A new multi-resolution based method for estimating local surface roughness from point clouds

Lei Fan[1*] and Peter M. Atkinson[2]

[1] The Department of Civil Engineering, Xi'an Jiaotong- Liverpool University, Suzhou, China (e-mail: Lei.Fan@XJTLU.edu.cn).
[2] The Faculty of Science and Technology, Lancaster University, Lancaster, LA1 4YR, U.K. (e-mail: pma@lancaster.ac.uk)
* corresponding author

Abstract—From some empirical and theoretical research on the digital elevation model (DEM) accuracy obtained for different source data densities, it can be observed that when the same degree of data reduction is applied to a whole area, the rate of change in the DEM error is statistically greater in local areas where the surface is rougher. Based on this observation, it is possible to characterize surface roughness or complexity from the differences between two digital elevation models (DEMs) built using point clouds that represent the same terrain surface but are of different spatial resolutions (or data spacings). Following this logic, a new approach for estimating surface roughness is proposed in this article. Numerical experiments are used to test the effectiveness of the approach. The study datasets considered in this article consist of four elevation point clouds obtained from terrestrial laser scanning (TLS) and airborne light detection and ranging (LiDAR). These types of topographical data are now used widely in Earth science and related disciplines. The method proposed was found to be an effective means of quantifying local terrain surface roughness.

Key words — laser scanning; point cloud; terrain surface roughness; DEM error

1. Introduction

Terrain surface roughness is an important parameter for describing terrain surface variability or complexity in Earth science. It is often used for investigating the DEM error and its spatial variation because DEM accuracy is affected by terrain surface complexity [1]-[5]. It is also commonly used for studying Earth surface processes and landforms [6]-[9]. In the literature, there exist a wide range of methods for estimating surface roughness, for example, the root mean square height (RMSH) or standard deviation of residual elevation over a particular scale ([6], [8]-[10]), standard deviation of slope or curvature ([9]-[10]), power spectrum [6], analysis of fractal dimension [11] and geostatistical analysis [12]-[13]. However, due to the wide variety of applications for this parameter, a single definition of terrain surface roughness may not be possible as its nature and calculation often depend on the type of analysis or application [9], and also the types of data used for calculating terrain surface roughness. This study concerns the quantification of local surface roughness for applications where the local DEM error is of interest. In recent years, scattering point cloud data obtained from terrestrial laser scanning (TLS) and airborne light detection and ranging (LiDAR) have been used widely for characterising terrain surfaces. For such data, the method commonly used for estimating terrain surface roughness is RMSH or standard deviation of residual elevation [8] [14]-[15], which is also a typical roughness descriptor for the cases where surface roughness needs to be quantified to investigate DEM accuracy.

To create a grid-based DEM or predict elevations at points where no source data are available, spatial interpolation is commonly used. At a predicted point, the elevation error (i.e., the difference between the predicted elevation and the reference elevation) consists of two components: (1) propagated measurement error and (2) interpolation error (or modelling error) [16]-[18]. The former error component is data-based, being strictly concerned with the source data used for interpolation [16]. The latter component is model-based, being concerned with how well the interpolation technique extends the data to the prediction point.

Data sampling interval is a critical factor for DEM accuracy [18]-[21]. For a given interpolation method, the local interpolation error is affected mainly by the local spacing of source data and the local roughness of a topographic surface [3], [22]-[24]. If one holds the spacing of source data constant over a whole area, the local interpolation error is then affected mainly by the local roughness. In other words, when the spacing of the source data used to build a DEM is the same, the local interpolation errors will be spatially variable and closely correlated to local terrain surface roughness (statistically the local interpolation error is expected to be greater at locations where the terrain surface is rougher or more complex). These are confirmed by various theoretical and empirical studies in the literature [25]-[32]. In applications where the DEM error is of interest, an effective descriptor of local terrain surface roughness should, therefore, be consistent with the aforementioned behaviour. In this context, the effectiveness of such a descriptor is determined by the magnitude of the correlation between the estimated local roughness and the local DEM interpolation error. In the literature, the correlation between the global surface roughness and the global DEM error was investigated using datasets that represent terrain surfaces of different terrain morphologies at different sites [3], [17]. In these cases, as the surface roughness of the DEMs considered varied significantly, commonly used surface roughness descriptors are sufficient to produce a large correlation between global roughness and global DEM error [17]. However, for a terrain surface of similar morphology (i.e. at the same site), it is not well understood quantitatively if commonly used error statistics (e.g. RMSH) are able to estimate the local surface roughness that is strongly correlated to the local DEM error.

As a widely used local surface roughness descriptor for scattered point cloud data, the RMSH was evaluated against local interpolation errors in this study. A conditional leave-one-out technique (i.e., conditioned by an equal data spacing of point cloud data) was used to estimate the local interpolation errors, which were then compared with the RMSH roughness values to investigate the effectiveness of the RMSH statistic. A comparatively small correlation between the local RMSH and the local interpolation error was observed. To develop a more effective alternative, a new approach was proposed for estimating local terrain surface roughness through the differences between two DEMs constructed using two point cloud datasets that represent the same terrain surface, but are of different spatial resolutions. This new approach is referred to as the multi-resolution method. The effectiveness of these methods (the RMSH and the multi-resolution approach) is compared.

2. Methods

2.1. Study data

The study datasets considered in this research include a TLS point cloud and three airborne LiDAR point clouds. These datasets represent a mix of different cases (different types of terrain, different roughness features, different scales of spatial variation in the elevation data). The TLS point cloud was used as the main dataset for demonstrating the multi-resolution method and the associated data processing. The airborne LiDAR point clouds were used to test the transferability of the multi-resolution method.

