WATER SURFACE ELEVATION CALCULATION USING LIDAR DATA

[1]**George J. Toscano***,
[1]**Partha P. Acharjee,**
[2]**Collin McCormick,**
[1]**Venkat Devarajan,**
[1]Dept. of Electrical Engineering, University of Texas at Arlington, Texas, USA,
[2]Natural Resources Conservation Service, Fort Worth, Texas, 76115
george.toscano@mavs.uta.edu,
partha.acharjee@mavs.uta.edu,
collin.mccormick@ftw.usda.gov,
venkat@uta.edu

ABSTRACT:

The spectral reflectance of water is very small in near infra-red range. Specular reflection from water surface occurs due to the smooth surface of water body. For these two reasons, high LiDAR data dropout rate is often observed from water bodies. Though DEMs or TINs created using LiDAR data have very good vertical accuracy, if any water body is present, those LiDAR deliverables sometimes become erroneous. If LiDAR data dropout rate is very high, the DEM or TIN generation system often interpolates the elevation obtained from the land edge of the water body. If the interpolated data is taken from land edge of the water body, water surface will no longer look flat. Therefore, in the DEMs and TINs, the presence of the water body boundary will not be well defined. In our previous paper on hydro break line generation, our proposed method required the accurate calculation of the surface elevation of water bodies, which could be problematic. In this paper, a solution to this problem is proposed and validated, which does not require pre-calculated break lines.

KEYWORDS: LiDAR, Water surface elevation, Point cloud, Hydro-flatten

1. INTRODUCTION

Hydro-flattening involves the process of levelling the water surface or giving a single elevation value to the water surface in case of still water body. In the case of streams and rivers, water surface must be levelled from bank to bank. In this paper, only still water bodies such as lakes and ponds are considered.

Though extensive research has been performed on detecting water bodies (Antonarakis et al., 2008; Brzank et al., 2008; Höfle et al., 2009, Smeeckaert et al. 2013, Toscano et al. 2014), no significant research is found in the literature on water surface elevation calculation. For example, Antonarakis et al. (2008) classified not only water but also gravel, bare earth and different types of forest. Brzank et al. (2008) classified water in Wadden sea area, based on supervised fuzzy classification using feature height, intensity and 2D point density. Höfle et al., (2009) modelled LiDAR data dropouts and used seeded region-growing segmentation to detect potential water regions.

Without enforcing hydro-breaklines, if a TIN is generated, the normal interpolation process might take data from the land edge of water body and create unrealistic TIN surfaces. An example is shown in figure 1 and 2. In figure 1, a point cloud is shown for a chosen area in the state of Oklahoma, USA. In figure 2, a TIN surface for the same area containing four water bodies (numbered 1 to 4) are shown. In this figure, large triangles are obtained on the water bodies due to LiDAR data voids (visible in zoomed-in mode). High data dropout rate or LiDAR data voids occur for the following two reasons:

a) very small spectral reflectance of water in near infra-red range and
b) issues with specular reflection from water surface.

Specular reflection or mirror like reflection from water surface occurs due to the smooth surface of a water body, resulting in very high intensity return if the deflection angle is small and, no return if the deflection angle is high. The surface elevation calculation becomes complex because of this high data dropout rate from a water surface.

This paper assumes the approximate location of water bodies based on the algorithm described in our previous paper (Toscano et al. 2014). It further improves upon our previous paper on hydro break line generation in two

ways: a) a significantly better approach to water body elevation calculation is presented and, b) the use of hard coded numbers in our previous algorithms is significantly reduced. Hence, the method is more robust and applicable to more geographic locations.
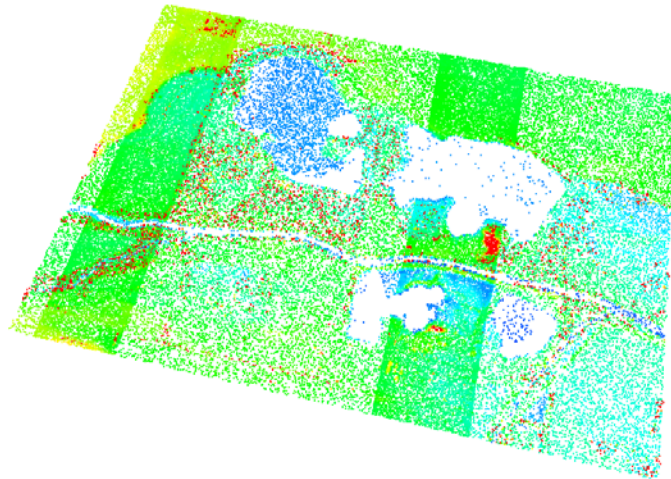

