A NEW FRAMEWORK FOR ACCURACY ASSESSMENT OF LIDAR-DERIVED DIGITAL ELEVATION MODELS

XiaoHang Liu

Department of Geography & Environment, San Francisco State University, U.S.A.

Commission IV, WG IV/3

KEY WORDS: LiDAR, DEM, Accuracy, Approximation Theory, Isomorphism

ABSTRACT:

Existing approaches to accuracy assessment of LiDAR-derived DEM are typically based on statistical methodology and focus on vertical accuracy. This paper presents a new framework which calls for assessment of not only a DEM's vertical accuracy, but also its ability to preserve a point's elevation rank in the bare earth topographic surface. New methods to assess each aspect are presented. For DEM's vertical accuracy, approximation theory from numerical analysis is used to quantify the total error at each DEM point as well as its three error components – sensor error, ground error, and interpolation error. For DEM's elevation order which is critical to model terrain structure, the concept of isomorphism in set theory is drawn as the mathematical rationale and Kendall's rank correlation efficient is used to quantify the accuracy. The new framework is illustrated using a DEM derived from LiDAR for sea-level rise vulnerability assessment of a tidal salt marsh. Compared to conventional methods based on statistics, the new framework and methods produce detailed mapping of error distribution thus enable the identification of the main sources of error and where improvement is needed most. Results call for re-evaluation of the current practice of assessing filtering accuracy in LiDAR data processing and further research on relative elevation and isomorphism.

1. INTRODUCTION

High-accuracy elevation data is essential to many applications such as flood risk mapping and sea-level rise vulnerability assessment. Light Detection and Ranging (LiDAR) is a standard source to obtain such data. Because of the high cost of LiDAR data acquisition and the steep learning curve of LiDAR data processing, few users fly LiDAR for their own projects but seek available LiDAR data and LiDAR-derived Digital Elevation Models (DEM). Different applications have different demands for elevation data accuracy, thus users of a LiDAR product not customized for their projects usually need an assessment of its accuracy to determine the product's appropriateness for their needs. For DEM producers, accuracy assessment is also important for the purpose of quality control and quality assurance. A sound framework for DEM accuracy assessment is thus necessary.

Many frameworks and methods to assess DEM accuracy have been reported. The majority focuses on the vertical accuracy aspect and resorts to statistical methodology for quantification (Höhle, Höhle, 2009; Maune, 2007). However, literature has pointed out the limitation of this methodology (Hu et al. 2009a). More importantly, increasing amount of research has realized that vertical accuracy is but *one* aspect of DEM accuracy; other aspects such as a DEM's ability to reproduce critical terrain features are also important. As illustrated by Chassereau et al. (2011) who compared a LiDAR-derived DEM with GPS-collected field data, a DEM of high vertical accuracy ended up modelling the overall shape of a salt marsh poorly because of its inability to depict terrain's microtopograhic variations. At the time being, there is little research beyond general discussions on how to assess this aspect of DEM accuracy, further research is thus necessary.

This paper strives to address the two challenges identified above. Given that terrain is a complex surface and DEM is meant to model it, it is imperative to understand terrain properties first. Only by doing so, it is possible to create a DEM to adequately account terrain properties. Liu et al. (2015) discussed terrain properties and their implications to DEM generation. This paper focuses on two of these properties which are based on absolute elevation and relative elevation respectively. The first is that the elevation at a location T is not random but has a single and determined true elevation denoted by z_T. While its exact value may never be known, z_T exists and has a single value. To account for this property of terrain, a DEM is expected to produce an accurate estimate of the absolute elevation at any points. This is the vertical accuracy aspect which has been examined intensively and extensively in the literature. The other property of terrain is that terrain is not a random collection of points but a surface resulting from the orderliness of point elevations. It is this elevation order that determines how water flows and what kind of landforms exist. For applications studying flood and landforms, a DEM's ability to preserve elevation order is very important, in fact more important than the vertical accuracy aspect because elevation order is about the *rank* of each point in the terrain in terms of elevation, i.e. the highest, the second highest, the third highest etc.. It is the relative elevation, not absolute elevation, that is of primary concern. Despite its importance, the implication of this property to DEM accuracy assessment and DEM generation is much less discussed in the literature. Researchers recognize it explicitly (Chassereau et al., 2011; Maune, 2007), but the mathematical rationale behind it is not articulated and methods to quantitatively assess it are yet to be reported.

