The effects of DEM interpolation on quantifying soil surface roughness using terrestrial LiDAR

Li Li,*, Mark A. Nearing, Mary H. Nichols, Viktor O. Polyakov, D. Phillip Guertin, Michelle L. Cavanaugh

School of Natural Resources and the Environment, University of Arizona, Tucson, AZ, 85705, USA
USDA-Agricultural Research Service, Southwest Watershed Research Center, Tucson, AZ, 85719, USA

ARTICLE INFO

Keywords:
LiDAR points
Interpolation
DEM error
Soil surface roughness

ABSTRACT

Soil surface roughness (SSR) is often calculated based on the Digital Elevation Models (DEMs) obtained by interpolating points from terrestrial LiDAR measurements. This study aimed to investigate the effects of DEM interpolation and interpolation methods on quantification of SSR. A series of rainfall at the intensity of 59 and 178 mm hr⁻¹ were simulated on a 2 by 6.1 m stony soil plot under 12% and 20% slope treatments. LiDAR measurements were conducted after each rainfall simulation at six positions around the soil plot. DEMs gridded at 5 and 10 mm were generated from LiDAR points using Inverse Distance Weighting, Natural Neighbors and Universal Kriging methods implemented in ESRI ArcGIS 10.5. Random roughness index (RR) was calculated based on the LiDAR-interpolated DEMs and LiDAR points directly and then compared. Results showed: 1) average values of RR from the DEMs were under-predicted compared to those directly calculated from LiDAR points by 7%–20%, indicating smoothing of the modeled surface during the interpolation process; 2) RR from LiDAR points indicated that soil surface became rougher as rainfall was progressed; 3) DEM errors increased as the surfaces evolved to rougher states and when 10 mm resolution was used; 4) the temporal SSR variations using the DEMs were -0.27 and -0.19 mm at 20% slope and 0.30 and 0.36 mm at 12% slope for the resolution 10 and 5 mm, respectively, while they were 0.22 mm at 20% slope and 0.65 mm at 12% slope when LiDAR points were directly used, the differences are due to the resulting increased DEM errors as the surface became rougher that masked the true changes in soil surface roughness. This study shows that at plot scale, DEMs generated through interpolations from LiDAR points underestimate soil surface roughness and are ineffective at tracking changes in soil surface roughness over time, and that LiDAR point data must be used instead at present scale.

1. Introduction

Soil surface roughness (SSR) describes the small-scale variations of surface elevation, which plays an important role in many hydrologic and erosion processes. A multitude of research studies have shown that SSR affects surface hydrologic behaviors, such as infiltration and runoff (Helming et al., 1993, 1998; Darboux et al., 2001), as well as hydraulic functions, such as retarding or localizing water velocity (Gómez and Nearing, 2005), diverting flow directions (Abrahams and Parsons, 1991; Weltz et al., 1992; Helming et al., 1993; Luo et al., 2017), and affecting the runoff sediment size distribution (Ding and Huang, 2017). The soil surface roughness status has also been shown to affect the evolution of drainage (or rill) networks (Gómez et al., 2003; Pelletier, 2003). The temporal variation of SSR, moreover, is considered as a descriptor that reflects how the soil surface behaves in the presence of erosive agents (Huang and Bradford, 1992). Unsurprisingly, SSR is incorporated as a crucial component in many process-based hydrologic erosion models. Thus, accurately quantifying the SSR is important for understanding erosion processes as well as for erosion modeling.

A wide range of techniques have been reported to sample surface elevation for assessing the SSR. The profile meter is one of the most commonly used apparatuses (Govers et al., 2000; Thomsen et al., 2015). A row of evenly spaced pins mounted on the meter are lowered onto the soil surface, the positions of pins are recorded, and then they are digitized either electronically or photographically for quantifying SSR. However, a profile meter can only capture limited transects of the surface, and the surface-contacted pins may disturb the soil surface for further measurements (Xingming et al., 2017) or penetrate the surface causing inaccurate measurements. Non-contact laser-scanning devices have also been deployed to scan the area of interest at high resolutions.

* Corresponding author.
E-mail address: lili2@email.arizona.edu (L. Li).

https://doi.org/10.1016/j.still.2019.104520
Received 28 March 2019; Received in revised form 8 October 2019; Accepted 23 November 2019
Available online 13 January 2020
0167-1987/ © 2019 Published by Elsevier B.V.
these DTMs. Shi et al. (2017) investigated the soil surface random rounding cells. The local roughness index was then estimated based on the minimum elevation contained within the defined search radius, and concentrated flow. The TLS points were first gridded to generate 1 by rangeland plots to study the effects of TLS-derived roughness on the DEMs. The main reason for generating DEMs from LiDAR points is their simple, efficient and organized formats for storing elevation information (Yilmaz and Uysal, 2016). For example, Eitel et al. (2011) applied TLS to quantify the SSR on 8.5 m² rangeland plots to study the effects of TLS-derived roughness on the concentrated flow. The TLS points were first gridded to generate 1 by 1 cm DTMs (digital terrain models) wherein the value of each cell was the minimum elevation contained within the defined search radius, and the values of void cells were linearly interpolated based on the surrounding cells. The local roughness index was then estimated based on these DTMs. Shi et al. (2017) investigated the soil surface random roughness on 1.4 by 0.65 m² plots before and after rainfall simulation by using the 1 by 1 cm DEMs derived from TLS point clouds, and each row of gridded points was then treated as an individual transect for assessing the random roughness index. Milenkovic et al. (2016) compared two mapping techniques using unmanned aerial vehicles and TLS for roughness assessments, and roughness analysis was performed on the DSMs (digital surface models) generated from unmanned aerial vehicles images and TLS point clouds. However, information about how to select the interpolation methods and the DEM resolutions is usually missing, leading to the implication that SSR quantification is independent of the selections of interpolation methods and resolutions.

