Propagation of vertical and horizontal source data errors into a TIN with linear interpolation

Lei Fan[a], Joel Smethurst[a], Peter M. Atkinson[b] William Powrie[a]

[a] Faculty of Engineering and the Environment, University of Southampton, University Road, Southampton, SO17 1BJ, UK. (Lei.Fan@soton.ac.uk; W.Powrie@soton.ac.uk; J.A.Smethurst@soton.ac.uk)

[b] Geography and Environment, University of Southampton, University Road, Southampton, SO17 1BJ, UK. (P.M.Atkinson@soton.ac.uk)

Digital elevation models (DEMs) have been widely used for a range of applications and form the basis of many GIS related tasks. An essential aspect of a DEM is its accuracy, which depends on a variety of factors, such as source data quality, interpolation methods, data sampling density and the surface topographical characteristics. In recent years, point measurements acquired directly from land surveying such as differential global positioning system (DGPS) and light detection and ranging (LiDAR) have become increasingly popular. These topographical data points can be used as the source data for the creation of DEMs at a local or regional scale. The errors in point measurements can be estimated in some cases. The focus of this article is on how the errors in the source data propagate into DEMs. The interpolation method considered is a triangulated irregular network (TIN) with linear interpolation. Both horizontal and vertical errors in source data points are considered in this study. An analytical method is derived for the error propagation into any particular point of interest within a TIN model. The solution is validated using Monte Carlo simulations and survey data obtained from a terrestrial laser scanner.

Keywords: Error propagation; DEM accuracy; Triangulated irregular network; Spatial interpolation; Topographical survey

1. Introduction

A digital elevation model (DEM) is a digital representation of a part of the Earth's surface. It forms a basic input to many GIS-related tasks and has a wide variety of applications, such as mapping, terrain analysis and environmental modelling (Webster *et al.*, 2006; Chen and Zhou, 2013; Brunori *et al.*, 2013). Two major categories of DEMs are used widely to represent a terrain surface, namely vector-based DEMs and grid-based DEMs (Zhu, *et al.*, 2005). The former usually refers to a triangulated irregular network (TIN). The grid-based DEMs represent elevations in a raster or grid framework, which can be used to calculate many elevation surface variables such as slope and flow direction. To create a grid-based DEM or predict the elevations at the points where no source data are available, spatial interpolation is commonly used. A wide variety of interpolation methods are available, which are either deterministic (e.g. TIN with linear interpolation, polynomial interpolation, inverse distance weighting) or geostatistical (e.g. kriging).

The interpolation method used in this article is a TIN with linear interpolation. It is common for GIS designers to use relatively simple and straightforward interpolation methods to speed up execution and simplify the decision process on the part of users (Kyriakidis and

Goodchild, 2006). For example, the Ordnance Survey in the UK uses TINs in the production process for orthorectified digital imagery and a range of height products. Researchers (Wheaton *et al.*, 2010; Day *et al.*, 2013) have used the TIN model derived from topographical survey data to detect surface deformations. Surveying strategies seem to be considered more important than the selection of interpolation methods (Bater and Coops, 2009; Heritage *et al.*, 2009). The reason for this is that when the density of the source data is high, there is little justification for employing a more complex interpolation method (Lloyd and Atkinson, 2006).

Historically, the source data used for the generation of DEMs have most frequently come from contour lines, paper maps and imagery such as stereo aerial photographs (Fisher and Tate, 2006; Zhu *et al.*, 2005). Imagery from various platforms still represents a popular source of data with which to construct DEMs (e.g. Gessesse *et al.*, 2010; Fonstad *et al.*, 2013). Point measurements acquired directly from land surveying were less frequently used to create DEMs in the past (Fisher and Tate, 2006). However, point measurement sources have become increasingly popular in recent years since differential global positioning system (DGPS) and laser scanning systems have come to be used more widely (e.g. Lloyd and Atkinson, 2006; Schürch *et al.*, 2011; Negishi *et al.*, 2012; Gallay *et al.*, 2013). Light detection and ranging (LiDAR) is now regarded as a standard tool for acquisition of DEM source data in many applications (Coveney, 2013).

Evaluation of the accuracy of a DEM is an essential part of DEM production. There are two error components associated with DEMs. These are: (i) errors in the individual data points used to build DEMs, and (ii) interpolation errors. A combination of these two components gives a measure of DEM accuracy that is influenced by several factors, such as the characteristics of the terrain surface and the density and distribution of survey points. Much research has been carried out to assess the quality of a DEM and investigate the effects of different factors on DEM accuracy (Kyriakidis *et al.*, 1999; Lloyd and Atkinson, 2002; Aguilar *et al.*, 2005; Erdogan, 2009; Bater and Coops, 2009; Hu *et al.*, 2009; Heritage *et al.*, 2009). The effects of DEM accuracy on other GIS variables (e.g. slope and aspect) have also been investigated (Oksanen and Sarjakoski, 2005; Wechsler and Kroll, 2006).

Error propagation in GIS is a broad topic, with the solution very often dependent on the nature of the dataset and its embedded error, the GIS operations and the GIS products the errors are propagated to. An excellent overview of error propagation in GIS can be found in Heuvelink (1998). Leung *et al.* (2004a, 2004b) also proposed a framework for error analysis and propagation in a measurement-based GIS (a concept proposed by Goodchild (1999)), and considered error analysis in several features such as intersections and overlays. Fisher and Tate (2006) gave a useful summary of previous research work on error propagation in relation to DEMs and DEM derivatives such as slope and visibility products. As there is a large literature on error propagation in GIS, the work most relevant to this study is reviewed in the following paragraph.

In this article, the propagated error variance from source data points is investigated. Previous studies on this have mainly focused on vertical error propagation. Zhu *et al.* (2005) reported that the average elevation error variance of a triangular surface was half the vertical error variance at node points, based on the assumption that the error variance of node points is identical and independent. Shi *et al.* (2005) estimated the propagated error variance from source data into DEMs created through higher-order interpolation (biquadratic and bicubic). Aguilar *et al.* (2006) studied the accuracy of grid DEMs linearly constructed from scattered sample data. Only the average error variance for a TIN model was evaluated in these studies (i.e. Zhu *et al.*, 2005; Shi *et al.*, 2005; Aguilar *et al.*, 2006). Kyriakidis and Goodchild (2006) also investigated the average error variance of three linear interpolation methods (line linear, TIN linear and bilinear), and the propagation of vertical error variance at node points into a

generic point using the geostatistics framework. Hu *et al.* (2009) approximated a complex function using simpler functions to assess the accuracy of DEMs, including TIN. They found that for each triangular surface the propagated error variance at any point within the triangular surface was bounded by the largest error at the triangle vertices. There may also be a lower bound dictating that the propagated error variance at any point is greater than a given value. In these studies, no survey data were used to validate the error variance propagated.

In previous research, the propagation of errors in the planimetric or horizontal direction was not considered. This may be reasonable for the assessment of an existing DEM since it is known that horizontal error will cause vertical error. The assessment can be carried out either by evaluating the elevation differences at a number of check points measured using another instrument, or by comparing the DEM with a benchmark DEM created using higher accuracy measurements. However, for most topographical point *measurement* data, the positional error has a horizontal component. The propagation of horizontal error variance directly from the topographical data points to DEMs has drawn little attention.

A propagation solution can be quite useful, in particular for topographical point measurements such as DGPS and LiDAR (including, terrestrial laser scanning (TLS)) data. In these cases, quantitative information on the accuracy in both the planimetric and elevation directions is usually available or can be assessed. For example, manufacturers of terrestrial laser scanners provide range, angular and positional accuracy information in the device specification. Some DGPS receivers are able to record horizontal and vertical data quality for each individual reading. The American Society for Photogrammetry and Remote Sensing (ASPRS) (2004) provided a set of guidelines for assessing and reporting vertical and horizontal accuracy of LiDAR data. When point measurements are used to create a DEM, the associated horizontal and vertical errors will be propagated into the DEM. An understanding of the propagated measurement errors is then important for further use of the DEM. The focus of this article is therefore on the propagation of error variance from individual source data points into any point of interest in a TIN model. Both horizontal and vertical errors in the source data are considered, and an analytical approach is introduced. To test the method, Monte Carlo simulations and laser scanning experiments were carried out. The survey experiments also demonstrate that the method is of practical application to topographical survey data.

2. TIN with linear interpolation

TIN is a popular means of representing a terrain surface in digital elevation modelling. In this method, the optimal Delaunay triangulation is commonly used for the selection of data points and formation of triangles. The triangles are created in such a way that no data point is inside the circumcircle of any triangle and their edges are not intersected by each other. All the triangle-shaped surfaces over the spatial extent of the data points constitute a complete and continuous surface. Within each triangle, the surface can be expressed by a plane (Equation (1)) determined from the three vertices of the triangle (Figure 1). The elevation of a particular point within a given triangle is determined by the linear function (i.e. linear interpolation):

$$z(x, y) = ax + by + c \tag{1}$$

Linear algebra gives the solutions for a, b and c in the format of determinants:

$$a = \frac{\begin{vmatrix} z_1 & y_1 & 1 \\ z_2 & y_2 & 1 \\ z_3 & y_3 & 1 \end{vmatrix}}{\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}} \qquad b = \frac{\begin{vmatrix} x_1 & z_1 & 1 \\ x_2 & z_2 & 1 \\ x_3 & z_3 & 1 \end{vmatrix}}{\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}} \qquad c = \frac{\begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix}}{\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}} \tag{2}$$

3. Error propagation

The law of propagation of uncertainty has been used widely in the field of measurement science, and is based on a Taylor series expansion. The error propagation law for a multivariate model is used in this study without proof. A detailed guide on uncertainty evaluation and propagation can be found in Taylor (1982) and Cox and Harris (2006).