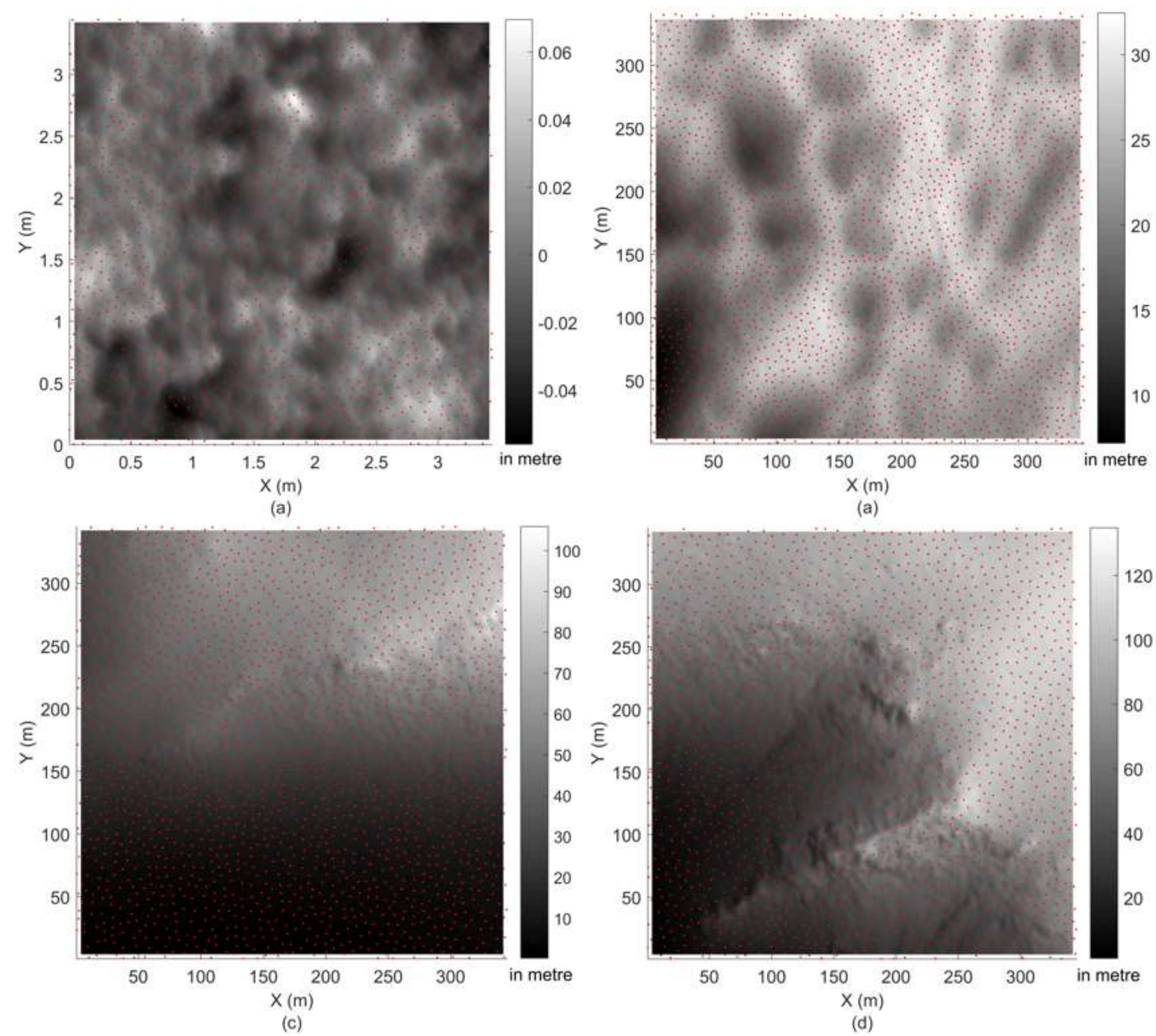

Fig. 1. The subsample point clouds and the shaded DEMs (built using the original point clouds): (a) TLS, (b) LiDAR-1, (c) LiDAR-2 and (d) LiDAR-3, in which the colour bar for TLS represents the residual elevation values after fitting a plane to remove linear trend in elevation and the colour bar for LiDAR-1, LiDAR-2 and LiDAR-3 depicts the elevation values.

The TLS point cloud represents a slightly rough clayey soil surface surveyed by a ScanStation C10. It consists of 34163 points over an area of approximately 3.4 m by 3.4 m. Fig. 1a shows a subsample (i.e., the dotted points) of the original TLS point cloud for clear visualization, and the DEM built using the original point cloud. The subsample was obtained in such a way that each individual data point has similar distances to its neighbouring data points. This is also a requirement of the multi-resolution method proposed, which is elaborated in a subsequent section. There is a global linear trend in elevation, which was removed to visualise the surface variability. In other words, the data shown in Fig. 1a represent residual elevations.

The spatial resolution (typically in metres) of data sampling by airborne LiDAR is much coarser than that (typically in several or tens of millimetres) by TLS. This leads to different scales of spatial variation in the elevation data measured. The airborne LiDAR point clouds considered represent three bare terrain surfaces of distinct characteristics: a comparatively flat and rough surface (Fig. 1b), a hilly and relatively smooth surface (Fig. 1c), and a hilly and relatively rough surface (Fig. 1d), which are referred to as LiDAR-1, LiDAR-2 and LiDAR-3, respectively. Each LiDAR dataset covers an area of approximately 340 m by 340 m, and consists of 81531 (LiDAR-1), 71467 (LiDAR-2) and 87275 (LiDAR-3) data points classified as bare ground. These LiDAR datasets were acquired by the National Centre for Airborne Laser Mapping, USA. LiDAR-1 comes from a large LiDAR dataset acquired along the north inner coastline of Cape Cod (Massachusetts, USA) with a total area of 37.106 km^2. LiDAR-2 and LiDAR-3 data were selected from a large LiDAR dataset (for a surface area of 366.64 km^2) acquired at a volcanic field in Central Nevada, USA.

2.2. Benchmark: The leave-one-out error for evenly distributed data points

In this Section, cross-validation is introduced, followed by an elaboration of the logic of the leave-one-out error (obtained under the condition of an equal data spacing for a point cloud) being used as a benchmark to evaluate the adequacy of a surface roughness measure.

In the literature concerning DEM accuracy, cross-validation is used widely for estimating how accurately a prediction model (or interpolation method) performs. In one round of cross-validation, a subset (validation dataset) of original data is held out and a model is fitted to the remaining data (training dataset). The model obtained is then used to predict the validation dataset. Much research work has been carried out to investigate the prediction accuracies of interpolation methods for building DEMs using cross-validation [3], [17], [29], [33]. In those research, to enable the analysis of the prediction errors, the validation dataset (i.e. check points) is often assumed to have negligible or zero measurement errors [3], [17], [29], [33]. This assumption is also adopted in this article but may be understood as follows. The topographic data contaminated by measurement errors are assumed to be new data that are free of measurement errors. In this case, the measurement errors associated with the true terrain surface are simply treated as a part of the natural spatial variability in the new data. In other words, the study object considered is the contaminated terrain surface rather than the true, error-free terrain surface. As the investigation reported in this article is on the adequacy of the methods considered for describing the roughness of a terrain surface, the error-contaminated surface represents an appropriate test dataset.

The leave-one-out method is a commonly used form of cross-validation for evaluating the interpolation error [29], [34]. In each iteration, a data point is removed from the original topographic dataset, and the remaining data points and a specified interpolation algorithm are used to predict the elevation value at the location where the data point is removed. The predicted value is compared to the observed value at the point removed, and their difference is referred to as the leave-one-out error. This process is repeated for all data points (note that the previously removed data point is replaced into the dataset at each iteration), leading to a set of leave-one-out errors (one for each data point location).

Given that the measurement errors in the test datasets are treated as natural spatial variation, the only error component for a predicted elevation value is the interpolation error. In this case, the leave-one-out errors represent the local DEM interpolation errors at all data point locations. The interpolation error is a function of the source data spacing and the terrain surface variability locally (as discussed in the Introduction), and so is the leave-one-out error. If one holds the data spacing fixed for a whole area, the spatial variation in the leave-one-out errors at different locations of the whole area must be due to only the variability of the terrain surface (or surface roughness). The leave-one-out error (for equally spaced data points) should statistically be greater at locations where the terrain surface is rougher, and *vice-versa*. Therefore, the leave-one-out error is a useful benchmark to test the adequacy of the surface roughness descriptors considered.

2.3. Root mean squared height

RMSH is the most commonly used measure of surface roughness, especially for scattering elevation data such as laser scanning point clouds [8], [14]-[15]. The definition of the RMSH is given by:

$$\mathrm{RMSH} = \sqrt{\frac{\sum_{i=1}^{n}(z_i - \bar{z})^2}{n-1}}$$

where n represents the number of data points, z_i is the elevation of the i^{th} data point and $\bar{z}$ is the mean elevation of all (n) data points.

The elevation data are often detrended by subtracting a best fit linear surface from the data, leaving a set of residual elevations with a mean value of zero. In this case, the RMSH is essentially the standard deviation of detrended residual elevations. To reflect the local surface roughness, the local RMSH is calculated over a particular spatial scale, which is often defined by a moving window size.