Spatial interpolation is the procedure of estimating values of variables at locations where those values are not measured using measured data from nearby locations. A great number of interpolation methods have been reported that can be divided into three groups according to Li and Heap (2014, 2011): 1) non-geostatistical methods that predict the unknown points based on the values of their neighborhood points. The non-geostatistical methods can be further distinguished by the methods used to calculate the weighting factor of each involving neighborhood point, such as Inverse Distance Weighting (IDW) or Natural Neighbors (NN); 2) geostatistical methods that utilize both the spatial structure and value of sample points, such as the Kriging method; 3) combined methods that generate the estimations of unknown points based on the non-geostatistical and geostatistical approaches or other statistical methods.

Due to the errors introduced by the interpolating process and the fact that natural terrain surfaces are continuous and comprised of infinite points while interpolation-generated terrain is represented in a discrete fashion, the quality of generated DEMs should be assessed before further applications. The accuracy of DEMs generated from different interpolation methods has been widely discussed and compared. Although a universal optimal interpolation method for generating DEMs has not been identified (Chaplot et al., 2006; Bater and Coops, 2009; Erdoğan, 2010), some consensus regarding the factors controlling the accuracies of interpolated DEMs has been reached, namely: 1) different interpolations applied over the same data may result in different DEMs, and the accuracy of DEMs are a function of the interpolation method (Arun, 2013); 2) the errors of interpolated DEMs depend on the surface characteristics and increase with increased terrain surface roughness regardless of the selections of interpolation method and resolutions (Tang et al., 2001; Heritage et al., 2009; Erdoğan, 2010); 3) the errors of interpolated DEMs vary for different resolutions, and the coarser the DEMs, the greater are the uncertainties in the values of the elevation points in the DEMs (Tang et al., 2001; Bater and Coops, 2009; Chen and Yue, 2010).

The above three considerations raise concerns about the reliability of estimating the SSR and SSR temporal variations using the LiDAR-interpolated DEMs. Since the interpolated DEMs and associated accuracies depend on factors associated with the interpolation methods, terrain surface characteristics, and DEM resolutions, it can be anticipated that the SSR estimations based on LiDAR-interpolated DEMs will be affected by these three factors as well. However, how these three factors ultimately affect the quantifications of SSR has not been studied. If such effects exist, the estimations of temporal variations in the SSR would likewise suffer uncertainties when the LiDAR-interpolated DEMs are utilized. For the dynamically evolving surface with temporally changing SSR, the temporal variations of DEM uncertainties (e.g. they increase with increase of surfaces roughness) due to the changes of surfaces could in turn mask the actual surface changes, resulting in misinterpretations of SSR variations. Those concerns are not trivial. First, unlike at the large scale, the magnitudes of SSR and SSR variations on small plots usually are at the order of centimeters or sub-centimeters (Darboux et al., 2001; Nouwakpo et al., 2016) and potentially affected by the qualities of DEMs if whose uncertainties are of the same order; Secondly, since SSR plays an important role in hydrologic-erosion models and the temporal variations of the SSR are usually related with rainfall amounts or tillage practices, the uncertainty of the estimations of SSR and temporal variations in SSR could impair model performances.

The main objective of this study was to investigate the effects of interpolation methods on the estimations of SSR and SSR temporal variations on small plots. We hypothesized that the interpolation-modeled surfaces will be smoother because of the interpolation process, leading to SSR estimations from LiDAR-interpolated DEMs that are lower than those directly calculated from LiDAR point cloud data. Also, as the actual surface changes (in this case the surface became rougher) due to erosion, the resultant change in DEM uncertainties may mask the true SSR changes, resulting in misinterpretation of temporal trends in SSR. We tested the hypotheses by using a rainfall simulator to apply rainfall events on 12.2 m² soil plots with two slope treatments (20 % and 12 %). LiDAR scans were conducted from six positions around the plots. For each surface, six scans were first registered to obtain a unified LiDAR point cloud and then generated into DEMs with resolution of 10 mm and 5 mm by using IDW, NN and Universal Kriging (UK) interpolators that were implemented in ESRI ArcGIS 10.5. Random roughness index was calculated based on the LiDAR-interpolated DEMs and LiDAR point clouds, and then compared. This study highlights the importance and effects of DEM uncertainties on SSR quantifications when LiDAR-interpolated DEMs are used.

2. Materials and methods

2.1. Rainfall simulations

Detailed descriptions of the rainfall simulation procedures used here are available in Nearing et al. (2017). Rainfall simulations were conducted on a 2 by 6.1 m long metal-bound box. The box was slope-adjustable, ranging from 0 to 20 %. Two slope gradients 12 % and 20 % were used. For each slope gradient, the box was first positioned horizontally and filled with soil.

The experimental soil was collected from the top layer (0–15 cm) of a Luckyhills-McNeal gravel sandy loam with 52 % sand, 26 % silt, 22 % clay and less than 1 % organic matter in the Lucky Hills area of the Walnut Gulch Experimental Watershed (31°44′ 34″N; 110°03′51″ W).
is the predicted value for unknown points, and is 

\[ Z_i^p = \sum_{j=1}^{n} w_i Z_j \]  


1 Trade names and company names, included for the benefit of the reader, do not imply endorsement or preferential treatment of the product listed by the USDA.