A DEM is created from a set of data points. In mathematical terms, the DEM is an output quantity while the data points are input quantities. The relationship between the output quantity and the input quantities is the model, which can be a mathematical function or a stepwise algorithm. For N input quantities denoted by $\mathbf{I} = (I_1, \cdots, I_N)$, the output quantity O can be expressed as:

$$O = f(\mathbf{I}) = f(I_1, \cdots, I_N) \tag{3}$$

where f is a mathematical function. There may be more than one output quantity (say M in number), $\mathbf{O} = (O_1, \cdots, O_M)$. In this case, the output quantities are:

$$\mathbf{O} = \mathbf{F}(\mathbf{I}) = \mathbf{F}(I_1, \cdots, I_N) = (f_1(I_1, \cdots, I_N), \cdots, f_M(I_1, \cdots, I_N)) \tag{4}$$

The law of uncertainty propagation for a multivariate model is expressed in Equation (5),

$$\mathbf{C}_O = \mathbf{J}_I \mathbf{C}_I \mathbf{J}_I^{\mathrm{T}} \tag{5}$$

or equivalently in matrix form,

$$
\begin{aligned}
&\begin{bmatrix}
\sigma_{O_1}^2 & \sigma_{O_1 O_2} & \cdots & \sigma_{O_1 O_M} \\
\sigma_{O_2 O_1} & \sigma_{O_2}^2 & \cdots & \sigma_{O_2 O_M} \\
\vdots & \vdots & \ddots & \vdots \\
\sigma_{O_M O_1} & \sigma_{O_M O_2} & \cdots & \sigma_{O_M}^2
\end{bmatrix} \\
&= \begin{bmatrix}
\frac{\partial f_1(\mathbf{I})}{\partial I_1} & \cdots & \frac{\partial f_1(\mathbf{I})}{\partial I_N} \\
\vdots & \ddots & \vdots \\
\frac{\partial f_M(\mathbf{I})}{\partial I_1} & \cdots & \frac{\partial f_M(\mathbf{I})}{\partial I_N}
\end{bmatrix}
\begin{bmatrix}
\sigma_{I_1}^2 & \sigma_{I_1 I_2} & \cdots & \sigma_{I_1 I_N} \\
\sigma_{I_2 I_1} & \sigma_{I_2}^2 & \cdots & \sigma_{I_2 I_N} \\
\vdots & \vdots & \ddots & \vdots \\
\sigma_{I_N I_1} & \sigma_{I_N I_2} & \cdots & \sigma_{I_N}^2
\end{bmatrix}
\begin{bmatrix}
\frac{\partial f_1(\mathbf{I})}{\partial I_1} & \cdots & \frac{\partial f_1(\mathbf{I})}{\partial I_N} \\
\vdots & \ddots & \vdots \\
\frac{\partial f_M(\mathbf{I})}{\partial I_1} & \cdots & \frac{\partial f_M(\mathbf{I})}{\partial I_N}
\end{bmatrix}^{\mathrm{T}}
\end{aligned} \tag{6}
$$

in which, $\sigma_{I_i}^2$ $(i = 1, \cdots, N)$ is the variance of input variables; $\sigma_{O_j}^2$ $(j = 1, \cdots, M)$ is the variance of output variables; $\sigma_{I_i I_k}$ $(i, k = 1, \cdots, N; i \neq k)$ is the covariance of input variables; $\sigma_{O_j O_l}$ $(j, l = 1, \cdots, M; j \neq l)$ is the covariance of output variables; $\frac{\partial f_i(\mathbf{I})}{\partial I_j}$ is the partial difference of the system of equations.

4. Propagated error variance of source data points

As it is difficult to consider the propagation of vertical and horizontal error variances simultaneously, they are considered separately and then combined in this research. In Section 4, the analytical solutions for propagation of error variance are derived, based on a location-dependent parameter *M*. This parameter is then quantified in Section 5. The analytical solutions are validated using Monte Carlo simulations and two simple survey experiments in Section 6.

4.1 Propagation of vertical error

It is assumed that the horizontal uncertainty of the source data points is zero (i.e. they are error free). Therefore, only the vertical errors are propagated. An analytical solution is derived to compute the error variance at any point of interest in a TIN model. The solutions given in Equation (2) can be written in a linear format and rearranged as shown in Equation (7). This operation aims to distinguish the elevation variable z from the location variables x and y.

$$a = \frac{(y_2-y_3)z_1 + (y_3-y_1)z_2 + (y_1-y_2)z_3}{x_1y_2 + x_3y_1 + x_2y_3 - x_3y_2 - x_1y_3 - x_2y_1} = a_1z_1 + a_2z_2 + a_3z_3$$

$$b = \frac{(x_3 - x_2)z_1 + (x_1-x_3)z_2 + (x_2 - x_1)z_3}{x_1y_2 + x_3y_1 + x_2y_3 - x_3y_2 - x_1y_3 - x_2y_1} = b_1z_1 + b_2z_2 + b_3z_3 \tag{7}$$

$$c = \frac{(x_2y_3 - x_3y_2)z_1 + (x_3y_1 - x_1y_3)z_2 + (x_1y_2 - x_2y_1)z_3}{x_1y_2 + x_3y_1 + x_2y_3 - x_3y_2 - x_1y_3 - x_2y_1} = c_1z_1 + c_2z_2 + c_3z_3$$

in which, $a_1 = \frac{y_2-y_3}{L}$; $a_2 = \frac{y_3-y_1}{L}$; $a_3 = \frac{y_1-y_2}{L}$, and $L = x_1y_2 + x_3y_1 + x_2y_3 - x_3y_2 - x_1y_3 - x_2y_1$; the same applies to b_1, b_2, b_3, c_1, c_2 and c_3, the expressions for which are not listed here.

By substituting a, b and c in Equation (1) with the expressions in Equation (7), the elevation z_P of the point $P(x_p , y_p)$ in the TIN model can be calculated (Equation (8)).

$$z_P = ax_P + by_P + c$$

$$= (a_1z_1 + a_2z_2 + a_3z_3)x_P + (b_1z_1 + b_2z_2 + b_3z_3)y_P + (c_1z_1 + c_2z_2 + c_3z_3) \tag{8}$$

$$= (a_1x_P + b_1y_P + c_1)z_1 + (a_2x_P + b_2y_P + c_2)z_2 + (a_3x_P + b_3y_P + c_3)z_3$$

Rewriting Equation (8) in matrix form:

$$z_P = \begin{bmatrix} x_P \\ y_P \\ 1 \end{bmatrix}^T \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} \tag{9}$$

On the right hand side of Equation (9), the first vector represents the location of the point P, the elevation of which is to be estimated (z_P). The second matrix is determined by the locations of the three vertices of a triangle. Recall that their positional errors are not considered at this stage. Therefore, the second term is a matrix of constants for a given triangle. The last vector represents the elevations of the three vertices of the triangle. Equation (9) shows that the elevation accuracy of the point P depends not only on the vertical accuracy of the survey data points but also its position (x_p, y_p).

Equation (9) can be expressed as:

$$z_P = m_1 z_1 + m_2 z_2 + m_3 z_3 \tag{10}$$

in which, $\begin{bmatrix} m_1 \\ m_2 \\ m_3 \end{bmatrix}^{\text{T}} = \begin{bmatrix} x_P \\ y_P \\ 1 \end{bmatrix}^{\text{T}} \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}$

For a given triangle and a known interpolation location (x_p, y_p) within the triangle, m_1, m_2 and m_3 can be calculated. It is known that there are errors associated with z_1, z_2 and z_3. The remaining question is how to propagate the vertical errors of data points into the elevation predicted at the location P.

Based on Equation (6) the error variance $\sigma_{z_p}^2$ of a particular point can be calculated as:

$$\sigma_{z_p}^2 = m_1^2 \sigma_{z_1}^2 + m_2^2 \sigma_{z_2}^2 + m_3^2 \sigma_{z_3}^2 + 2m_1 m_2 \sigma_{z_1 z_2} + 2m_1 m_3 \sigma_{z_1 z_3} + 2m_2 m_3 \sigma_{z_2 z_3} \tag{11}$$

in which, $\sigma_{z_1}^2$, $\sigma_{z_2}^2$ and $\sigma_{z_3}^2$ are the error variances of the three node points; $\sigma_{z_1 z_2}$, $\sigma_{z_1 z_3}$ and $\sigma_{z_2 z_3}$ are the covariances between the three node points.

It should be noted that Kyriakidis and Goodchild (2006) obtained a solution similar to Equation (11), based on the proportion (weight) of the areas of sub-triangles formed by triangle vertices and a prediction location, to the total area of a TIN element. The weight was derived using a geostatistical formulation. Our article shows that the solution can also be obtained with error propagation directly through the linear function for a triangular surface. In addition, the dependence of the weight (m_i) on the prediction location and the spatial distribution of propagated measurement error variance are both investigated. The solution is validated using Monte Carlo simulations and laser scanning survey data. The propagation of horizontal measurement error is also investigated since most topographical survey data have horizontal measurement uncertainty.

If the random errors in the elevations of node points (z_1, z_2 and z_3) are independent of each other, Equation (12) can be used to propagate the error variances into the point P.

$$\sigma_{z_p}^2 = m_1^2 \sigma_{z_1}^2 + m_2^2 \sigma_{z_2}^2 + m_3^2 \sigma_{z_3}^2 \tag{12}$$

If the random errors are identical ($\sigma_{z_{\text{node}}}$) for all node points, Equation (12) can be further simplified to Equation (13).

6

$$\sigma_{z_p}^2 = (m_1^2 + m_2^2 + m_3^2)\sigma_{z_{node}}^2$$

$$= M\sigma_{z_{node}}^2 \tag{13}$$

M is a location-dependent parameter. The value of M depends on the actual location of the point of interest. This is demonstrated in Section 5.

4.2 Propagation of horizontal error

The horizontal error of a measurement does not affect its elevation component. However, if the terrain surface is inclined, the horizontal error will generate an elevation error during digital elevation modelling. Several researchers (Hodgson and Bresnahan, 2004; Su and Bork, 2006; Spaete *et al.*, 2011) have reported that terrain slope influenced the accuracy of DEMs created from topographical survey data.

4.2.1 Geometric relation method

A straightforward way of investigating the effect of horizontal errors on elevation is to use the geometric relation. To simplify the demonstration, a 2D case is illustrated in Figure 2. It is also valid for a triangular surface in a TIN model. The horizontal error will generate an elevation error (from a DEM accuracy perspective) if the surface is inclined. The effect of the horizontal error (line AC) on DEM accuracy is equivalent to that of a vertical error (line AD), $|AD| = |AC| \tan \beta$. The angle β can be estimated from a slope map using a variety of means suggested in the literature. Alternatively, $|AD|$ can be approximated by $|EC|$ ($|EC| = |AC| \tan \alpha$), if the horizontal error $|AC|$ is small compared with the data point spacing. α represents the angle of line BA, and is a known variable (because points B and A are measured and their coordinates are known). In the context of a triangular surface in a TIN model, α (in both X and Y axes) can be derived directly from Equation (1). In this approach, a separate slope map is not required, hence simplifying the work process. The following discussion is based on the second approach.

Based on the above, Equation (14) can be used to estimate the equivalent elevation error caused by the horizontal measurement error. The slope angle is derived from individual triangular surfaces in a TIN model, and may vary from one triangular surface to another.

$$\delta_{z_x} = \delta_x \tan \alpha_x \tag{14}$$

in which, δ_x is the horizontal error; α_x is the angle of inclination of the surface with respect to the X axis; δ_{z_x} is the equivalent vertical error caused by δ_x.

