

Article

Accuracy Assessment of LiDAR-Derived Digital Elevation Models Based on Approximation Theory

XiaoHang Liu ¹, Hai Hu ^{2,*} and Peng Hu ²

¹ Department of Geography and Environment, San Francisco State University, 1600 Holloway Avenue, San Francisco, CA 94132, USA; E-Mail: xhliu@sfsu.edu

² School of Resource and Environment Science, Wuhan University, Wuhan 430079, China; E-Mail: penghu@whu.edu.cn

* Author to whom correspondence should be addressed; E-Mail: huhai@whu.edu.cn; Tel.: +86-27-6877-8319; Fax: +86-27-6877-8893.

Academic Editors: Richard Müller and Prasad S. Thenkabail

Received: 16 February 2015 / Accepted: 7 May 2015 / Published: 29 May 2015

Abstract: The cumulative error at a point in a LiDAR-derived DEM consists of three components: propagated LiDAR-sensor error, propagated ground error, and interpolation error. To combine these error components so as to assess the vertical accuracy of a LiDAR-derived DEM, statistical methods based on the error propagation theory are often used. Due to the existence of systematic error, statistical methods are only effective if a large number of checkpoints are available, which may not be affordable in many practical applications. This paper presents approximation theory as an alternative methodology that departs from error propagation theory in fundamental ways. Using approximation theory, an error bound of the cumulative error at any point in the study site can be obtained, thus informing users conservatively of the spatial variation of DEM accuracy and pointing out the weakly determined areas. The new method is illustrated from DEM users' perspective by assessing whether a publicly available LiDAR-derived DEM meets FEMA's accuracy standard for flood risk mapping. The paper calls for a change in the existing methods of assessing and reporting the errors in a LiDAR-derived DEM, in particular those introduced during the ground filtering process.

Keywords: LiDAR; DEM; accuracy; approximation theory; error

1. Introduction

High-density, high-accuracy elevation data are essential for myriad applications such as flood mitigation, coastal management, and topographical mapping [1–3]. To acquire such data, Light Detection and Ranging (LiDAR) has become the established technology in the aeroservice community. LiDAR is an optical remote sensing system that actively illuminates objects with laser pulses. By measuring the time it takes for the pulses to travel through space, interact with the object, and return to the detector, the distance between the object and the sensor can be calculated, thus enabling inferences about elevation and height information [4]. Compared to other elevation sources such as topographic maps, photogrammetry, and satellite remote sensing, LiDAR-acquired elevation has unprecedented resolution and accuracy. LiDAR's ranging information can also be augmented with GPS-based location and sensor-based orientation to enable three-dimensional representation [5]. Because LiDAR has the potential to benefit so many applications, a national LiDAR dataset for the entire U.S. has been proposed [2].

While LiDAR brings exciting new opportunities, the high cost of collecting LiDAR data prevents most projects from conducting their own LiDAR flight. Even if a LiDAR flight is affordable, the direct output from LiDAR is only a point cloud, which has to be processed to scatter points of elevations. The scatter points serve as a topographical model, but most LiDAR users would like this to be deliverable in grid format so that existing cartography and terrain analysis methods can be applied [6]. To this end, the scatter points must be converted to a grid DEM through filtering, interpolation, and accuracy assessment. The steep learning curve for understanding and processing the raw point cloud makes it difficult for most DEM users to produce their own LiDAR-derived DEM. Instead, most DEM users resort to a LiDAR-derived DEM already available for their study area. In the United States, such LiDAR products are often provided by federal agencies such as the Federal Emergency Management Administration (FEMA) and United States Geological Survey (USGS), along with county and state agencies.

While publicly available LiDAR-derived DEMs relieve DEM users from LiDAR acquisition and processing, such DEMs were usually collected under the general goal of supporting a wide variety of non-DEM science and mapping applications that are in need of LiDAR technology. The producers acquire, process, and deliver the LiDAR data following general guidelines and standards such as the U.S. National Standard of Spatial Data Accuracy (NSSDA). However, real-world applications have different accuracy criteria. The accuracy reported by DEM producers does not always translate to the accuracy standard of DEM users. Moreover, most LiDAR producers only provide a global accuracy for the entire project area, masking the spatial variation of accuracy. Whether the LiDAR product suits a DEM user's specific need is therefore unknown.

In this paper, we present a new method to assess a LiDAR-derived DEM. The new method is different from the widely used statistical methodology in that it is based on approximation theory in numerical analysis. In the rest of the paper, we first discuss the limitations of statistical approaches, then present the theoretical framework and rationale behind our new method. The new method is illustrated from DEM users' perspective by assessing whether a USGS LiDAR dataset can be used for flood risk mapping. Along the way, we demonstrate the need for a change in assessing and reporting DEM error, in particular that introduced during the ground filtering process.

2. Error Components at a LiDAR-Derived DEM Point

To explain how our new method departs from statistical methods, we start with a clarification of the error components at a point in a LiDAR-derived DEM. While this subject has been probed by several previous studies [7,8], this paper is the first to define each error component mathematically instead of verbally. The rigor of mathematical definition enables the articulation of the type of each error component, which turns out to be critical in understanding the limitations in existing methods of DEM accuracy assessment. We start with a discussion of the errors introduced during LiDAR acquisition and processing, then show how these errors are propagated to a DEM point.

2.1. LiDAR Sensor Error

The generation of a LiDAR-derived DEM starts with the point cloud acquired by LiDAR sensors. LiDAR sensors, though much more accurate compared to other techniques, are not error-free. An elevation value returned by LiDAR thus has errors. Hodgson and Bresnen [7] explained that this error is due to two sources: the vertical error of the LiDAR sensor/platform and the impact of LiDAR horizontal error. Because the two errors are both due to LiDAR sensor system and often assessed together, they are referred to as LiDAR-sensor error in this research. Let a raw LiDAR return at point a be Z_a . If an error-free LiDAR sensor is used, the return at a would be z_a . Z_a and z_a are related by

$$Z_a = z_a + \varepsilon_a, \quad (1)$$

Where ε_a is the LiDAR-sensor error at point a . Note that the value of ε_a may be positive or negative.

2.2. Ground Error

The next step in processing a LiDAR point cloud is to label each raw LiDAR return as “ground,” “vegetation,” “building,” or other using automated and manual methods. This step, often called a labeling process [7] or the filtering process [8], is considered the most critical step for DEM generation from LiDAR data [9]. The goal is to identify bare-earth points so that a grid DEM can be created through interpolation. The exact method and parameters used in this step are often proprietary and dependent on the researchers or LiDAR mapping vendors. A review of the critical issues of ground filtering algorithms for airborne LiDAR data can be found in [10]. During the processing, some returns may be incorrectly labeled as “ground” while some true “ground” returns may be omitted. The “ground” return set thus contains errors. This error, called filtering error or classification error, is typically measured using methods developed for the accuracy assessment of categorical data, e.g., overall accuracy, Kappa statistics [10], and type I and type II errors [11].