A programmable Walnut Gulch Rainfall Simulator (Paige et al., 2003) was used to apply rainfall. This rainfall simulator consisted of a single oscillating boom with four VectJet 80100 nozzles. Nozzles were mounted on the boom at a spacing of 1.52 m apart. By adjusting the oscillation frequency of nozzles, rainfall intensity from the simulator may range from 13 to 190 mm h\(^{-1}\) with a kinetic energy of 204 kJ ha\(^{-1}\) mm\(^{-1}\) obtained over the 2 by 6.1 m area. Detailed information about Walnut Gulch Rainfall Simulator may be found in Paige et al. (2003). Three windscreens were set up around the rainfall simulator to minimize unwanted wind disturbance.

Before applying rainfall, the soil surface was covered with a porous cloth, and a 30-minute pre-wetting rainfall with intensity of 35 mm h\(^{-1}\) was applied to create a relatively consistent initial moisture condition for each experiment. The box was then adjusted to the designated slope (12 % or 20 %) and the protective cloth cover was removed.

Rainfall was simulated three times on 20 % slope and four times on the 12 % slope. That difference was due to the fact that the surface of 20 % slope changed more quickly over time, due to a faster erosion rate, than at 12 % slope. Each simulation was a continuous rainfall application with two intensities. The duration of each simulation varied from 1.5 to 5 h. For each slope gradient, the first simulation started from low intensity (59 mm h\(^{-1}\)) until steady runoff rate was reached, followed by approximate one-hour of rainfall at high intensity (178 mm h\(^{-1}\)), ending with approximately 15 min of rainfall at low intensity (59 mm h\(^{-1}\)). The subsequent simulations started from high intensity (178 mm h\(^{-1}\)) and then decreased to low intensity (59 mm h\(^{-1}\)) and lasted for up to 5 h.

2.2. Surface elevation measurement

The soil surface was scanned with a RIEGL VZ400 terrestrial LiDAR scanner (www.riegl.com) after the first and last rainfall applications. The scanner was configured to emit a collimated beam with a nominal divergence of 0.3 mrad for sampling the surface at high speed and high resolution (0.3 mrad corresponds to an increase beam diameter by 3 mm per 10 m distance) (Rieg GmbH, 2012). The range-dependent nature of the TLS measurements results in the fact that soil surfaces farther from the scanner were sampled at lower density than those closer to the scanner. Six scan positions surrounding the soil box were selected to maximize the capture of the soil surface morphology. The scan positions are schematically shown in Fig. 1. Position #1, #5 and #6 were set up with the mounting tripod on the ground while #2, #3, #4 were from platforms 4 m above the ground. Scan positions #2, #3 and #4 were elevated to obtain more favorable viewing angles. Each position was located approximately 2 m from the soil box boundaries to minimize the shadow produced by the instrument. The soil surface on the box was 1.5 m above the ground at its center pivot point, hence the relative heights of the scanner in each position to the center of the soil box were approximately 0.5 m for position #1, #5 and #6, and 1.5 m for positions #2, #3 and #4. During the course of measurements, the scanner sampled the surface at angular increments of 0.04° in the azimuth (0-360°) and zenith (30-130°) direction, resulting in one point every 1 mm at a distance of 2 m and every 6 mm at a distance of 8 m.

2.3. LiDAR data processing

The surfaces after the first and the last simulation for each slope gradient were used for analyses. Thus, there were two surfaces with two sets of scans on each for a total of four scan sets. Here, we designate the initial and final LiDAR measurements as A1 and A2 for the 20 % slope, B1 and B2 for 12 % slope.

For each surface, scans from the six positions were transformed into a common external coordinate system from the different internal spherical coordinate systems to obtain a composite point cloud. Scan to scan registrations were accomplished in two steps. The first step involved matching the corresponding control point targets among the scans. In this study eight control points were established around the soil box and their coordinates were surveyed with a Trimble R8 robotic total station. Eight cylindrical reflectors with diameters of 10 cm were set up over the control points on survey tripods (Fig. 1). In that step, at least three reflector targets were required to match the scans. After the coarse match in the first step, the iterative closest point (ICP) (Besl and McKay, 1992) algorithm implemented in the Riegli RiScan Pro® software was used to further register the six scans to obtain a composite LiDAR point cloud.

To avoid the boundary effects of the soil box borders, each composite point cloud was clipped to within approximately 25 cm from the soil box edges, with resulting LiDAR points covering an area of approximate 5.5 by 1.5 m. Prior to the generation of DEMs, the resultant LiDAR points were partitioned evenly into three sections of the plot: upper, middle and lower. In each section, five percent of the LiDAR points within each of the three sections were randomly selected as check points for subsequent error analyses. This was done to ensure each section of the plot had check points when random selections were conducted. The remaining 95 % of the LiDAR points were used for generating DEMs.

2.4. DEM generations

Three interpolation methods were selected to generate DEMs: Inverse Distance Weighting (IDW), Natural Neighbors (NN), and Universal Kriging (UK). These three methods have been widely employed to create raster DEMs from LiDAR data and implementation tools are available in Spatial Analyst Tools of ESRI ArcGIS 10.5. In this study, all DEMs were generated with 10 mm and 5 mm resolution.

2.4.1. Inverse distance weighting (IDW)

IDW is one of simplest interpolation methods (El-Shemy et al., 2005; Mitas and Mitasova, 1999; Li and Heap, 2008, 2011; Huang et al., 2011). The basic assumption is that the value of an unknown point can be predicted as a weighted average of values at several nearby known points, and the weights are inversely proportional to a power or exponential function of distances between the unknown point to the given points. The guiding principle is that the closer a given point is to the unknown point, the greater influence the given point imposes. IDW method can be represented by the equation:

\[ Z_p = \sum_{i=1}^{n} w_i Z_i \]  

where \( Z_p \) is the predicted value for unknown points, \( Z_i \) is the value of a given point of measured elevation; \( w_i \) is the weight assigned to \( Z_i \); and is equal to an inverse function of the distance (d\(^{-\alpha}\), where \( d \) is distance and \( \alpha \) is the user-designated exponent of the power function) between the known and unknown points. The numbers of sample points and their corresponding weights are shown to affect IDW interpolation results. In this study, 12 neighboring sample points were used, and the power of distances was set as 2.

