Grand Lake Meadows: Accuracy Comparison of Elevation Providers

Undergrad NSERC Research

Michel Leger

Supervisor: Dr. Stefanakis, Dr. McGrath

Abstract

Grand Lake Meadows (GLM), the largest wetland in New Brunswick, is a historically and ecologically significant area. GLM is ecologically diverse and home to an abundance of wildlife due in part to the extensive floodplains, presiding water levels and the presence of Grand Lake. Records of human settlement in this area date back 6000 years. Reasons for settlement in GLM include: (i) the extensive network of rivers and lakes for ease of transportation, fresh water and fishing, (ii) the rich, moist soil being excellent for planting crops, and (iii) the moderated climate and rich ecosystem provide a suitable habitat for a variety of animals, providing both meat and leathers. Low elevations are found in this wetland, with ranges between 0 and 14m above sea level. In this paper, we discuss the topography of GLM, specifically elevation data from a variety of providers and how the reported elevation data of each of these providers: Google, Bing and Esri compare to New Brunswick LiDAR data collected in 2014.

1. Introduction

The specific study area of GLM in this research is bounded, to the south by the Saint John River. On the eastern boundary, the area is north of Gagetown Island, and cuts up Jemseg River. The boundary crosses across the lower portions of Grand, Maquapit and French Lakes, passing through the Lower and Main Thoroughfares between these waterbodies. In the west, the northern border follows the Portobello Creek National Wildlife Area [Washburn & Gillis, 1996; Paponnet-Cantat and Black, 2003], Figure 1.

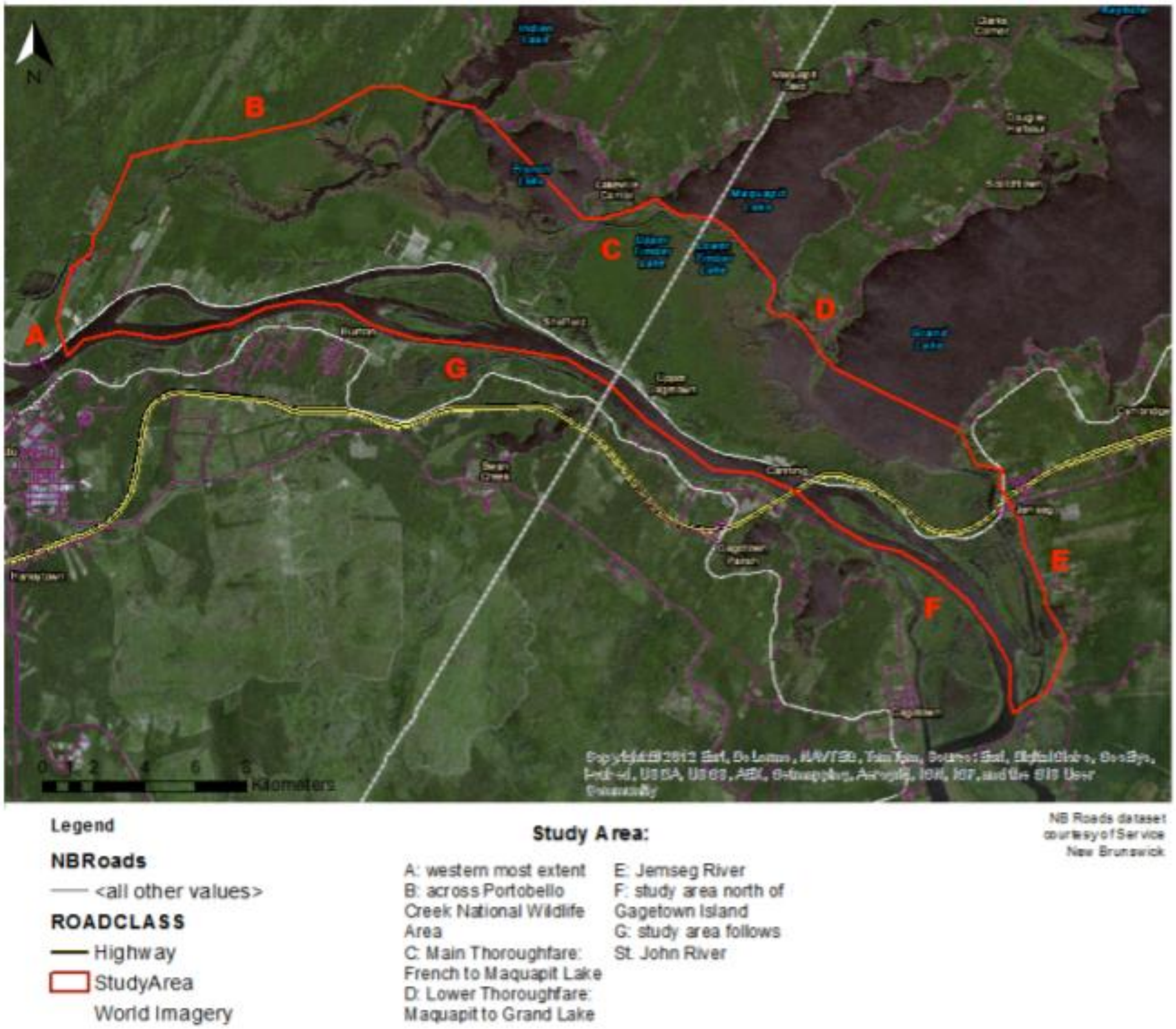


Figure 1 (from McGrath, 2014)

2. Elevation Datasets

Elevation data for this area has been collected and is available from a variety of providers: Provincial government LiDAR, and REST API Data sources: Google, Bing and Esri.