This paper presents a new examination of the above two aspects in DEM accuracy assessment. For each aspect, a new method will be given; the mathematical rationale behind each method will also be articulated – approximation theory for vertical accuracy and isomorphism for elevation order. A LiDAR-derived DEM, employed to assess the vulnerability of a salt marsh to sea-level rise, is used to illustrate the methods.

2. VERTICAL ACCURACY

2.1 Error Component

Generally speaking, there are two main steps when generating a DEM from a LiDAR point cloud: filtering and interpolation. During filtering, each LiDAR point is labelled as "ground", "vegetation", "building" etc. During interpolation, bare-earth ground points are used to create a Triangulated Irregular Network (TIN) to which interpolation can be applied to creating a raster DEM. Because LiDAR sensor/platform has inherent errors, the filtering algorithm is rarely 100% accurate, meaning that there are omission and commission errors in LiDAR points classified as bare-earth ground. When TIN is constructed and converted to raster DEM, these errors are propagated through the interpolation function. As a result, a point in a LiDAR-derived DEM has three error components: error due to LiDAR sensor/platform, error due to filtering/classification, and interpolation error. The vertical error at a point T covered by a DEM, denoted as ΔZ_T, is the sum of these three components:

$$\Delta Z_T = Z_T - z_T = \mathcal{E}_T + G_T + R_T \tag{1}$$

where Z_T = T's elevation in a LiDAR-derived DEM
z_T = T's true bare-earth elevation
$\mathcal{E}_T$ = error due to LiDAR sensor or platform
G_T = ground error due to imperfect filtering
R_T = interpolation error

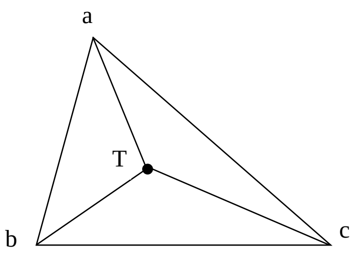

Figure 1. Triangular interpolation.

When triangular interpolation (Figure 1) is used, there are

$$\mathcal{E}_T = w_a \mathcal{E}_a + w_b \mathcal{E}_b + w_c \mathcal{E}_c$$

$$R_T = w_a z_a + w_b z_b + w_c z_c - z_T$$

$$G_T = w_a G_a + w_b G_b + w_c G_c$$

Where w_a, w_b, w_c = weight of point a, b, and c respectively; $w_a + w_b + w_c = 1$
$\mathcal{E}_a, \mathcal{E}_b, \mathcal{E}_c$ = LiDAR-senor error at a, b, and c respectively
z_a, z_b, z_c = true bare-earth elevations at a, b, and c respectively
G_a, G_b, G_c = ground errors at a, b, and c respectively

Ground error is the difference between a point's elevation reported by an error-free LiDAR sensor and the point's true bare-earth elevation. This error is introduced by filtering, labelling, and/or classification and has nothing to do with the LiDAR sensor or platform. In the literature, filtering accuracy is widely reported as a percentage value such as overall accuracy, Kappa statistics, or type I/ II error (Meng et al., 2010; Sithole, Vosselman, 2004; Yan et al., 2012). While these statistics summarize the percentage of points classified correctly, they do not carry the critical information on the error occurred at each misclassified point. For example, two points both have true bare-earth elevation of 10 cm but LiDAR-reported elevations are 9 cm and 13 cm respectively. Both are misclassified as bare-earth ground thus both are included in TIN interpolation and DEM generation. However, the 9-cm point only introduces an error of -1 cm while the 13-cm point introduces an error of 3 cm. If only summary statistics is used to report filtering accuracy, the important information that one misclassified point carries less error than the other misclassified point will be lost.

To stress the importance of retaining such information, we use the term ground error, not filtering error, in this paper and define it as follows: the ground error at a LiDAR point classified as bare-earth ground, e.g. point a in Figure 1, is

$$G_a = LIDAR_a - z_a \tag{2}$$

where z_a = true bare-earth elevation at a
$LIDAR_a$ = elevation at a reported by an error-free LiDAR

In typical TIN construction and triangular interpolation, only points classified as bare-earth ground are included thus ground error is concerned mostly at these points. However, the concept of ground error is applicable to *any* points. For example, if vegetation heights are known, the bare-earth elevation of points classified as vegetation can be inferred by subtracting vegetation heights from LiDAR-reported heights. These vegetation points can then be included in DEM generation. Ground error as defined in Equation 2 can be easily modified to apply to such situations. This flexibility is important for circumstances where sparse bare-earth ground points are available, e.g. non-open space. It enables the employment of non-bare-earth LiDAR points in DEM construction after their above-ground elevations are shaved.