Once δ_{z_x} has been obtained, the problem is converted into the case discussed in Section 4.1, i.e. the propagation of vertical error variances. Therefore, Equation (11) can also be used to propagate the horizontal error. In the special case of the random horizontal error being independent and identical ($\sigma_{x_{node}}^2$) for all three node points of a triangular surface, the vertical error variance $\sigma_{z_p}^2$ of a particular point due to the horizontal error is given by:

$$\sigma_{z_p}^2 = M\sigma_{z_x}^2 = M(\sigma_{x_{node}} \tan \alpha_x)^2$$

$$= M(\tan \alpha_x)^2 \sigma_{x_{node}}^2 \tag{15}$$

in which $\sigma_{x_{node}}^2$ is the error variance of node points in the X axis.

4.2.2 Error propagation method

It is also possible to use the law of uncertainty propagation (Equation (6)) to solve the horizontal error propagation problem. To use this method, the measurement error was assumed to be small compared with the data point spacing (i.e. the second approach to derive slope angles in Section 4.2.1). Only the horizontal error component in the X axis is considered (it is the same for the Y axis). Equation (2) is rearranged so that the $x_i (i = 1,2,3)$ are separated from the other components y_i and $z_i (i = 1,2,3)$. Equation (16) is then obtained, which aims to emphasize that only $x_i (i = 1,2,3)$ are random variables having uncertainty and the other components are constants if the three node points are known.

$$a = \frac{(z_3 - z_2)y_1 + (z_1 - z_3)y_2 + (z_2 - z_1)y_3}{(y_2 - y_3)x_1 + (y_3 - y_1)x_2 + (y_1 - y_2)x_3}$$

$$b = \frac{(z_2 - z_3)x_1 + (z_3 - z_1)x_2 + (z_1 - z_2)x_3}{(y_2 - y_3)x_1 + (y_3 - y_1)x_2 + (y_1 - y_2)x_3} \tag{16}$$

$$c = \frac{(y_2 z_3 - y_3 z_2)x_1 + (y_3 z_1 - y_1 z_3)x_2 + (y_1 z_2 - y_2 z_1)x_3}{(y_2 - y_3)x_1 + (y_3 - y_1)x_2 + (y_1 - y_2)x_3}$$

The expressions for a, b and c in Equation (16) are substituted into Equation (1). As this will make Equation (1) complex, the following expressions are used to simplify the equation after substitution.

$$r_0 = x_p[(z_3 - z_2)y_1 + (z_1 - z_3)y_2 + (z_2 - z_1)y_3]$$
$$r_1 = y_p z_2 - y_p z_3 + y_2 z_3 - y_3 z_2$$
$$r_2 = y_p z_3 - y_p z_1 + y_3 z_1 - y_1 z_3$$
$$r_3 = y_p z_1 - y_p z_2 + y_1 z_2 - y_2 z_1$$
$$r_4 = y_2 - y_3$$
$$r_5 = y_3 - y_1$$
$$r_6 = y_1 - y_2$$

Giving:

$$z_p = \frac{r_1 x_1 + r_2 x_2 + r_3 x_3 + r_0}{r_4 x_1 + r_5 x_2 + r_6 x_3} \tag{17}$$

As only the horizontal error in the X direction is considered, it is assumed that there are no uncertainties associated with y_i and $z_i (i = 1,2,3)$. The point (x_p, y_p) is an interpolation location and does not have any uncertainty. Therefore, for a given triangle, $r_i (i = 4,5,6)$ are constants across the whole triangle while $r_i (i = 0,1,2,3)$ are location-dependent constants (i.e. they depend on the location of the point P).

Now the propagation of measurement errors in the X direction has been converted into a mathematical problem. Based on Equation (17), the original error propagation problem can be stated mathematically. It is known that $r_i (i = 1:7)$ are constants and $x_i (i = 1,2,3)$ are random variables with uncertainty. If the variances of $x_i (i = 1,2,3)$ and their covariances are known, what is the variance of z_p?

It is not easy to calculate the partial differences for x_1, x_2 and x_3 in Equation (17). To avoid this, Equation (17) is further simplified as Equation (18).

8

$$z_p = \frac{n_1}{n_2} \quad : \boldsymbol{F}(n_1, n_2) \tag{18}$$

in which,

$$n_1 = r_1 x_1 + r_2 x_2 + r_3 x_3 + r_0 \quad : f_1(x_1, x_2, x_3)$$
$$n_2 = r_4 x_1 + r_5 x_2 + r_6 x_3 \quad\quad : f_2(x_1, x_2, x_3) \tag{19}$$

n_1 and n_2 are two new input variables and z_p is the output variable. This simplification will make the partial differences easier to calculate, but introduces two new variables that are correlated with each other, owing to the presence of common variables in the expressions for n_1 and n_2. This should not be a problem since the covariance of n_1 and n_2 can be calculated using Equation (6).

Applying Equation (6) for this case:

$$\begin{bmatrix} \sigma_{n_1}^2 & \sigma_{n_1 n_2} \\ \sigma_{n_2 n_1} & \sigma_{n_2}^2 \end{bmatrix} = \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1} & \dfrac{\partial f_1}{\partial x_2} & \dfrac{\partial f_1}{\partial x_3} \\ \dfrac{\partial f_2}{\partial x_1} & \dfrac{\partial f_2}{\partial x_2} & \dfrac{\partial f_2}{\partial x_3} \end{bmatrix} \begin{bmatrix} \sigma_{x_1}^2 & \sigma_{x_1 x_2} & \sigma_{x_1 x_3} \\ \sigma_{x_2 x_1} & \sigma_{x_2}^2 & \sigma_{x_2 x_3} \\ \sigma_{x_3 x_1} & \sigma_{x_3 x_2} & \sigma_{x_3}^2 \end{bmatrix} \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1} & \dfrac{\partial f_2}{\partial x_1} \\ \dfrac{\partial f_1}{\partial x_2} & \dfrac{\partial f_2}{\partial x_2} \\ \dfrac{\partial f_1}{\partial x_3} & \dfrac{\partial f_2}{\partial x_3} \end{bmatrix} \tag{20}$$

Equation (20) is a generic solution, which can account for any given error variances (e.g. different error variances: $\sigma_{x_1}^2 \neq \sigma_{x_2}^2 \neq \sigma_{x_3}^2$) and correlations (i.e. $\sigma_{x_i x_j}$ $(i, j = 1,2,3 \ \& \ i \neq j) \neq 0$) in the error propagation. However, in the following discussion, the special case of zero covariance and equal variances ($\sigma_{x_1}^2 = \sigma_{x_2}^2 = \sigma_{x_3}^2 = \sigma_x^2$) is considered in detail. There are cases where covariances are assumed to be zero, either because the error in source data is believed to be independent, or because no information on covariances is available (Kyriakidis and Goodchild, 2006). Some researchers (Li, 1993; Huang, 2000; Aguilar *et al.*, 2006) have adopted this assumption (independent and identical) for propagating error variances of node points into DEM. In some cases, error in data points acquired by some instruments may be independent. For example, Hodge (2010) investigated random positional error of TLS data points.

The partial differences with respect to $x_i (i = 1,2,3)$ are:

$$\frac{\partial f_1}{\partial x_1} = r_1 \quad \frac{\partial f_1}{\partial x_2} = r_2 \quad \frac{\partial f_1}{\partial x_3} = r_3$$
$$\frac{\partial f_2}{\partial x_1} = r_4 \quad \frac{\partial f_2}{\partial x_2} = r_5 \quad \frac{\partial f_2}{\partial x_3} = r_6 \tag{21}$$

Substituting these partial differences into Equation (20):

$$\begin{bmatrix} \sigma_{n_1}^2 & \sigma_{n_1 n_2} \\ \sigma_{n_2 n_1} & \sigma_{n_2}^2 \end{bmatrix} = \begin{bmatrix} r_1 & r_2 & r_3 \\ r_4 & r_5 & r_6 \end{bmatrix} \begin{bmatrix} \sigma_x^2 & 0 & 0 \\ 0 & \sigma_x^2 & 0 \\ 0 & 0 & \sigma_x^2 \end{bmatrix} \begin{bmatrix} r_1 & r_4 \\ r_2 & r_5 \\ r_3 & r_6 \end{bmatrix} \tag{22}$$

$$= \begin{bmatrix} (r_1^2 + r_2^2 + r_3^2)\sigma_x^2 & (r_1\,r_4 + r_2 r_5 + r_3\,r_6)\sigma_x^2 \\ (r_1\,r_4 + r_2 r_5 + r_3\,r_6)\sigma_x^2 & (r_4^2 + r_5^2 + r_6^2)\sigma_x^2 \end{bmatrix}$$

The variances and covariance of n_1 and n_2 are now obtained.

$$\sigma_{n_1}^2 = (r_1^2 + r_2^2 + r_3^2)\sigma_x^2$$
$$\sigma_{n_2}^2 = (r_4^2 + r_5^2 + r_6^2)\sigma_x^2 \tag{23}$$
$$cov(n_1, n_2) = \sigma_{n_1 n_2} \; or \; \sigma_{n_2 n_1} = (r_1\,r_4 + r_2 r_5 + r_3\,r_6)\sigma_x^2$$

Recall that the problem to be solved is Equation (18). This is relatively simple since the variances and covariance of the input variables n_1 and n_2 are already known.

