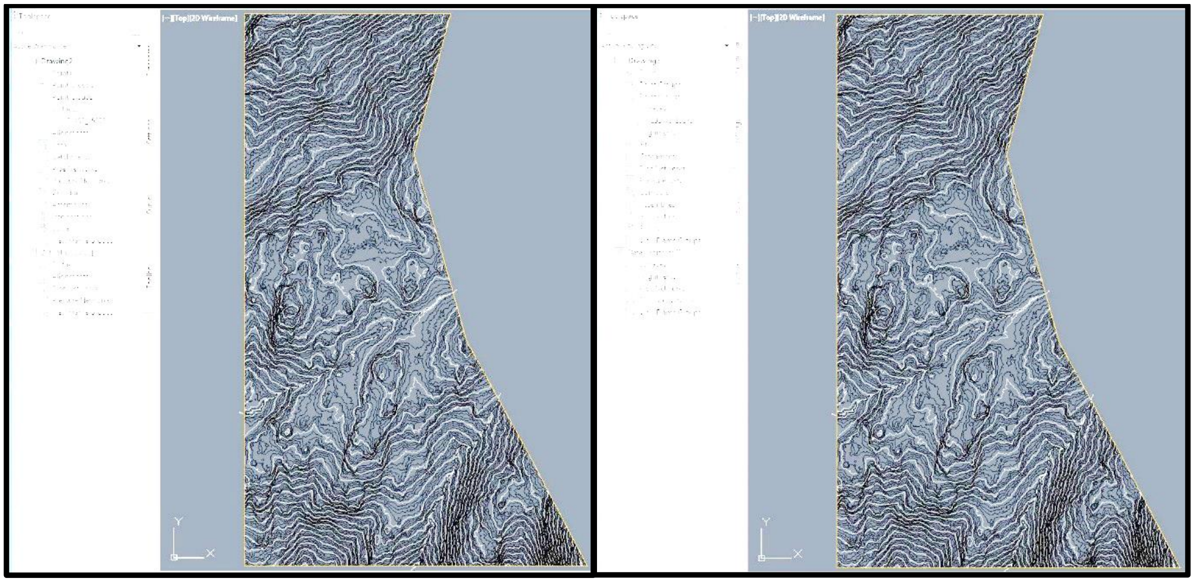


Figure 34, L31475_15275 - Incorrect drawing name of L31465_15275. This causes a loading conflict with tiles from adjacent data. Please correct drawing name.

Figure 35, Haze bleaching can be observed in small pockets in the imagery.

Figure 36, Cut line contrast can be observed in the imagery.