In this article, we use ground error instead of filtering error to describe the error introduced by ground filtering. Unlike filtering error, which only concerns classification accuracy, ground error also concerns the magnitude of an error. For example, if a misclassified LiDAR return reads 20 m but the actual bare-earth elevation at the point is 15 m, the corresponding ground error is 5 m. To define ground error mathematically, let a be a point classified as ground. Let the LiDAR return at point a be Z_a . Z_a could be the first return, the second return, the last return, or a combination of them depending on the filtering algorithm [10]. According to Equation (1), $Z_a = z_a + \varepsilon_a$. Let the true bare-earth elevation at point a be

H_a . The ground error at a , denoted as G_a , is the difference between z_a and H_a , *i.e.*, the difference between an error-free LiDAR return (z_a) and the true bare-earth elevation (H_a). In other words, G_a is the above-the-ground height of the LiDAR return at a . If point a is indeed a bare-earth point, G_a is 0. Otherwise, G_a is a positive value. Equation (2) provides the definition of G_a :

$$G_a = z_a - H_a = Z_a - \varepsilon_a - H_a, \quad (2)$$

which is equivalent to

$$Z_a = H_a + \varepsilon_a + G_a. \quad (3)$$

2.3. Interpolation Error

Once the bare-earth points are identified through ground filtering, the next step is to apply an interpolation method to create a grid DEM. Many interpolation methods have been used in the literature, *e.g.*, nearest neighbor, kriging, and inverse distance weighting. This paper uses Triangulated Irregular Network (TIN) interpolation as an example, though the discussions can be easily extended to other interpolation methods (Figure 1). TIN is the *de facto* standard method used by FEMA to create LiDAR-derived DEMs because of its simplicity and effectiveness [12]. Most importantly, TIN interpolation is proven to be isomorphic [13]. This means that if point a is higher than point b on the ground, the TIN-interpolated elevation of point a is also guaranteed to be higher. Similarly, if point a is higher than point b in a TIN-interpolated DEM, a is guaranteed to indeed be higher in the field. Isomorphism is important because it preserves the elevation order and sequence among terrain points. An isomorphic interpolation function is necessary if a DEM is to be used for flood control, hydrological modeling, and geomorphology analysis [13,14].



Figure 1. Triangulated Irregular Network (TIN) interpolation. a , b , and c have known elevation values. The elevation of T is to be interpolated.

In Figure 1, T is a point whose bare-earth elevation is to be interpolated based on ground points a , b , and c . The interpolated elevation of T , denoted as Z_T , is a weighted sum of these elevations, *i.e.*,

$$Z_T = \omega_a Z_a + \omega_b Z_b + \omega_c Z_c \quad \omega_a + \omega_b + \omega_c = 1, \quad (4)$$

where ω_a , ω_b , ω_c are the areal proportions of the sub-triangles and they add up to 1. Interpolation error refers to the error introduced by the interpolation function even if error-free source data is used. Its mathematical definition will be derived in Section 2.4. The existence of interpolation error is illustrated in Figure 2. It can be seen that interpolation error is inevitable except in flat terrain because

TIN always approximates a terrain patch using a flat triangle no matter of the curvature. TIN introduces an overestimation error in concave terrain (Figure 2a) and an underestimation error in convex terrain (Figure 2b). The sign of an interpolation error is thus not random but depends on terrain curvature.

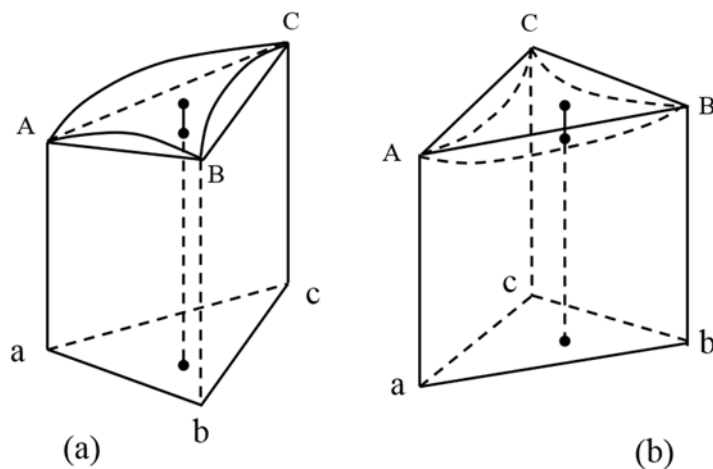


Figure 2. Interpolation using TIN. The interpolation always results in (a) an underestimation error in a convex terrain, and (b) an overestimation error in a concave terrain.

2.4. Error Propagation to a DEM Point

With the errors introduced in each step defined, the total error at a DEM point can be derived as follows. Let the LiDAR-based elevation at a , b , and c be Z_a , Z_b , and Z_c respectively. Per Equations (1–4), the interpolated elevation of T , denoted as Z_T , can be written as

$$Z_T = \omega_a Z_a + \omega_b Z_b + \omega_c Z_c = \omega_a (H_a + G_a + \epsilon_a) + \omega_b (H_b + G_b + \epsilon_b) + \omega_c (H_c + G_c + \epsilon_c). \tag{5}$$

Let the true error-free bare-earth elevation at T be H_T . The vertical error at point T , denoted as Δz_T , is the difference between Z_T and H_T , i.e.,

$$\Delta z_T = Z_T - H_T = (\omega_a H_a + \omega_b H_b + \omega_c H_c - H_T) + (\omega_a \epsilon_a + \omega_b \epsilon_b + \omega_c \epsilon_c) + (\omega_a G_a + \omega_b G_b + \omega_c G_c),$$

or, equivalently,

$$\Delta z_T = R_T + \epsilon_T + G_T \tag{6}$$

$$R_T = \omega_a H_a + \omega_b H_b + \omega_c H_c - H_T \tag{7}$$

$$\epsilon_T = \omega_a \epsilon_a + \omega_b \epsilon_b + \omega_c \epsilon_c \tag{8}$$

$$G_T = \omega_a G_a + \omega_b G_b + \omega_c G_c. \tag{9}$$

Equation (6) suggests that the total error at a point is the sum of three components. The first component is R_T , which is the interpolation error explained in Section 2.3. This error has nothing to do with LiDAR sensor error or ground filtering. It is entirely due to the imperfect nature of TIN interpolation. The sign of interpolation error is not random (Figure 2). The magnitude of interpolation error is also fixed because every point on terrain must have a true bare-earth elevation. Thus, even though the values of H_a , H_b , H_c , and H_T are unknown, each of them is a single fixed value. As the arithmetic

combination of them, interpolation error (R_T) must also be a single fixed value. One can thus conclude that interpolation error is not random error but systematic error.