2.2 Approximation Theory

Previous research has shown that, while LiDAR-senor error ($\mathcal{E}_T$) may be random error, interpolation error (R_T) and ground error (G_T) are not, they are systematic error instead (Liu et al., 2015). Since the total vertical error at a point is the sum of these three error components, vertical error cannot be random. As a result, the errors cannot be assumed to follow normal distributions as in the case of random error. This challenges the validity of using parametric statistics such as 95% confidence interval or Root Mean Squared Error (RMSE) to describe a DEM's vertical accuracy. A viable alternative is approximation theory which is widely used in numeric analysis in computational science to control the errors introduced when approximating a function by simpler functions (Atkinson, K. and Han, W., 2004). For example, the function $f(x)$ in Figure 2 is to be approximated by $F(x)$ and a set of reference points are provided. According to approximation theory, the accuracy of the overall approximation is determined by the largest error of *any* point in the entire domain, i.e., max $|F(x) - f(x)|$. The rationale is simple: If the largest error is acceptable, the error at *any* point must also be acceptable, hence the overall approximation is acceptable.

In the case of DEM research, terrain is the function $f(x)$ to be approximated by a simplifier function which is the DEM. By approximation theory, the vertical accuracy of a DEM is determined by the largest error at *any* point in the terrain. Based on Equation 1, the vertical accuracy of a DEM is controlled by

$$\max|\Delta Z_T| = \max|\mathcal{E}_T + G_T + R_T| \leq \max|\mathcal{E}_T + G_T| + \max|R_T| \quad (3)$$

After rearrangement of the terms, it can be seen that the largest error at a point is capped by the sum of the largest interpolation error, the largest ground error, and the largest sensor error. In this paper, sensor and ground error are grouped together because of the difficulty to separate them by users who did not produce the LiDAR dataset by themselves.

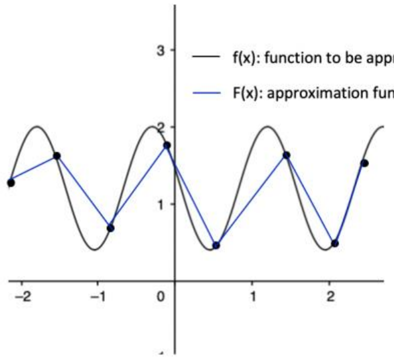

Figure 2. Approximation theory to estimate accuracy

2.3 Sensor and Ground Error ($\mathcal{E}_T + G_T$)

Sensor error is typically provided by the metadata of a LiDAR-acquisition project. LiDAR sensor error is usually believed to be random error, hence it is routinely reported using RMSE. Currently, many LiDAR sensors have a RMSE around 9 cm or better.

Ground error is ideally measured when evaluating filtering accuracy. If a is a LiDAR point classified as bare-earth ground, and if the LiDAR sensor used is error free, then a's ground error is the difference between these two elevations. In reality, it is difficult to separate sensor error from ground error, therefore they are assessed together, i.e.

$$\Delta Z_a = Z_a - z_a = \mathcal{E}_a + G_a \quad (4)$$

where Z_a = LiDAR-reported elevation at a,
z_a = a's true bare-earth elevation.

Note a in Equation 4 refers to a LiDAR point while T in Equation 3 refers to any point on the terrain covered by the DEM.

2.4 Interpolation Error

Hu et al. (2009a) shows that the interpolation error in triangulated interpolation is bounded by

$$|R_T| \leq \frac{3}{8} M_2 h^2 \quad (5)$$

where R_T = interpolation error at point T
M_2 = the maximum norm of the second-order derivative,
h = the longest edge of the triangle containing T.

h can be calculated based on the coordinates of triangle vertices. M_2 is the maximum norm of the second-order derivative of the triangle containing T. Since the first-order derivate is slope, the second-order derivative is slope of slope, i.e. curvature. M_2 can be calculated using the method described in Liu et al. (2015). Equation 5 articulates why large errors in DEM tend to occur in steep areas or where points density is low. For flat areas where curvature is near 0, even if only sparse points are available, the interpolation error will be small. On the other hand, if an area is curvy, dense points are needed in order to reduce interpolation error.