Figure 1: LiDAR point cloud for an area in the state of Oklahoma, USA.
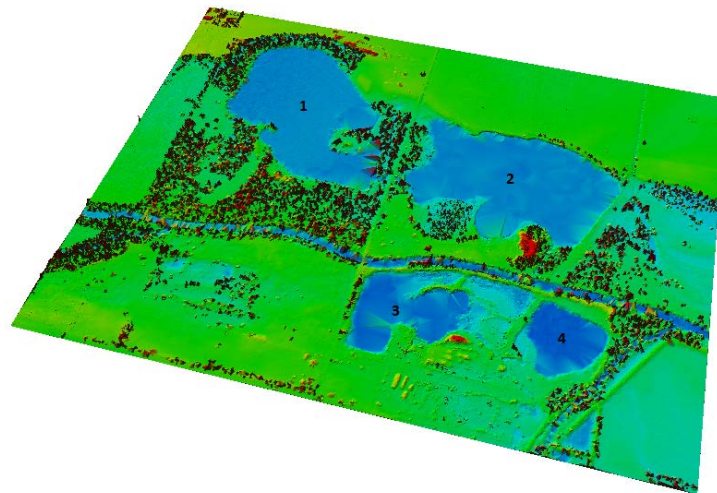

Figure 2: TIN of the test area in Oklahoma

2. OVERVIEW OF OUR METHOD

In order to calculate the surface elevation of a water body by the method proposed in this paper, the area is first rasterized and then its elevation histogram is generated. The surface of a water body has the same nominal elevation. So, if a histogram is created based on elevation, a sharp peak should be obtained at the corresponding water surface elevation. However, sometimes a very small number of LiDAR returns is obtained from water bodies (example: test area water body 2, 3 and 4) and in those cases, a sharp peak might not be obtained in the elevation histogram. To solve this problem, a method is proposed to modify the histogram.

Along with a sharp peak for a water body, many peaks are created which do not represent any water body. To remove those false peaks from the histogram, a Gaussian filter is designed using water surface elevation distribution.

After Gaussian filtering, a spline interpolation is used to determine the peak location and from that surface elevation of the water body is determined. The block diagram of the proposed method is shown in figure 3.
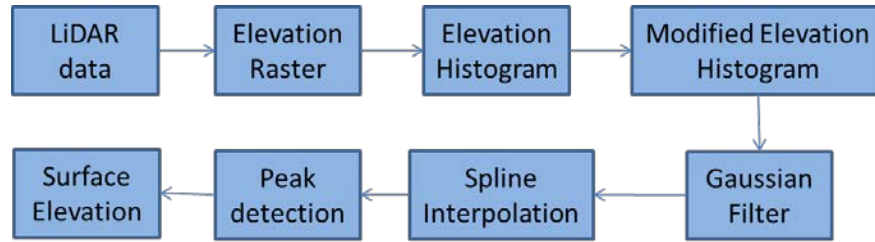

Figure 3: Block diagram of the proposed method

3.   DETAILED DESCRIPTION OF OUR METHOD

In this section, our proposed method is described step by step for a test area. For better representation of an elevation histogram all the elevation values are given from a reference elevation of 369.18 meters (the lowest elevation of all the LiDAR returns in the .las file of the test area) in this paper.

3.1 Creating 2m Elevation Raster

LiDAR elevation data is rasterized at the very beginning. The area is divided into 2m x 2m square grid. Now to create a 2m elevation raster, in each 2m x 2m square, all the single returns and, all the last returns (in case of multiple returns) are stored. Thereafter, the median elevation value of those LiDAR returns are calculated and assigned as the elevation of each 2m x 2m square. Interpolation methods such as the inverse distance weighted method, kriging etc., are generally used for DEM generation. However, for our algorithm, the median elevation value is taken to avoid some high elevation last returns which might not be from the ground. Zero elevation and zero intensity values are assigned to squares, from which no LiDAR return is recorded. The median elevation raster for the test area is shown in figure 3.
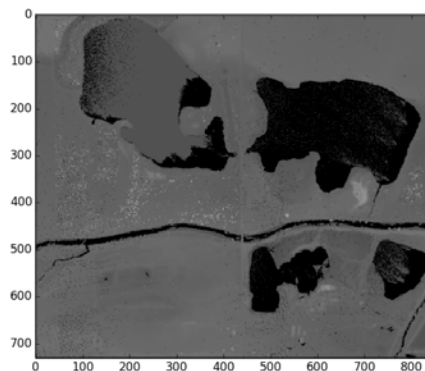