Another error component is ε_T , which is called *propagated sensor error*. This is the LiDAR sensor error in a , b , and c (ε_a , ε_b , and ε_c) that is propagated to T through TIN interpolation. ε_T may be random error or a mixture of random and systematic error depending on whether there are systematic errors in the LiDAR sensor system. The last error component is G_T , which is the ground error at a , b , and c (G_a , G_b , and G_c) that is propagated to T during interpolation. Recall that ground error is the above-the-ground height of a LiDAR return that is classified as ground. Its value is 0 if this LiDAR return point is indeed bare earth; otherwise it is positive. G_T is a weighted sum of G_a , G_b , and G_c , per Equation (9); its value is therefore non-negative. Furthermore, because G_a , G_b , and G_c , are single fixed values, G_T must be also a single fixed value. G_T is thus systematic error, not random error. ε_T , G_T , and R_T are similar to the source data error, filtering error, and gridding error used in Aguilar *et al.* [8], though it is difficult to be certain because the errors in Aguilar *et al.* [8] were described only verbally, not mathematically.

The above discussion establishes two points. One is that the vertical error at point (T) in a LiDAR-derived DEM is the sum of three components defined in Equations (7–9). The other is that R_T and G_T are both systematic error, while ε_T may be random error or a mixture of random and systematic error. In either case, it suffices to say that the total error at a point in a LiDAR-derived DEM is not random, but a mixture of random and systematic error.

3. Problems in Applying Error Propagation Theory

With the error components at a DEM point clarified, the next question is how to assess the overall DEM, which consists of numerous points. In the literature, the dominant methodology is to use the error propagation theory in statistics as the theoretical framework (e.g., [7,8,15]). This theory says that if an error is the sum of several error components, e.g., $z = x \pm y$, each error can be described by its variance, e.g., σ_x^2 , σ_y^2 , σ_z^2 . Better yet, the variance of the cumulative error is the sum of the variances of each error, i.e., $\sigma_z^2 = \sigma_x^2 + \sigma_y^2$ or equivalently $\sigma_z = \sqrt{\sigma_x^2 + \sigma_y^2}$. An introduction to error propagation theory and how it has been used by the National Mapping Accuracy Standard (NMAS) in the United States can be found in [16]. Based on this theory, variance and Root Mean Square Error (RMSE) are routinely used. RMSE is considered an equivalent of variance, because under the assumption that errors are normally distributed, RMSE can be interpreted as standard deviation, which is the square root of variance [17]. Error propagation theory and RMSE are very popular in the literature. For example, Hodgson and Bresnahan [7] modeled the cumulative error at a DEM point as the sum of four error components. Each error component was described by its RMSE, and the RMSE of the cumulative error was modeled as the square root of the sum of RMSEs. Aguilar *et al.* [8] used a similar framework but modeled the cumulative error as the sum of three components; each component is quantified using variance estimated through a hybrid of theoretical and empirical methods. Even the U.S. National Standard of Spatial Data Accuracy [18] requires using the 95% confidence interval by calculating 1.96(RMSE), though the 95th percentile was introduced later to address RMSE inadequacy.

While the error propagation theory is applied widely in DEM accuracy assessment, the critical assumption behind this error is often neglected. Error propagation theory requires that all errors involved, including the cumulative error as well as each error component, must be random errors, independent of

each other, and identically distributed. The reason is because random error is usually normally distributed with a mean of 0, making it possible to describe the error using variance alone. Independence between error components makes covariances negligible so that the total variance can be written as the sum of individual variances [9]. In reality, numerous research has observed that DEM vertical errors are not random, not normally distributed, and not even from a spatially stationary process [15,17]. In fact, DEM errors are found to be correlated with terrain structure, land cover type, sampling density, and interpolation method [19–24]. The observations of spatial autocorrelation among DEM errors further suggest that vertical errors are not independent of each other. As derived in Section 2, both interpolation error and propagated ground error are systematic errors; even propagated sensor error may contain systematic error sometimes. These characteristics of DEM errors present serious challenges to the application of error propagation theory. Unless a large number of checkpoints are available, the effectiveness of RMSE and other statistical methods will be compromised [7,10,12,13]. Yet another problem with RMSE or 95th percentile is that it only provides a single global accuracy assessment, masking out the spatial variation of DEM quality [25]. For a DEM user, a DEM with acceptable global accuracy does not guarantee that the same accuracy holds for his/her specific need. Neither does a global accuracy point out weekly determined areas. To address these problems, other statistical methods have been explored such as quantile method [15], spatial regression [8,25], and geostatistics [26,27].

4. Approximation Theory: A Departure from Error Propagation Theory

In this paper, we present approximation theory as an alternative framework to assess DEM accuracy. Approximation theory is widely used in numerical analysis in computer science to evaluate the quality of approximating a complex function by simpler functions and quantify the errors introduced therein [28]. In Figure 3, $f(x)$ is a complex function whose exact form is unknown, but some samples are available and piecewise linear interpolation $F(x)$ is chosen to approximate $f(x)$. $f(x)$ is divided into segments based on the sample points. Each segment is approximated by the simpler linear function $F(x)$. 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 must be acceptable.

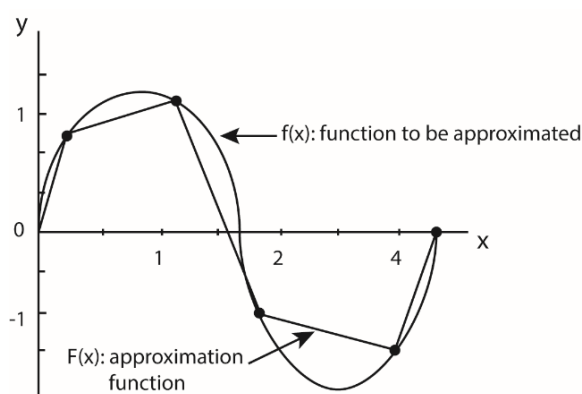


Figure 3. Illustration of approximation theory. A complex function $f(x)$ is approximated by a simpler function $F(x)$ using sample points.