3. ISOMORPHISM

While a DEM's ability to estimate the absolute elevation at a point is important, its ability to reproduce the order and rank of terrain points according to their elevations is equally important if not more. Taking flow direction as an example: whether water flows from point a to point b or vice versa depends little on their absolute elevations but whose elevation is higher. Thus for a DEM to be useful for applications such as watershed management and landform analysis, its ability to preserve elevation order is critical, more critical than vertical accuracy.

Literature recognizes the importance of this aspect of DEM and had approached it from relative elevation perspective (Dakowicz and Gold, 2003). While relative elevation is very relevant to elevation order, the two are not exactly the same. Relative elevation, like absolute elevation, is about the distance above or below a reference. While absolute elevation uses the mean sea level as the reference, relative elevation uses another point. For example, if the absolute elevation of point a and b are 10m and 30m respectively, the relative elevation of point b with reference to point a would be 20m. As it can be seen, the value of a relative elevation has both magnitude and direction. For the purpose of identifying the higher point, however, magnitude is irrelevant; direction or sign of the value is sufficient. As long as the relative elevation of point b is positive – it does not matter whether the magnitude is accurately-estimated as 30m or poorly-estimated as 1m, one can conclude that point b is higher.

Like relative elevation, elevation order involves comparing points according to their elevations. In fact, elevation order of a terrain can only be determined by comparing every possible pair of points. However, elevation order is about ranking points according to their elevations, i.e. finding the highest point in the terrain, the second highest point, the third highest point etc. How much a point a higher than the other points is not of interest. It is from this perspective that this paper uses elevation order or elevation ranking instead of relative elevation.

Sparse discussion is available in the literature on how to evaluate a DEM's ability to preserve elevation order. In this paper, we draw on the concept of isomorphism in set theory. Let a and b be two points whose true elevations are z_a and z_b, respectively; their elevations in the DEM are Z_a and Z_b correspondingly. The ability of a DEM to preserve elevation order can be described mathematically as follows: If $z_a \leq z_b$, will $Z_a \leq Z_b$? Similarly, if $Z_a \leq Z_b$, will $z_a \leq z_b$? The first question examines whether a lower point in terrain will remain lower in a DEM. The latter question examines whether a lower point in a DEM is indeed lower in the field. If a DEM can answer yes to both questions, it is called an isomorphic DEM. In reality, the latter question is especially important because most often a DEM user cannot visit the field in person but relies on the DEM to understand the terrain of his/her study site.

An isomorphic DEM guarantees the preservation of elevation order, thus appropriate for applications involving flow directions and terrain structures. Hu et al. (2009b) discussed the necessary conditions to create an isomorphic DEM and applied them to the examination of three interpolation methods. They found that, when input data is error free, linear interpolation and triangular interpolation can generate isomorphic DEM but bilinear interpolation cannot. While their mathematical proofs are valuable, how to assess whether a DEM is isomorphic remains a challenge because, in reality, LiDAR data is never error free. Even if an isomorphic method is used to generate a DEM, e.g. triangular interpolation as is often used in TIN to Raster DEM conversion, filtering error and ground error discussed previously in Section 2.1 may still render the resultant DEM non-isomorphic.

In this paper, we explore Kendall's rank correlation coefficient as a first solution. Kendall's rank correlation coefficient is a non-parametric method that compares the ranks of two sets by measuring their ordinal association. Let $\{ Z_i \}$ be the set of points in a DEM, $\{ z_i \}$ be their corresponding true elevations. Any pair of observant (z_i, Z_i) and (z_j, Z_j) are said to be concordant if both $z_i < z_j$ and $Z_i < Z_j$, or if $z_i > z_j$ and $Z_i > Z_j$. They are said to be discordant if $z_i < z_j$ but $Z_i > Z_j$, or if $z_i > z_j$ but $Z_i < Z_j$. The Kendall's rank correlation efficient τ measures the difference between concordant pairs and discordant pairs, and normalizes the difference by the number of possible pairs. Kendall's τ value ranges between -1 and 1: 1 means the two rankings agree perfectly; -1 means one ranking is the reverse of the other; 0 means the two sets are independent with no correlation. Thus, an isomorphic DEM would have Kendall's τ of 1. The higher a DEM's Kendall's τ is, the better terrain's elevation order is preserved.

The next section illustrates the ideas and methods discussed in Section 2 and Section 3 using a case study.