Figure 3. 2m median Elevation Raster

3.2 Creating elevation histogram

In this step, the elevation histogram is created from the elevation raster. For each of the water bodies, a rectangular elevation raster containing a water body is chosen and, an elevation histogram is generated from it. In figure 4 and 5, rectangular elevation raster for water body 1 and its corresponding elevation histogram are shown.

It is found that on the one hand, a smaller bin size for the histogram unnecessarily increases the complexity of the overall algorithm and, on the other hand a larger bin size lacks precision in water body detection. The bin size for the elevation raster is chosen to be one inch as a reasonable compromise.

From the surface of the water body 1, a large number of LiDAR returns are obtained. Since water surface is nominally flat, many LiDAR returns must come from the same elevation, creating a peak at that elevation. Therefore, the highest peak of the histogram in figure 5 represents the surface elevation of the water body. The highest peak is found at 608 inches elevation from the reference elevation of 369.18 meters.
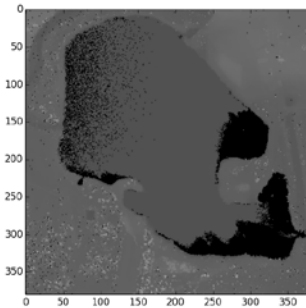

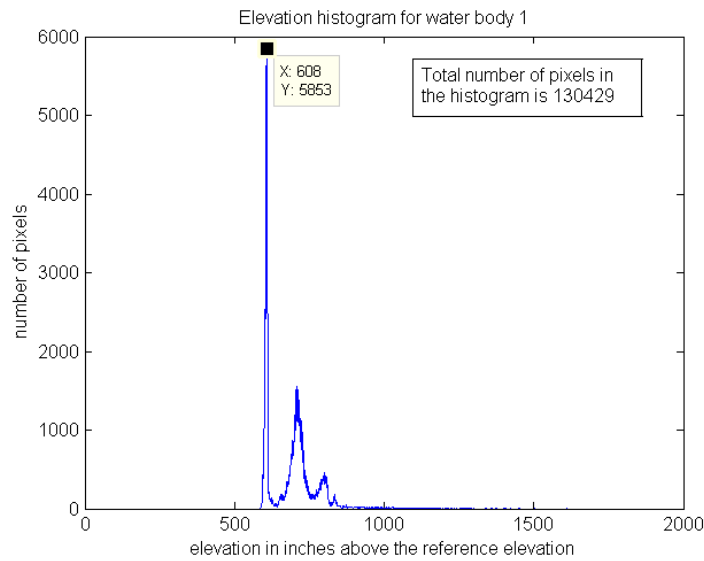Figure 4: Elevation raster for water body 1.

Figure 5. Elevation histogram for water body1.

For water body 2, the elevation histogram is shown in figure 6. The highest peak in the histogram is found at elevation of 768 inches. However, a manual approach to determine the water surface elevation found it to be approximately 602 inches. This discrepancy is caused by the high LiDAR data dropout rate from within the water body causing a large data void in it, which results in a very small peak at actual water surface elevation. In the next section, a method is discussed to solve this large data void problem by creating a modified elevation histogram.
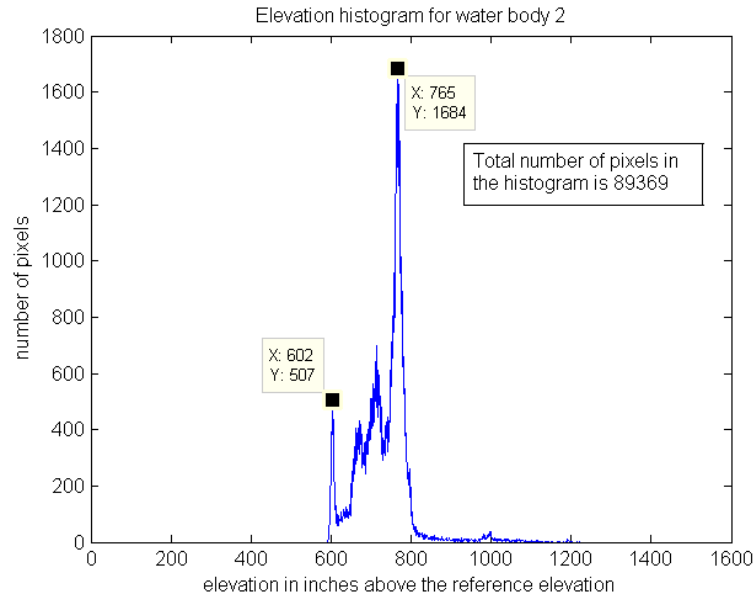
Figure 6. Elevation histogram for water body 2.