In the context of LiDAR-derived DEM, terrain is the complex function to be approximated by a simpler function. If TIN is used, the terrain is divided into facets and each facet is approximated by a triangle. According to approximation theory, the approximation accuracy of each facet is determined by the largest error at *any* point in that facet, and the overall accuracy of the TIN-interpolated DEM is determined by the largest error at *any* point in the entire study site. Since the total error at a point T in a TIN-interpolated DEM, per Equation (6), is

$$\Delta z_T = R_T + \varepsilon_T + G_T,$$

it follows that

$$\max |\Delta z_T| = \max |R_T + \varepsilon_T + G_T| \leq \max |R_T| + \max |\varepsilon_T| + \max |G_T|. \quad (10)$$

$\max |\varepsilon_T|$, which is the largest propagated sensor error at a point, can be obtained via Equation (8):

$$\max |\varepsilon_T| = \max |\omega_a \varepsilon_a + \omega_b \varepsilon_b + \omega_c \varepsilon_c| \leq (\omega_a + \omega_b + \omega_c) |\varepsilon_{\max}| \leq |\varepsilon_{\max}|, \quad (11)$$

where $|\varepsilon_{\max}| = \max\{\varepsilon_a, \varepsilon_b, \varepsilon_c\}$, *i.e.*, the largest error due to LiDAR sensor system at a triangle vertex. Similarly, it can be shown that

$$\max |G_T| = \max |\omega_a G_a + \omega_b G_b + \omega_c G_c| \leq (\omega_a + \omega_b + \omega_c) |G_{\max}| \leq |G_{\max}|, \quad (12)$$

where $|G_{\max}| = \max\{G_a, G_b, G_c\}$ is the largest ground error among a , b , and c .

For interpolation error, previous research [29] has derived that the interpolation error of TIN is

$$|R_T| \leq \frac{3}{8} M_2 h^2, \quad (13)$$

where M_2 is the maximum norm of the second-order derivative of the triangle containing T . M_2 describes how fast slope changes, and is therefore equivalent to curvature. h is the longest edge of a triangle and is a description of the density of ground points used in interpolation. Combining Equations (11–13), the error at any point T is bounded by:

$$\max |\Delta z_T| \leq |\varepsilon_{\max}| + |G_{\max}| + \frac{3}{8} M_2 h^2. \quad (14)$$

From Equation (14), it can be seen that approximation theory departs from error propagation theory in fundamental ways. Approximation theory is not based on the statistics paradigm, thus it does not make any assumption about error distribution. More importantly, it does not use variances, which are known to be inappropriate for systematic errors. Approximation theory also enables the derivation of an error for any point in the entire domain, thus creating an opportunity to depict a DEM's accuracy with local details. On the other hand, the use of an error bound instead of an error may lead to overemphasis of large errors. From this perspective, an accuracy obtained through approximation theory is a conservative assessment of the DEM's quality. However, if triangles are sufficiently small, which can be achieved by using high-density points, the largest error in a triangle will become close to the error at a point in the same triangle.

5. Study Area and LiDAR Data

In this section, we illustrate approximation theory through a case study. Our approach is from DEM users' perspective, though the proposed methodology can be employed by both DEM producers and users. The task of the case study is to assess whether an existing LiDAR product is suitable for flood risk analysis. Our study site, located in UTM zone 10, is a densely populated area in the San Francisco Peninsula in California. It has about 30% urban or built-up area, 20% annual grassland, 20% chaparral, and the rest is coastal scrub. Chaparral ranges from 1 to 4 m in height, while coastal scrub rarely exceeds 2 m in height. Like most DEM users, budget constraints prevented us from acquiring and processing LiDAR data directly. Instead, we found an existing LiDAR dataset acquired by the Golden Gate LiDAR project (GGLP) [30]. The Golden Gate LiDAR project was contracted by USGS. The project was conducted in 2010 under the goal to acquire LiDAR data for the San Francisco Peninsula in California. According to the project's final report, the Earth Eye and Galileo acquisition team collected the LiDAR data using a Leica ALS60 MPiA LiDAR system mounted on a Cessna 207 aircraft between the months of April and July 2010 during the lowest tides possible. Calibration was done using data acquired from GPS and IMU attuned to sensor and flight line data. All horizontal data collected are in North American Datum 1983 and all vertical data collected are in North American Vertical Datum of 1988. The horizontal accuracy of the LiDAR dataset is stated as < 1-meter RMSE and the vertical accuracy is stated as less than or equal to 9.25 cm. No systematic error was detected.

The raw LiDAR data were processed by Earth-Eye LLC using a minimum block mean algorithm [31]. Because our LiDAR data was acquired in a dry season, LiDAR's difficulty in penetrating vegetation is expected to be reduced. The overall accuracy of this classification was reported to be 95%. Breaklines including lakes, coast, streams, rivers, islands, and above-ground features allowing water to pass underneath were digitized manually using orthoimages and USGS topographical maps through utilities in Earth-Eye Viewer. A bare-earth DEM was created at a grid resolution of 1.0 meter by incorporating all breaklines. The 1-meter bare-earth DEM is the final product delivered to USGS, and will be added to the National Map, which is a set of national geospatial datasets available to the public at no cost. The raw LiDAR point cloud, together with filtered mass points, breaklines, and grid DEM is available through the USGS Center for LiDAR Information Coordination and Knowledge (CLICK). The technical report of the project [31] was obtained and became our only source of information about this LiDAR dataset.

For the purpose of flood risk mapping, we assessed whether the DEM created by the Golden Gate LiDAR project in our study site met the FEMA Guidelines for Regions and Mapping Partners concerning the accuracy and processing of high-quality topographic data including LiDAR [12]. FEMA maintains a dataset that estimates flood risk nationwide. This dataset assigns a risk rank to each area. The risk rank is generated by grouping areas into 10 risk classes with an equal number of members. Based on risk rank and terrain type, FEMA created specification levels that provide the minimum elevation standards for empirical analysis. In an area with rolling to hilly terrain, as is the case with our study site (Figure 4), areas with a high level of flood risk (deciles 1–5) may have a specification level of “high” or “medium.” If the specification level is “high,” a LiDAR-derived DEM must have an accuracy equivalent to 1.2192 m (4 feet) contours [12]. This means the 95th percentile or 95% of the errors must be less than 72.6 cm [21]. If the specification level is lowered to “medium,” the accuracy requirement is 2.4384 m

(8 feet) contour equivalency, where the corresponding 95th percentile becomes 1.45 m. A look-up table provided by FEMA to translate contour equivalency to 95th percentile can be found in [12].