4. A CASE STUDY

4.1 Data and Study Area

As an illustration, we assess the accuracy of a LiDAR-derived DEM. The study site is a tidal salt marsh in China Camp State Park located approximately 25 km north of the city of San Francisco in California, U.S.A. Tidal salt marsh and surrounding estuaries are among the World's most biologically productive ecosystems, providing commercial harvesting and recreational fishing opportunities that produce tremendous economic benefits (Allen, J. et al., 1992). Marsh elevation relative to the local tidal range, however, must be monitored closely, as deviations from established tidal marsh elevations as small as 10 cm can change salt marsh ecology, erosion and accretion (Silvestri et al., 2005; Vanderzee, M., 1988). An accurate DEM is thus critical.

To obtain an accurate DEM at low cost, LiDAR data from the Golden Gate LiDAR Project was used. The Golden Gate LiDAR Project was commissioned by the United States Geological Survey; its dataset including a DEM is publicly available. The minimum point density is 2 points/m^2 and the vertical accuracy is less than or equal to 9.25 cm when measured as root mean squared error (RMSE) (Hines, E., 2011). The producers used TerraScan, a proprietary LiDAR processing software to filter the dataset; the result is a set of points classified as bare-earth ground (Hines, 2011). For our study site of China Camp salt tidal marsh, over 18 million points were extracted to cover it; bare-earth points within the study site boundary were used to build a TIN. Triangular interpolation was then conducted to generate a 1-m DEM.

In addition to LiDAR data, a set of 753 points collected by the United States Geological Survey (USGS) in China Camp using Real Time Kinetic (RTK) Global Position System (GPS) during the same time when LiDAR was flied was also available and obtained. The horizontal and vertical accuracy of these points are reported as ± 1 cm and ± 2 cm respectively (Takekawa, et al. 2013). A vegetation classification map, created from a 0.25-m resolution colour-infrared aerial image of the study site, was also obtained from the United States National Oceanic and Atmospheric Administration. The overall accuracy of the classification was 91% and the Kappa statistics was 82% for six vegetation species.

4.2 Vertical Accuracy

4.2.1. Sensor and Ground error

LiDAR data filtering was conducted by the Golden Gate LiDAR Project using a proprietary software; the only publicly-available information is that the overall accuracy of filtering process was 95%. Since this value does not tell ground error, this research did a separate assessment of it. In theory, ground error calculation requires the LiDAR sensor to be error-free (Equation 2) but this is not feasible, therefore sensor error and ground error were assessed together using Equation 3. Among the 753 RTK GPS points, 733 had a bare-earth ground LiDAR point within 1 m. The LiDAR-reported elevation of these 733 points were compared with the RTK-GPS elevation which was used as the true bare-earth elevation in this research; the difference is the sum of sensor and ground error.

Figure 3. Absolute error at RTK GPS points.

Results show that the vertical difference between the elevation values reported by RTK GPS and LiDAR ranged between -51.5 cm to 47.0 cm; the median is 18.1 cm, the mean is 17.9 cm, and the standard deviation is 8.6 cm. Histogram shows that the errors, when taking absolute values as needed by approximation theory, are only slightly positively skewed with a skewness of 0.37. The 95th percentile is 31.3 cm. Figure 3 shows the spatial distribution of these absolute errors at RTK GPS points.

4.2.2. Interpolation Error

To calculate interpolation error, triangles in the TIN were extracted so as to calculate h and M_2. Triangles near site boundary were excluded – many such triangles are erroneous because LiDAR bare-earth points outside our study site were not included in TIN construction. Overall, the study area was covered by over 4.9 million triangles.

The longest edge of these triangles (h) ranged between 0.9 cm to over 2 m. The median value of h was 84 cm, the mean was 86 cm, and the standard deviation was 57 cm. Histogram showed that the distribution was highly positively skewed. 19% triangles had their longest edge less than 50 cm, 50% between 50 cm and

1 m, 30% between 1 cm and 2 cm. Only about 1% triangles had its longest edge greater than 2 m.