2.4.2. Natural neighbors (NN)

NN also uses Eq. (1) as the estimation formula with IDW while the calculation algorithm for the weight for each sample point differs. Once the DEM resolution is defined, sampled points that have fallen into a
cell are first used to construct a Voronoi diagram (Thiessen polygon) V’. Then another Voronoi diagram V around the unknown point is also created. From those, the corresponding weight of each sample point is determined by the proportion of overlap between the V and V’ (Li and Heap, 2008). In ESRI ArcGIS 10.5, the only input parameter required for this method was the designed DEM resolutions, which were 10 and 5 mm in this study.

2.4.3. Universal Kriging (UK)

Kriging is a generic name for a family of generalized least squares regression algorithms (Li and Heap, 2008). Kriging assigns weights according to a data-derived weighting function, such as the semivariogram function (Anderson et al., 2005). For a large number of points, such as are common with a LiDAR point cloud, the semivariance γ(h) can be estimated as:

\[ γ(h) = \frac{1}{2n} \sum_{i=1}^{n} (Z(x_i) - Z(x_i + h))^2 \]  

(2)

Where \( x \) is a point in the cloud, \( Z(x) \) is its elevation, and \( n \) is the number of pairs of sample points separated by a distance \( h \). All kriging estimators are variants of the basic equation:

\[ \hat{Z}(x_0) - \mu = \sum_{i=1}^{n} \lambda_i [Z(x_i) - \mu(x_i)] \]  

(3)

Where \( \mu \) is a known stationary mean of the whole domain while \( \mu(x) \) is the mean elevation of points in the search windows, \( \lambda_i \) is the kriging weight linked to \( γ(h) \) (Li and Heap, 2008) and \( n \) is the number of sampled points included in the estimation.

In the Universal Kriging algorithm, three simplifications are made based on the Eq. (3): 1) replace the \( \mu \) with \( \mu(x_0) \); 2) forcing \( \sum \lambda_i = 1 \); 3) \( \mu(x_0) \) is the trend component of each point in the neighborhood search window. In this study, UK with linear drift was selected and 12 neighboring sample points were used in ESRI ArcGIS 10.5.

2.5. DEM accuracy estimation

2.5.1. Root mean square errors

DEM accuracy defines how reliably the interpolated DEMs represent the true terrain surface, and is assessed based on the values of DEM errors (Tang et al., 2001). Hu et al. (2009) analyzed DEM error components and concluded that two primary sources that contributed to DEM uncertainties were: a) the error from the data source due to inaccuracy of the surveying techniques and b) error from the interpolation process in which true elevation is approximated from the surrounding points. Root Mean Square Error (hereafter denoted by RMSE_EI) is most frequently used for quantifying the overall DEM accuracy including both sources (Tang et al., 2001; Heritage et al., 2009; Hu et al., 2009; Yilmaz and Uysal, 2016) and its formula was given as:

\[ RMSE_{EI} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (E_{DEM} - E_{CHECK})^2} \]  

(4)

Where \( E_{DEM} \) is the elevation extracted from DEM, \( E_{CHECK} \) is the measured elevation of the correspondent check point, and \( n \) is the number of check points.

2.5.2. Terrain representation error

Because a natural terrain surface is a continuous surface comprised of infinite points while interpolation-generated terrain is represented in a discrete fashion, the interpolated-DEMs cannot completely represent the true terrain surface. Limited by the resolution of DEMs, no matter how accurate surveying and interpolating procedures are, the interpolated-DEMs only approximate to the true terrain surface, therefore a terrain representation discrepancy, or accuracy loss is unavoidable (Tang et al., 2001; Chen and Yue, 2010). However, \( RMSE_{EI} \) only reflects the vertical differences between the interpolated values with the true values (in our case it is the LiDAR measured value) and is insufficient to describe how completely the DEM represents the true surface. An alternative error term was put forward and estimated by Tang et al. (2001), designated as terrain representation error (ETR), which may be used to describe the discrepancy between the DEM and the true terrain surface. The basic algorithm was: for the DEM with resolution of \( R \), moving windows with size \( k \times k \) (\( k = 2, 4, 6 \ldots \), e.g., 2R, was used to cover the DEM, and then five points with one center (i, j) and four corners ((i-1, j-1), (i-1, j+1), (i+1, j-1), (i+1, j+1)) of each window were found. The maximum ETR at the position (i, j) was then defined as the elevation of center \( h(i, j) \) minus the average elevation of those four corners:

\[ ETR(i, j) = Z(i, j) - \frac{1}{5} \sum_{c} Z(c(i, j)) \]  

Where \( Z(i, j) \) is the elevation of the correspondent check point, \( ETR(i, j) \) is the maximum terrain representation error, \( c(i, j) \) is the maximum ETR at the position (i, j) and \( n \) is the number of check points.
\[ ETR_{ij} = h_{ij} - \frac{h_{i-1,j-1} + h_{i-1,j+1} + h_{i+1,j-1} + h_{i+1,j+1}}{4} \]  

(5)

Using this algorithm, ETR can be calculated for moving windows with sizes of \(2R\), to \(4R\), to \(6R\), etc. More details about the ETR concept and algorithm are available in Chen and Yue (2010) and Tang et al. (2001). As described by Tang et al. (2001), once ETR is calculated for each center point \((i, j)\), the ETR matrix can be obtained for the entire area of interest. In our study, the moving window size was set as two times the resolution (for 10 and 5 mm resolution, the window size was 20 and 10 mm, respectively), and the root mean square of ETRs (hereafter denoted by \(RMSE_{ET}\)) was calculated based on the following equation

\[ RMSE_{ET} = \sqrt{\sum_{i,j}^{M,N} ETR_{ij}^2} / M \times N \]  

(6)

where, \(M\) and \(N\) represent the rows and columns of the ETR matrix, respectively.