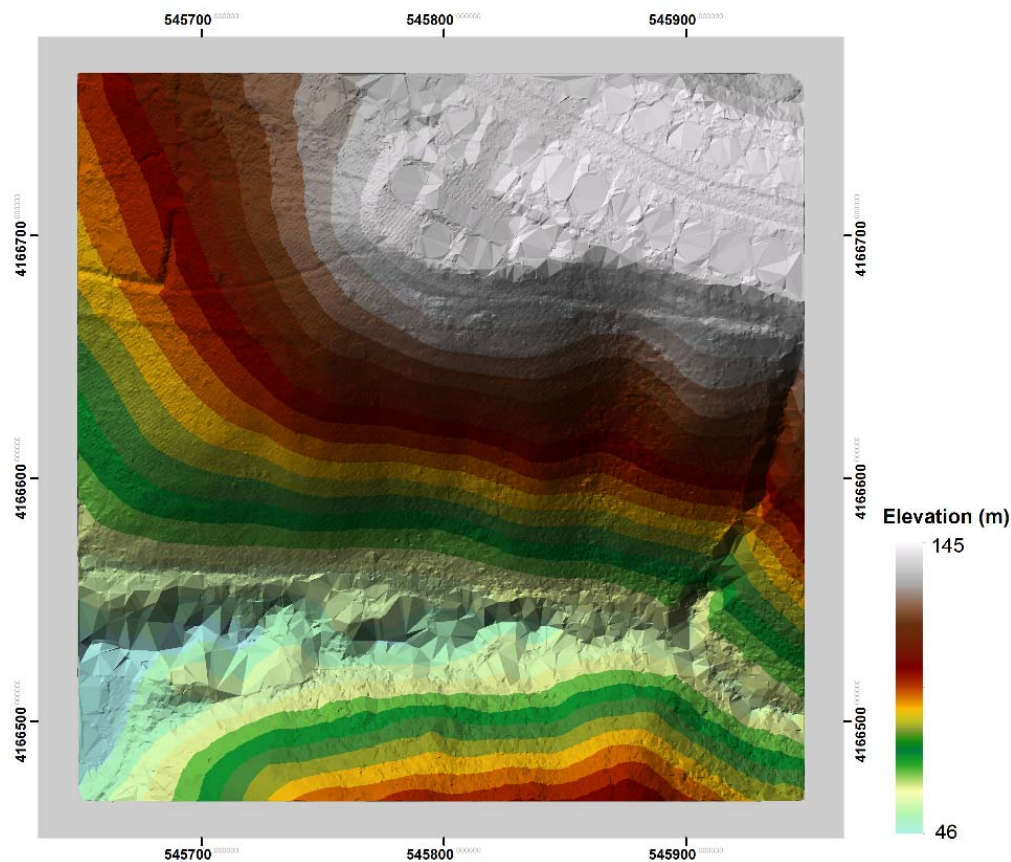


Figure 4. An elevation map of the study area, which is located in UTM zone 10.

6. Error Assessment

6.1. Propagated Sensor Error

Per Equation (10), the total error at a DEM point is bounded by the sum of three error components: propagated sensor error, propagated ground error, and interpolation error. This section explains how to infer the value of each. For propagated sensor error, Equation (11) shows that it is bounded by the largest LiDAR-sensor error in the source data. In our case study, the LiDAR acquisition team reported a LiDAR-sensor accuracy of 9.25 cm in RMSE with no systematic error detected. This accuracy was obtained through rigorous LiDAR calibration where all inherent errors such as that in the sensor's ranging and torsion parameters, GPS, and IMU were considered. To ensure compliance with the accuracy standard, a set of 47 ground control points were collected during the acquisition of LiDAR. Each control point was collected at a place of constant slope and at moderate intensity surfaces to avoid any intensity-based anomalies. Absolute accuracy was achieved by matching flight line to flight line data and subsequent adjustment to the 47 ground survey control points [31]. Because all errors, both horizontal and vertical in nature, have been addressed, and no systematic error was detected, it is reasonable to assume that only random error exists. Random error is typically normally distributed with a mean near 0. The RMSE of 9.25 cm can therefore be interpreted as standard deviation. Based on this

information, we infer that 95% of the points in the study area have a LiDAR sensor error less than or equal to 1.96(RMSE), *i.e.*, 18.1 cm. These LiDAR sensor errors will be propagated to a point during DEM generation. However, Equation (11) shows that LiDAR sensor errors are not amplified during their propagation. Therefore, it can be inferred that the propagated sensor error ($|\delta_T|$) in 95% of the points in the study area is no more than 18.1 cm.

While our inference of propagated LiDAR-sensor error in this case study is relatively straightforward, it has to be clarified that the above inference is made under the circumstance that blunders and systematic errors have been removed. To assist DEM users to make a judgement whether such a circumstance exists, when DEM producers report their empirical assessment of LiDAR sensor error, they should include information on sampling method, sample size, and error distribution in particular. The GGLP project report does not include plots of error distribution. However, a separate evaluation of GGLP data in [32] provided a histogram of LiDAR elevation errors in open or sparsely-vegetated terrain, and that histogram conformed to normal distribution. In the case that systematic errors still exist after blunders and systematic errors have been removed as much as possible, *i.e.*, LiDAR-sensor errors are in the form in Equation (15), each systematic error should be identified and quantified separately, and then combined using approximation theory, not RMSE as in [7] because variance is not appropriate to describe systematic errors. If a systematic error varies spatially, the study area will need to be stratified so that each stratum can be assessed separately. This suggests that the current method on reporting LiDAR sensor errors needs to be changed:

$$\varepsilon = \text{random error} + \text{systematic error}_1 + \text{systematic error}_2 + \dots \quad (15)$$

6.2. Propagated Ground Error

The propagated ground error is potentially a main impactor on DEM accuracy. Like LiDAR sensor error, the ground errors in points classified as bare earth by ground filtering are also not amplified during their propagation to a DEM point (Equation (12)); therefore, propagated ground error is bounded by the largest ground error in the area. As defined in Section 2.2, ground error is the height of a laser return used in ground filtering. It has a value equal to 0 for correctly-classified points and larger than 0 in the case of misclassifications. Interestingly, the current practice on reporting filtering accuracy is to use *filtering error*, not *ground error*. The typical approach is to select some reference points, and then classify every point into ground and non-ground based on image segmentation, field survey, and manual editing. The accuracy of ground filtering is obtained by cross-tabulating the reference data and classified data, similar to the accuracy assessment in remote sensing. Filtering error is reported as a percentage value such as overall accuracy, Kappa statistics, and type I or type II error. This approach is so widely applied that even the International Society for Photogrammetry and Remote Sensing (ISPRS) uses it to compare the accuracies of various filtering algorithms [10,11,33].