2.1. LiDAR

Light Detection and Ranging (LIDAR) data for majority of the Province of New Brunswick is available through the open source GeoNB data sets. The 1m resolution Lidar data can be downloaded through the GeoNB application, as a .laz file that will then need to be converted into a .las file for most software to read. The LIDAR data is reference in the grid horizontal coordinate system: NB Double Stereographic Projection (EPSG 2953), and a vertical coordinate system of Canadian Geodetic Vertical Datum 1928 (CGVD28) or Canadian Geodetic Vertical Datum 2013 (CGVD2013). All the GLM LIDAR vertical for this study is referenced in the CGVD28 system. As reported by Paponnet-Cantat and Black (2003), the elevation range in this area is between 0m to 16 meters above sea level. Recent 2014 elevation data collected by LiDAR, was processed to generate a Digital Elevation Model (DEM) in ArcGIS and the DEM reported the highest elevation to be 13.83m, referenced to the Canadian Geodetic Vertical Datum 1928 (CGVD28), (SNB, 2016).

2.2. REST API Elevation Data

Other than the LiDAR data, available through GeoNB, multiple providers offer elevation datasets covering this area, including: Google, Esri and Bing,

. The resolution of the datasets in this area, Google, Bing and Esri, is 1 arc-second, approximately 30m, while the LiDAR is approximately 1 meter resolution.

2.2.1. REST API

A REST API is a tool used were the user inputs requests and receives a response to those requests. (Deering, 2012) This project contained many Elevation APIs to request elevation values at various geographic coordinates. These Elevation APIs are offered through Google, Bing, and Esri web services and many others. Research focused on three of the API services, Google, Bing, and Esri. Elevation dataset requests are achieved through individual requested values, or by a batch request of requested of values. For this research batch request was used through the python computing language that requested elevation values within a ~10m x 10m bounding box of the GLM study area. The three services are varying in user friendliness and ease of accessing the data. The Bing and Google service provided adequate and helpful documentation on how to create request in the API. Esri API was difficult to access data, and retrieve helpful documentation covering how to create a successful request.

The elevation datasets from Esri, Google and Bing were accessed via REST API services. The python scripts takes the boundaries of the GLM – extended to a rectangular bounding box and requests elevation data with 30m spacing. This downloaded point data was then converted to raster, with 30 m cell size.

2.3. Point to Raster

The LiDAR point data was converted to raster via natural neighborhood interpolation technique with a cell size of 30 m, using the average value for cell assignment – only those points with returns classified as ground or bare earth were included. While this increase of cell size of the LiDAR data reduces accuracy, it is preferable than increasing cell size of the other datasets to 1m – which would introduce false accuracy statistics. The LiDAR DEM for the GLM study area is shown in Figure 2. The roads are clearly visible as the higher elevation data in each of the three section maps of Figure 2. A slope raster, computed from this DEM indicates that 80% of the GLM area has a slope less than or equal to 1° and 98.6% of the area has slopes less than 5°.

The REST API data, once converted to a comparable vertical datum (section 3 of report) with the LIDAR data (CGVD28) was imported as a point shapefile into the ArcMap software where a point interpolation processes was conducted to determine the cell assignment for the creation of the DEM raster files. Included in the ArcMap tool library is a Point-To-Raster tool that uses various cell assignment methods including: most frequent, mean, standard deviation, etc…, to create a raster from a given point file. The option determined to best represent unbiased elevations in the GLM study area was to us the most frequent cell assignment option as the parameter for the interpolation.

3. Methods

In order to assess the vertical accuracy of the elevation datasets, they needed to be transformed to the same horizontal and vertical projection.

3.1. Vertical/Horizontal Transformation

In this case, to the Canadian Geodetic Vertical Datum of 1928 (CGVD28). The National Oceanic and Atmospheric Administration (NOAA) application, VDatum (https://vdatum.noaa.gov/) was used to transform each of the datasets to WGS84(ITRF2008). VDatum is a free software, developed jointly between three U.S government organizations: NOAA's National Geodetic Survey (NGS), Office of Coast Survey (OCS), and the Center for Operational Oceanographic Products and Services (CO-OPS) (NOAA 2016). These converted x,y.z text files, referenced to WGS84 were then uploaded to the GPS-H application to convert to HTv2.0 heights. GPS-H allows users to convert their GNSS ellipsoidal heights (h) that are in either NAD83 (CSRS) or ITRF, to orthometric heights (H, heights above mean sea level) through the application of a gravimetric or hybrid geoid model (N)" [Natural Resources Canada, 2016]. The hybrid model HTv2.0 provides heights that are compatible with the Canadian Geodetic Vertical Datum of 1928 (CGVD28).