3.3 Creating and applying modified elevation histograms

It is clear that the peak corresponding to water body 2 is smaller because the void pixels failed to contribute to the peak. The peak is reduced in size roughly by a factor of v/n where v is the void data size and n is actual pixels from the water body (scattered pixels) that contributed to the histogram. Based on this logic, a modified histogram is calculated as follows.

1. Large data voids are identified using connected component analysis.
2. Let the number of *scattered* pixels inside a large data void be n. It should be noted that pixels in large clusters might represent an island inside the water body and will not be scattered.
3. Let the distribution of LiDAR data return from such pixels be $h_n(x)$.
4. Let the data void size be v
5. Let the unmodified histogram containing the actual water surface returns be $h_u(x)$.
6. A modified elevation histogram $h_v(x)$ can then be expressed as:

$$h_v(x) = h_u(x) + (v/n)*h_n(x) \ldots \quad \ldots \quad \ldots (1)$$

Application of this equation has the effect of accentuating the impact of the waterbody in the overall histogram.

Analysis shows four large data voids can be identified (per step1 above) for water body 1 and, one large data void each is obtained for water body 2, 3 and 4. Next the number of scattered pixels within each void was identified per step 2 above. The total number of pixels in the elevation histogram was also calculated in each water body. For all the water bodies the data void size (v) and the number of scattered pixels (n) inside the voids is shown in table 1. All of this information is used together with equation 1 above to modify the original histograms to obtain the modified histograms.

Modified histograms for water body 1 and 2 are shown in figure 7 and 8 respectively. It can be seen that for water body 1 the elevation peak at 608 inches is shifted by 2 inches to 606 inches and for water body 2 the highest peak is now obtained at 602 inches elevation from the previous 765 inches elevation.

Table 1: Size of data voids inside water body and number of scattered pixels inside the void

Water body	Total number of pixels in the elevation histogram $h_u(x)$	Size of data voids in pixels (v)	Number of scattered pixels inside data void (n)
1 (top left)	139420	270	130
		3455	84
		645	72
		7001	317
2 (top right)	89369	55131	3804
3 (bottom left)	35946	13335	372
4 (bottom right)	20472	7688	1343
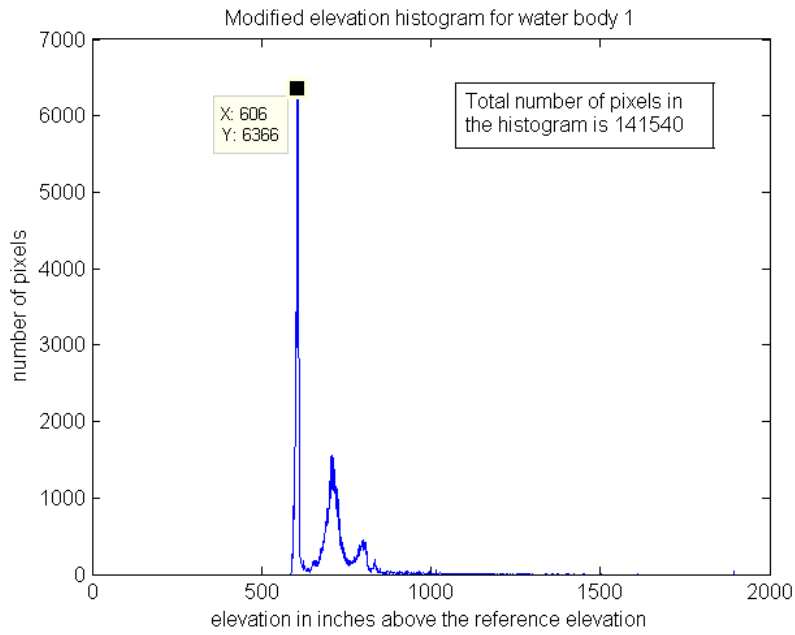
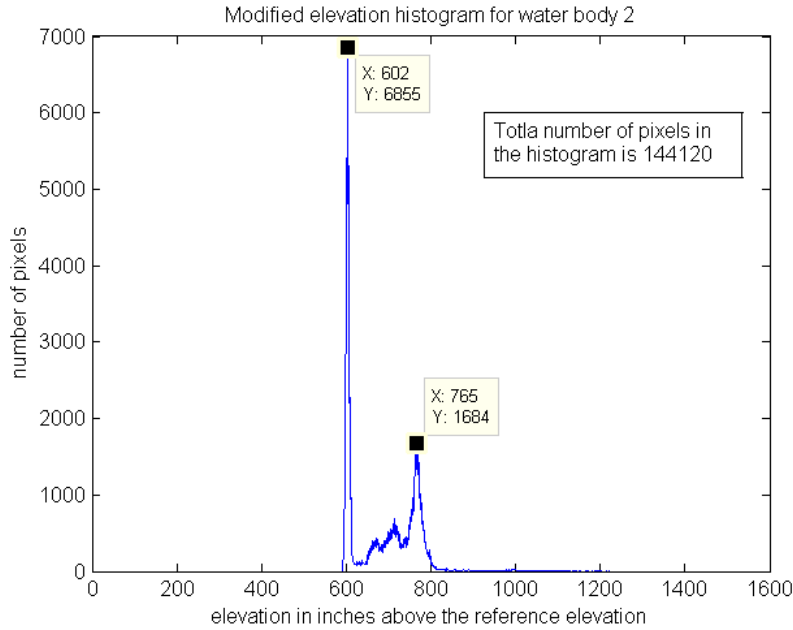Figure 7: Modified elevation histogram for water body 1