The maximum norm of the second-order derivative (M_2) of these triangles can be calculated using the method in Liu et al. (2012) where contour lines are generated from TIN and M_2 is calculated as the slope change along the flow line of waterdrops between adjacent contour lines. This method is rigorous but involves complex computations. Since second-order derivative is essentially curvature, this study calculated curvature using a GIS software and M_2 is the largest curvature values. Results show that M_2 in the study area varied between -0.1 and 0.1 with a mean of nearly 0. This is expected since our study area is a tidal salt marsh. Using Equation 5 to combined M_2 and h , interpolation errors are calculated. It is found that 88.7% of the area had an interpolation error less than 10 cm and 4.7% had an interpolation error between 10 cm and 20 cm, the rest was over 20 cm. In other words, more than 95% area had an interpolation error less than 20 cm.

Figure 4 shows the spatial distribution of interpolation errors. It can be seen that nearly all large errors (> 20 cm) are found along tidal creeks where LiDAR bare-earth points are much sparser than other areas. In fact, the longest edge in triangles in these areas was 105 cm on average with a standard deviation of 56 cm. In contrast, the triangles in areas whose interpolation error was less than 20 cm had their longest edges much shorter, about 26.8 cm. The curvature of the triangles where large interpolation errors occurred turned out also higher than other areas. These two factors combined explain the large interpolation error found along tidal creeks. This result echoes the observation by Chassereau et al. (2011) that LiDAR-derived DEM performed best on the marsh platform but poorly along tidal creeks and creek heads.

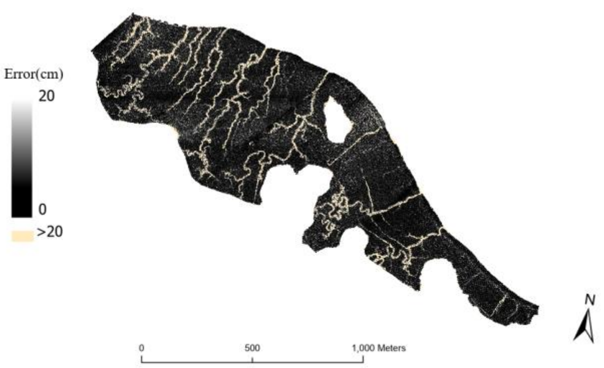

Figure 4. Spatial distribution of interpolation errors

In the NOAA vegetation map of the study site, some areas do not have vegetation information. These areas turned out to overlap significantly with areas whose interpolation error was over 20 cm. Field knowledge suggests that these are the most challenging areas for data collection, for both LiDAR and airborne colour-infrared imagery. LiDAR point density is much lower in these areas. Thus, future efforts to create a more accurate DEM for China Camp should focus on these areas.

4.2.3. Overall vertical error

As shown in Equation 1, the total error at a DEM point is bounded by the sum of sensor error, ground error, and interpolation error. Due to the limited number of RTK GPS points available to assess sensor and ground error, a surface showing the spatial distribution of each error cannot be generated. However, since the sum of the two errors were bounded by 51.5 cm in our samples, we will use 51.5 cm as the error bound for sensor and ground error combined. If 95th percentile of each types of errors were used – 31.3 cm in ground and sensor error combined, 20 cm in interpolation error, it is safe to say that quite some of the study area had a vertical error exceeding 51.3 cm. Such vertical accuracy renders the DEM not suitable for applications such as vulnerability to sea-level rise. The California Climate Action Team (2013) estimates that sea-level rise for coastal areas in California including our study site is 15 cm between the years of 2000 and 2030, 30 cm between 2000 and 2050, and 83 cm between 2000 and 2100. To use a DEM to study the vulnerability of China Camp tidal salt marsh to these projected sea-level rises, the DEM's vertical accuracy must be at least twice as certain as the sea-level rise increment (NOAA 2010). This means that the corresponding DEM must have a vertical error no more than 7.5 cm, 15 cm, and 41.5 cm in order to assess the impact in 30, 50, and 100 years respectively. Since the vertical error of the tidal marsh DEM in this research clearly exceeds the expectation, it is not useful to study the impact of sea-level rise to China Camp salt marsh.

It has to be pointed out that the examination of vertical accuracy did not take horizontal accuracy, i.e. LiDAR points' positional accuracy, into account. As noted by ASPRS (2015), while horizontal errors in elevation data do not always impact vertical accuracy, they normally contribute significantly to the error detected in vertical accuracy tests. The Golden Gate LiDAR dataset did not report its horizontal accuracy but it must have positional errors. Thus the true overall vertical error in the LiDAR-derived DEM is likely to be even higher than that reported in this paper.