Table 1 Metadata of available elevation data available in the Grand Lake Meadows study area

	Esri[1]	**Google**[2]	**Bing**[3]	**Lidar**
Resolution/ horizontal vertical	10m, 30m, or 90m, or 1000m. (30m in GLM)	10 – 900 m (depending on location)	10 – 900 m (depending on location) *rounded to nearest metre	1m horizontal
Source Data	USGS National Elevation Dataset	Unknown	unknown	Leading Edge Geomatics
Vertical Reference	NAVD88	WGS84 EGM96 Geoid	EGM2008 2.5' or Height Above Ellipsoid (HAE) WGS84	CGVD28
Date	4/26/2012	Unknown	2008	2014
Coverage Area	Global, 30m (U.S and Canada)	Global	Global: Only latitude in range: -85 to +85	Study Area

[1] USGS Metadata, National Elevation Dataset - NAVD88 Meters - 1/3rd-Arc-Second: https://gisdata.nd.gov/Metadata/ISO/html/metadata_DEM_NED_10m.html
[2] Google Rest API metadata source: https://developers.google.com/maps/documentation/elevation/intro
[3] Microsoft Bing Rest API metadata source: https://msdn.microsoft.com/en-ca/library/jj158959.aspx

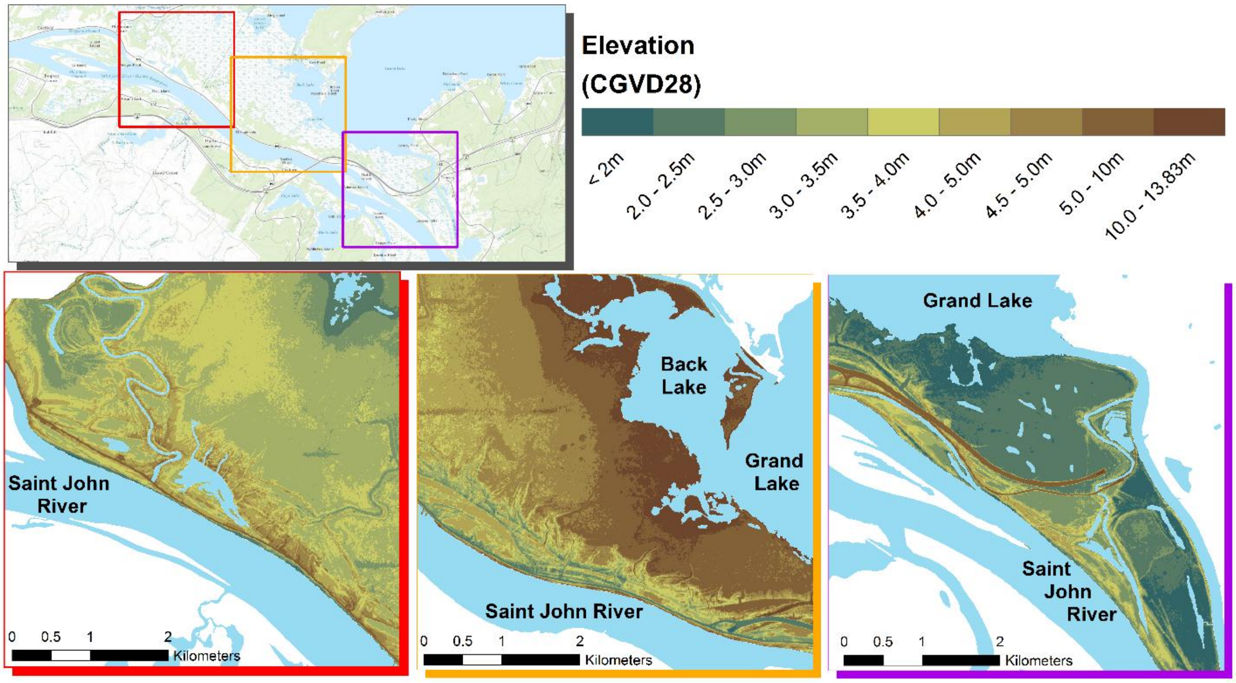


Figure 2 DEM from LiDAR data – bare earth returns

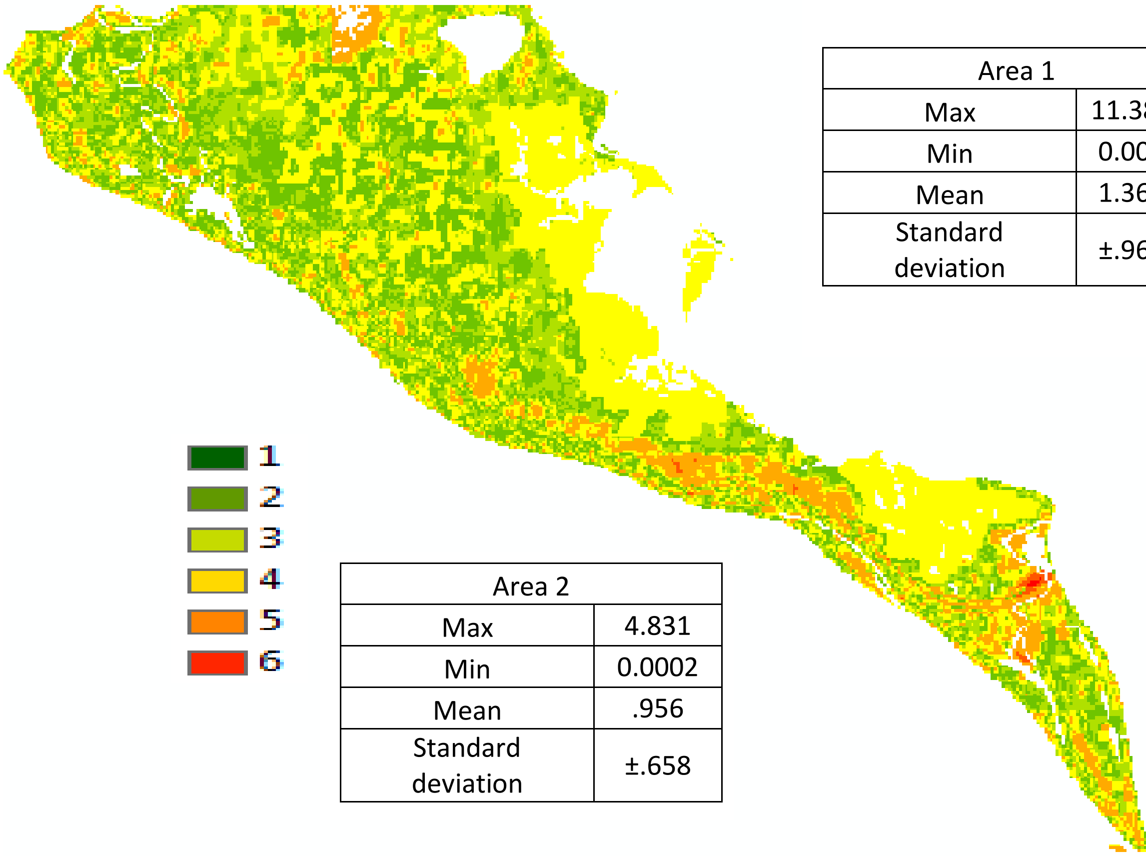
3.2. Raster Comparison

Once all DEMs were created with 30m resolution in the same horizontal and vertical reference datum, each was subtracted from the LiDAR to create a difference DEM (dDEM) were the statistics from these dDEMs were analyzed. The absolute difference surfaces that just concentrate solely on classification on land cover of the contained in the boundary of GLM study area are shown in Figures 3 to 5. The comparison between absolute high elevation differences in the Google rasters compared to location of high vegetation is illustrated and described in Figure 4, Figure 5 and Table 3.

4. Results

4.1. -difference DEMs Surrounding Area

BING



Area 1	
Max	11.380m
Min	0.001m
Mean	1.364m
Standard deviation	±.966m

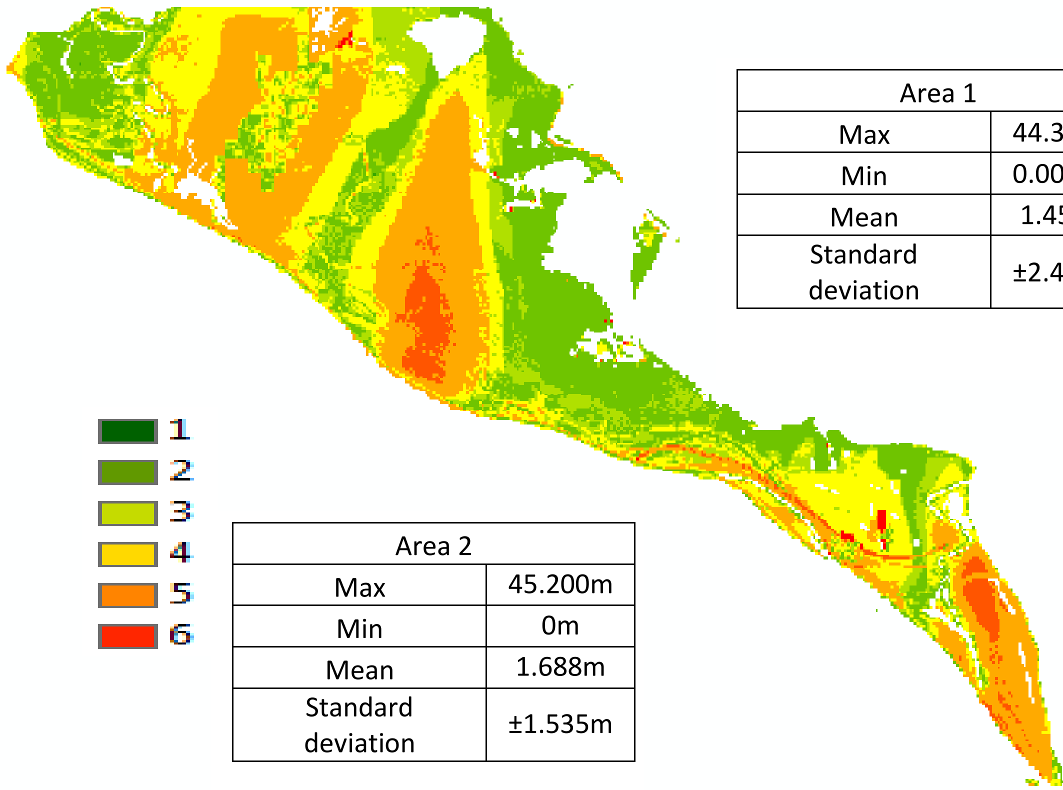
Area 2	
Max	4.831
Min	0.0002
Mean	.956
Standard deviation	±.658