Figure 8: Modified elevation histogram for water body 2

3.4 Low pass filtering the modified elevation histogram:

After modifying the elevation histogram, it is low pass filtered to remove spurious peaks and to more precisely determine the water surface elevation. The correlation between two random variables *X* and *Y* is maximized if their density functions have similar shapes. This logic motivates us to calculate water surface elevation histogram from some test area containing a water body and use that histogram to design a Gaussian filter. Now if elevation histogram of an area is correlated with the Gaussian filter, only the peaks due to water surface will be significantly amplified compared with other spurious peaks. A one dimensional Gaussian filter is used as the low pass filter. The impulse response of the Gaussian filter is given in equation (2)

$$g(x) = \frac{1}{\sqrt{2\pi}.\sigma} e^{-\frac{x^2}{2\sigma^2}} \quad \dots \quad \dots \quad \dots \quad (2)$$

The parameters of this filter are determined by statistical analysis of several water body returns. The standard deviation of the Gaussian filter is calculated from three selected 200m*200m areas from three different water bodies. One of the water bodies and the selected area are shown in figure 9. The standard deviation of the elevation in this area is found to be 1.6721 inches. Similarly, for two more water bodies, the standard deviation of water surface elevation is calculated and found to be 1.9233 inches and 1.3482 inches. So the overall standard deviation is

$$\sigma = \sqrt{\left(\frac{\sigma_1^2 + \sigma_2^2 + \sigma_3^2}{3}\right)}$$
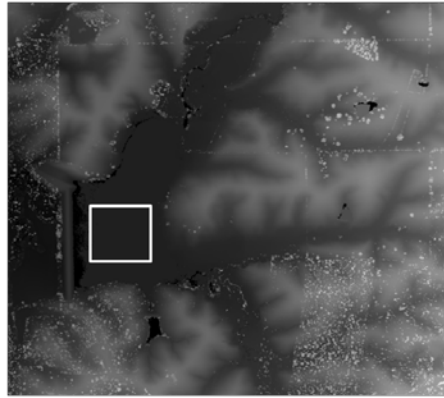$$= 1.6646 \text{ inches}$$

7

Figure 9. 200m*200m area selected for water surface elevation standard deviation calculation