2.4. The multi-resolution method

An empirical study [35] on the DEM error using TLS point cloud data suggested that the DEM interpolation error increases more quickly in areas of greater surface roughness when the same degree of data density reduction is applied to point clouds of different surface complexities. This observation is consistent with earlier studies. For example, an experimental study [1] suggests that the interpolation error might be approximated by RD (the original form was expressed in terms of error variance: $(RD)^2$), where R is a parameter describing the terrain surface roughness and D represents the spacing of source data. This simple model was also considered in [17] and [28]. In a theoretical study on the DEM error [24], it was demonstrated that the interpolation error for a predicted elevation point using a TIN with linear interpolation can be expressed as 3/8 Mh^2, where the parameter M is a terrain morphology roughness parameter and h represents data spacing. The aforementioned studies suggest that when the same level of data reduction is applied to a whole area, *the rate of change* in the DEM interpolation error with data spacing is statistically greater in local areas where the surface is rougher. This forms the basis of the multi-resolution method introduced below for inferring surface roughness or complexity.

Fig. 2 illustrates the main steps involved in the multi-resolution method. First, the raw point cloud (the data spacing of which often varies spatially) is filtered to produce a subsample, aiming to generate a point cloud of similar data spacing between adjacent data points. Uniform sampling can be implemented readily in various point cloud processing software packages (e.g. CloudCompare, Leica Cyclone). Typically, these software packages can produce only one subsample for a given data spacing requirement. To produce different subsamples under a specified minimum spacing between neighbouring data points, the data filtering algorithm used in this research is continuous random selection constrained by a minimum distance between selected points, and it involves the following steps: (1) select randomly one data point from the point cloud, (2) create a circular buffer zone with the data point newly selected being its centre and the minimum distance being its radius, (3) remove all the data points inside the buffer zone, which results in a new point cloud, (4) select randomly the next data point from the new point cloud obtained in Step 3, (5) repeat Step 2 to Step 4 until there are no data points left. When the point cloud is adequately dense, this filtering process leads to a subsample point cloud of very similar data spacing S_1, which may be checked by visual inspection. This first point cloud obtained is referred to as the fine-resolution point cloud. Using the same data filtering method but with a larger minimum distance, the fine-resolution cloud is thinned further to generate a coarse-resolution point cloud of very similar data spacing S_2 ($S_2 > S_1$). Ideally, the data spacing should be identical. However, this is practically unlikely due to the scattered nature of the point clouds obtained using laser scanners. As such, both S_1 and S_2 represent the average data spacing of the subsample point clouds, and the calculation of their values should be based on the actual subsamples selected. In this study, the average data spacing (S) is estimated using $S = \sqrt{A}/\left(\sqrt{N} - 1\right)$, where A and N represent the whole planar (i.e. XY) area and the total number of data points, respectively.

The fine-resolution and the coarse-resolution point clouds are used to produce a DEM of the same grid resolution r and grid locations, respectively, and this leads to two DEMs (DEM_1 and DEM_2). The interpolation method used for constructing these DEMs was the TIN with linear interpolation for two main reasons. First, it does not require any input modelling parameters, which simplifies the modelling process and can produce more consistent results. Second, as a local interpolation method using only three adjacent data points, it can reflect the realised local variation of the interpolation error.

Based on the aforementioned discussions, the local differences between the two DEMs (DEM of Difference or DoD in abbreviation) is a function of the data spacing difference and local variability. As the data spacing difference is fixed spatially, this can be regarded as being controlled for, such that in a relative sense the values in the DoD should be a function of local variability only. Therefore, the DoD can be used as an index for inferring surface roughness.

Although a single DoD obtained using the aforementioned approach gives an approximate estimate of surface roughness, it was found in this research that such results show variability when different coarse-resolution subsamples were used. More specifically, if the same fine-resolution dataset was used to compare several different coarse-resolution point clouds of the same data spacing, the resulting DoD varied from one coarse-resolution point cloud to another. Such variation is likely caused by the nature of randomness in the filtering method used for producing subsamples. To reduce this variability, multiple rounds (n) of coarse-resolution subsample selection were performed, which leads to n corresponding DoD maps. In these rounds, the same data spacing S_2 was used for all coarse-resolution subsamples. The DoD maps were averaged over all rounds to produce a mean DoD map, which can be used to measure terrain surface roughness or complexity. The method was validated by comparing the mean DoD map to the leave-one-out errors.

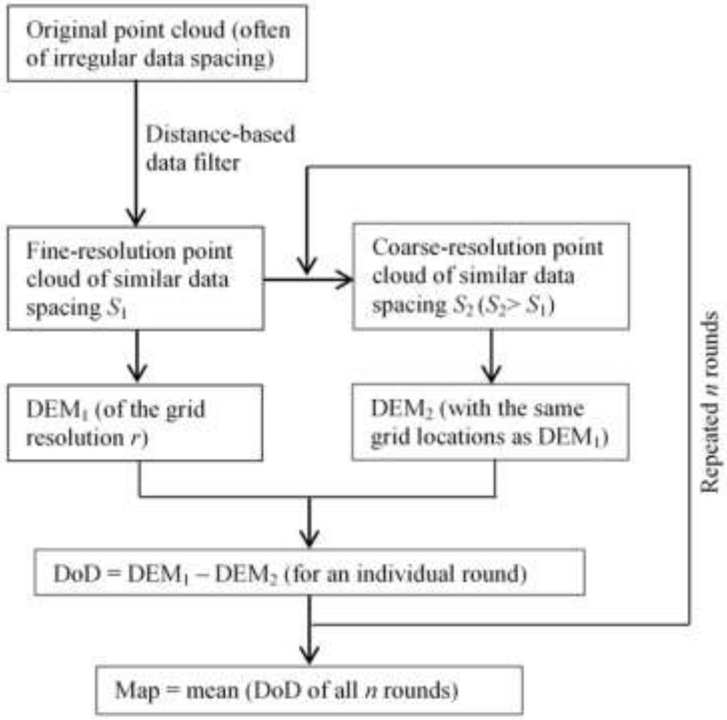

Fig. 2. The main steps of the multi-resolution method.

2.5. Data processing

All the data processing was implemented in MATLAB. To demonstrate the data processing, only the TLS point cloud was used. As discussed in Section 2.3, to obtain a leave-one-out error dataset that represents the local interpolation error, it is necessary to use a point cloud that has equally spaced data points. Due to the effects of scanning parameters (such as incidence angles and scanning distance), the data points in the original point cloud are often not equally spaced. Therefore, a subsample of the original TLS point cloud was selected using the filter (i.e. constrained by a minimum distance) introduced for the multi-resolution method. This subsample has a similar data spacing of approximately 0.037 m and was used as the test dataset for calculating (1) the leave-one-out error at each individual data point location, (2) RMSH and (3) the mean DoD. Spatial resolutions finer than that (i.e. 0.037 m) were not considered, due to the limitation of the data density of the raw point cloud used. However, TLS laser scanning surveys can be carried out with a fine or very fine sampling resolution. Therefore, if oversampling is undertaken for data collection in a survey, a very fine-resolution point cloud can be obtained even after the data filtering for similar data spacings is carried out.

To calculate the RMSH, a moving window size needs to be determined. The literature shows that the RMSH often varies with the window size used [9]. A typical window size suggested in the literature is five times the data spacing [9], [15]. Based on this, a square window with a size of 0.19 m was considered. To remove any elevation trend, a local fitting plane was fitted to the data point within each moving window to derive the residual elevations that were used for calculating the local RMSH. A direct comparison between the RMSH and the leave-one-out error is not feasible because the former was calculated over a given scale (i.e. the window size) while the latter was calculated at individual data point locations. To enable the comparison, the arithmetical mean absolute values of the leave-one-out errors over the same window size were calculated and then compared with the RMSH.