Bing GLM Area 1 Classes	Classes Elevation Differences	Count of Cells in Class	Percentage of cells in Class (%)
1	0 m -.5 m	1696	15.20
2	.5 m – 1 m	1809	16.21
3	1 m – 2 m	5845	52.39
4	2 m – 5 m	1731	15.51
5	5 m – 10 m	69	0.62
6	> 10 m	7	0.06
Sums		11157	~100

Bing GLM Area 2 Classes	Classes Elevation Differences	Count of Cells in Class	Percentage of cells in Class (%)
1	0 m -.5 m	6153	28.93
2	.5 m – 1 m	5552	26.11
3	1 m – 2 m	8222	38.66
4	2 m – 5 m	1338	6.29
5	5 m – 10 m	0	0
6	> 10 m	0	0
Sums		21265	~100

Figure 3

ESRI



Area 1	
Max	44.327m
Min	0.0002m
Mean	1.459m
Standard deviation	±2.489m

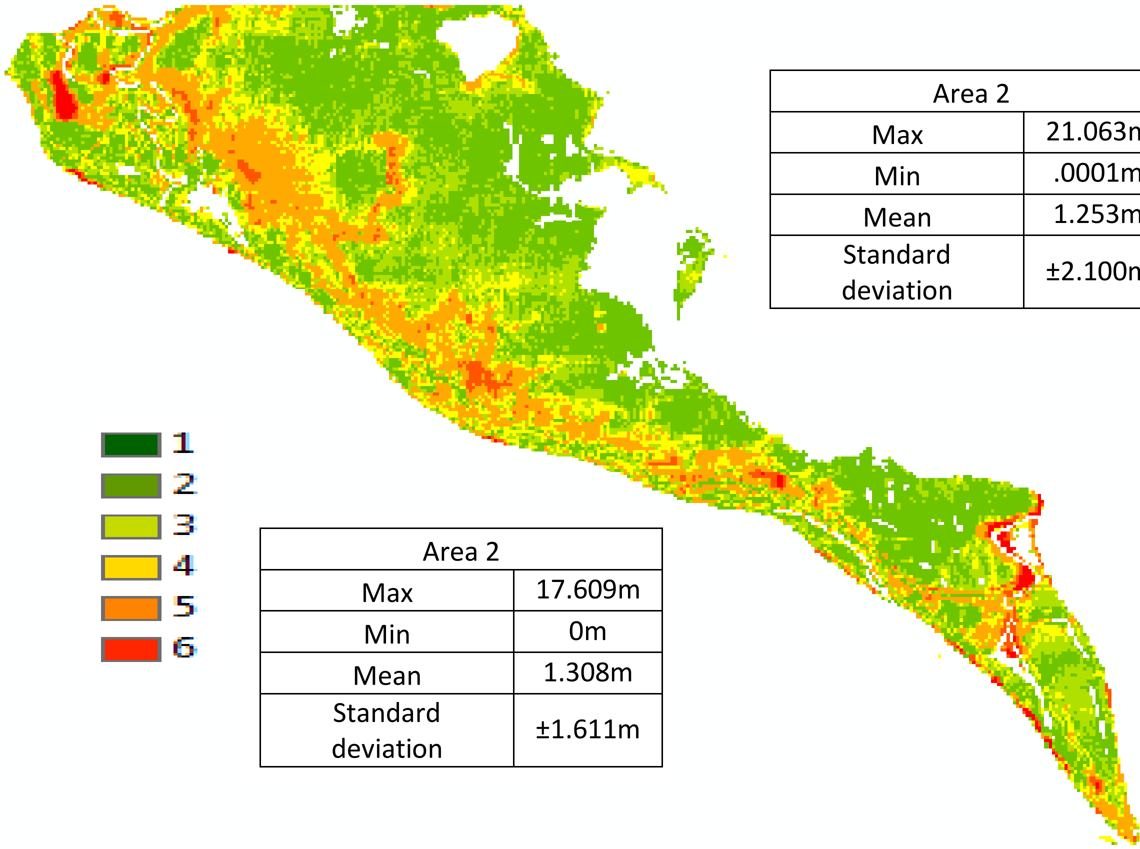
Area 2	
Max	45.200m
Min	0m
Mean	1.688m
Standard deviation	±1.535m

Esri GLM Area 1 Classes	Classes Elevation Differences	Count of Cells in Class	Percentage of cells in Class (%)
1	0 m -.5 m	4782	42.36
2	.5 m – 1 m	1502	13.30
3	1 m – 2 m	2328	20.62
4	2 m – 5 m	2186	19.36
5	5 m – 10 m	443	3.92
6	> 10 m	48	0.43
Sums		11289	~100