2.5.3. Total DEM error

The total DEM errors consist of the errors that are directly introduced by the interpolations (the vertical differences between the interpolated values with the true values) and the accuracy loss during interpolation process. Based on the error propagation theorem, total DEM error is given by Chen and Yue (2010) as:

\[ RMSE_{total} = \sqrt{RMSE_{EI}^2 + RMSE_{ET}^2} \]  

(7)

2.6. Roughness index estimations

A great number of SSR parameters have been reported and can be divided into three types following descriptions of Martinez-Agirre et al. (2016); Smith (2014) and Govers et al. (2000): single parameters, hybrid parameters and multi-scale parameters. The most widely used parameter is the Random Roughness index (RR) which is a length scale that is equal to the standard deviation of elevations from a mean surface. Moreover, RR was shown to be the best estimator in distinguishing soils with different surface roughness levels (Vermang et al., 2013). Therefore, in this study RR values were calculated using the following formula and then applied to test the interpolation effects on roughness quantifications.

\[ RR = \frac{1}{N} \sum_{j=1}^{N} (Z_j - \bar{Z})^2 \]  

(8)

where \(Z_j\) is the elevation measurement at point \(j\), \(N\) is the number of the measurements, and \(\bar{Z}\) is the average value of those \(N\) measurements.

For each surface, grid networks with 10 and 5 mm resolution were created. Then the elevation value of each grid point was extracted from the DEMs generated by IDW, NN and UK with the aid of Sample tool in ESRI ArcGIS 10.5. Each row of the grid networks was considered as single transect that was across the plot, and subsequently a linear de-trending method was utilized to remove the effect of slope. RR index was then calculated for each row.

When the LiDAR points were directly utilized, each row of the above created grid network was treated as a center line for a band with 2 mm width. Since the average point spacing was 3 mm, the width with a value of 2 mm ensured that only the points (at least 100 points) near to the center line were involved in the calculations. The number of the bands was the same as the number of rows of the grid networks. LiDAR points that fell into each band were firstly linearly de-trended, and then the RR index was calculated for each band. In this study, in total 553 transects were defined for each set of elevation points on the plots at 10 mm resolution and 1106 transects at 5 mm resolution. The averaged value of RR of those 553 (or 1106) transects was used for further analysis.

2.7. Data analyses

For each surface, the relative RR, which is the ratio of the average RR calculated from the interpolated DEM to that from the LiDAR points directly, was used to illustrate the accuracy of interpolated DEM in terms of quantification of surface roughness. The roughness changes were calculated as the RR values of A2 (B2) compared to those of A1 (B1), and positive variations indicated that the surface has become rougher. A paired t-test was performed using Excel to compare the differences among interpolation methods. In all statistical tests \(P = 0.05\) was used, unless indicated otherwise.

The relationships between DEM errors and surface roughness were developed using the RMSE\(_{EI}\) of each interpolator and the average RR calculated from the LiDAR points directly, to illustrate the effects of surface roughness on the DEM errors; RMSE\(_{total}\) was also related to the relative RR to determine the effects of DEM errors on SSR quantifications.

3. Results

3.1. DEM errors

The RMSE\(_{EI}\), RMSE\(_{ET}\) and total DEM error for each interpolation method, surface and resolution were computed and shown in Table 1. For all eight sets of DEMs (four surfaces at two resolutions), RMSE\(_{EI}\) values of IDW and NN were similar (\(p > 0.05\)), while RMSE\(_{EI}\) values of UK was lower than those of IDW and NN (\(p < 0.05\)); both RMSE\(_{ET}\) and total DEM error showed a consistent trend in the values with NN being the greatest, followed by IDW and finally UK (\(p < 0.05\)). Values of RMSE\(_{EI}\), RMSE\(_{ET}\) and total DEM error were less at finer resolution (5 mm) DEMs than at 10 mm resolution (\(p < 0.05\)). The RMSE\(_{EI}\), RMSE\(_{ET}\) and total DEM error were greater on the second surfaces (A2 and B2) than on the first (A1 and B1) (\(p < 0.05\)) for every combination of slope, resolution, and interpolation method. Also, the RMSE\(_{EI}\) and RMSE\(_{ET}\) were both consistently greater at 20 % (A1 and A2) than 12 % (B1 and B2), that is, for A1 compared to B1 and for A2 compared to B2.

3.2. Roughness quantifications

The RR values for each transect (or band) are shown in Fig. 2 using surface A1 as an example. The RR values of a majority of the transects estimated from the LiDAR points directly were greater than those calculated from interpolated DEMs, regardless of interpolation method.

Average Random Roughness (RR) calculated based on the LiDAR points directly and the interpolated DEMs using IDW, NN and UK interpolator are shown in Table 1. For each surface with the same resolution, the average RR values computed from interpolated DEMs were different from each other, and always lower (\(p < 0.05\)) than those calculated from the LiDAR points directly, irrespective of interpolation method. The average RR calculated using the 5 mm resolution DEMs was always greater than that calculated using the 10 mm resolution DEMs (\(p < 0.05\)).