In the multi-resolution method, the aforementioned test dataset (i.e. the point cloud with the data spacing of approximately 0.037 m) was used as the fine resolution point cloud. By applying the minimum distance filter introduced in Section 2.4, coarse-resolution point clouds with data spacing of approximately 0.07 m were obtained. The DEM grid resolution used was r = 0.02 m. To determine the number of rounds (n) for the calculation of the mean DoD, two independent realisations of the multi-resolution method were produced. Take n = 50 as an example. In one realisation, 50 different DoD maps were produced for calculating a mean DoD map. In the second realisation, another 50 DoD maps were obtained for the mean DoD map. The mean DoD map from the first realisation was compared to that from the second realisation to check if their results (i.e. the mean DoD values) were consistent.

To enable a more direct comparison between the leave-one-out errors (at individual data point locations) and those obtained using the multi-resolution method, the TIN with linear interpolation was used to provide the leave-one-out results for interpolated values at the same grid locations as those used in the multi-resolution method.

3. Results

3.1 The TLS data

For the point cloud with the data spacing of 0.037 m, the RMSH and the mean absolute values of the leave-one-out errors are shown in Fig. 3(a) and Fig. 3(b), respectively. In Fig. 3(c), the grid values in Fig. 3(a) are plotted against those at the same grid locations in Fig. 3(b), enabling a more direct comparison between the results. There is a positive correlation between the RMSH and the mean absolute leave-one-out errors, confirming expectations about greater interpolation errors in rougher areas. However, the correlation coefficient was relatively small. These results were based on a moving window of 0.13 m (i.e. five times the data spacing). Some other window sizes were also tested and small coefficient of determination values (similar values or smaller than 0.49) were observed.

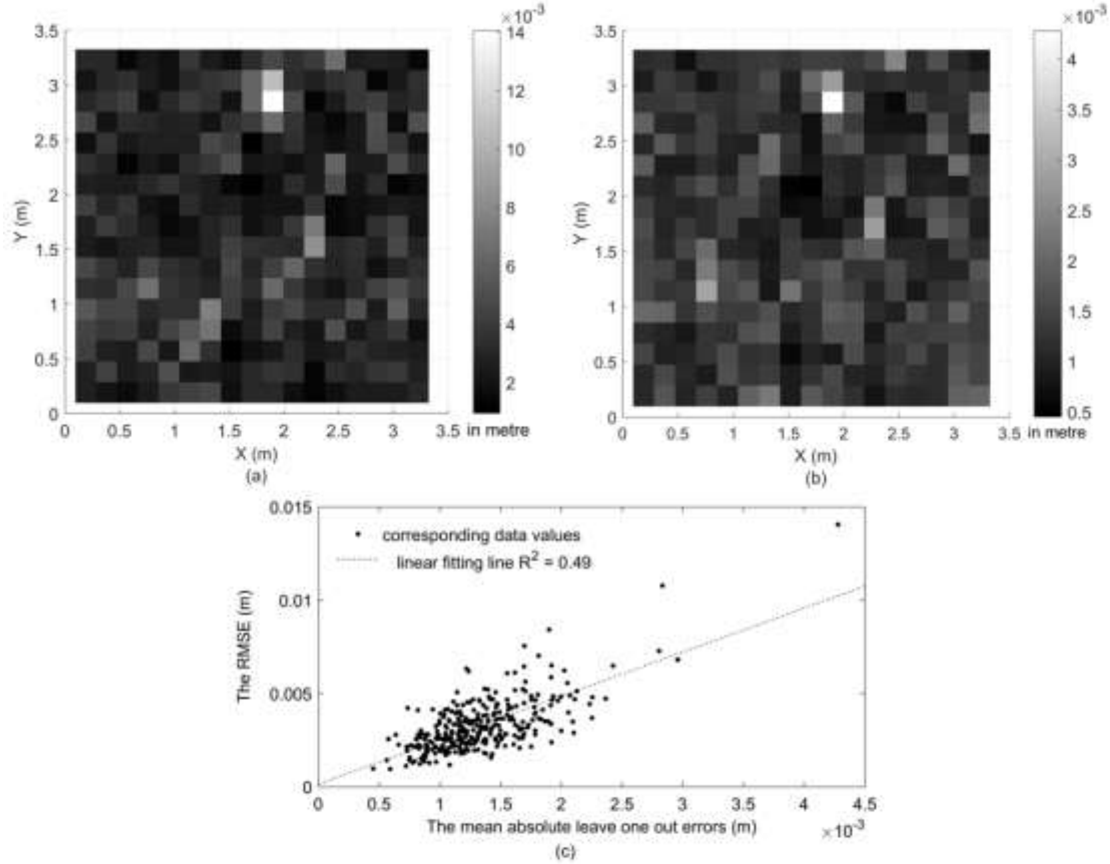

Fig. 3. (a) RMSH, (b) arithmetical mean absolute values of the leave-one-out errors, (c) the grid values in (a) plotted against the corresponding grid values in (b).

Fig. 4 shows the multi-resolution results for several different repeated rounds of subsample selection. For example, the results (i.e. the mean DoD map) of two independent realizations of the multi-resolution method at n = 50 (i.e. 50 DoD maps based on 50 different coarse-resolution point clouds) are shown in Fig. 4(a) and Fig. 4(b), respectively. As the values of the mean DoD vary with the amount of difference in the data spacings between the fine-resolution and coarse-resolution point clouds, these values were normalised. The normalisation was applied in such a way that the maximal absolute value after normalisation was 1. The grid values in Fig. 4(a) and Fig. 4(b) were plotted against each other for better visualisation of their correlation. The correlation was large as indicated by a large coefficient of determination. This suggests that as long as a reasonable number of rounds (e.g. n = 50) of subsample selections is used, the multi-resolution results obtained lead to consistent results. A range of other n values were investigated and some of those are reported in Fig. 4(d)-(f). Large variability for n = 1 (i.e. only one coarse-resolution point cloud used) was observed. It was also found that using more repetition rounds than n =50 did not lead to an obvious improvement.