Esri GLM Area 2 Classes	Classes Elevation Differences	Count of Cells in Class	Percentage of cells in Class (%)
1	0 m -.5 m	5424	25.41
2	.5 m – 1 m	3035	14.21
3	1 m – 2 m	5280	24.75
4	2 m – 5 m	6946	32.56
5	5 m – 10 m	620	2.91
6	> 10 m	25	0.17
Sums		21330	~100

Figure 4

Google



Area 2	
Max	21.063m
Min	.0001m
Mean	1.253m
Standard deviation	±2.100m

Area 2	
Max	17.609m
Min	0m
Mean	1.308m
Standard deviation	±1.611m

Google GLM Area 1 Classes	Classes Elevation Differences	Count of Cells in Class	Percentage of cells in Class (%)
1	0 m -.5 m	4161	46.93
2	.5 m – 1 m	1896	21.38
3	1 m – 2 m	1235	13.99
4	2 m – 5 m	1220	13.76
5	5 m – 10 m	233	2.63
6	> 10 m	122	1.38
Sums		8867	~100

Google GLM Area 2 Classes	Classes Elevation Differences	Count of Cells in Class	Percentage of cells in Class (%)
1	0 m -.5 m	6496	38.66
2	.5 m – 1 m	3843	22.87
3	1 m – 2 m	2681	15.95
4	2 m – 5 m	3367	20.04
5	5 m – 10 m	333	1.98
6	> 10 m	84	0.50
Sums		16804	~100

Figure 5

The Bing difference raster comprises the least amount of cells in the >10m classification (class 6), with only .06% cells in Area 1 and 0% cells of Area 2 belonging to this classification. From the Bing difference raster the majority (84% in Area 1 and 93% in Area 2) of cells are classified falling between 0m and 2m (or classification 1-3). The Bing difference raster's contain 99% to 100% of all cells between 0m and 5m (or between class 1 and 4).

The Esri difference rasters contains small percentages of the cells that are classified under the >10m classification (class 6) with only .43% cells in Area 1 and .12% in Area 2. In Area 1, 42% of the cells are between 0m and .5m, and over 95% of cells are classified to measuring between 0m and 5m in both Area 1 and 2 signifying that Esri elevations match the closets to the LIDAR in this study area according these cell percentages. It is important to note that the Esri raster also contains the highest max difference value with magnitudes close to 45m, however these high classification are only represented in a small percentage of the total cells with less than 0.4%. A spot check was done in the location with the reported high difference value by creating a smaller bounding request and creating a raster around the location. This check proved to show that a possible error is present in the Esri elevation data, as the two different requests achieved similar results.

The Google difference rasters include the highest percentage of cells classified as >10m when compared to the other services. Almost 1.4% of the total 8867 cells in Area 1 and almost 0.5% of the total cells in Area 2 are classified as >10m. The Google raster statistics also have a pattern of more cells being classified in higher elevation classes. This has led to more investigation into where the high elevation cells in the Google difference raster location with respect to high vegetation classified LIDAR data.

4.2. Differences based on Land

4.3.