The relative RR indicated average values of RR from the DEMs were under-predicted compared to those directly calculated from LiDAR points by 7%–20%. As the surface evolved from A1 to A2, and from B1 to B2, the underestimations of SSR using interpolated DEMs were more pronounced, indicated by the decreased relative RR values. The relative RR of DEMs with 5 mm resolution were always greater (\(p < 0.05\)) than those of DEMs with 10 mm resolution.

3.3. Roughness changes

The estimated changes in average RR for each slope treatment are shown in Table 2. At 20 % slope, results using the interpolated-DEM showed a decrease in surface roughness irrespective of the interpolation method and DEM resolution. However, when LiDAR points were used...
The decreased EDEM ECHECK terrain surface characteristics on DEM quality which was previously aforementioned effects of interpolation method, DEM resolution and ported in this study. Our results from small plots reaffirm the solutions, and terrain surface characteristics, all of which were sup-
terpolated-DEM would be affected by interpolation method, DEM re-

4.1. The errors of interpolated DEMs

<table>
<thead>
<tr>
<th>surface</th>
<th>interpolator</th>
<th>RMSE_EI (mm)</th>
<th>RMSE_ET (mm)</th>
<th>total DEM error (mm)</th>
<th>RR (mm)</th>
<th>Relative RR</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>IDW</td>
<td>2.20</td>
<td>0.94</td>
<td>2.39</td>
<td>4.77</td>
<td>92.7 %</td>
</tr>
<tr>
<td></td>
<td>NN</td>
<td>2.20</td>
<td>1.00</td>
<td>2.42</td>
<td>4.80</td>
<td>93.2 %</td>
</tr>
<tr>
<td></td>
<td>UK</td>
<td>2.10</td>
<td>0.87</td>
<td>2.27</td>
<td>4.75</td>
<td>92.3 %</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>2.17</td>
<td>0.94</td>
<td>2.36</td>
<td>4.77</td>
<td>92.7 %</td>
</tr>
<tr>
<td></td>
<td>LiDAR</td>
<td>2.07</td>
<td>0.62</td>
<td>2.19</td>
<td>4.87</td>
<td>94.7 %</td>
</tr>
<tr>
<td>A2</td>
<td>IDW</td>
<td>2.90</td>
<td>1.20</td>
<td>3.14</td>
<td>4.50</td>
<td>83.8 %</td>
</tr>
<tr>
<td></td>
<td>NN</td>
<td>2.90</td>
<td>1.30</td>
<td>3.18</td>
<td>4.54</td>
<td>84.6 %</td>
</tr>
<tr>
<td></td>
<td>UK</td>
<td>2.80</td>
<td>1.10</td>
<td>3.01</td>
<td>4.46</td>
<td>83.1 %</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>2.87</td>
<td>1.20</td>
<td>3.11</td>
<td>4.50</td>
<td>84.1 %</td>
</tr>
<tr>
<td></td>
<td>LiDAR</td>
<td>2.80</td>
<td>0.79</td>
<td>2.91</td>
<td>4.67</td>
<td>86.9 %</td>
</tr>
<tr>
<td>B1</td>
<td>IDW</td>
<td>1.60</td>
<td>0.68</td>
<td>1.74</td>
<td>3.11</td>
<td>87.6 %</td>
</tr>
<tr>
<td></td>
<td>NN</td>
<td>1.60</td>
<td>0.73</td>
<td>1.76</td>
<td>3.13</td>
<td>88.2 %</td>
</tr>
<tr>
<td></td>
<td>UK</td>
<td>1.60</td>
<td>0.60</td>
<td>1.71</td>
<td>3.08</td>
<td>86.8 %</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>1.60</td>
<td>0.67</td>
<td>1.73</td>
<td>3.11</td>
<td>87.6 %</td>
</tr>
<tr>
<td></td>
<td>LiDAR</td>
<td>1.57</td>
<td>0.47</td>
<td>1.64</td>
<td>3.55</td>
<td>89.8 %</td>
</tr>
<tr>
<td>B2</td>
<td>IDW</td>
<td>2.30</td>
<td>0.96</td>
<td>2.49</td>
<td>3.41</td>
<td>81.3 %</td>
</tr>
<tr>
<td></td>
<td>NN</td>
<td>2.30</td>
<td>1.05</td>
<td>2.53</td>
<td>3.45</td>
<td>82.3 %</td>
</tr>
<tr>
<td></td>
<td>UK</td>
<td>2.20</td>
<td>0.88</td>
<td>2.37</td>
<td>3.36</td>
<td>80.2 %</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>2.27</td>
<td>0.97</td>
<td>2.46</td>
<td>3.41</td>
<td>80.2 %</td>
</tr>
<tr>
<td></td>
<td>LiDAR</td>
<td>2.17</td>
<td>0.63</td>
<td>2.26</td>
<td>4.19</td>
<td>83.4 %</td>
</tr>
</tbody>
</table>

Table 1
The Root Mean Square Error of the elevation differences, RMSE_EI, Root Mean Square Error of the terrain representation error, RMSE_ET and total DEM error of each interpolation method, and the calculated average Random Roughness from interpolated DEMs and LiDAR point clouds for each surface with 10 mm and 5 mm resolutions (Note: A1 and A2 represented the soil surface after the first and the third rainfall simulation under 20 % slope; B1 and B2 represented the soil surface after the first and the fourth rainfall simulation under 12 % slope).

4. Discussion

4.1. The errors of interpolated DEMs

Our hypotheses were that the quantification of SSR based on interpolated-DEM would be affected by interpolation method, DEM resolution, and terrain surface characteristics, all of which were supported in this study. Our results from small plots reaffirm the aforementioned effects of interpolation method, DEM resolution and terrain surface characteristics on DEM quality which was previously reported solely on a large scale.