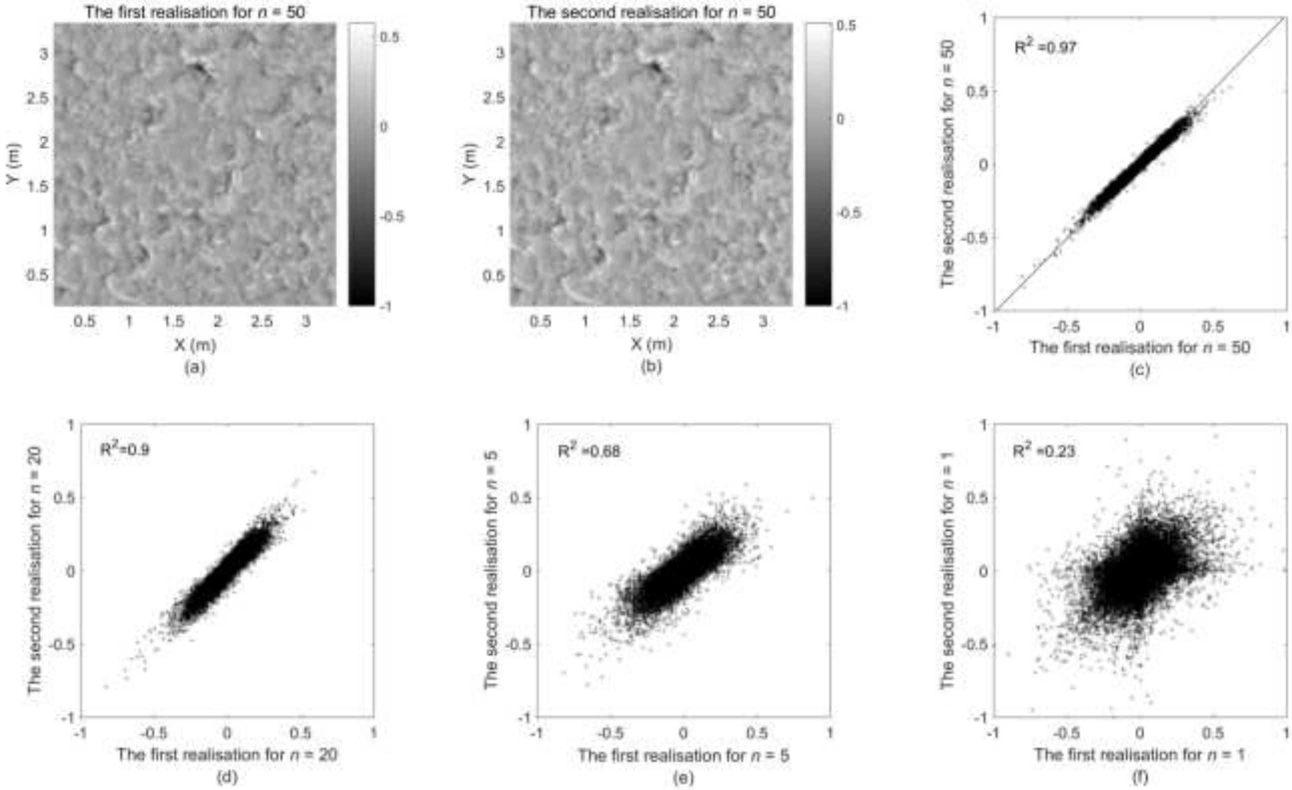

Fig. 4. Results of some trial tests (using the TLS point cloud) for determining the number of rounds for the subsample selection.

Based on n = 50 rounds of coarse-resolution point cloud selection and the data spacing pairing (0.037m and 0.071 m), the mean DoD map obtained using the multi-resolution method is shown in Fig. 5(a) and the leave-one-out results are shown in Fig. 5(b). The values of the leave-one-out results were not the same as those derived using the multi-resolution method as they achieve different things. The former represents the local prediction errors while the latter are the differences in prediction values from two DEMs created using point-clouds of different data spacings. There was a large positive correlation between the two datasets (Fig. 5(a) and (b)) in most areas. To confirm this, the grid values shown in Fig. 5(a) were plotted against those at the same grid locations in Fig. 5 (b), and these grid values are shown in Fig. 5(c). This gives a line of least-squares fit with a large coefficient of determination.

Compared to the RMSH method, some advantages of the multi-resolution method were observed. The surface roughness using the multi-resolution method has a larger correlation with the DEM interpolation errors, and it has both positive and negative values that can be used to differentiate areas where the DEM prediction error is positive or negative. As the RMSH is calculated over a particular scale (often defined by a non-overlapping window size), there is only one roughness value per window size (e.g. 0.19 m in this study) and, therefore, a loss of more detailed roughness information within the window size. However, in the multi-resolution method, a fine DEM grid resolution can be used for a fine-resolution roughness map.

The multi-resolution method requires a difference between the data spacing of the fine-resolution point cloud and that of the coarse-resolution point clouds. Several data spacings (S_2) for the coarse-resolution point cloud were investigated to understand their effects on the multi-resolution results. The analysis procedure used for that shown in Fig. 5(c) was repeated for several data spacing S_2, and the results are shown in Fig. 5(d). There exists an optimal data spacing (S_2 = 0.07 m) for the coarse-resolution point cloud, at which the largest correlation between the multi-resolution results and the leave-one-out results is observed. For the TLS data considered, the optimal data spacing is approximately 1.9 times the data spacing of the fine-resolution point cloud.

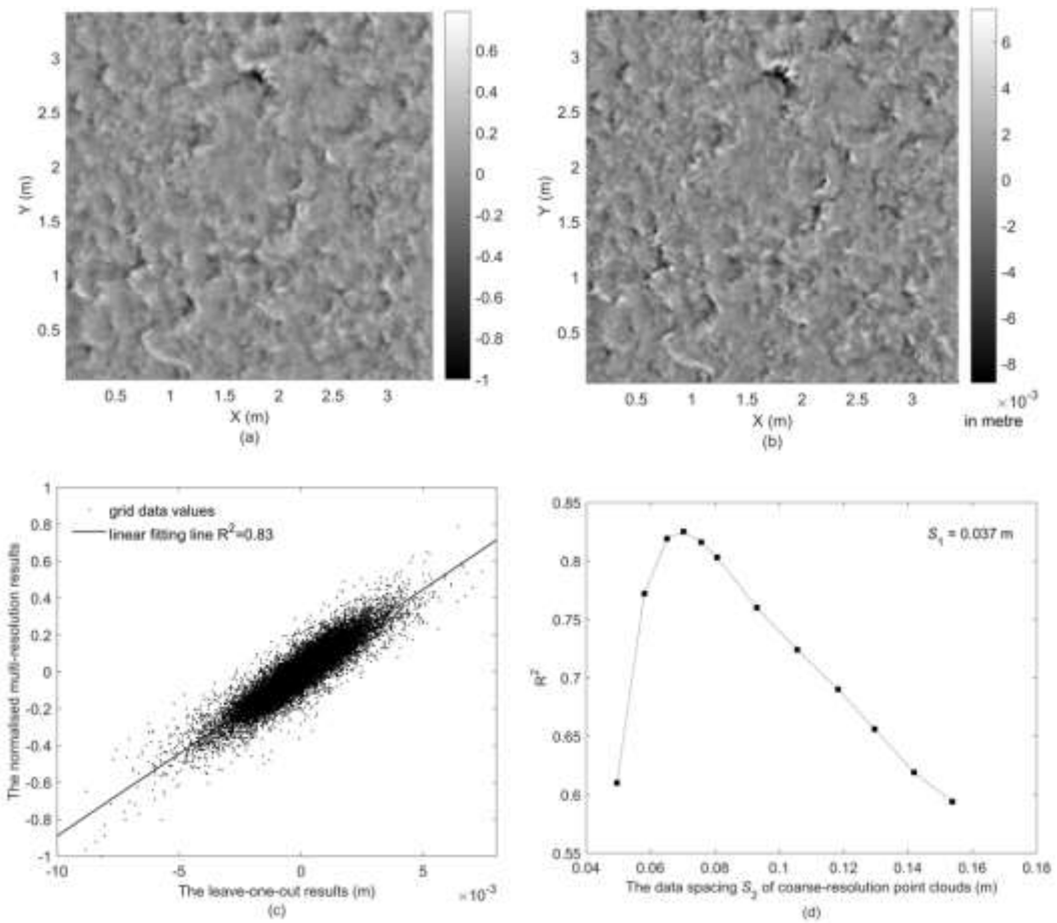

Fig. 5. Results for the TLS point cloud: (a) The multi-resolution method based on $S_1 = 0.037$ m, $S_2 = 0.07$ m, $r = 0.02$ m and $n = 50$, (b) the leave-one-out method, (c) the grid values in (a) plotted against the corresponding grid values in (b), (d) the change of coefficient of determination R^2 (refers to that shown in (c)) with the data spacing S_2 of the coarse-resolution point cloud.