Now, applying Equation (6) to Equation (18) (there are two input variables and only one output variable) to obtain:

$$\sigma_{z_p}^2 = \left(\frac{\partial F}{\partial n_1}\right)^2 \sigma_{n_1}^2 + \left(\frac{\partial F}{\partial n_2}\right)^2 \sigma_{n_2}^2 + 2\frac{\partial F}{\partial n_1}\frac{\partial F}{\partial n_2} cov(n_1, n_2) \tag{24}$$

Calculating the partial differences:

$$\frac{\partial F}{\partial n_1} = \frac{1}{n_2}; \qquad \frac{\partial F}{\partial n_2} = -\frac{n_1}{n_2^2}$$

Substituting the partial differences in Equation (24) and, rearranging, to have:

$$\sigma_{z_p}^2 = \frac{n_1^2}{n_2^2}\left(\left(\frac{\sigma_{n_1}}{n_1}\right)^2 + \left(\frac{\sigma_{n_2}}{n_2}\right)^2 - \frac{2cov(n_1, n_2)}{n_1 n_2}\right) \tag{25}$$

Substituting the variances and covariances (in Equation (23)) into Equation (25):

$$\begin{aligned} \sigma_{z_p}^2 &= \frac{n_1^2}{n_2^2}\left(\frac{(r_1^2 + r_2^2 + r_3^2)\sigma_x^2}{n_1^2} + \frac{(r_4^2 + r_5^2 + r_6^2)\sigma_x^2}{n_2^2}\right. \\ &\quad \left. - \frac{2(r_1\,r_4 + r_2 r_5 + r_3\,r_6)\sigma_x^2}{n_1 n_2}\right) \\ &= \frac{n_1^2}{n_2^2}\underbrace{\left(\left(\frac{r_1}{n_1} - \frac{r_4}{n_2}\right)^2 + \left(\frac{r_2}{n_1} - \frac{r_5}{n_2}\right)^2 + \left(\frac{r_3}{n_1} - \frac{r_6}{n_2}\right)^2\right)}_{M_x}\sigma_x^2 \end{aligned} \tag{26}$$

Based on the assumption that the variances of the node points in the X direction are identical and independent of each other (although any correlation can be included in Equation (20)), the propagated error variance at the point P can be expressed in a similar way as for Equation (13). By considering several triangular surfaces, it is found that $M_x = M(\tan \alpha_x)^2$. This is exactly the same as the solution (Equation (15)) derived from the geometric relation.

$$\sigma_{z_p}^2 = M_x \sigma_{x_{node}}^2 = M(\tan \alpha_x)^2 \sigma_{x_{node}}^2 = M(\tan \alpha_x \, \sigma_{x_{node}})^2 \tag{27}$$

5. The parameter *M*

It is known from Section 4.1 that the propagation of error variance depends on the parameters m_i ($i = 1,2,3$), which represent the influence of the error variances at each node point on the propagated error variance. For independent and identical error variances at node points, a location-dependent parameter M ($= m_1{}^2 + m_2{}^2 + m_3{}^2$) was derived, which combines the effects of error variances at three node points. These parameters are further explored in this section. To understand how the propagated error variance changes with location, a single triangular surface is considered. Only the propagation of vertical error variance at node points is considered since the same M is used for the propagation of horizontal error variance.

Suppose that three data points being used to build a triangular surface and their coordinates are (1,1,1), (3,1,3) and (4,4,4) (Figure 3(a)). The units are not given because it is not relevant to the analysis. The relative location of a particular point with respect to node points is of interest. About 10000 points, regularly distributed on the triangle surface, were considered in this analysis. The method introduced in Sections (4.1) was used to propagate the error variances at the nodes into points at different locations within the triangle. A single node point (i.e. (1,1,1)) was considered first. Figure 3(b) shows that the value of m_1 was maximal (=1) at this node point, and gradually decreased to zero with a pattern parallel to the opposite side of the triangle. It also shows the dependence of the weight m_1 on the prediction location. The actual error variance propagated into a particular point was a combination of the error variances at three node points. In the case of independent and different error variances at nodes, Equation (12) can be used. For a one-unit error variance at two nodes and a two-unit error variance at the third, the propagated error variance is shown in Figure 3(c), and is bounded by the maximum error variance at the node points (consistent with Hu *et al.*, 2009).

In the case of independent and identical error variance, the parameter M was derived. The value of M is shown in Figure 3(d). The analysis showed that the minimum value of M was about 1/3 and its maximum value was about 1. The value of M reached a maximum when the point of interest was located at a node point, and was minimal when the point was at the geometric centre of the triangle. It was thought that 10000 points can reasonably represent the whole triangular surface. On this basis, the mean value of M for the whole triangular surface was calculated as 1/2, which is consistent with the value given by Zhu *et al.* (2005). This suggested that the mean error variance ($0.5\sigma_{z_{node}}^2$) of a triangle surface could also be derived through the simple experimental method presented here. The average error variance is likely to underestimate or overestimate the error of a particular point. For a simple estimation, it is more conservative to assume that the propagated elevation error at any point in a TIN model is the same as that of the node points, namely $M = 1$. Alternatively, the procedure introduced in Section 4 should be used to find the exact value of M for each point of interest.

6. Validation

6.1 Monte Carlo simulation

Monte Carlo simulations were carried out to validate the results given in Sections 4 and 5. Only the cases of independent and identical error variance (represented by Equations (13) and (27)) were simulated. The 10000 points considered in Section 5, regularly spaced over the triangular surface, were used. In the Monte Carlo simulation, a realization of propagated errors at 10000 check points was created, by adding at each node (of a triangular surface) a random error drawn from a normal distribution. This realization was stored and this process

was repeated 10000 times. For each check point, an error distribution (10000 realizations) was formed, which was then used to calculate error variances at each check point. These were carried out separately for the propagation of vertical and horizontal error variances.

It was found that the distribution of M was the same as that shown in Figure 3(d), with minimum, average and maximum values of about 1/3, 1/2 and 1, respectively. These results validated Equation (13). The simulation results also showed that the propagated horizontal error variance was almost the same as that predicted by Equation (27).

6.2 Experiments

Two experiments were carried out to further support the analytical solutions. In these experiments, the vertical and horizontal errors in the actual survey data, obtained using a terrestrial laser scanner, were considered separately. The scanner used was a Leica ScanStation C10. The data processing and visualisation were carried out in MATLAB® and ArcMap®.

A clarification concerning the use of the survey data obtained is made here. The Leica ScanStation C10 is a short-range scanner with a high positional accuracy of a few millimetres within its useful scan range. In addition, the survey conditions were favourable (i.e. a short scan distance, sound reflectance and smooth surface of the object scanned). Therefore, the survey data obtained had an accuracy of better than 2 mm at 1σ (detailed in the following sub-sections). The survey data obtained in other cases (e.g. a long-range scan over 1 km) or from other devices (e.g. airborne LiDAR and DGPS) may have much larger positional errors. For example, the positional accuracy of individual data points obtained using long-range terrestrial scanners (over 1km) is in the magnitude of several centimetres, varying with the actual scanning distance (Leica HDS8810 and RIEGEL VZ-6000 datasheets). The accuracy of DGPS measurements is usually about 10 - 20 mm. In this study, the ScanStation C10 was used, partly because the instrument was available.

6.2.1 Experiment 1: propagation of vertical error variance

A thick wooden plate was placed over an area of flat ground and levelled to avoid the effect of the horizontal measurement error on error propagation. The scanner was used to scan the plate. The data points (1237 points) acquired are shown in Figure 4(a). The top surface of the plate was taken as the datum-plane, that is zero elevation. The distribution of the elevations (i.e. vertical error) of the data points is shown in Figure 4(b), with one standard deviation of 1.48 mm and variance of 2.19 mm². It approximately follows the fitted normal distribution. This suggests that the vertical errors in these data points were likely to be random.

The TLS data points were used to generate a TIN (Figure 5(a)), which was then used to derive a fine resolution raster (2 mm) by linear interpolation. Such a fine resolution was used to show the spatial pattern of propagated error in this study. The DEMs created for practical use would not have so fine a resolution. There was no modelling error because the plate scanned was a plane. Also, as the plate was levelled, the horizontal measurement error did not affect the elevation in the DEM. Thus, the raster (Figure 5(b)) was a representation of propagated vertical measurement error.

Figure 6 shows the spatial characteristics of the propagated vertical error. To better visualise the data, only a sub-area is shown in detail. The propagated vertical error varied and tended to be larger around node points. This suggests that for a grid-DEM a grid point closer to node points had a larger propagated measurement error variance. The mean error and variance of propagated error at all grid points were 0 and 1.14 mm², respectively. Given that the variance of node points was 2.19 mm², the average value of $M = 1.14/2.19 = 0.52$,

almost the same as that (0.5) derived in Section 5. It was also known from Section 5 that the value of *M* was 1/3 at the geometric centres of triangular surfaces. The propagated vertical errors at all centroids (black dots in Figure 6) were extracted from the raster. The distribution of these is shown in Figure 7, with the standard deviation of 0.89 mm. The actual error variance and that derived using the analytical method are compared in Table 1. The small difference in 1σ (between the analytical method and the experiment) was probably a result of the limited number of data points used.

6.2.2 Experiment 2: propagation of horizontal error variance

The horizontal measurement error will cause an elevation error when the terrain surface surveyed is inclined. To investigate this, a plate with an inclination angle should be scanned. In this case, however, it is impossible to obtain the horizontal error since it cannot be separated from the vertical error. A different approach was adopted in this experiment (Figure 8(a)). A plate was held vertically at about the same height as the scanner. The *Y* axis of the laser scanner was set up by orientating the scanner to the centre of the plate. As the scanner was levelled, the vertical direction was the *Z* axis. The plate was then scanned head-on. As the plate was vertical, the vertical measurement error would not cause data points to move off the plate surface in the *Y* direction. A plane was fitted to all data points. The distances (residuals) from the data points to the fitting plane represent the horizontal measurement error in the *Y* direction.

In the next step, the best fitting plane was artificially rotated to an angle of 30° with the *Y* axis in MATLAB®. The residuals (i.e. the horizontal errors) were then added back to the projected locations of the data points on the fitting plane. This process generated a new data set, which had error in the *Y* direction only. This new data set was used to create a TIN model, which was then interpolated into a raster (Figure 8(b)). Subsequent data processing was the same as for Experiment 1. It was found that $M_y = M(\tan \alpha_y)^2$. The minimum, mean and maximum values of *M* were 1/3, 1/2 and 1. Some results are shown in Figure 8 and Table 2. The interpretation of these results is not repeated here as it is similar to that for Experiment 1.

The results from Monte Carlo simulations and survey experiments were the same as those derived from the analytical solutions. Therefore, the analytical method proposed for vertical and horizontal error propagation was validated.

7. Combination of vertical and horizontal error variance

In previous sections, the propagation of vertical and horizontal error variances from node points into a TIN model was studied separately. In this section, the combination of these effects is discussed.

Based on the discussion in Sections 4.1 and 4.2, the propagated error variances on a TIN model are:

Effect of vertical error variance: $\qquad\qquad\qquad\qquad M\sigma^2_{z_{\text{node}}}$

Effect of horizontal error variance in *X* direction: $\qquad M(\tan \alpha_x)^2 \sigma^2_{x_{\text{node}}}$

Effect of horizontal error variance in *Y* direction: $\qquad M(\tan \alpha_y)^2 \sigma^2_{y_{\text{node}}}$

These results are based on identical and independent vertical or horizontal error variances at three node points. General cases involving correlation between node points can also be considered (using Equations (11) and (20)), but are not detailed here. However, it is important to appreciate that the errors in the X, Y and Z coordinates of a single data point are likely to be correlated. This is because the coordinates (X, Y, Z) of an individual measurement point are derived from common variables, i.e. distance and angle.