When DEM resolution was fixed on a surface, the RMSE_EI values were similar between the NN and IDW methods. These results are not unusual. For example, Bater and Coops (2009) used the TLS data collected on Vancouver Island, Canada, to investigate the errors associated with interpolated DEMs, and found NN and IDW were similar with respect to the RMS and mean absolute errors. This was explainable in that when the sample density (e.g. LiDAR points) was great enough, the resultant small distance between known values would dominate over the impacts of differences of interpolation techniques (Chaplot et al., 2006). However, in our study, both values of errors generated from UK (RMSE_EI, RMSE_ET and total DEM error) were all less than those of IDW and NN. This is likely explainable, because UK takes spatial correlations into consideration while IDW and NN do not. The different RMSE_ET values also implied the potential of using RMSE_ET to provide complementary information for selecting the appropriate interpolators and resolutions (Chen and Yue, 2010; Tang et al., 2001), especially in the case where RMSE_EI are similar (e.g. the values of IDW and NN in present study).

When finer resolution (5 mm) DEMs were used, all interpolation methods resulted in better quality DEMs with lower RMSE_EI and RMSE_ET values (Table 1). The fact that the finer DEMs produce the lower interpolation error (RMSE_EI) has been reported previously (Bater and Coops, 2009). This occurs probably because the distance from the check points to the cell centers of the coarser grids was greater on average than for the finer grids, leading to the greater difference between ECHECK and EDEM. The decreased RMSE_ET as resolution increased from 10 to 5 mm is because more gridded points were used to represent the terrain surface, hence the information loss (the discrepancy between DEM and surface it presented) decreased (Tang et al., 2001). It is also worthy to note that the RMSE_ET values were similar between the NN and IDW method while different when various resolutions were selected, implying that the spatial resolution in some cases may be more important than the choice of interpolation method when LiDAR points are interpolated for DEMs (Bater and Coops, 2009).

The positive relationship between average RR and total DEM error (Fig. 3) highlighted that DEM quality is impaired on complex and rough terrain, namely at the same resolution, a rough surface is less accurately
approximated by DEM than a simple or flat surface. The results are consistent with previous research (Tang et al., 2001; Heritage et al., 2009; Erdoğan, 2010). In the case when terrain is represented by a DEM, the natural continuous terrain of each cell is approximated by a

### Table 2

Changes in average roughness over time for the 20% and 12% slope from A1 to A2, and from B1 to B2, respectively.

<table>
<thead>
<tr>
<th>Roughness changes (mm)</th>
<th>10 mm resolution</th>
<th>5 mm resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>20% surface (A1 to A2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IDW</td>
<td>−0.27**a</td>
<td>−0.20a</td>
</tr>
<tr>
<td>NN</td>
<td>−0.25a</td>
<td>−0.17a</td>
</tr>
<tr>
<td>UK</td>
<td>−0.29a</td>
<td>−0.22a</td>
</tr>
<tr>
<td>LiDAR points</td>
<td>0.22b</td>
<td>0.23b</td>
</tr>
<tr>
<td>12% surface (B1 to B2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IDW</td>
<td>0.30a</td>
<td>0.35a</td>
</tr>
<tr>
<td>NN</td>
<td>0.32a</td>
<td>0.38a</td>
</tr>
<tr>
<td>UK</td>
<td>0.28a</td>
<td>0.34a</td>
</tr>
<tr>
<td>LiDAR points</td>
<td>0.65b</td>
<td>0.64b</td>
</tr>
</tbody>
</table>

**Value with the same letters are similar, for each set of slope percentage and resolution.

Fig. 2. Random roughness of each transect of interpolated DEMs at a resolution of 10 mm and the corresponding LiDAR point band on 20% slope (“LiDAR points” refer to the random roughness calculated from the LiDAR points directly, “IDW”, “NN” and “UK” represent the random roughness calculated from the interpolated DEMs using Inverse Distance Weighting, Natural Neighbors and Universal Kriging respectively).

Fig. 3. Total DEM error at resolution of 10 and 5 mm as a function of average random roughness which estimated from LiDAR points directly for the four surfaces (A1 and A2 for 20% slope; B1 and B2 for 12% slope).

Fig. 4. Relative Random Roughness, which is the ratio of average RR calculated from the interpolated DEM to that from LiDAR points directly, as functions of total DEM error at 20% and 12% slope.
flat plane with a single interpolated elevation value. The discrepancy between the flat plane and natural surface it approximated is expected to be greater when the terrain surface is rougher and more complex. Our results showed that, in order to achieve the same accuracy level, rougher surface requires a finer DEM resolution.

4.2. The effects of interpolation on the quantification of SSR

The overall result was that SSR is underpredicted by the interpolated-DEM, and the extent of underestimation (which is indicated by the relative RR) was dependent on, at least, the DEM resolution, and terrain surface characteristics. More specifically, when finer resolution was applied, and when the terrain surface was smoother, it would be better approximated by the DEM, and the values of RR were closer to those from LiDAR points directly (Table 1, Fig. 3). Such findings came from a basic fact that the elevation variations of natural surface covered by each cell can be captured by LiDAR points, while this variation information is lost when a smooth flat plane is assigned for that cell during DEM generations. The smoothing effects were more pronounced when the sampled natural surface was rougher and DEM cell size was larger. This would seem to have apparent importance when using LiDAR technology to characterize surface microtopography. It is not only because random roughness is frequently used to estimate the friction factor which is then related with the velocity prediction in the erosion models (Welz et al., 1992), but also because random roughness is found to control many transfer processes on and across the surface boundary (Huang and Bradford, 1992), even related with the tillage practices (Allmaras et al., 1996).