3.2 The LiDAR data

The data analysis procedure for all the LiDAR point clouds is the same as that for the TLS point cloud, and is not repeated here. The analysis results for LiDAR-1, LiDAR-2 and LiDAR-3 are shown in Fig. 6, Fig. 7 and Fig. 8, respectively. The values of the parameters used in the multi-resolution method for each LiDAR point cloud are given in the corresponding figure caption. For all the LiDAR data considered, a large correlation between the local surface roughness (represented by the multi-resolution results) and the local DEM interpolation error (represented by the leave-one-out results) was observed. The results confirm the existence of an optimal data spacing for the coarse-resolution point cloud. The multi-resolution results shown in Fig. 6 – Fig.8 were based on the corresponding optimal data spacing for each dense-resolution point cloud.

It was found that the optimal data spacing of the coarse-resolution subsample ranged from 1.8 – 2.0 times the data spacing of the associated fine-resolution point cloud, which is similar to that for the TLS dataset considered. This finding can serve as a basis for determining the optimal data spacing when the multi-resolution method is applied to a new point cloud. The results obtained using the TLS and the LiDAR datasets confirm the effectiveness and the transferability of the multi-resolution method.

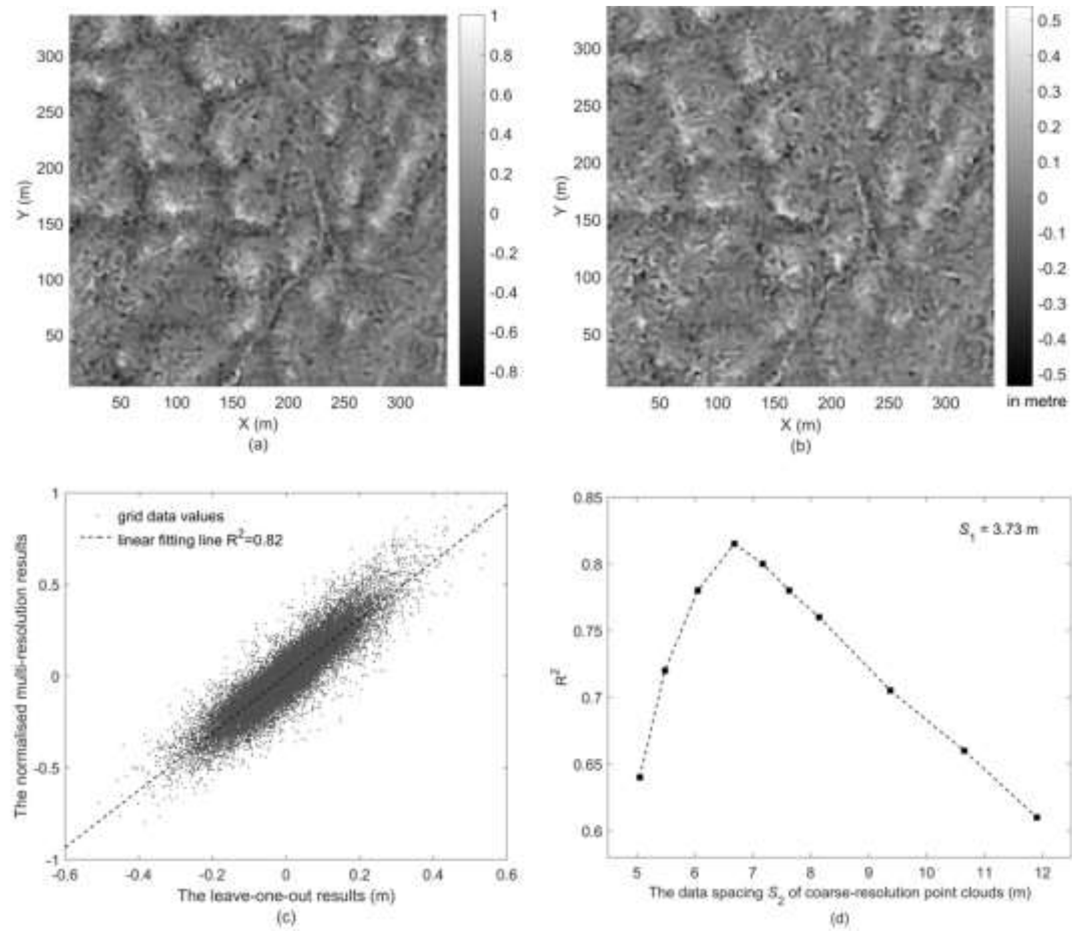

Fig. 6. Results for the LiDAR-1 point cloud: (a) The multi-resolution method based on $S_1 = 3.73$ m, $S_2 = 6.69$ m, $r = 2$ m and $n = 50$, (b) the leave-one-out method, (c) the grid values in (a) plotted against the corresponding grid values in (b), (d) the change of coefficient of determination R^2 (refers to the one shown in (c)) with the data spacing S_2 of the coarse-resolution point cloud.

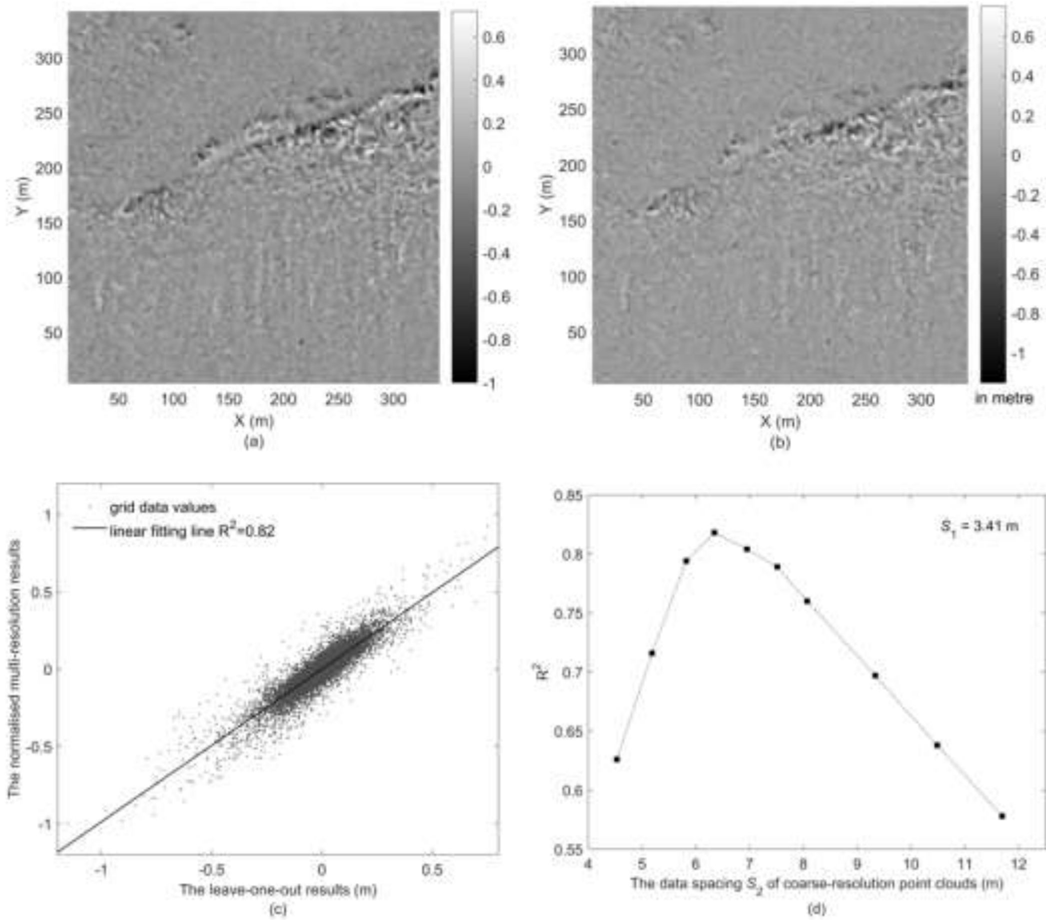

Fig. 7. Results for the LiDAR-2 point cloud: (a) The multi-resolution method based on S_1 = 3.41 m, S_2 = 6.35 m, r = 2 m and n = 50, (b) the leave-one-out method, (c) the grid values in (a) plotted against the corresponding grid values in (b), (d) the change of coefficient of determination R^2 (refers to the one shown in (c)) with the data spacing S_2 of the coarse-resolution point cloud.