There are two different ways of combining the effects of vertical and horizontal source data errors on a TIN model, depending on how these errors are assessed. An empirical approach is to compare the elevation data obtained with benchmark elevations of higher accuracy at check points. The benchmark elevations are usually acquired with another instrument of higher accuracy. A distribution of the elevation differences at these check points can then be established, which is a measure of the elevation errors at the data points. Such errors have already combined the effects of vertical and horizontal measurement errors. In this case, these errors can be regarded as "vertical" errors, and propagated into a TIN model using the propagation of vertical error variance method presented in Section 4.1. This approach avoids the consideration of the correlation between X, Y and Z coordinates of a single data point.

Another possibility is that the positional (vertical and horizontal) errors of measurements are already known from a technical specification, theoretical calculation or control experiments. In this case, the propagation of horizontal and vertical source data errors is required. The correlation between X, Y and Z coordinates of a single data point needs to be assessed. The actual correlation may vary from one case to the other, depending on how a data point is derived, and may not be zero. For example, the X, Y and Z coordinates of a data point obtained from a terrestrial laser scanner are calculated from the range and angle measurements. Therefore, these coordinates are correlated. If the error variances are assumed to be independent of each other (unlikely to be the case in practice), the error variance at a particular point P can be expressed as:

$$\sigma_{z_p}^2 = M \left(\sigma_{z_{\text{node}}}^2 + (\tan \alpha_x)^2 \sigma_{x_{\text{node}}}^2 + \left(\tan \alpha_y \right)^2 \sigma_{y_{\text{node}}}^2 \right) \tag{28}$$

$$(1/3 \le M \le 1)$$

If the correlation is known or can be assessed, Equation (28) should be modified to include the influence of the correlation. A detailed discussion of this is outside the scope of this article as the nature of the correlation depends on the techniques used to acquire data points. However, it is possible to determine the correlation between X, Y and Z coordinates if the algorithm used to derive these coordinates is known. Alternatively, the error propagation can be based on the worst case scenario, which depends on the significance of the influence of the error variance in each direction.

8. Discussion

Error propagation reflects the effect of uncertainties in input variables on the uncertainty of an output variable based on some functions of the inputs. The well-known error propagation method has been used to derive the analytical solution. The weight m_i represents the influence of the error variance of an individual node point on the propagated error variance at a particular point. Its value gradually decreases from one at the ith node point to zero on the opposite side of the triangle surface. In the case of identical and independent error variances at nodes, Equations (13) and (27) were derived to calculate the propagated error variance at

any particular points. It was found that the propagated error variance of a particular point varied with its relative location to the node points enclosing it. This variation depended on the parameter M introduced, which was essentially the location-dependent weighting for error variances at node points. The value of M for a point other than the geometric centre and nodes lay between 1/3 and 1. A greater propagated measurement error variance resulted when the point was closer to either of the three node points. This is consistent with the classical statistical theory, in which the propagation of measurement error at all points is a weighted average of the errors at node points.

Some GIS operations (e.g. a comparison of multi-temporal DEMs) require interpolation at particular points. In such cases, the uncertainty of a particular point should be considered. The method introduced in this study provides a means of assessing the elevation error of any point propagated from node points. On this basis, the elevation error variances at the grid points of a raster DEM can be evaluated. The analysis in this study suggests that the average error variance is likely to be an under- or over-estimate of the actual error variance at a particular point. If an error propagation analysis is not carried out, it is more conservative to assume that the error variance in a TIN model is the maximum error variance of node points.

The horizontal errors at node points were studied. Their effects on the elevation at a particular point were found to depend on the slope angles with respect to the X and Y axes. The larger the slope angles, the greater the propagated error variances. When the slope angle is greater than 45°, the effect of the horizontal errors of source data points on the predicted elevation of a particular point P is greater than that of a vertical error of the same magnitude. The slope angle can either be determined by a slope map or from triangular surfaces directly.

In cases where data points are extremely dense and the point spacing approaches the horizontal error of data points, the interpolation point is relatively likely to lie within a different triangular facet. Because of discontinuities in slope along the sides of these facets, error in the interpolated height could be greater than if the interpolation point occurs within the same facet. As the horizontal error in real survey data points is usually much smaller than the point spacing, this is considered a minor issue in practice, and is beyond the scope of this article.

Although a special case of equal error variances and zero covariance is discussed in detail, a generic solution (Equation (11)) has been derived for propagation of vertical error variance. Also, a generic solution for propagation of horizontal error variance can readily be derived from Equation (20) (simply following the same procedure as Equations (22)-(25)), by using appropriately different error variances and covariances between node points. Therefore, the solutions provided in this article can be used to assess the propagated source data errors at a particular point when the uncertainties at node points are different and dependent.

Two approaches to combine the effects of vertical and horizontal measurement errors were presented. These errors can be combined as an elevation error through empirical assessment at check points. If these elevation errors are regarded as a representative of the measurement error at node points, only the method proposed for propagation of vertical error is required. If the horizontal and vertical errors at node points can be evaluated separately or are known from specifications, an understanding of their correlation is required.

The results of the survey experiments carried out in this study show that the method is of practical application to topographical survey data. If the error variances at data points are known, the error propagation can be implemented in the following main steps: (1) generate a TIN with a reference number assigned to each triangle formed in the triangulation process; (2) identify points of interest within the TIN model, that is the particular points described in earlier sections. These can be the grid points of a raster-based DEM; (3) match each point of interest to the relevant triangle which encloses the point (there are several ways (e.g.

Hormann and Agathos, 2001) to determine whether a point is within a triangle); and (4) apply the analytical solution.

This article, as its title suggests, addresses the propagation of error variances from source data points; it does not address interpolation uncertainty. However, the accuracy of a DEM also depends on interpolation uncertainty, which reflects how well the interpolated elevation model approximates the real physiography (Shortridge, 2001; Fisher and Tate, 2006). Several researchers (e.g. Li, 1993; Huang, 2000; Bruin, 2008; Erdogan 2009; Bater and Coops, 2009; Hu *et al.*, 2009) have investigated the uncertainty associated with interpolation algorithms for approximating features of various dimension. For example, using approximation theory from computational science, Hu *et al.* (2009) derived a solution for estimating the interpolation error associated with a TIN with linear interpolation. In combining propagated error from source data points with interpolation error, it is commonly assumed that they are independent (e.g. Li, 1993; Huang, 2000; Aguilar *et al.*, 2006).

9. Conclusion

In this article, an analytical solution was derived to calculate the error variance of any point within a linearly interpolated TIN model due to vertical and horizontal error variances in the source data (node points). The solution was validated against real survey data obtained using a terrestrial laser scanner. The propagated error variance depended on its relative location with respect to the relevant three node points that form a triangular surface. For the error variance of a single node point, its effect was maximal at this node point and gradually decreased to zero with a unique pattern (parallel to the opposite side of the triangle surface).

The error propagation solutions in Equations (13) and (27) were based on identical and independent error variances at node points. A location-dependent parameter M was derived. The maximum, average and minimum values of M were 1 (when the point was at one of node points), 1/2 and 1/3 (when the point was at the geometric centre of a triangle). The average error variance ($0.5\sigma_{Z_{node}}^2$) overestimated or underestimated the actual propagated error variance at a particular point (except for the case where the error variance at some points was just half that at node points). The propagated error variances were greater when the particular point was closer to one of the node points. This was consistent with the expectation using classical statistics. More general cases can be considered, in which the error variances at node points are different and dependent (see discussion in Section 8).

The effect of horizontal error at node points depended on the actual slope angles with respect to the X and Y axes. The propagation of horizontal error variance can be carried out in a similar way to the vertical error variance.

References

Aguilar, F.J., Aguera, F., Agullar, M. and Carvajal F., 2005. Effects of terrain morphology, sampling density, and interpolation methods on grid DEM accuracy. Photogrammetric Engineering & Remote Sensing, 71(7), 805–816.

Aguilar, F.J., Aguilar, M.A., Agüera, F. and Sánchez, J., 2006. The accuracy of grid digital elevation models linearly constructed from scattered sample data. International journal of geographical information science, 20(2), 169-192.

American Society for Photogrammetry and Remote Sensing, ASPRS Guidelines: Vertical Accuracy Reporting for Lidar Data [online], edited by Flood, M. (2004); Draft ASPRS Guidelines: Horizontal Accuracy Reporting for Lidar Data [online] http://www.asprs.org/a/society/committees/standards/Vertical_Accuracy_Reporting_for_Lidar_Data.pdf [Accessed 16 Sep 2013]; http://www.asprs.org/a/society/committees/standards/Horizontal_Accuracy_Reporting_for_Lidar_Data.pdf [Accessed 16 Sep 2013].

Bater C.W. and Coops N.C., 2009. Evaluating error associated with lidar-derived DEM interpolation. Computers and Geosciences, 35, 289-300.

Bruin, S.D., 2008. Modelling positional uncertainty of line features by accounting for stochastic deviations from straight line segments. Transactions in GIS, 12(2), 165-177.

Brunori, C.A., Civico, R., Cinti F.R. and Ventura G., 2013. Characterization of active fault scarps from LiDAR data: a case study from Central Apennines (Italy). International Journal of Geographical Information Science, 27(7), 1405-1416.

Chen Y.M. and Zhou Q.M., 2013. A scale-adaptive DEM for multi-scale terrain analysis, International Journal of Geographical Information Science, 27(7), 1329-1348.

Cox M.G. and Harris P.M., 2006. Software support for metrology best practice guide no. 6: uncertainty evaluation. National Physical Laboratory UK.

Coveney S., 2013. Association of elevation error with surface type, vegetation class and data origin in discrete-returns airborne LiDAR. International Journal of Geographical Information Science, 27(3), 467-483.

Day, S.S., Gran, K.B., Belmont, P. and Wawrzyniec, T., 2013. Measuring bluff erosion part 1: terrestrial laser scanning methods for change detection. Earth Surf. Process. Landforms, 38, 1055–1067.

Erdogan, S., 2009. A comparison of interpolation methods for producing digital elevation models at the field scale. Earth Surface Process and Landforms, 34, 366–376.

Fisher P.F. and Tate N.J., 2006. Causes and consequences of error in digital elevation models. Progress in physical geography, 30(4), 467-489.

Fonstad, M. A., Dietrich, J. T., Courville, B. C., Jensen, J. L. and Carbonneau, P. E., 2013. Topographic structure from motion: a new development in photogrammetric measurement. Earth Surf. Process. Landforms, 38, 421–430.