The kernel length required for the FIR Gaussian filter is $\lceil 6\sigma - 1 \rceil$ or 9. The coefficients for the Gaussian filter are determined to be [0.0134, 0.0472, 0.1164, 0.2001, 0.2397, 0.2001, 0.1164, 0.0472, 0.0134] using equation (2). After convolving with the Gaussian filter, resultant histograms for water body 1, 2, 3 and 4 are shown in figure 10, 11, 12 and 13 respectively.
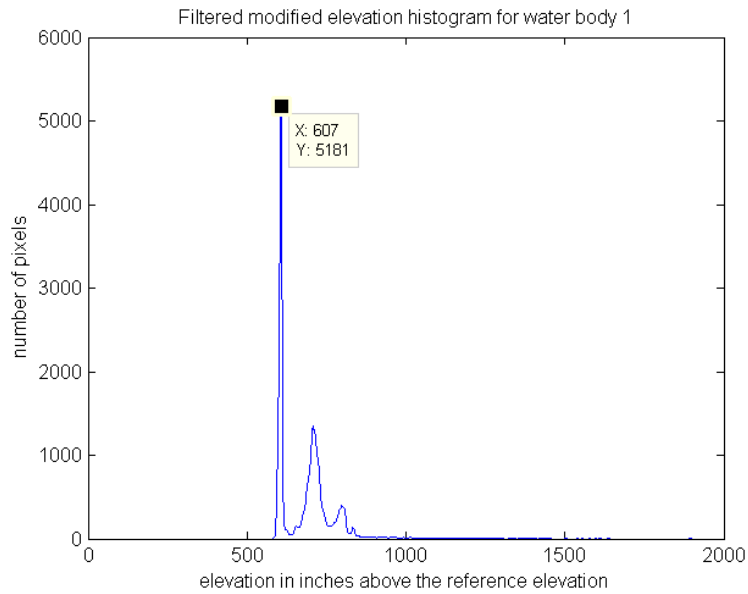

Figure 10: Filtered, modified elevation histogram for water body 1
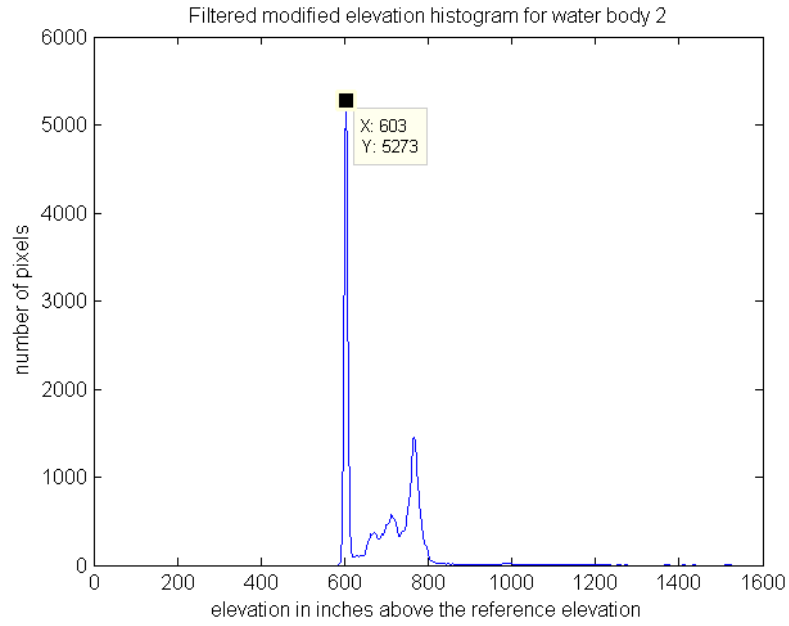
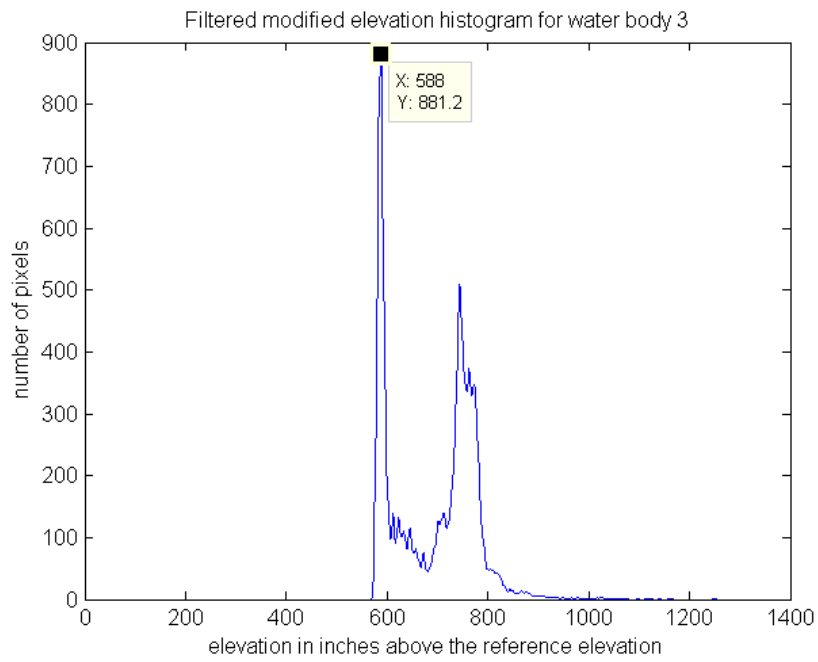Figure 11: Filtered, modified histogram for water body 2