4.3 Isomorphism

Kendall's τ rank correlation was calculated by comparing the elevation values at RTK GPS points: One reported by the RTK GPS, the other by the LiDAR-derived DEM. Kendall's τ rank correlation between the two sets of values was 0.38, far from the ideal value of 1, suggesting that the two datasets' rank orders do not agree well. We also calculated Pearson's correlation coefficient between the two datasets and its value was 0.6. Pearson's correlation efficient measures how well LiDAR-derived DEM elevation agrees with RTK-GPS reported elevation; it provides a summary of the vertical accuracy of the LIDAR-derived DEM. Kendall's τ correlation efficient, on the other hand, evaluates how they agree in terms of the elevation ranks of the points. The fact that Pearson's correlation efficient is much higher than Kendall's rank correlation coefficient tells that, between vertical accuracy and isomorphism, our LiDAR-derived DEM had a better performance in vertical accuracy. The exact elevation at a point may be estimated quite accurately by the LiDAR-derived DEM, but the elevation order of the tidal salt marsh was not reproduced well. This explains the DEM's poor ability to reflect the micro topology of the marsh especially on area along small-scale tidal creek networks.

5. DISCUSSION AND CONCLUION

In this paper, we presented a new framework to assess the accuracy of LiDAR-derived DEM, and illustrated it using a case study. Compared to existing approaches which often focus exclusively on vertical accuracy, the new framework takes into account isomorphism which is another critical aspect in DEM generation. Furthermore, for vertical accuracy assessment, a new methodology based on approximation theory from numerical analysis is introduced. Unlike statistical methodology where DEM's accuracy is typically described using one or a few summary statistics, the approximation-theory-based approach creates a map showing the error bound at each point. Such a map is valuable as it points out where large errors occur; subsequent efforts can thus focus on these areas to effectively and efficiently improve the overall quality of the DEM.

The separation of sensor error, ground error, and interpolation error also helps identify the main source of error. In the case that interpolation error is higher than the combined sensor and ground error, further efforts to improve the DEM should focus on the interpolation phase in DEM generation. Recall that interpolation error depends on the second-order derivative of terrain (M_2) and the longest edge of a triangle (h^2). Terrain curvature is fixed, but the longest edge of a triangle can be reduced by having denser points during triangulation. In our study, we limited triangulation based on bare-earth points only. In the case an area has few bare-earth points, non-bare-earth points may also be used provided their ground error is carefully assessed so that their bare-earth elevations are reliably estimated. As high-density LiDAR points are increasingly available, it can be expected that h^2 will reduce significantly, thus interpolation error will be significantly smaller.

On the other hand, if combined ground error and sensor error are found higher than interpolation error, further efforts on improving DEM should focus on the filtering stage since little can be done on LiDAR sensor/platform error. Filtering is a critical step in LiDAR data processing. However, existing methods focus on classification and describes filtering error in terms of summary statistics similar to those used in assessing the accuracy of remote sensing image classification. Our research found that such a method, though serves the purpose of describing classification accuracy, does not help DEM accuracy assessment and quality control. Instead, we propose the concept of ground error which is the elevation difference between LiDAR-reported elevation and the true bare-earth elevation. Measurement of ground error should be conducted during filtering. For example, during the ground truthing phase which is usually required to assess filtering accuracy, one collects not only the class/label of a point but also measures its true bare-earth elevation, especially at points that are classified as bare-earth. In this way, whether a point is indeed bare-earth is known as well as the error in its LiDAR-reported elevation.

Another implication of this research is isomorphism. As pointed out already, DEM's ability to preserve elevation order is an aspect whose importance is widely recognized but no established method exists yet to assess it. This research introduced isomorphism as the mathematical rationale and applied Kendall's rank correlation coefficient to quantify it. Though Kendall's rank correlation coefficient provides a summary on how well elevation ranking and orders are preserved, it is far from satisfactory because it does not provide any information on the spatial distribution of the errors. Many questions remain for future research on isomorphism – should it be assessed using statistical methodology as we did in this research? Or there are better methods? In the case of statistical methodology, how many samples should be used and where to collect these samples? How to ensure that results based on sampling can be generalized to the entire DEM? Admittedly, these are grand challenges for future DEM research, but they are extremely important.