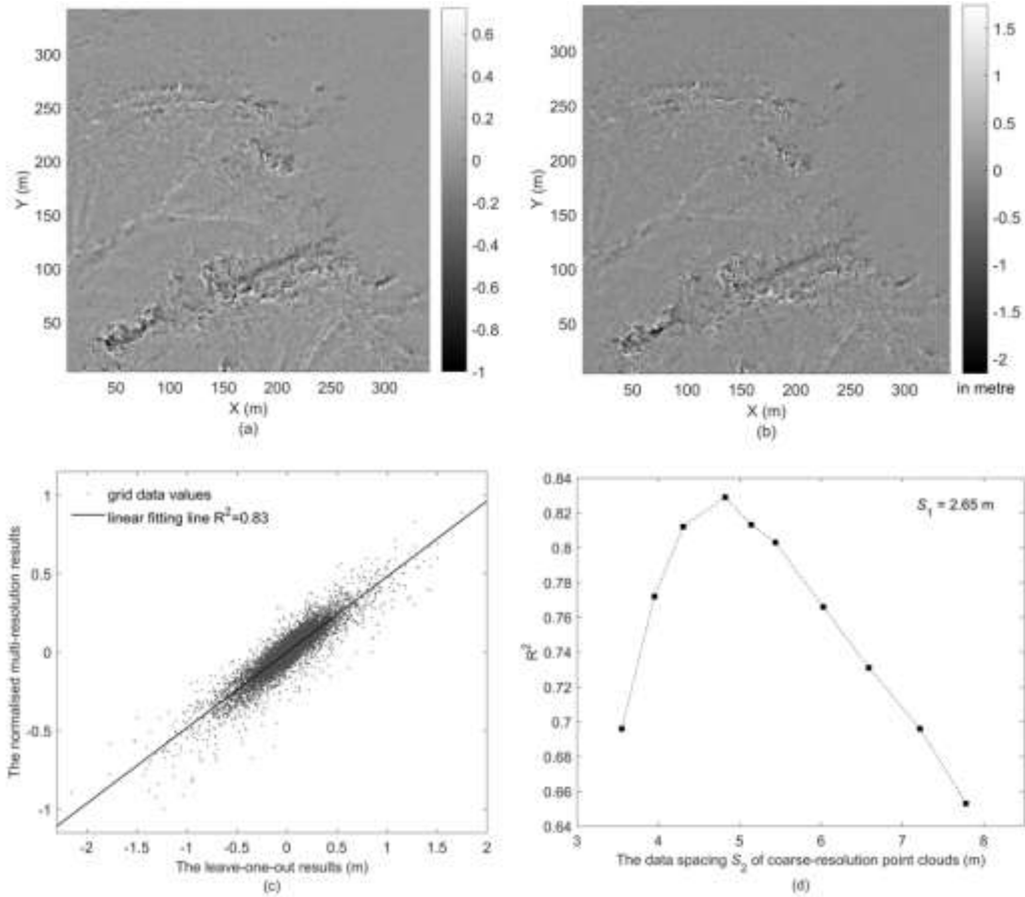

Fig. 8. Results for the LiDAR-3 point cloud: (a) The multi-resolution method base on S_1 = 2.65 m, S_2 = 4.83 m, r = 2 m and n = 50, (b) the leave-one-out method, (c) the grid values in (a) plotted against the corresponding grid values in (b), (d) the change of coefficient of determination R^2 (refers to the one shown in (c)) with the data spacing S_2 of the coarse-resolution point cloud.

4. Discussion

Although the correlation in the scatterplots shown in Fig. 5(c) is obvious, there are clear deviations of data from the fitted line. The likely sources of uncertainty causing those deviations may include the following. Firstly, the spacing of the filtered point cloud used in this investigation is not perfectly fixed and equal, which may lead to either larger or smaller local residuals (because the multi-resolution residual is affected by the difference in data spacing between the coarse-resolution and fine-resolution data, as discussed in Section 2.4). Secondly, to enable a more direct comparison, interpolation was used for the leave-one-out results. This second source of interpolation error may be another factor explaining the deviations. Thirdly, the instrumental measurement error can result in some spatial variabilities that are not detectable in the point clouds acquired, which may contribute to the deviations. For the TIN with linear interpolation, the measurement error component in the DEM error would be independent of data spacing. In an attempt to reduce the deviations, some trial tests were carried out where a moving average window or a Gaussian filter was applied to the grid maps in Fig. 5. It was found that the correlation (similar to that shown in Fig. 5(c)) can be improved and the value of coefficient of determination R^2 increased to 0.9. However, this additional filtering process should be considered with caution as it smooths the maps locally. For clarity, no filter was applied to the results shown in Fig. 5 – Fig. 8.

In the multi-resolution method, the number of rounds n = 50 was found to be adequate for all the point clouds considered. For a different dataset, the evaluation of the adequacy of a specific value of n can be achieved readily using the approach introduced in Section 2.5 (i.e. generate two realizations of multi-resolution results for n and compare the consistency of the results).

As demonstrated in Section 2.2, the leave-one-out errors conditioned by an equal data spacing represent the local interpolation errors. Therefore, the conditioned leave-one-out errors can also be used for inferring local surface roughness. However, the leave-one-out technique is often slow because it needs to sweep through each and every data point. For a point cloud of m data points, interpolation has to be carried out m times. Although the multi-resolution method involves certain rounds of interpolation (for the DEM construction using n different coarse-resolution point clouds), the number of rounds (e.g. n=50) is much smaller than the number (m) of data points. Therefore, the multi-resolution method is more computationally efficient. For the LiDAR-3 point cloud considered in this research using an average office computer (4 cores of i5-3470 CPU @ 3.2 GHz and 8 GB RAM), the time taken for the leave-one-out process was approximately 1003 seconds, which is much slower than that (approximately 13 seconds for n = 50 and 8 seconds n = 30) taken for the multi-resolution method. The leave-one-out process could be very time-confusing if there is a very large number of data points in a point cloud.

The method proposed in this article was found to be effective for describing local surface roughness or complexity and is of use in applications where the local DEM error is of interest. Future research will investigate its effectiveness for estimating surface roughness in other applications.

5. Conclusion

In this article, a simple approach for describing local terrain surface roughness or complexity was investigated. It was found that the roughness map obtained using the multi-resolution method was highly consistent with the spatial patterns of the local DEM error, and performed better than the RMSH descriptor. These suggest that the method is an effective means of quantifying local surface roughness or complexity, at least for applications where the DEM accuracy is of interest.

Acknowledgment:

The authors acknowledge funding for the research from Natural Science Foundation of Jiangsu Province under grant number BK20160393 and the University internal support under RDF-15-01-52. LiDAR data access is based on [LiDAR, ground] services provided by the OpenTopography Facility with support from the National Science Foundation under NSF Award Numbers 1226353 & 1225810. Lidar data acquisition completed by the National Center for Airborne Laser Mapping (NCALM - http://www.ncalm.org). NCALM funding provided by NSF's Division of Earth Sciences, Instrumentation and Facilities Program. EAR-1043051. https://doi.org/10.5069/G9PR7SX0.

