Accuracy Assessment of Digital Elevation Models Based on Approximation Theory

XiaoHang Liu[1], Peng Hu[2]

[1]San Francisco State University

1600 Holloway Avenue HSS 279, Department of Geography, San Francisco, CA 94132
Telephone: 1-415-338-7509
Fax: 1-415-338-6243
Email: xhliu@sfsu.edu

[2]Wu Han University

School of Resource and Environment Science, 129 LuoYu Road, WuHan, P.R.China 430079

Telephone: (86) 027-68778381
Fax: (86) 027-68778893
Email: penghu@whu.edu.cn

1. Introduction

Digital elevation models (DEM) are essential to various applications such as terrain modeling and watershed delineation. Since the 1980s, error propagation theory has been the dominant framework to assess DEM accuracy, which leads to the current practice on using error variance or Root Mean Squared Error (RMSE) to describe DEM quality. However, error propagation theory is found incapable to explain some empirical observations such as the correlation between DEM error and terrain complexity (e.g. (Wise 2000; Oksanen and Sarjakoski 2006). This paper presents approximation theory in computational science as an alternative framework. Specifically, the paper clarifies the nature and composition of DEM error; discusses the challenges facing the error propagation theory; and illustrates how approximation theory can be applied to assess a DEM generated from linear polynomial interpolation such as linear interpolation in 1D, Triangulated Irregular Network (TIN) interpolation, and bilinear interpolation in a rectangle.

2. Errors in a DEM

To assess DEM error, an understanding of its nature and composition is necessary. In the literature, some scholars have made explicit assumption that DEM error is randomly distributed (Li 1993; Huang 2000; Weng 2002). Others expressed reservations because of findings from empirical analysis (Oksanen and Sarjakoski 2006). Suppose T is a DEM point whose true elevation is Z_T . Given an interpolation method, the interpolated elevation using source data with and without error are denoted by $Z_T^{'}$ and H_T respectively. It can be seen that

$$Z_T^{'} = H_T \pm \delta_T \qquad (1)$$

where δ_T the impact of the errors in the source data on point T. The total error of T, denoted by ΔZ_T is:

$$\Delta Z_T = Z_T - Z_T^{'} = Z_T - (H_T \pm \delta_T) = (Z_T - H_T) \pm \delta_T = R_T \pm \delta_T \qquad (2)$$

where $R_T = Z_T - H_T$ is the interpolation error and can be shown to be systematic error. Under the assumption that systematic error and gross error in the source data are removed, δ_T can be shown random error. Equation 1 shows that the total error in a DEM point is not random but a mixture of random and systematic error.

3. Error propagation theory

3.1. Error propagation theory in its general form

Since 1980, error propagation theory has been used to assess DEM accuracy. The general form of error propagation theory is as follows. Suppose error ε is the sum of two errors x and y, i.e.

$$\varepsilon = x + y \qquad (3)$$

if ε is observed repeatedly for n times, there is:

$$\varepsilon_i = x_i \pm y_i, \quad i = 1, ..., n$$

or equivalently

$$\varepsilon_i^2 = x_i^2 + y_i^2 \pm 2x_i y_i$$

The average of these observations is:

$$\frac{\varepsilon_i^2}{n} = \frac{x_i^2}{n} + \frac{y_i^2}{n} \pm \frac{2x_i y_i}{n}$$

Under the assumption that x and y are random errors which typically have normal distribution with mean of 0, the above equation can be rewritten as

$$\sigma_\varepsilon^2 = \sigma_x^2 + \sigma_y^2 \pm 2\sigma_{xy}^2 \qquad (4)$$

where $\sigma_\varepsilon^2 = \sum \varepsilon_i^2 / n$, $\sigma_x^2 = \sum x_i^2 / n$, $\sigma_y^2 = \sum y_i^2 / n$, and $\sigma_{xy}^2 = \sum x_i y_i / n$. Furthermore, if x and y are independent of each other, equation 4 can be simplified to

$$\sigma_\varepsilon^2 = \sigma_x^2 + \sigma_y^2 \qquad (5)$$

3.2. Application of Error propagation theory in DEM

As shown in equation 1, the total error of a DEM point can be written as

$$\Delta Z_T = R_T \pm \delta_T$$

which is in the same form as equation 2, therefore error propagation theory has been applied to study DEM accuracy. Per equation 4,

$$\sigma_{\Delta Z_T}^2 = \sigma_{R_T}^2 + \sigma_{\delta_T}^2 \pm 2\sigma_{R_T \delta_T}^2 \qquad (6)$$