While filtering error helps to identify the percentage of points classified correctly, it does not carry any magnitude information needed to assess propagated ground error. In the case that a vegetation point is misclassified as ground, meaning that its ground error must be bigger than 0, filtering error does not provide the value of this overestimation error. It would be very beneficial to DEM users if DEM producers could provide detailed ground error data. This request does not add a burden to DEM

producers, because the bare-earth elevations of the reference points can be collected concurrently during accuracy assessment. Ground error at a point is a simple calculation using Equation (5) by subtracting the field-surveyed bare-earth elevation from the LiDAR return used for filtering. For example, if the LiDAR return used for filtering vegetation is 5 meters, and the true ground elevation (reference) is 4 m, then the ground error at this vegetation point is 1 m. In the case that filtering accuracy depends on topography and land cover, the study area can be stratified first. The ground error of each stratum can then be reported in terms of 95th percentile.

For our case study, the LIDAR acquisition and processing team reported the ground filtering accuracy in filtering error, as done in most studies. The filtering algorithm considered slope, angle, relationships, and distances, and classified 95% of the points automatically. Further processing was done on more than 10% of the points through manual inspection. Points were classified into five categories including processed but unclassified, bare-earth ground, vegetation and all above-ground objects, noise, and water. During manual editing, editors used a combination of imagery, intensity of the LiDAR reflection, profiles, and TIN-editing software to assess points. The overall classification accuracy was reported to be 95%. In the absence of more details on ground errors, a DEM user of this LiDAR dataset can only assume that 95% of the points have a ground error of 0. The other 5% were misclassified and thus their ground errors are greater than 0. Since propagated ground error does not exceed ground error, we infer that the 95th percentile of ground error is 0. It has to be noted, though, that this inference is a best-case scenario in that noticeable filtering errors usually exist in vegetated areas.

6.3. Interpolation Error

The interpolation error at a point is bounded by $\left| \frac{3}{8} M_2 h^2 \right|$, which is the largest interpolation error of any point within the same triangle patch. Its calculation requires the value of the longest edge of a triangle (h) and the maximum norm of second-order derivative $|M_2|$. h can be calculated based on the coordinates of triangle vertices. In this case study, 298,761 triangles were built based on 224,686 points classified as bare earth. After removing the triangles near the boundary, which can be erroneous because points outside of the study area were not incorporated in TIN construction, 290,792 triangles were included in the accuracy assessment. It was found that some triangle edges were as long as 27.6 m, while some others were only 0.01 m. On average, the longest edge of a triangle is about 1.0 m, with a standard deviation of 0.79 m.

The maximum norm of second-order derivative (M_2) can be obtained by using the contour lines generated directly from TIN [17]. In Figure 5, let the contour interval be cl . A drop of water flows downhill by crossing a number of contour lines Z_a , Z_b , and Z_c . Let T be a point on the flow path between Z_a and Z_b , M_2 is the slope change rate at Z_b , *i.e.*,

$$M_2 \cong \frac{\frac{Z_c - Z_b}{d_{bc}} - \frac{Z_b - Z_a}{d_{ab}}}{d_{ac}} = cl \frac{d_{ab} - d_{bc}}{d_{ab} \cdot d_{bc} \cdot d_{ac}}. \quad (16)$$

The value of d_{01} and d_{12} are the horizontal distances between adjacent contour lines. Their ratio d_{01}/d_{12} describes how fast the slope changes. In this research, we generated contours of 1.2192 m (4 feet) interval because the goal is to assess whether the DEM accuracy is equivalent to 1.2192 m (4 feet) contour lines.

It was found that $|M_2|$ values ranged between 0 and 1.2. The majority of the area has a rather small $|M_2|$ value. In fact, 99.7% of the triangles have an M_2 of less than 0.5/m. The mean of M_2 is 0.118/m and the standard deviation is 0.08/m. Combining this information with h , the interpolation error bound of $3/8M_2h^2$ is obtained.

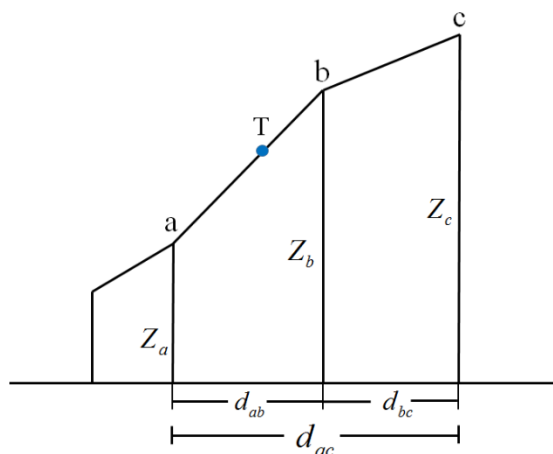


Figure 5. Estimation of M_2 at T using the contour line approach. a , b , and c are points on the contour lines of Z_a , Z_b , and Z_c respectively. T is a point in between Z_a and Z_b .

7. Results and Discussion

FEMA standards for floodplain mapping in rolling to hilly terrain require that the DEM accuracy should be equivalent to 1.2192 m (4 feet) contours if the specification level is “high.” As discussed previously, this means that 95% of the area must have an error of no more than 72.6 cm. In our case study, the interpolation error bound varied from 0 to 5.846 m. Ninety-five percent of the propagated LiDAR-sensor error is bounded by 18.1 cm and 95% of the propagated ground errors are 0. This suggests that, under the best case scenario, 95% of the interpolation errors must be bounded by 54.5 cm in order to meet the 72.6 cm threshold. In reality, 98% of the triangles covering the area are bounded by such an interpolation error, meaning that all points in these triangles have an interpolation error of no more than 54.5 cm. However, these points accounted for only 72% of the study area. In fact, 68% of the area has an interpolation error of less than 30 cm and 58% of the area has an interpolation error of less than 18 cm. This result suggests that the LiDAR-derived DEM cannot reach 1.2192 m (4 feet) contour equivalency in our study area, thus it is not suitable if the FEMA specification level is “high.” However, if the specification level is lowered to “medium,” the accuracy requirement would be lowered to 2.438 m (8 feet) contour equivalency, which means that 95% of the area must have a vertical error bounded by 1.45 m. In order to meet this accuracy requirement, 95% of the area must have an interpolation error of less than 1.27 m. In reality, 99.4% of the triangles have an interpolation error bounded by this value and these triangles accounted for 94.6% of the total area. This suggests that the LiDAR-derived DEM has a vertical accuracy nearly equivalent to 8-foot contours and thus can be used where the FEMA specification level is “medium.” Figure 6 shows how interpolation error varied in the study area. If propagated sensor error and propagated ground error also vary spatially, they can be combined with interpolation error bounds to create an error bound for each point in the study area. It has

to be stressed, though, that the above conclusion was obtained under the best-case scenario that 95% of the area has a ground error of 0. If the 95th percentile of ground error turns out to be higher than 0, the interpolation error threshold has to be adjusted accordingly.