Gallay, M., Lloyd, C.D., McKinley, J. and Barry, L., 2013. Assessing modern ground survey methods and airborne laser scanning for digital terrain modelling: A case study from the Lake District, England. Computers & Geosciences, 51, 216-227.

Gessesse, G.D., Fuchs, H., Mansberger, R., Klik, A. and Rieke-Zapp, D.H., 2010. Assessment of erosion, deposition and rill development on irregular soil surfaces using close range digital photogrammetry. The Photogrammetric Record, 25, 299–318.

Goodchild M.F., 1999. Measurement-based GIS. In: Shi W.Z. et al. (eds) Proceedings of the International Symposium on Spatial Data Quality '99. Hong Kong Polytechnic University, Hong Kong, pp. 1–9.

Heritage, G. L., Milan, D. J., Large, A. R. G. and Fuller, I. C., 2009. Influence of survey strategy and interpolation model on DEM quality. Geomorphology, 112(3-4), 334-344.

Heuvelink G.B.M., 1998. Error propagation in environmental modelling with GIS. Taylor & Francis, London.

Hormann, K. and Agathos, A., 2001. The point in polygon problem for arbitrary polygons. Computational Geometry: Theory and Applications, 20(3), 131 - 144 .

Hodge, R.A., 2010. Using simulated terrestrial laser scanning to analyse errors in high-resolution scan data of irregular surfaces. ISPRS Journal of Photogrammetry and Remote Sensing 65(2), 227-240.

Hodgson, M.E. and Bresnahan, P., 2004. Accuracy of airborne LiDAR-derived elevation: empirical assessment and error budget. Photogrammetric Engineering & Remote Sensing, 70, 331–339.

Hu, P., Liu X.H., Hu, H., 2009. Accuracy assessment of digital elevation models based on approximation theory. Photogrammetric Engineering & Remote Sensing, 75(1), 49–56.

Huang, Y.D., 2000. Evaluation of information loss in digital elevation models with digital photogrammetric systems. Photogrammetric Record, 16(95), 781-791.

Kyriakidis P.C., Shortridge, A.M., and Goodchild M.F., 1999. Geostatistics for conflation and accuracy assessment of digital elevation models. International Journal of Geographical Information System, 13(7), 677-707.

Kyriakidis P.C. and Goodchild M.F., 2006. On the prediction error variance of three common spatial interpolation schemes. International Journal of Geographical Information System, 20(8), 823-855.

Leica HDS8810 datasheet [online]: http://www.leica-geosystems.co.uk/downloads123/zz/hds/HDS8800/brochures/HDS8800_BRO_en.pdf [Accessed 23 Sep 2013].

Leung, Y., Ma, J.H. and Goodchild, M.F., 2004a. A general framework for error analysis in measurement-based GIS Part 1: The basic measurement-error model and related concepts. Journal of Geographical Systems, 6, 325-354.

Leung, Y., Ma, J.H. and Goodchild, M.F., 2004b. A general framework for error analysis in measurement-based GIS Part 3: Error analysis in intersections and overlays. Journal of Geographical Systems, 6, 381-402.

Li, Z.L., 1993. Mathematical models of the accuracy of digital terrain model surfaces linearly constructed from square gridded data. Photogrammetric Record, 14(82), 661-674.

Lloyd C.D. and Atkinson P.M., 2002. Non-stationary approaches for mapping terrain and assessing prediction uncertainty. Transactions in GIS, 6(1) 17-30.

Lloyd C.D. and Atkinson P.M., 2006. Deriving ground surface digital elevation models from LiDAR data with geostatistics. International Journal of Geographical Information Science, 20(5), 535-563.

Spaete L.P., Glenn N.F., Derryberry D.R., Sankey T.T., Mitchell J.J. and Hardegree S.P., 2011. Vegetation and slope effects on accuracy of a LiDAR-derived DEM in the sagebrush steppe. Remote Sensing Letters, 2(4), 317-326.

Negishi, J. N., Sagawa, S., Sanada, S., Kume, M., Ohmori, T., Miyashita, T. and Kayaba, Y., 2012. Using airborne scanning laser altimetry (LiDAR) to estimate surface connectivity of floodplain water bodies. River Res. Applic., 28, 258–267.

Oksanen J. and Sarjakoski T., 2005. Error propagation of DEM-based surface derivatives. Computers and Geosciences, 31, 1015-1027.

RIEGEL VZ-6000 datasheet [online]: http://www.riegl.com/uploads/tx_pxpriegldownloads/DataSheet_VZ-6000_23-09-2013_PRELIMINARY.pdf [Accessed 23 Sep 2013].

Schürch, P., Densmore, A. L., Rosser, N. J., Lim, M. and McArdell, B. W., 2011. Detection of surface change in complex topography using terrestrial laser scanning: application to the Illgraben debris-flow channel. Earth Surf. Process. Landforms, 36, 1847–1859.

Shi, W. Z., Li, Q. Q. and Zhu, C. Q., 2005. Estimating the propagation error of DEM from higher-order interpolation algorithms. International Journal of Remote Sensing, 26(14), 3069-3084.

Shortridge, A., 2001. Characterizing uncertainty in digital elevation models. In Hunsaker, C.T., Goodchild, M.F., Friedl, M.F. and Case, T.J., editors, Spatial uncertainty in ecology: implications for remote sensing and GIS applications, New York: Springer, 238–257.

Su, J.G. and Bork, E.W., 2006. Influence of vegetation, slope and LiDAR sampling angle on DEM accuracy. Photogrammetric Engineering and Remote Sensing, 72, 1265–1274.

Taylor, J.R., 1982. An introduction to error analysis: the study of uncertainties in physical measurements (1st edition).

Webster, T.L., Forbes, D.L., MacKinnon, E. and Roberts, D., 2006. Flood-risk mapping for storm-surge events and sea-level rise using LIDAR for southeast New Brunswick. Canadian Journal of Remote Sensing, 32 (2), 194–211.

Wechsler S.P., and Kroll C.N., 2006. Quantifying DEM uncertainty and its effect on topographic parameters. Photogrammetric Engineering & Remote Sensing, 72(9), 1081-1090.

Wheaton, J. M., Brasington, J., Darby, S. E. and Sear, D. A., 2010. Accounting for uncertainty in DEMs from repeat topographic surveys: improved sediment budgets. Earth Surf. Process. Landforms, 35, 136–156.

Zhu, C., Shi, W., Li, Q., Wang, G., Cheung, T. C. K., Dai, E., and Shea, G. Y. K., 2005. Estimation of average DEM accuracy under linear interpolation considering random errors at the nodes of tin model. International Journal of Remote Sensing, 26(24), 5509-5523.

1σ of vertical error at all node points ($1\sigma_{z_{node}}$):	1.48 mm (Figure 4(b))
1σ of elevation error at triangle centroids ($1\sigma_{cen}$) derived using the analytical solution (Equation (13)):	$\sqrt{M\sigma_{z_{node}}^2} = \sqrt{\frac{1}{3}1.48^2} = 0.86$ mm ($M = 1/3$ for triangle centroids)
1σ of elevation error at all triangle geometric centres ($1\sigma'_{cen}$) in Experiment 1:	0.89 mm (Figure 7)
The percentage that the values at triangle centroids in Experiment 1 lie between confidence intervals:	63.9% $[-1\sqrt{M}\sigma_{z_{node}} , 1\sqrt{M}\sigma_{z_{node}}]$; 95.4% $[-2\sqrt{M}\sigma_{z_{node}} , 2\sqrt{M}\sigma_{z_{node}}]$

Table 1: Propagated vertical error variance (standard deviation) at geometric centres

1σ of horizontal (Y axis) error at all node points ($1\sigma_y$):	1.76 mm (Figure 8(c))
1σ of elevation error at triangle centroids ($1\sigma_{cen}$) derived using the analytical solution (Equation (27)):	$= \sqrt{M_y \sigma_y^2} = \sqrt{M(\tan\alpha_y\, \sigma_{y_{node}})^2} =$ $\sqrt{\frac{1}{3} \times (\tan 30° \times 1.76)^2} = 0.59$ mm
1σ of elevation error at all triangle geometric centres ($1\sigma'_{cen}$) in Experiment 2:	0.6 mm (Figure 8(d))
The percentage that the values at triangle centroids in Experiment 2 lie between confidence intervals:	67.4% [$-1\sqrt{M_y}\sigma_y$, $1\sqrt{M_y}\sigma_y$]; 95.1% [$-2\sqrt{M_y}\sigma_y$, $2\sqrt{M_y}\sigma_y$]