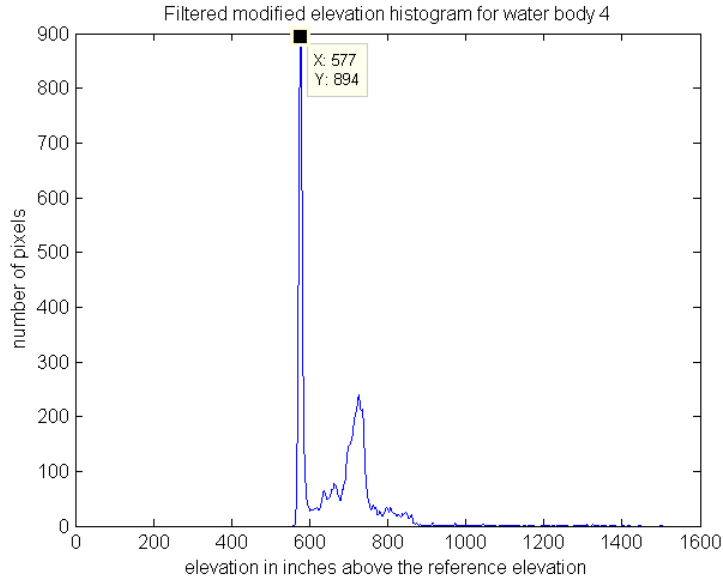Figure 12: Filtered, modified histogram for water body 3

Figure 13: Filtered, modified histogram for water body 4

3.5 Cubic Spline Interpolation of the modified and low pass filtered histogram data points:

After smoothing the histogram using the Gaussian filter, cubic spline interpolation is performed to determine the redefined location of peaks. Cubic spline data interpolation connects data points in adjacent bin locations of the histogram. Each spline (between two adjacent bins with elevation x_i and x_{i+1}) can be represented by equation (3)

$$f_i(x) = a_i x^3 + b_i x^2 + c_i x + d_i \qquad \dots \dots \dots (3)$$

Part of interpolated histograms of both water body 1 and 2 is shown in figure 14 and 15, from which it can be seen that surface elevation of water body 1 is most likely at 606.7938 inches and surface elevation of water body 2 is similarly at 602.9388 inches. The highest peak elevation/location change with each stage of histogram operation is shown in table 2.
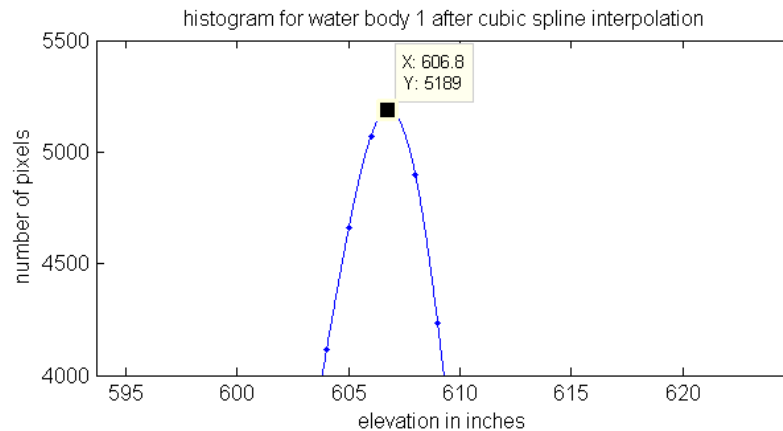