Given the fact that relative RR decreased with increased total DEM errors (Fig. 4), it is interesting that, for each individual surface, the UK, which had the lowest values of total DEM error, displayed the lowest values of relative RR (Table 1). That is probably due to the selection of linear drift for considering the slope effects. The linear drift actually is a process used to determine the trend component in the search radius (see Method section). Bryant et al. (2007) have reported the effects of detrending method on the roughness quantification. Whether the selection of linear drift imposes additional smoothing effects on DEMs requires further study. However, the unexpected results that UK when compared to NN and IDW, has a lower value of relative RR, does not alter our interpretations. For all the eight sets of DEMs generated from UK, the increasing trend of total DEM error with RR (Fig. 3), and the decreased trend of relative RR with the total DEM error (Fig. 4), are still reasonable and valid.

4.3. The effects of interpolation on the quantification of SSR changes

Another hypothesis was that estimations of temporal RR variations were unreliable if interpolated-DEMs were used due to the changes of DEM uncertainties associated with the changes of surface roughness, which was also supported here. The average RR directly estimated from LiDAR points indicated that A2 and B2 was rougher than A1 and B1, respectively (This trend of increased roughness was also supported by the three transects measurements of laser device as reported in the study of Nearing et al. (2017), and by visual observations during experiments). As the surface evolved into rougher stages, for example from A1 to A2, and B1 to B2, the increased total DEM error due to the increased surface roughness (Fig. 3) imposed greater smoothing effects, consequently increasing the underestimations of RR (Fig. 4) when using the interpolated DEMs. As shown in Table 2, in the case of surface A (20 % slope), the greater roughness which was tracked by the LiDAR points, however exhibited as a decrease using the interpolated DEMs. In the case of surface B (12 % slope), the increase of RR captured using the interpolated DEM was only approximately half of that calculated from the LiDAR points. We therefore speculate, if these RR variation estimations from the interpolated DEMs were further related with rainfall, water flow velocity and sediment transport, different conclusions would be made, leading to a misunderstanding of the erosion process.

4.4. Further considerations

With the rapid development of new terrain measurement techniques, accurate and highly dense elevation data are available and increasingly utilized in geoscience. DEM offers a simple format for storing terrain information, and they are widely used as inputs to many GIS-supported models, such as the Topography-based hydrological Model (TOPMODEL) and the Soil and Water Assessment Tool (SWAT) (Yilmaz and Uysal, 2016). Generally, the fashions of generating DEMs from LiDAR points come from either computability with other datasets or tradeoff for ease of use. However, interpolation errors are unavoidable during the process of DEM establishment. Even if the interpolation processes are free of errors, the finite grid points that make up a DEM, cannot represent the true continuous surface without any discrepancy. The inherent errors have the potential to negatively affect the DEM applications (Tang et al., 2001). For example, the estimated surface temporal elevation changes by directly comparing two interpolated DEMs include uncertainties and hence assessing the DEM errors and their propagated errors is a crucial step in estimating surface spatio-temporal changes (Goodwin et al., 2016). Estimations of DEM accuracy should, therefore, be fundamental preliminary work necessary for preparing LiDAR data for DEM modeling and associated analyses.

In our study, only the random roughness index was calculated, the effects of interpolation process on other roughness indices (e.g., fractal dimension and multiple fractals) which describe the spatial heterogeneity of surface roughness has not been tested, this is worthy further investigations. This is because the direct use of LiDAR points to conduct fractal analysis remains with some challenges. For one, fractal dimensions and multiple fractals are more easily derived from gridded format data instead of unevenly spaced LiDAR points. For two, due to the huge amounts of data, some practical compromises of reducing the original LiDAR point cloud density or selecting smaller patches from areas of interest have to be made when LiDAR points are used directly (Nouwakpo et al., 2016; Abban et al., 2017). If the interpolation processes mislead the interpretation of fractal analysis and LiDAR points must be used instead, new methods for fractal analysis should be developed.

The above findings might not be limited to the LiDAR measurements, but could be extended to other techniques, such as Structure from Motion (SFM) in which a huge number of points are generated from series of images. For example, Prodocimi et al. (2017) used the SFM to obtain a 3D point cloud, and then the point cloud was generated into DEMs with 1 cm resolution to investigate the soil water erosion in Mediterranean vineyards. Though the mechanism of obtaining point clouds differs between LiDAR and SFM, once these points are utilized for developing DEMs, the quality of interpolated-DEMs should be carefully considered and the DEM-derived parameters should be treated with caution.

Our study was limited to the effect of interpolation on the quantification of roughness, but not extended to its propagated effects on the hydrological modeling. Whether or how the roughness parameter (or other hydrographic features) extracted from interpolated-DEMs influences the performances of hydrological modeling also need to be investigated. We also note that the finer resolutions produce better DEM quality in terms of roughness quantification. Whether is necessary to take attentional efforts to generate high resolution for estimating roughness parameter in order to significantly improve the model results, also need to be addressed.
5. Conclusions

In the present study, the random roughness indices were computed based on the interpolated DEMs and directly on LiDAR points. The results showed that the quality of interpolated DEMs is dependent on the selection of interpolation method, surface roughness values and choice of resolution. Due to the smoothing effects of interpolation method and the loss of elevation information between the grid points, the interpolated DEMs underestimated the RR, and those underestimations were greater when the surface evolved into rougher stages, masking the true surface elevation changes and leading to underestimated or wrong temporal RR variation estimations. Therefore, once the point cloud is utilized for generating DEMs, the quality of interpolated DEMs and associated DEM-derived parameters should be treated with cautions in order to avoid misinterpretation of the surface processes.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of Competing Interest

The authors declare that they have no conflict of interest.

Acknowledgements

We thank John Smith for the soil box preparation and soil collection. The Southwest Watershed Research Center hosted a number of workshops.

References