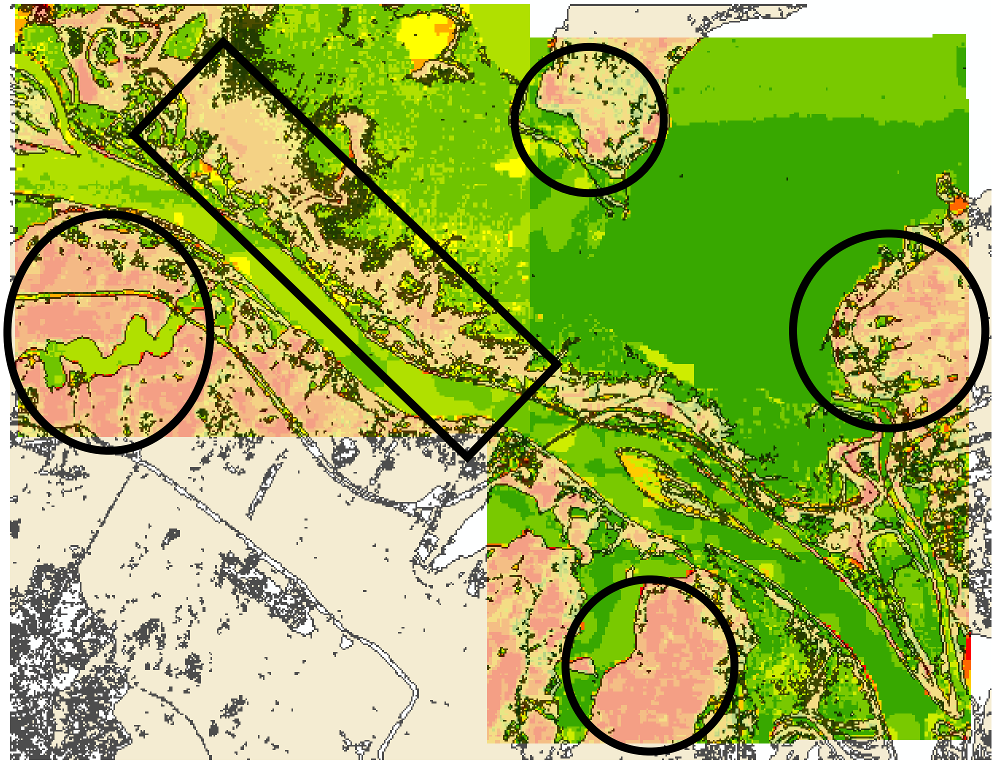


Figure 6

The figure above (Figure 6) illustrates the absolute Google difference rasters overlaid on a polygon layer that represents the high vegetation classified areas of GLM according to the class 5 filter of LIDAR data. By inspection the red areas (the areas that represent a higher difference value from the Google and the Ground LIDAR), are commonly located within the boundaries of the high vegetation polygon, leading to a probable conclusion that it is highly likely that the Google elevations service could represent more than just ground elevation values. An extraction by polygon was conducted to confirm the location of the high Google elevations with the alignment with the high vegetation polygon form LIDAR.

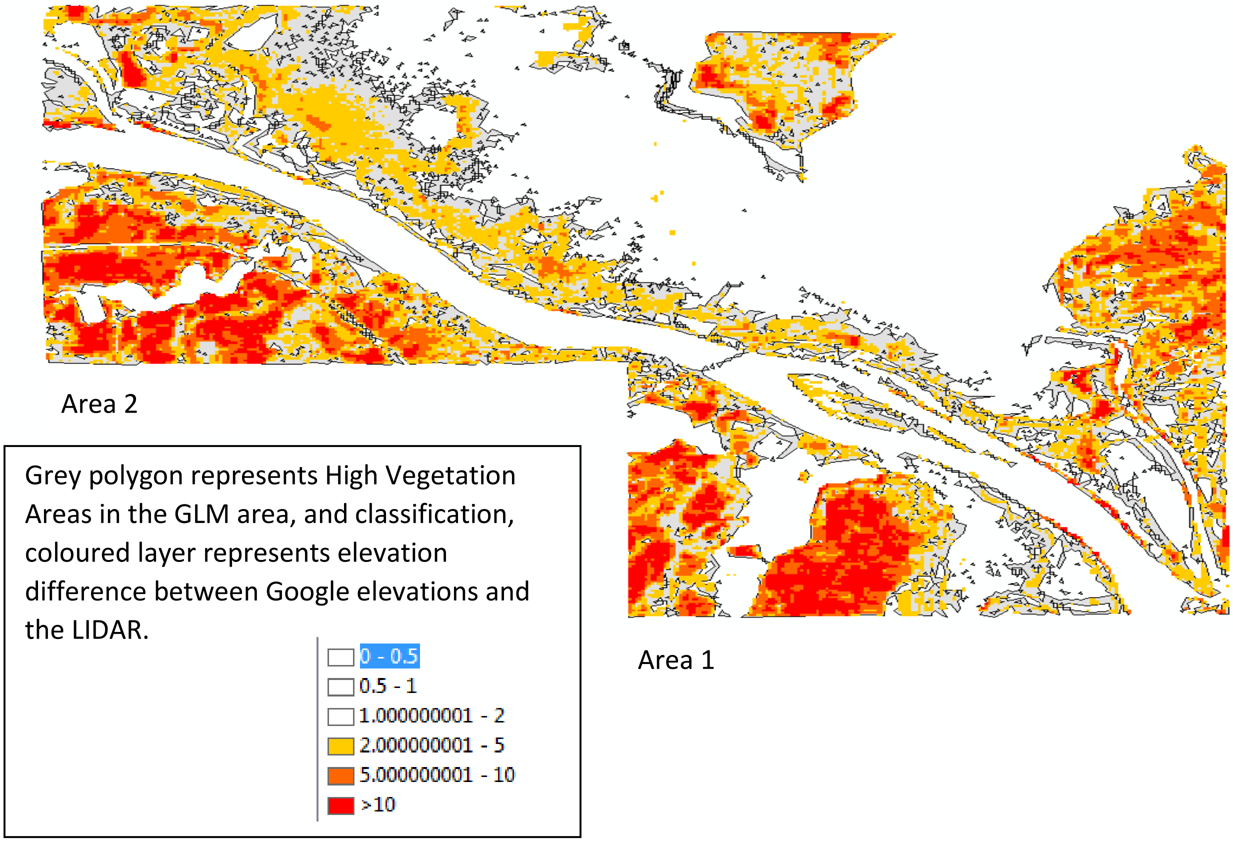


Figure 7 Google Elevations compared to High Vegetation Polygon

Table 2 Percentage of each Elevation Classification within the High Vegetation Polygon

Google Area 1				Percentage of cells classified as 2->10m	Google Area 2			
1	0 -.5	1444	8.60%		1	0 -.5	1300	7.94%
2	.5 - 1	1589	9.47%	Google Area 1	2	.5 - 1	1760	10.74%
3	1 - 2	2839	16.92%	65%	3	1 - 2	3065	18.71%
4	2 - 5	4902	29.21%	Google Area 2	4	2 - 5	5643	34.44%
5	5 - 10	3663	21.83%		5	5 - 10	2701	16.49%
6	> 10	2344	13.97%	63%	6	> 10	1913	11.68%
Total Cells in Area		16781			Total Cells in Area		16382	

The figures and table above provide information that confirms that over 60% of the raster cells lie within the high Vegetation (class 5 filter) polygon and are classified as falling between the 2m and 10m elvation differences.