Figure 14: Partial histogram for water body 1 after cubic spline interpolation
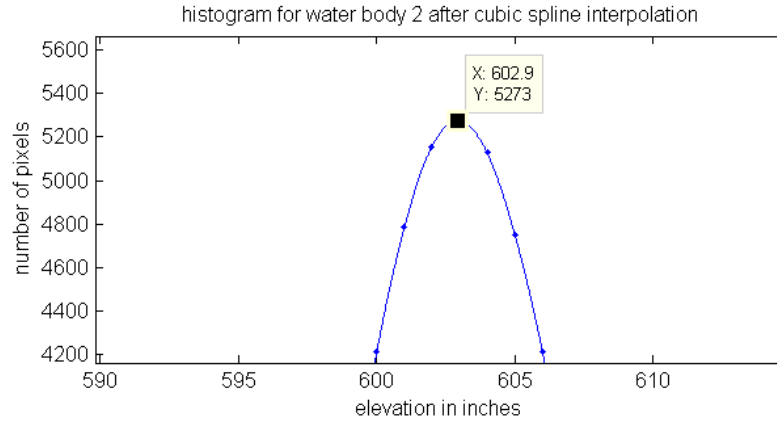
Figure 15: Partial histogram for water body 2 after cubic spline interpolation

Table 2: Highest peak elevation/ location of histograms for water bodies

Water body	Initial elevation histogram	Modified elevation histogram	Filtered elevation histogram	Cubic spline interpolation
1 (top left)	608	606	607	606.7938
2 (top right)	765	602	603	602.9388
3 (bottom left)	745	589	588	588.0868
4 (bottom right)	725	578	577	576.5470

4. RESULTS AND DISCUSSIONS

In this section, a method is described to validate the results obtained using our automated approach to elevation calculation. The correct answer for the water body elevation or ground truth is calculated using a manual approach. It is known that if the break lines enclosing the water body were known, the elevations within the break lines can be averaged to determine the correct elevation. Therefore, using ArcGIS a polygon feature was manually created inside the water body 1 as shown in figure 16. All the LiDAR returns are extracted from the polygon and the mean is found to be 384.5482 meters. Subtracting the reference elevation, the mean elevation is found to be (384.7482-369.18) meters or 605.04 inches, which is close to the elevation (606.7938 inches) obtained by the method proposed in this paper. In a similar way, for all the other water bodies, the water surface elevation is calculated in the proposed method and then verified manually by drawing polygons inside the actual hydro break lines. The comparison is shown in table 3.

Figure 16: Polygon feature drawn manually inside the hydro break line to calculate water surface elevation

Table 3: Comparison between proposed method and manually extracting data method

Water body number	Surface elevation in proposed method (in inches)	Surface elevation by manually extracting data (in inches)	Difference (in inches)
1	606.7938	605.04	1.7538
2	602.9388	602.924	0.0148
3	588.0868	586.2271	1.8579
4	576.5470	576.40	0.147

5. CONCLUSIONS

A new method is proposed and validated to accurately calculate the elevation of still water bodies from LiDAR point clouds. The advantage of the proposed method is, it does not require that the hydro break lines be pre-calculated. Only the LiDAR data of a rectangular area around the water body is required. Therefore, if hydro break lines are not available with the LiDAR data, the proposed method is very useful. On the other hand, if the water surface elevation is calculated using this approach, together with the approach that we have described in a previous paper (Toscano et al., 2014) to identify the boundaries of a water body, accurate hydro break line generation can essentially be automated.

REFERENCES

Antonarakis, A.S., Richards, K.S., Brasington, J. 2008. Objects-based land cover classification using airborne LiDAR. *Remote Sensing of Environment,* 112(6), pp. 2988–2998.

Brzank, A., Heipke, C., Goepfert, J., Soergel, U. 2008. Aspects of generating precise digital terrain models in the Wadden Sea from lidar-water classification and structure line extraction. *ISPRS Journal of Photogrammetry and Remote Sensing,* 63(5), pp. 510–528.

Höfle, B., Vetter, M., Pfeifer, N., Mandlburger, G., & Stötter, J., 2009. Water surface mapping from airborne laser scanning using signal intensity and elevation data. *Earth Surface Processes and Landforms,* 34(12), pp. 1635-1649

Maune, David F., February 2010. Digital Elevation Model (DEM) Whitepaper, NRCS High Resolution Elevation Data.

Smeeckaert, J., Mallet, C., David, N., Chehata, N., & Ferraz, A. 2013. Large-scale classification of water areas using airborne topographic lidar data. *Remote Sensing of Environment*, 138, pp. 134-148.

Toscano, G. J., Gopalam, U. K., & Devarajan, V., March 2014. Auto Hydro Break Line Generation Using LiDAR Elevation And Intensity Data. Proceedings, *ASPRS Conference*

ACKNOWLEDGEMENT