References

[1] F. Ackermann, “The accuracy of digital terrain models,” in Proceedings of 37th Photogrammetric Week (University of Stuttgart), pages 113-143.

[2] B. H. Carlisle, “Modelling the spatial distribution of DEM error”, *Transaction in GIS*, vol.9, no.4, pp. 521-540, 2005.

[3] F. J. Aguilar, F. Agüera, M.A. Aguilar, and Carvajal F., “Effects of terrain morphology, sampling density, and interpolation methods on grid DEM accuracy,” *Photogrammetric Engineering & Remote Sensing*, vol. 71, no. 7, pp. 805-816, Jul. 2005.

[4] F.J. Aguilar, J.P. Mills, J. Delgado, M.A. Aguilar, J.G. Negreiros and J.L. Pérez, “Modelling vertical error in LiDAR-derived digital elevation models,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 65, no. 1, pp. 103-110, Oct. 2010.

[5] K. Kraus, W. Karel, C. Briese and G. Mandlburger, “Local accuracy measures for digital terrain models,” *The Photogrammetric Record*, vol. 21, no. 116, pp. 342-354, Dec. 2006.

[6] M. Milenković, N. Pfeifer and P. Glira, “Applying Terrestrial Laser Scanning for Soil Surface Roughness Assessment,” Remote Sensing, vol. 7, pp. 2007-2045, 2015.

[7] R. D. Hobson, “Surface roughness in topography: Quantitative approach,” in *Spatial Analysis in Geomorphology*, R. J. Chorley, Ed. London, U.K.: Methuer, 1972, pp. 225–245.

[8] J. M. Nield and G. F. S. Wiggs, “The application of terrestrial laser scanning to aeolian saltation cloud measurement and its response to changing surface moisture,” *Earth Surf. Processes Landforms*, vol. 36, no. 2, pp. 273–278, 2011.

[9] C. H. Grohmann, M. J. Smith, and C. Riccomini, “Multiscale analysis of topographic surface roughness in the Midland Valley, Scotland,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 4, pp. 1200–1213, 2011.

[10] K. L. Frankel and J. F. Dolan, “Characterizing arid region alluvial fan surface roughnesswith airborne laser swathmapping digital topographic data,” *Journal of Geophysical Research F*, vol. 112, no.2, Article ID F02025, 2007

[11] R. Andrle and A. D. Abrahams, “Fractal techniques and the surface roughness of talus slopes,” *Earth Surface Processes & Landforms*, vol. 14, no. 3, pp. 197–209, 1989.

[12] C. H. Huang and J. M. Bradford, “Applications of a laser scanner to quantify soil microtopography,” *Soil Science Society of America Journal*, vol. 56, no. 1, pp. 14–21, 1992.

[13] U.C. Herzfeld, H. Mayer, W. Feller, and M. Mimler, “Geostatistical analysis of glacier-roughness data,” *Ann. Glaciol.*, vol. 30, no. 1, pp. 235– 242, Jan. 2000.

[14] C.H. Hugenholtz, O.W. Brown and T.E. Barchyn, “Estimating aerodynamic roughness (z_0) from terrestrial laser scanning point cloud data over un-vegetated surfaces,” *Aeolian Research*, vol. 10, pp.161-169, 2013.

[15] K.M. Brubaker, W.L. Myers, P.J. Drohan, D.A. Miller and E.W. Boyer, “The Use of LiDAR Terrain Data in Characterizing Surface Roughness and Microtopography,” *Applied and Environmental Soil Science*, vol. 2013, 13 pages, 2013.

[16] P. F. Fisher, and N.J. Tate, “Causes and consequences of error in digital elevation models,” *Progress in physical geography*, vol. 30, no. 4, pp. 467-489, Aug. 2006.

[17] F.J. Aguilar, M.A. Aguilar, F. Agüera, and J. Sánchez, "The accuracy of grid digital elevation models linearly constructed from scattered sample data," *International Journal of Geographical Information Science*, vol. 20, no. 2, pp. 169-192, 2006.
[18] B. Makarovic, "Progressive sampling for DTMs", ITC Journal, 1973-4: 397-416. 1973
[19] P. Frederiksen, "Terrain analysis and accuracy prediction by means of the Fourier transformation," *Photogrammetria*, vol. 36, pp. 145-157, 1981.
[20] P. Frederiksen, O. Jacobi and K. Bubik, "Optimum sampling spacing in digital elevation models," *International Archives of Photogrammetry and Remote Sensing*, vol. 26(3/1), pp. 252-259, 1986.
[21] Z.L. Li, "A comparative study of the accuracy of digital terrain models based on various data models," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 49, no. 1, pp. 2-11, 1994.
[22] L. Fan, J. A. Smethurst, P. M. Atkinson and W. Powrie, " Propagation of vertical and horizontal source data errors into a TIN with linear interpolation," *International Journal of Geographical Information Science*, 28(7), 1378-1400, 2014.
[23] Q. Guo, W. Li, H. Yu and O. Alvarez. "Effects of topographic variability and lidar sampling density on several DEM interpolation methods," *Photogrammetric Engineering & Remote Sensing*, vol.76, pp.1-12. 2010.
[24] P. Hu, X.H. Liu and H. Hu, "Accuracy Assessment of Digital Elevation Models based on Approximation Theory," *Photogrammetric Engineering & Remote Sensing*, vol. 75, no. 1, pp. 49–56, January 2009.
[25] C. H. Davis, H. Jiang and X. Wang, "Modeling and estimation of the spatial variation of elevation error in high resolution DEMs from stereo image processing", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, pp. 2483–89, 2001.
[26] H. Liu and K.C. Jezek, "Investigating DEM error patterns by directional variograms and Fourier analysis", *Geographical Analysis*, vol.31, pp. 249–66, 1999.
[27] B.H. Carlisle, "Modelling the spatial distribution of DEM error", *Transactions in GIS*, vol. 9, no. 4, pp. 521-540, Sep. 2005.
[28] Z.L. Li, "Theoretical models of the accuracy of digital terrain models: an evaluation and some observations", *Photogrammetric Record*, vol. 14, no. 82, pp. 651-660, October 1993.
[29] S. Erdogan, "A comparison of interpolation methods for producing digital elevation models at the field scale," *Earth Surface Process and Landforms*, vol. 34, no. 3, pp. 366-376, Mar. 2009.
[30] X.H. Liu, H. Hu, and P. Hu, "Accuracy Assessment of LiDAR-Derived Digital Elevation Models Based on Approximation Theory", *Remote Sensing*, vol.7, no.6, pp. 7062-7079, 2015.
[31] M.E. Hodgson and P. Bresnahan, "Accuracy of airborne lidar-derived elevation: empirical assessment and error budget", *Photogrammetric engineering and remote sensing*, vol. 70, no. 3, pp. 331-340, 2004.
[32] F.J. Aguilar and J.P. Mills, "Accuracy assessment of lidar-derived digital elevation models, *Photogrammetric Record*, vol. 23, no.122, pp. 148-169, Jun. 2008.
[33] S. Wise, "Cross-validation as a means of investigating DEM interpolation error," *Computers & Geosciences*, vol. 37, no. 8, pp. 978-991, 2011.
[34] S.L. Smith, D.A. Holland and P.A. Longley, "Quantifying interpolation errors in urban airborne laser scanning models," *Geographical Analysis*, vol. 37, no. 2, pp. 200–224, 2005.
[35] L. Fan and P. M. Atkinson, "Accuracy of Digital Elevation Models Derived From Terrestrial Laser Scanning Data," *IEEE Geoscience and Remote Sensing Letters*, vol. 12, no. 9, pp. 1923-1927, Sept. 2015.