Below shows the comparison by percentage, the number of cells in each classification of the total Google difference raster, with the amount of cells in each class that lie within the high vegetation polygon, as well as the total percentage of each class that fall within the polygon compared to the total number of cells for each classification class.

Table 3 Percentages of cells inside the High Vegetation Polygon compared with total cell count for each classification of the whole difference raster

<table>
<tr><td colspan="2">Total cells classified in Google difference raster</td></tr>
<tr><td>Area 1 =
2-5m - 5769
5-10m – 3870
>10m – 2383
Total # of cells = 49539</td><td>Area 2 =
2-5m - 6215
5-10m – 2808
>10m – 1933
Total # of cells = 33394</td></tr>
<tr><td colspan="2">Number of cells inside High Vegetation Polygon</td></tr>
<tr><td>Area 1 =
2-5m - 4902
5-10m – 3663
>10m – 2344
Total # of cells = 16781 or 33% of total cells</td><td>Area 2 =
2-5m - 5643
5-10m – 2701
>10m – 1913
Total # of cells = 16382 or 49% of total cells</td></tr>
<tr><td colspan="2">Percentages of cells inside the High Vegetation Polygon compared with total cell count for each classification of the total difference raster</td></tr>
<tr><td>Area 1 =
2-5m – 85%
5-10m –95%
>10m – 99%</td><td>Area 2 =
2-5m – 91%
5-10m –96%
>10m – 99%</td></tr>
</table>

5. Conclusion

The lowest elevation differences between the LIDAR and API elevations were determined to be in the LIDAR, Esri comparison. The Esri difference raster contained 95% of cells being classified as a difference between 0m to 2m, and only a fraction of a percent to be classified having a difference greater than 10m. This comparison also contained the largest difference absolute magnitude of 45m. These high magnitude elevations appeared during a check to see if the same result was determined in this location and the results were similar. This led to the conclusion that the Esri elevation data has a small errors in these locations and reports possible incorrect elevations of those locations.

The second lowest elevation differences appear in the LIDAR, Bing comparison. The LIDAR, Bing comparison has a lower percentage of cells appearing in the 0m to 2m range (84% for Area 1 and 93% for area 2) than the LIDAR Esri comparison with 95% of cells appearing in the 0m to 2m range. The LIDAR, Bing comparison however has the lowest absolute difference elevation with the maximum at about 12m. By containing the lowest maximum value and having over 80% of cells classified under the 0 to 2m difference range signifies that Bing is considered the closest services to the LIDAR.

The LIDAR, Google comparison contained the highest percentages of cells in the >10m range with 1.4% in Area 1 and .5 % in Area 2. When the larger 10m x 10m area was compared with LIDAR only 60% of difference cells were classified as being in the 2m to 10 m range. During inspection the Google raster showed that majority of the Google high elevations difference were located in the boundaries of the LIDAR high vegetation areas with 99% of total cells greater the 10m to be located in this boundary. This information supports that the Google elevations contain a combination of ground and vegetation elevation values and caution is needed when interpreting results from this service.

The ease of access and the user friendly aspect of the services are varying for each API service. The two that were very straight forward and contained helpful documentation were the Bing and Google services. The Esri service was very difficult to comprehend how to create a request and the lack of helpful documentation made Esri less user friendly compared to the other services.

Final conclusion remarks that the most suitable API elevation service when considering both user friendliness and representing ground LIDAR data would be the Bing API Elevation Service.

References

Deering, S. (2012). Do you Know What a REST API is?, retrieved: http://www.sitepoint.com/developers-rest-api/

McGrath, H. (2014). Historical Maps of Grand Lake Meadows, M. Sc. E. University of New Brunswick, 2014 Department of Geodesy and Geomatics Engineering

Natural Resources Canada (2016). Earth Sciences, Geomatics, Geodetic Reference Systems, Tools and Applications, retrieved: http://www.nrcan.gc.ca/earth-sciences/geomatics/geodetic-reference-systems/tools-applications/10925#gpsh

NOAA (2016). Vertical Datum Transformation, Integrating Americas Elevation Data, retrieved: https://vdatum.noaa.gov/

Paponnet-Cantat, C., & Black, W. (2006). 20th century land use mapping report. (pdf). Fredericton, NB: UNB.

SNB (2017). GeoNB Data Catalogue, Lidar Data, retrieved: http://www.snb.ca/geonb1/e/DC/lidarData.asp

Stefanakis (2015, Nov2). Elevation Web Services: Limitations and prospects. GoGeomatics Canada, online Magazine

Washburn and Gillis Associates Ltd. (1996). Environmental impact assessment, TransCanada Highway Fredericton to Salisbury, component study report Grand Lake Meadows Wetland. Report prepared for the New Brunswick Department of Transportation, Fredericton, New Brunswick.