Isomorphism relates to another important issue in DEM research, namely DEM generalization. As high-density point clouds are increasingly available, point reduction has become necessary in generating LiDAR-derived DEM. However, mechanistic reduction will result in the loss of important points which defines the terrain structure. Generalization which involves selection, reduction, aggregation, and even exaggeration is thus necessary. In fact, even before LiDAR, generalization is necessary for *any* DEM generation because terrain is made up by infinite number of points but a DEM can only comprise a limited number of them. As discussed in Section 3, a DEM can preserve terrain structure only if it is created by an isomorphic process which ensures high points remain high and low points remain low in the output DEM. Some DEM generalization algorithms are available in the literature, but few has studied whether they are isomorphism or not. Meanwhile, errors introduced by LiDAR sensor and filtering algorithm may result in non-isomorphic DEM even if a DEM generation method is proved to be isomorphism in theory. Intentional errors introduced by exaggerations and displacements during generalization may further complicate the issue. More advanced research on isomorphism is necessary in order to create DEMs effectively accounting all terrain properties.

ACKNOWLEDGEMENTS

The author thanks Adam McClure for sharing the LiDAR and GPS dataset.

REFERENCES

Allen, J., Cunningham, M., Greenwood, A., Rosenthal, L., 1992: *The Value of California Wetlands: An Analysis of their Economic Benefits*. Oakland, California: The Campaign to Save California Wetlands.

American Society for Photogrammetry and Remote Sensing (ASPRS), 2015: ASPRS positional accuracy standards for digital geospatial data. *Photogrammetry Engineering and Remote Sensing* 81(3), A1-A26.

Atkinson, K., Han, W., 2004: *Elementary Numerical Analysis*. John Wiley & Sons, Hoboken, NJ, USA.

Chassereau, J.E., Bell, J.M., Torres, R., 2011. A comparison of GPS and Lidar salt marsh DEMs. *Earth Surface Processes and Landforms* 36, 1770–1775.

Dakowicz, M., Gold C., 2003. Extracting meaningful slopes from terrain contours. *International Journal of Computational Geometry & Applications* 13, 339-357.

Hines, E., 2011: *Final Report: Golden Gate LiDAR Project*. San Francisco, California: San Francisco State University, 12p.

Höhle, J., Höhle, M., 2009. Accuracy assessment of digital elevation models by means of robust statistical methods. *ISPRS Journal of Photogrammetry and Remote Sensing* 64, 398–406.

Hu, P., Liu, X., Hu, H., 2009a. Accuracy assessment of digital elevation models based on approximation theory. *Photogrammetry Engineering and Remote Sensing* 75, 49–56.

Hu, P., Liu, X., Hu, H., 2009b. Isomorphism in digital elevation models and its implication to interpolation functions. *Photogrammetry Engineering and Remote Sensing* 6, 713-721.

Liu, X., Hu, H., Hu, P., 2015. The "M" in digital elevation models. *Cartography and Geographic Information Science* 42, 235–243.

Liu, X., Hu, P., Hu, H., Sherba, J., 2012. Approximation theory applied to DEM vertical accuracy assessment. *Transaction in GIS* 16, 397-410.

Maune, D.F (Ed.), 2007: *Digital Elevation Model Technologies and Applications: the DEM Users Manual*, 2nd ed. American Society of Photogrammetry and Remote Sensing, Bethesda, MD, USA.

Meng, X., Currit, N., Zhao, K., 2010. Ground filtering algorithms for airborne LiDAR data: A review of critical issues. *Remote Sensing* 2, 833–860.

Silvestri, S., Defina, A., Marani, M., 2005. Tidal regime, salinity and salt marsh plant zonation. *Estuarine, Coastal and Shelf Science* 62, 119–130.

Sithole, G., Vosselman, G., 2004. Experimental comparison of filter algorithms for bare-earth extraction from airborne laser scanning point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing* 59, 85–101.

Takekawa, J., Thorne, K., Buffington, K., Spragens, K., Swanson, K., Drexler, J., Schoellhamer, D., Overton, C., Casazza, M., 2013: *Final Report for Sea-Level Rise Response Modeling for San Francisco Bay Estuary Tidal Marshes*. U.S. Geological Survey Open File Report 2013-1081, 161p.

Vanderzee, M., G.I., 1988. Changes in saltmarsh vegetation as an early indication of sea-level rise. In: Pearman, G. (Ed.), *Greenhouse: Planning for Climate Change*. Brill Archive, pp. 147–160.

Yan, M., Blaschke, T., Liu, Y., Wu, L., 2012. An object-based analysis filtering algorithm for airborne laser scanning. *International Journal of Remote Sensing* 33, 7099–7116.