Table 2: Propagated horizontal error variance (standard deviation) at geometric centres
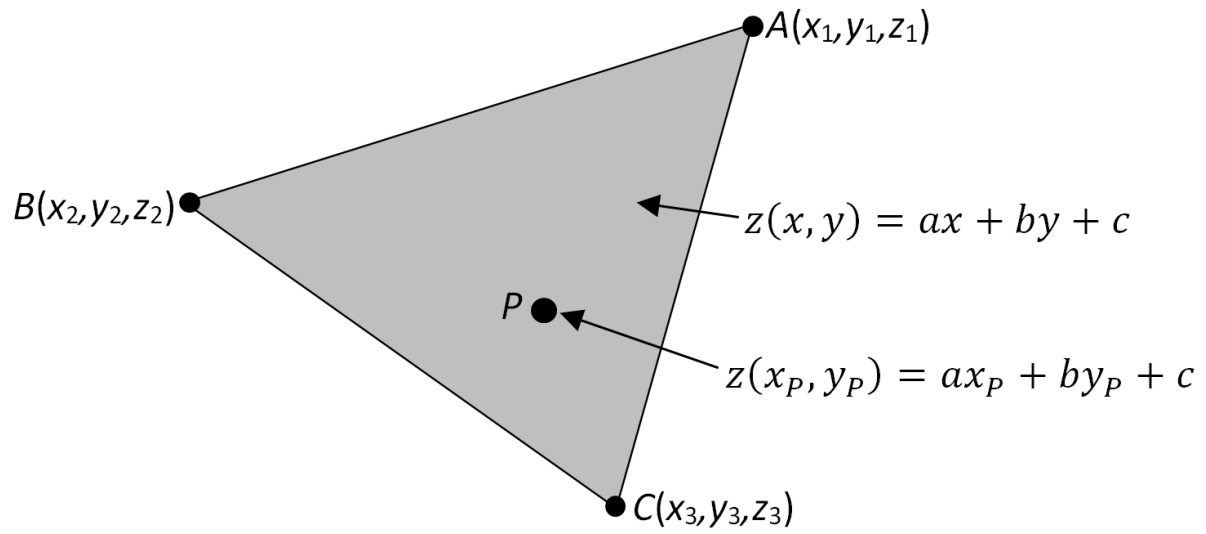
Figure 1: TIN model with linear interpolation: survey data points A, B and C are three vertices of a triangle in a TIN model and their coordinates are (x_1, y_1, z_1), (x_2, y_2, z_2) and (x_3, y_3, z_3), respectively; Point P is located anywhere within the triangle and its elevation can be calculated by the linear function given in Equation 1.
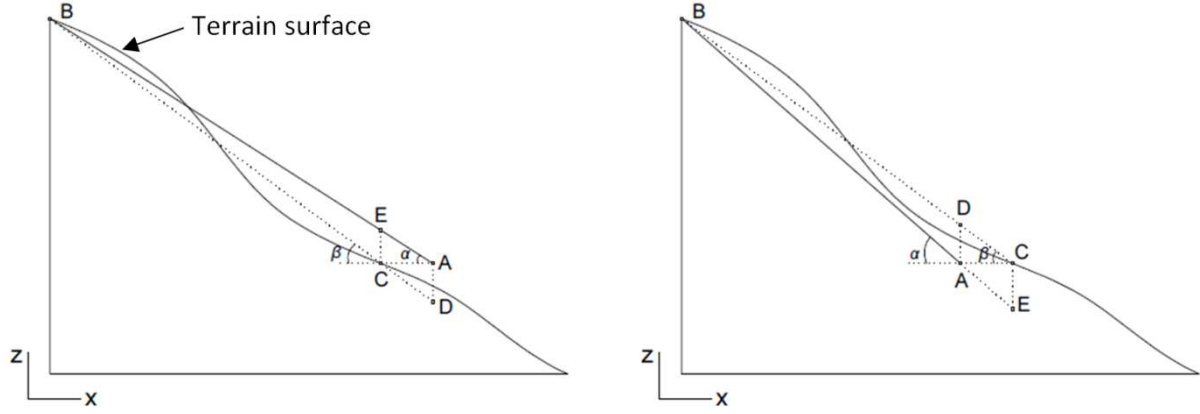
Figure 2: Elevation error caused by horizontal error of source data points: Points *B* and *C* are two terrain surface points to be measured. Assume the measurement of point *B* is free of error, and the measurement of point *C* has a horizontal error represented by dotted line *CA*; Line *BC* represents the terrain surface without any measurement error (note that modelling error is not considered in this study). Points *B* and *A* are actual survey points acquired. Line *BA* is the actual representation of the terrain surface due to the horizontal measurement error from point *C*. The rise or fall of point *E* with respect to point *C* (the imaginary true terrain point) is equivalent to the rise or fall of point *A* (the actual survey point) with respect to point *D*, causing an elevation error (i.e. the elevation difference between line *BE* and *BC*). α is the slope angle of line *BA* and β is the angle of inclination of the terrain surface.
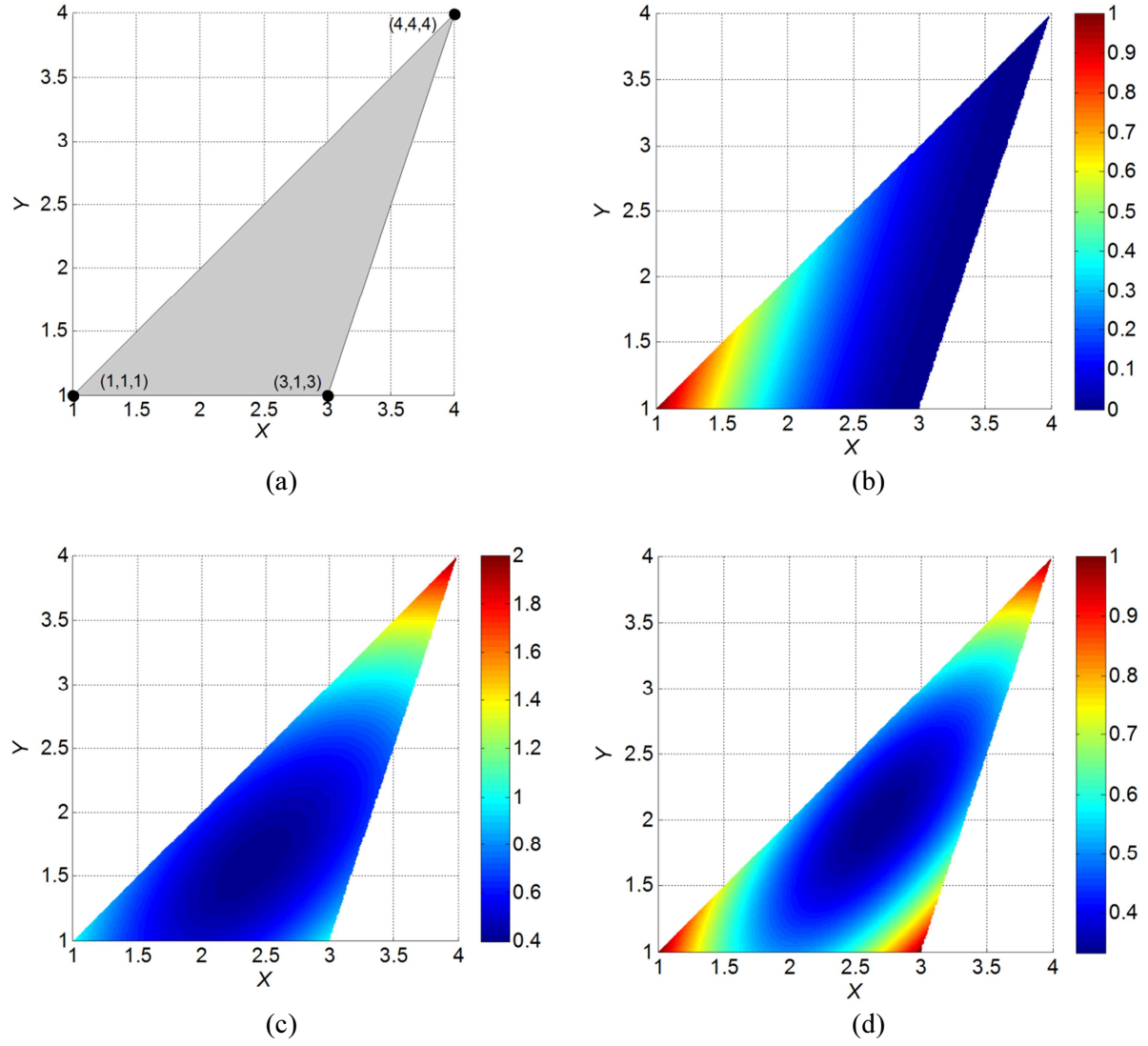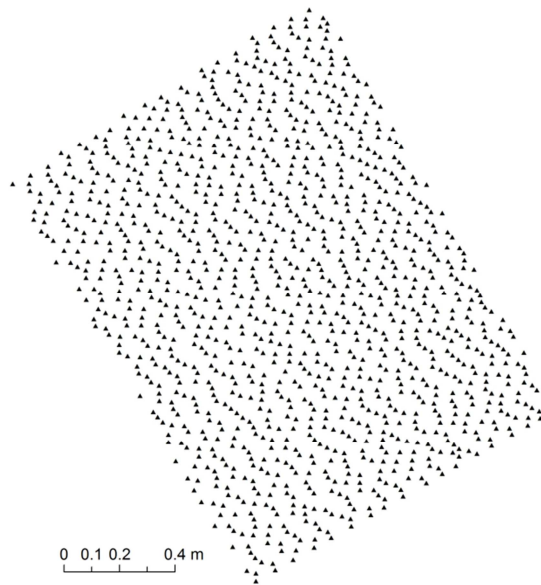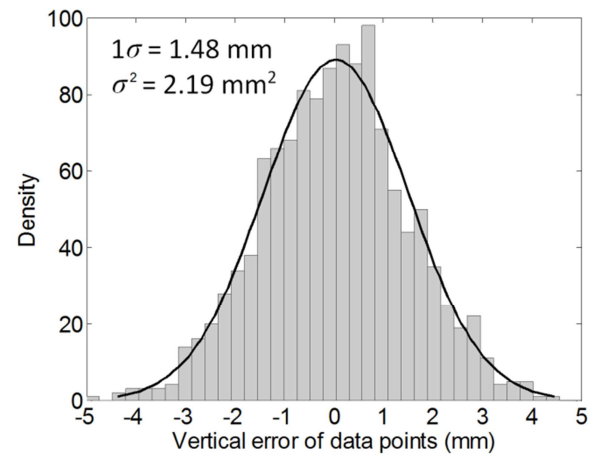
Figure 3: A single triangle: (a) a triangle surface created using TIN with linear interpolation based on the three node points; (b) values of m_1; (c) The propagated error variance in the case of different error variances at nodes ($\sigma^2_{node\_1} = 1$; $\sigma^2_{node\_2} = 1$; $\sigma^2_{node\_3} = 2$); (d) values of M at different locations.

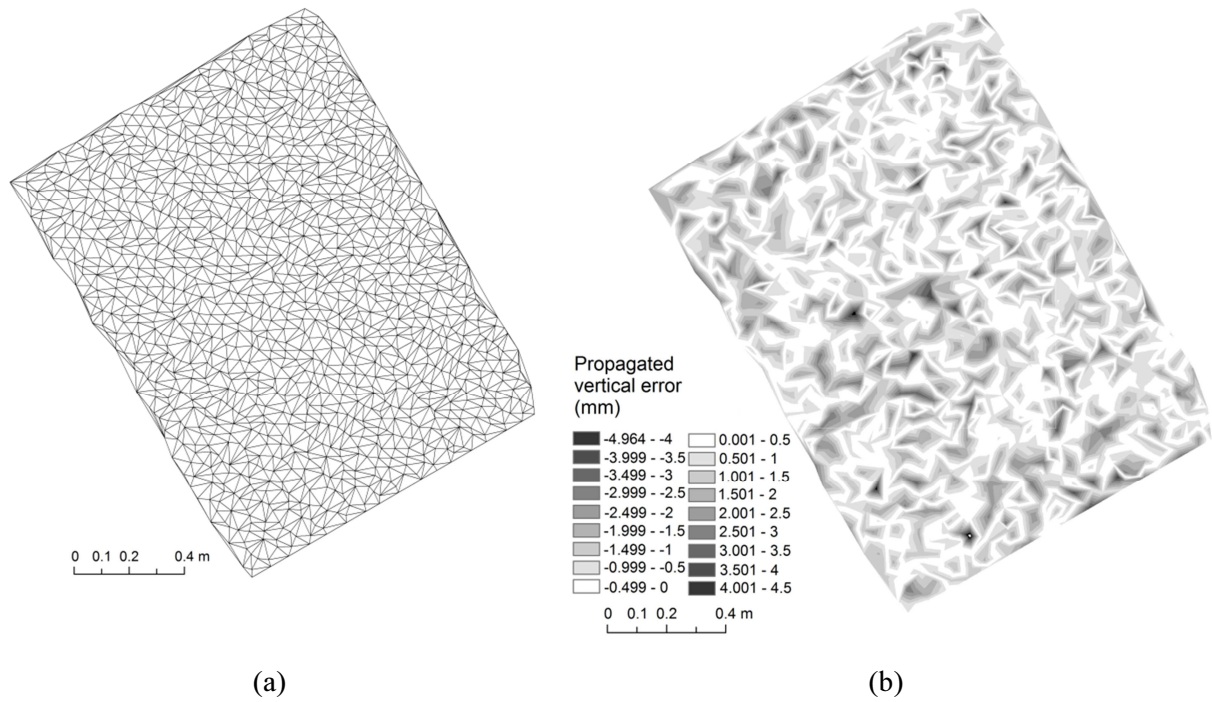Figure 4: (a) TLS data points; (b) histogram of vertical error of TLS data points with normal fit