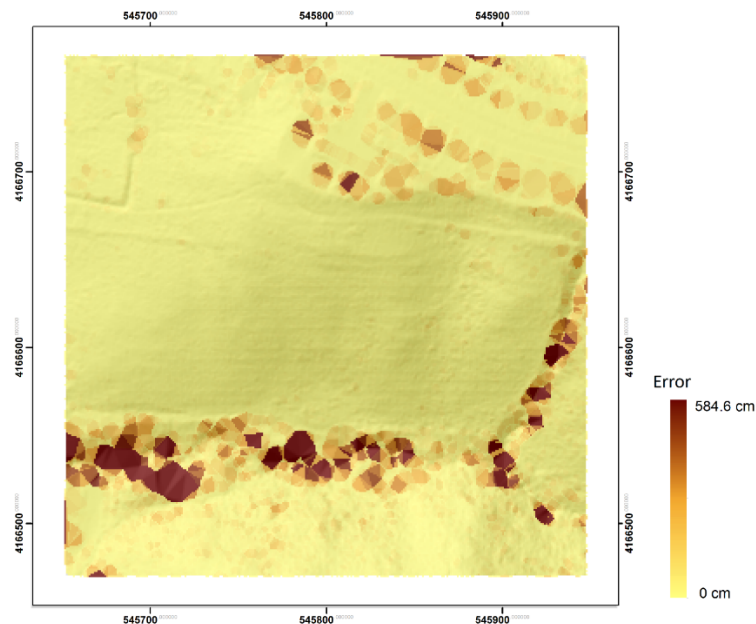


Figure 6. The spatial variation of interpolation errors.

A close examination of the interpolation errors shows that large errors appear mainly in areas of few ground points or areas of high curvature. This is expected because, as suggested by Equation (7), interpolation error depends on point density (h) as well as slope change rate (M_2). M_2 is not modifiable because its value is determined by topography. However, h can be modified by adding additional bare-earth points. In the case that no additional ground points can be obtained, one may consider adding some non-bare-earth points. By doing so, ground error at the non-bare-earth points will be propagated during interpolation. However, as shown in Equation (12), ground error is not amplified. If the added ground error can effectively offset interpolation error, then the overall vertical error at a point is still smaller. This returns us to the discussions in Section 6.2, which called for using ground error instead of filtering error to evaluate the effectiveness of a filtering algorithm. With both ground error and interpolation error available, it is possible to assess their potential to offset each other.

The implication of M_2 to DEM error assessment is interesting. In the literature, slope has often been picked as a primary derivative to correlate with DEM error [15,20,22,34]. However, Equation (7) reveals that it is not slope but curvature (the slope of slope) that correlates with DEM error. If a DEM's accuracy is to be assessed by sampling a set of test points, it is important that the test points are collected from locations of different curvatures so as to truly represent the entire terrain. In fact, large errors often occur in areas with large curvature and sparse ground points. Failure to include such area in accuracy assessment is likely to overestimate DEM accuracy. FEMA guidelines for assessing LiDAR-derived DEMs specify that test points should be selected in terrain that is flat or uniformly sloped within 5 m in all directions [12]. By this standard, the M_2 of all checkpoints can be expected to be nearly 0. Consequently, interpolation errors at these checkpoints would be nearly 0. Yet the underestimation of interpolation error is only one aspect. As Sithole and Vosselman found in their testing of ground filtering

algorithms [33], noticeable errors are usually found in regions with urban structures, abrupt slope changes, or a complex vegetation covering. By only collecting checkpoints from relatively flat or uniformly sloped areas, the ground error is also likely to be underestimated. The RMSE or 95th percentile calculated from these checkpoints is therefore likely to be small, resulting in an overestimation of the DEM accuracy.

8. Conclusions

The advancement of LiDAR technology has opened up many exciting opportunities. However, budget constraints and the steep learning curve associated with LiDAR acquisition and processing force most projects in need of a LiDAR-derived DEM to use an existing product instead of creating their own. Because most DEM producers only report a global accuracy, which may or may not be applicable to a DEM user's study site and specific need, there exists a vital need for developing a practical and cost-effective method to assess a LiDAR-derived DEM. In the literature, statistical methods based on the error propagation theory are routinely applied. These methods may not be effective because of the existence of systematic error. In this paper, we presented approximation theory as a new viable alternative. By separating the total error at a DEM point into three error components and assessing each separately, a bound can be obtained for the total error as well as each error component. Better yet, approximation theory points out where the desired accuracy is or is not met. For points that have excessive error, approximation theory further points out that these errors can be effectively reduced by adding additional bare-earth points before gridding. On the other hand, it has to be noted that, because of the use of error bound instead of error, conclusions drawn from approximation theory tend to be conservative thus may lead to underestimation of DEM accuracy. However, one can expect the difference between error and error bound to shrink once triangles become smaller, which can be achieved by using a high density of points during interpolation.

The application of approximation theory calls for a change on the existing approach to assessing and reporting the accuracy of a LiDAR-derived DEM. Contrary to the current method, which lumps all errors together and summarizes them by sample statistics, we believe that errors introduced in different stages should be separated and assessed individually. To combine these error components, approximation theory is more appropriate than error propagation theory. Currently, LiDAR sensor error is often tested and documented. However, ground error and interpolation error are reported in much less detail, as in the case of GGLP. In this research, we call for a change on how to report these errors. For ground error, it should not be reported as filtering error, which is only a percentage value. Instead, the 95th percentile or 95% confidence level of the heights of misclassified LiDAR returns used for filtering should be reported. This ground error information is very beneficial because it provides an opportunity to offset interpolation error. In the current DEM generation, triangulation is done using bare-earth points only. Non-ground points are avoided so that propagated ground error can be minimized. Our research finds that such an approach may increase the chance of higher interpolation error, especially in areas where slope changes fast. If detailed data on ground error are available and one uses the method provided this research to assess interpolation error, one can weigh the impact of having a larger ground error *versus* reducing the interpolation error. Wherever it turns out to be beneficial, it is possible to include some non-bare-earth points yet still end up with a DEM of better vertical accuracy.

Acknowledgments

This paper was funded by the San Francisco State University and the Chinese National Natural Science Foundation. We thank the three anonymous reviewers whose thorough and thoughtful reviews made the manuscript much stronger.