In the literature, the covariance term is typically omitted to result in (Tempfli 1980; Li 1993; Aguilar *et al.* 2006):

$$\sigma_{\Delta Z_T}^2 = \sigma_{R_T}^2 + \sigma_{\delta_T}^2 \qquad (7)$$

While the error propagation theory in equation 6 and 7 provide an elegant framework to study DEM, a few challenges exist: (1) error propagation theory assumes both errors are random error. However, the interpolation error R_T is not random error but systematic error. Whether error propagation theory is applicable in the context of DEM error is thus worth of consideration. Furthermore, the application of equation 7 requires that R_T and δ_T are independent of each other. In reality, R_T and δ_T both dependant on the interpolation function. Their independence is thus difficult to determine. Considering these challenges, error propagation theory which is the root of using error variance and RMSE to describe DEM accuracy seem to deserve a second thought.

4. Accuracy assessment based on approximation theory

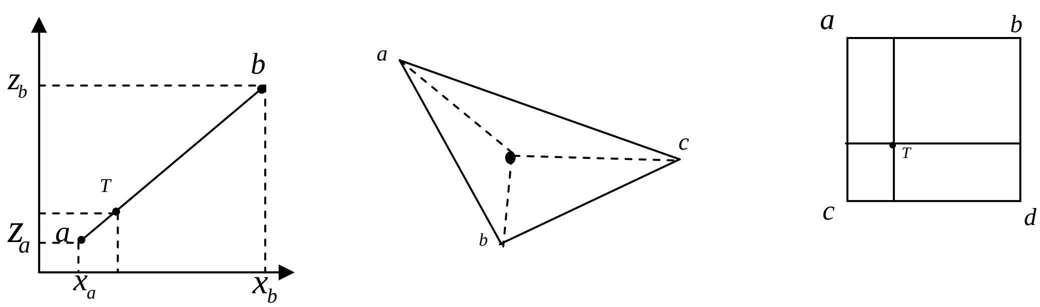
An alternative framework is approximation theory which concerns how to approximate an actual complex function $f(x, y)$ using simpler functions $F(x, y)$ and quantitatively characterize the errors introduced thereby. In the context of DEM research, $f(x, y)$ is the terrain, $F(x, y)$ is a DEM. According to approximation theory, the goodness of approximation is measured by truncation error which is the largest difference between the two functions, i.e. $\max|f(x, y) - F(x, y)|$. Let $T(x_T, y_T)$ a DEM point. Its error is

$$\Delta z_T = f(x_T, y_T) - F_1(x_T, y_T) = R_T \pm \delta_T$$

The truncation error of the DEM is $\max\{|\Delta z_T|\}$, i.e.

$$\max|\Delta z_T| = \max|R_T \pm \delta_T| \le \max|R_T| + \max|\delta_T| \tag{8}$$

In the following, we derive $\max|R_T|$ and $\max|\delta_T|$ for the three linear polynomial interpolation methods in Figure 1.



(a) (b) (c)

Figure 1: Three linear polynomial interpolation methods to generate a DEM: (a) linear interpolation, (b) TIN interpolation, and (c) bilinear interpolation in a rectangle.

It can be shown that the functions of the three interpolation methods are as follows:
Linear interpolation:

$$Z_T^{'} = \omega_1 Z_b + \omega_2 Z_a, \omega_1 + \omega_2 = 1, \omega_1, \omega_2 > 0 \tag{8}$$

TIN interpolation:

$$Z_T^{'} = \omega_1 Z_a + \omega_2 Z_b + \omega_3 Z_c, \omega_1 + \omega_2 + \omega_3 = 1, \omega_1, \omega_2, \omega_3 > 0 \tag{9}$$

Bilinear interpolation in a rectangle:

$$Z_T^{'} = \omega_1 Z_a + \omega_2 Z_b + \omega_3 Z_c + \omega_4 Z_d \;, \omega_1 + \omega_2 + \omega_3 + \omega_4 = 1,\; \omega_1, \omega_2, \omega_3, \omega_4 > 0 \qquad (10)$$

where Z_a, Z_b, Z_c, Z_d are the elevations of the corresponding vertices as shown in Figure 1.