Propagated
vertical error
(mm)

-4.964 - -4	0.001 - 0.5
-3.999 - -3.5	0.501 - 1
-3.499 - -3	1.001 - 1.5
-2.999 - -2.5	1.501 - 2
-2.499 - -2	2.001 - 2.5
-1.999 - -1.5	2.501 - 3
-1.499 - -1	3.001 - 3.5
-0.999 - -0.5	3.501 - 4
-0.499 - 0	4.001 - 4.5

0   0.1 0.2   0.4 m

(a)   (b)

Figure 5: (a) The TIN created from the TLS survey data; (b) a fine-resolution raster interpolated linearly from the TIN.
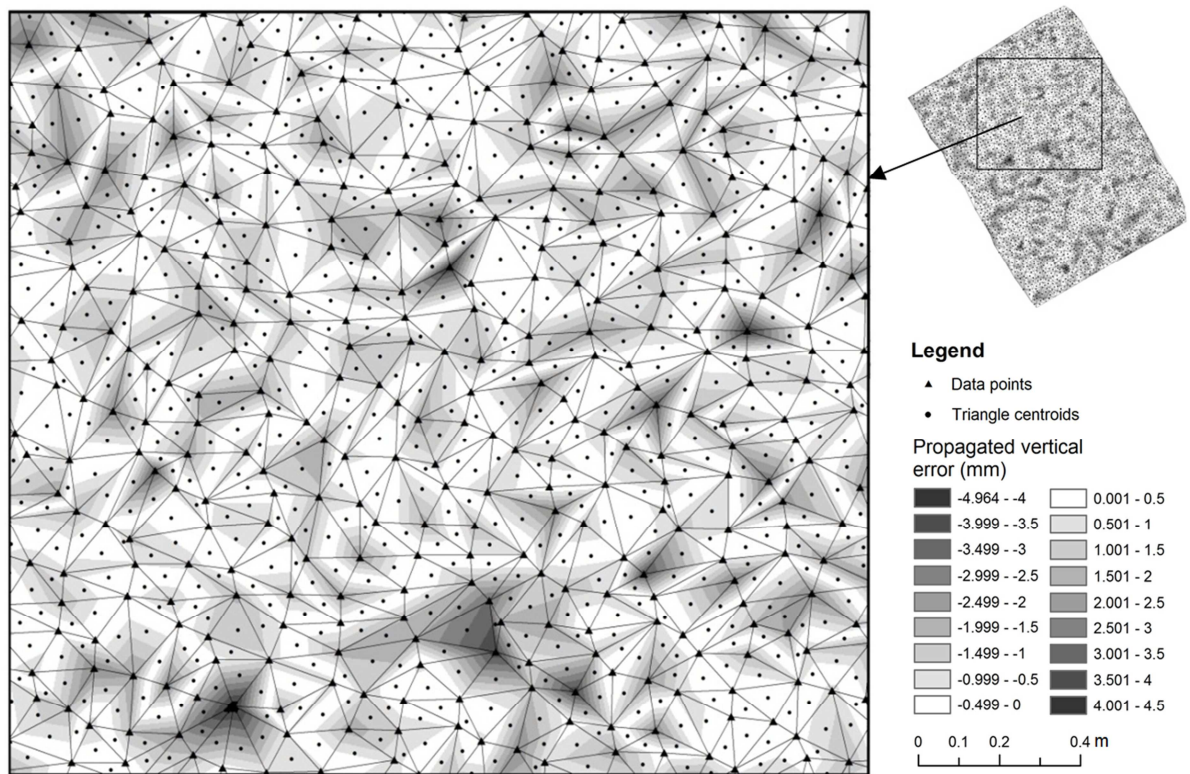
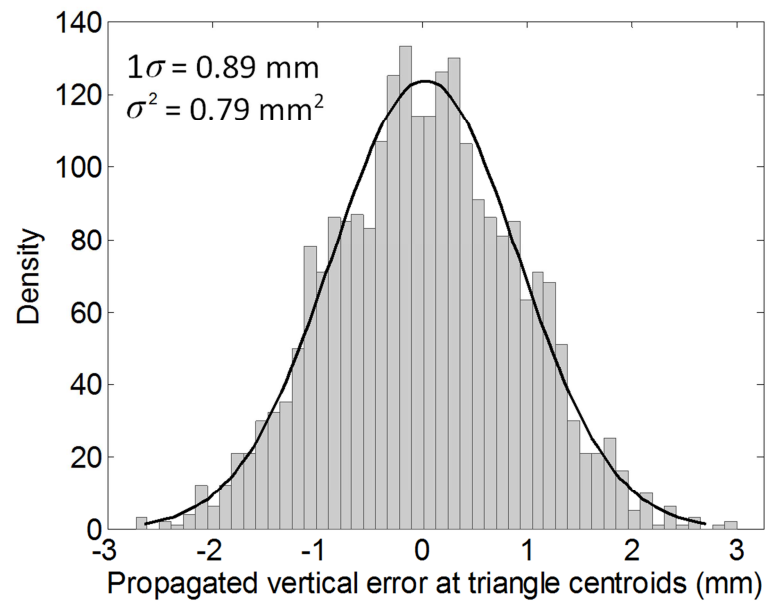Figure 6: Spatial characteristics of propagated vertical error.

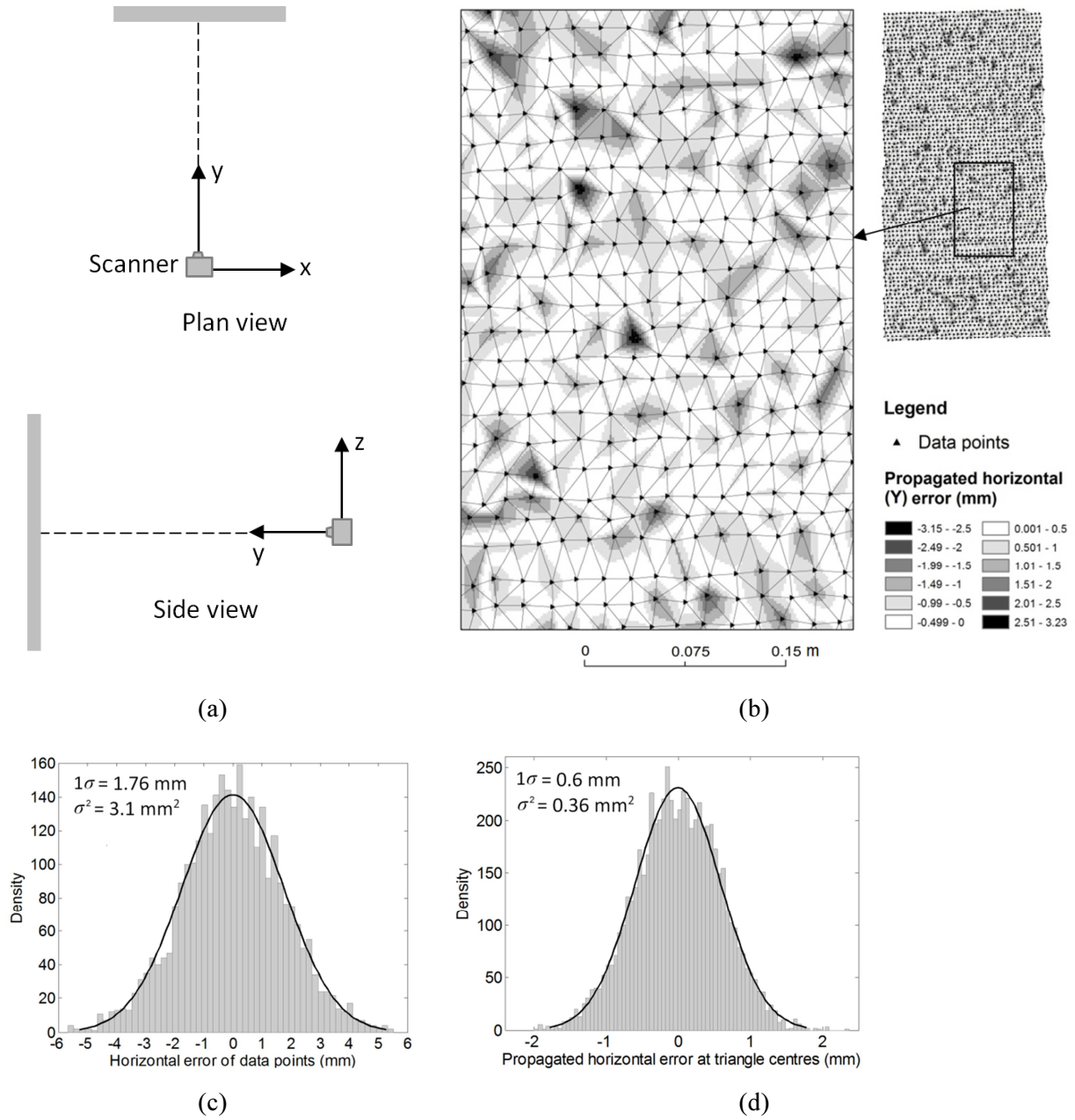Figure 7: Histogram of propagated vertical error at triangle centroids with normal fit

Figure 8: (a) experiment layout; (b) spatial characteristics of propagated vertical error; (c) histogram of vertical error of TLS data points with normal fit; (d) histogram of propagated vertical error at triangle centroids with normal fit.