Author Contributions

XiaoHang Liu, Hai Hu, and Peng Hu designed the research and contributed all stages of the analysis. XiaoHang Liu wrote the manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Bitenc, M.; Lindenbergh, R.; Khoshelham, K.; van Waarden, A.P. Evaluation of a LiDAR land-based mobile mapping system for monitoring sandy coasts. *Remote Sens.* **2011**, *3*, 1472–1491.
2. Stoker, J.; Harding, D.; Parrish, J. The need for a national Lidar dataset. *Photogramm. Eng. Remote Sens.* **2008**, *74*, 1066–1068.
3. Webster, T.L. Flood risk mapping using LiDAR for Annapolis Royal, Nova Scotia, Canada. *Remote Sens.* **2010**, *2*, 2060–2082.
4. Wehr, A.; Lohr, U. Airborne laser scanning—An introduction and overview. *ISPRS J. Photogramm. Remote Sens.* **1999**, *54*, 68–82.
5. Wang, R. 3D building modeling using images and LiDAR: A review. *Int. J. Image Data Fusion* **2013**, *4*, 273–292.
6. *Digital Elevation Model Technologies and Applications: The DEM Users Manual*, 2nd ed.; Maune, D., Ed.; ASPRS: Bethesda, MD, USA, 2007.
7. Hodgson, M.E.; Bresnahan, P. Accuracy of airborne LIDAR-derived elevation: Empirical assessment and error budget. *Photogramm. Eng. Remote Sens.* **2004**, *70*, 331–339.
8. Aguilar, F.J.; Mills, J.P.; Delgado, J.; Aguilar, M.A.; Negreiros, J.G.; Pérez, J.L. Modelling vertical error in LiDAR-derived digital elevation models. *ISPRS J. Photogramm. Remote Sens.* **2010**, *65*, 103–110.
9. Liu, X. Airborne LiDAR for DEM generation: Some critical issues. *Progr. Phys. Geogr.* **2008**, *32*, 31–49.
10. Meng, X.; Currit, N.; Zhao, K. Ground filtering algorithms for airborne LiDAR data: A review of critical issues. *Remote Sens.* **2010**, *2*, 833–860.
11. Yan, M.; Blaschke, T.; Liu, Y.; Wu, L. An object-based analysis filtering algorithm for airborne laser scanning. *Int. J. Remote Sens.* **2012**, *33*, 7099–7116.
12. Federal Emergency Management Agency (FEMA). *Standards for Lidar and Other High Quality Digital Topography*; Federal Emergency Management Agency: Boise, U.S.A., 2010.

13. Hu, P.; Liu, X.; Hu, H. Isomorphism in digital elevation models and its implication to interpolation functions. *Photogramm. Eng. Remote Sens.* **2009**, *75*, 713–721.
14. Liu, X.; Hu, H.; Hu, P. The “M” in digital elevation models. *Cartogr. Geogr. Inf. Sci.* **2015**, *42*, 235–243.
15. Höhle, J.; Höhle, M. Accuracy assessment of digital elevation models by means of robust statistical methods. *ISPRS J. Photogramm. Remote Sens.* **2009**, *64*, 398–406.
16. Greenwalt, C.R.; Schultz, M.E. *Principles of Error Theory and Cartographic Applications*; Aeronautical Chart and Information Center: St. Louis, MO, USA, 1962.
17. Liu, X.; Hu, P.; Hu, H.; Sherba, J. Approximation theory applied to DEM vertical accuracy assessment. *Trans. GIS* **2012**, *16*, 397–410.
18. Federal Geographic Data Committee. *National Standard for Spatial Data Accuracy*; Federal Geographic Data Committee: Reston, USA, 1998.
19. Godone, D. The role of morphometric parameters in Digital Terrain Models interpolation accuracy: A case study. *Eur. J. Remote Sens.* **2013**, *46*, 198–214.
20. Aguilar, F.J.; Agüera, F.; Aguilar, M.A.; Carvajal, F. Effects of terrain morphology, sampling density, and interpolation methods on grid DEM accuracy. *Photogramm. Eng. Remote Sens.* **2005**, *71*, 805–816.
21. Shortridge, A.; Messina, J. Spatial structure and landscape associations of SRTM error. *Remote Sens. Environ.* **2011**, *115*, 1576–1587.
22. Su, J.; Bork, E. Influence of vegetation, slope, and LiDAR sampling angle on DEM accuracy. *Photogramm. Eng. Remote Sens.* **2006**, *72*, 1265–1274.
23. Bater, C.W.; Coops, N.C. Evaluating error associated with LiDAR-derived DEM interpolation. *Comput. Geosci.* **2009**, *35*, 289–300.
24. Sharma, A.; Tiwari, K.N.; Bhadoria, P.B.S. Determining the optimum cell size of digital elevation model for hydrologic application. *J. Earth Syst. Sci.* **2011**, *120*, 573–582.
25. Kraus, K.; Karel, W.; Briese, C.; Mandlbürger, G. Local accuracy measures for digital terrain models. *Photogramm. Rec.* **2006**, *21*, 342–354.
26. Gong-Saholiariliva, N.; Gunnell, Y.; Petit, C.; Mering, C. Techniques for quantifying the accuracy of gridded elevation models and for mapping uncertainty in digital terrain analysis. *Progr. Phys. Geogr.* **2011**, *35*, 739–764.
27. Lloyd, C.D.; Atkinson, P.M. Non-stationary approaches for mapping terrain and assessing prediction uncertainty. *Trans. GIS* **2002**, *6*, 17–30.
28. Atkinson, K.; Han, W. *Elementary Numerical Analysis*; John Wiley & Sons: Hoboken, NJ, USA, 2004.
29. Hu, P.; Liu, X.; Hu, H. Accuracy assessment of digital elevation models based on approximation theory. *Photogramm. Eng. Remote Sens.* **2009**, *75*, 49–56.
30. AARA Golden Gate Lidar Project. Available online: http://online.sfsu.edu/ehines/arra_golden_gate_lidar_project.htm (accessed on 11 April 2015).
31. Hines, E. *Final Report: Golden Gate LiDAR Project*; San Francisco State University: San Francisco, CA, USA, 2011; p. 12.

32. O'Connor, M. *Easkoot Creek Hydrology & Hydraulics Study Task 2—Evaluation of Golden Gate LiDAR Data for Use in Hydrologic and Hydraulic Modeling of Easkoot Creek*; O'Connor Environmental, Inc.: Healdsburg, CA, USA, 2012; p. 8.
33. Sithole, G.; Vosselman, G. Experimental comparison of filter algorithms for bare-earth extraction from airborne laser scanning point clouds. *ISPRS J. Photogramm. Remote Sens.* **2004**, *59*, 85–101.
34. Sharma, A.; Tiwari, K.N.; Bhadoria, P.B.S. Vertical accuracy of digital elevation model from Shuttle Radar Topographic Mission—A case study. *Geocarto Int.* **2010**, *25*, 257–267.

© 2015 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).