4.1. Impact of source-data error

The impact of the errors in the source data on a DEM point T is as follows:

Linear interpolation: $\delta_T = \omega_1 \delta_b + \omega_2 \delta_a$

TIN interpolation: $\delta_T = \omega_1 \delta_a + \omega_2 \delta_b + \omega_3 \delta_c$

Bilinear interpolation in a rectangle: $\delta_T = \omega_1 \delta_a + \omega_2 \delta_b + \omega_3 \delta_c + \omega_4 \delta_d$

The bound of the impact can be shown to be

$$\max \left| \delta_T \right| \leq \left| \delta \right|$$

where $\left| \delta \right|$ is the largest error in the source points.

4.2. Impact of interpolation error

All of the three interpolation methods are piecewise polynomial interpolation, i.e. dividing the terrain into patches and conduct DEM interpolation patch by patch. For each patch, the corresponding terrain $f(x, y)$ is assumed to be twice or more continuously differentiable. It can be derived that the interpolation error bounds are

Linear interpolation in 1D: $\left| R_T \right| \leq \frac{1}{8} M_2 h^2$

TIN interpolation: $\left| R_{TIN} \right| \leq \frac{3}{8} M_2 h^2$

Bilinear interpolation in a rectangle: $\left| R_T \right| \leq \frac{1}{4} M_2 h^2 + \frac{1}{64} M_4 h^4$

where M_2, M_4 are the second and fourth order maximum norm over the entire terrain, h is the largest distance between two source points. M_2, M_4 are essentially indicators of terrain complexity, and h is the sampling density.

Combing the bounds of the interpolation error and the impact of the source-data error, the total error of a DEM point generated by each of the three linear polynomial interpolation methods is summarized in Table 1.

Table 1. DEM error of three linear polynomial interpolation methods.

	δ_T *: random error*	R_T : interpolation error
Linear interpolation in 1D	$\lvert\delta_T\rvert \le \lvert\delta\rvert$	$\lvert R_T\rvert \le \frac{1}{8} M_2 h^2$
TIN interpolation	$\lvert\delta_T\rvert \le \lvert\delta\rvert$	$\lvert R_T\rvert \le \frac{3}{8} M_2 h^2$
Bilinear interpolation in a rectangle	$\lvert\delta_T\rvert \le \lvert\delta\rvert$	$\lvert R_T\rvert \le \frac{1}{4} M_2 h^2 + \frac{1}{64} M_2 h^4$
Total error bound	$\max\lvert R_T\rvert + \lvert\delta_T\rvert \le \max\lvert R_T\rvert + \max\lvert\delta_T\rvert$	

5. Discussion and Conclusion

Many research on DEM accuracy observed that DEM accuracy is related to terrain complexity and sampling density (Wood 1994; Aguilar *et al.* 2006). The research in this paper based on approximation theory offers a theoretical explanation for this observation. It can be seen that DEM error is a mixture of the interpolation error and random error. In the case that terrain is very complex and source point density is low, $M_2 h^2$ will be large which may render all three interpolation methods ineffective. Increasing sample density will decrease the interpolation error. In the case that very high density of source data is available and the source data is highly accurate, as in the case of LiDAR, high accuracy DEM is almost guaranteed regardless of the interpolation method used.

6. References

Aguilar, F. J., Aguilar, M., Aguera, A.F. and Sanchez, 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.

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

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

Oksanen, J. and Sarjakoski, T. (2006). Uncovering the statistical and spatial characteristics of fine toposcale DEM error, *International Journal of Geographical Information Science*,**20**(4), 345–369.

Tempfli, K. (1980). Spectral analysis of terrain relief for the accuracy estimation of digital terrain models, *ITC Journal*, **1980**(3), 478-510.

Weng, Q., 2002. Quantifying uncertainty of digital elevation models derived from topographic maps. *Advances in Spatial Data Handling* (D. Richardson and P. van Oosterom, editors).Springer, New York, pages 403-418.

Wise, S., 2000. Assessing the quality for hydrological applications of digital elevation models derived from contours,*Hydrological Process*,**14**: 1909-1929.

Wood, J. D., 1994. Visualising contour interpolation accuracy in digital elevation models *Visualization in Geographical Information Systems* (H. M. Hearnshaw and D. J. Unwin, editors).John Wiley and Sons, Chichester, pages168–80.

Biography

XiaoHang Liu is an Assistant Professor in Geography at San Francisco State University. Her research interest is in spatial interpolation, uncertainty in geographic data, and remote sensing image analysis.

Peng Hu is a distinguished Professor at school of Resource and Environment Science of WuHan University. His research has focused on digital elevation model generation and accuracy assessment and computational geometry.