Determination of Complexity Factor and Its Relationship with Accuracy of Representation for DEM Terrain

LIU Chun1,2, SUN Weiwei1, WU Hangbin1
1. Department of Surveying and Geo-informatics, Tongji University, 1239 Siping Road, Shanghai 200092, China
2. Key Laboratory of Advanced Engineering Survey of SBSM, 1239 Siping Road, Shanghai 200092, China

Abstract Based on the estimating rule of the normal vector angles between two adjacent terrain units, we use the concept of terrain complexity factor to quantify the terrain complexity of DEM, and then the formula of terrain complexity factor in Raster DEM and TIN DEM is deduced theoretically. In order to make clear how the terrain complexity factor \( \text{CFE} \) and the average elevation \( h \) affect the accuracy of DEM terrain representation RMSE, the formula of Gauss synthetical surface is applied to simulate several real terrain surfaces, each of which has different terrain complexity. Through the statistical analysis of linear regression in simulation data, the linear equation between accuracy of DEM terrain representation RMSE, terrain complexity factor \( \text{CFE} \) and the average elevation \( h \) is achieved. A new method is provided to estimate the accuracy of DEM terrain representation RMSE with a certain terrain complexity and it gives convincing theoretical evidence for DEM production and the corresponding error research in the future.

Keywords terrain complexity factor; Gauss synthetical surface; accuracy; DEM terrain representation

CLC number P231.5

Introduction

DEM is one of the most important spatial data in the spatial database for a GIS application. It now becomes the elementary data during three dimension representation and terrain analysis in 3D GIS.\(^1\) However, the error of DEM may result from various errors during the modeling stage when building the DEM. The influencing factors include:\(^1\) (1) characteristic of terrain relief, (2) attribute of original DEM data, and (3) modeling algorithm of DEM.

So far, many scholars have performed their research on complexity representation of a terrain and accuracy determination of a DEM. The relationship among the error of DEM, spatial resolution and mean profile curvature was described using terrain representation error indicator.\(^2\) Later, the model of terrain representation error is improved, and the exact association is found with the spatial resolution and terrain slope acting on the error of DEM.\(^3\) Also, research work can be found focused on the relationship be-
 tween terrain complexity degree and terrain slope or aspect. A valid scale range is given to represent the slope and aspect of DEM using composition of surface vector. Spatial self-relevance analysis is also performed to explain the spatial distribution feature of elevation error in DEM. For quantization of terrain complexity, the concept and calculation of the terrain complexity indicator are put forward based on grid DEM. Meanwhile, fractal geometry is adopted to give a fractal analysis to describe the terrain complexity. With respect to DEM accuracy, the mathematic transform model under diverse spatial scale is discussed as DEM representation error, while the gully region of the Loess Plateau in Northern Shanxi Province is used as a case study area. After selecting six typical Loess Plateau in the Southern Shanxi Province profile as the case study area, this paper discusses the fractal terrain complexity and spatial isomers feature aimed at the typical Loess Plateau. Zhu put forward an accuracy assessment model on DEM by reconstructing the contour, and discussed the node error of TIN (triangulate irregular network) and gave a mean accuracy determination method.

Besides, many researches can be found abroad surrounding the terrain factors representation for a DEM. A soil-landscape interference model is used to explain the influence of spatial resolution of DEM on real surface attribute. Grid DEM is used to determine the error of the river length, and some other work was given to explain the vibration of DEM accuracy when the terrain location changed.

However, lots of existing works are performed on the determination of DEM accuracy and relationship with its relative terrain. It is still difficult to depict the terrain representation error. Actually, the error of terrain representation has a relationship with spatial resolution of DEM. Therefore, it is significant that we put forward a statistical description indicator to explain the influence of terrain complexity on the error of terrain representation. Focusing on the relationship between terrain complexity and the error of terrain representation, this paper will give a quantization terrain factor to represent the terrain complexity. Afterward, the influence of the diverse terrain feature on the error of terrain representation is discussed. Significance role is obviously found when a representation accuracy standard is requested according to real complex terrain relief.

1 Terrain complexity factor

The accuracy of DEM is a precision indicator when DEM is used to depict a real terrain. Many researchers had performed their theoretical research on DEM representation accuracy, and had achieved a mature theory system. Therefore, this paper emphasizes on involving a terrain complexity factor to quantify the quality when representing a real terrain relief. Actually, the spatial convergence angle between positive normal vector and adjacent positive normal vector for a grid or triangles polygon can be used as the factor to depict terrain complexity in a certain region.

For a DEM with rolling topographical relief, a single grid or triangle can be regarded as a relative consistent terrain unit. Terrain shape generally changes at the boundary of adjacent terrain unit. As in Fig.1, A and B are two terrain units in a DEM data set, and they can be a grid or triangle unit. When $\alpha = 0$, the two adjacent terrain units have no distinct change in terrain shape; when $0^\circ < \alpha < 180^\circ$, terrain shape between two adjacent terrain units changes more significantly while the value of $\alpha$ becomes bigger. Obviously, convergence angle between two adjacent terrain units can be used to describe the change of the terrain shape in a certain region.

![Fig.1 Normal angle of two adjacent terrain units](image)

However, the value $\alpha$ has the unit of angle, so adjacent terrain unit complexity factor $cf$ is used to replace $\alpha$ in order to unify the dimension for convenient calculation. The formula is as:
\[ cf_\alpha = e^{-\cos \alpha} = e^{-(a_1x_2 + a_2x_3 + a_3x_4)} \]  

(1)

The involved value of \( cf \) is a monotone increasing function with the independent variable \( \alpha \) inside the range of \( 0^\circ < \alpha < 180^\circ \), and the values always bigger than 0. Therefore, the \( cf \) can depict and quantify the representation of terrain shape change.

1.1 Terrain complexity factor for grid DEM

For grid DEM, each grid has four adjacent grids. The four terrain units’ complexity factors are calculated by using the positive normal vector of each grid and its four adjacent grids. The mean value of the four complexity factors is determined as the terrain complexity value \( CF \). Actually, each adjacent grid shares two grid nodes, so total least squares is used to perform the adjustment. The obtained result can fit a plane for an adjacent grid unit and simulate the real terrain relief effectively. As in Fig.2, \( A, B, C, D \) and \( E \) represent grid planes in a certain region. \( A \) is the center grid, \( a, b, c, d, e \) are positive normal vectors for each grid plane, and \( \alpha, \beta, \gamma, \delta \) are the convergence angles between \( a \) and \( b, c, d, e \). Therefore, the terrain complexity factor for non-boundary grid \( A \) is calculated as:

\[ CF = (cf_a + cf_b + cf_c + cf_d)/4 \]  

(2)

However, a boundary grid can locate alternately at up, down, right or left. Under boundary circumstance, adjacent terrain grids have a terrain complexity factor value of 0; the terrain grid has a terrain complexity factor as the mean value of all non-zero adjacent grids. For a grid DEM with the size of \( m \times n \), terrain complexity factor for each grid is expressed as \( CF_{(i,j)} \), where \( i = 1, \ldots, m; j = 1, \ldots, n \).

Actually, because of the limitation of local representation, terrain factor \( CF_{(i,j)} \) can only express a local terrain complexity degree, overall terrain complexity degree is determined by factor of \( E_{CF} \), which is calculated as:

\[ E_{CF} = \frac{\sum_{i=1}^{m} \sum_{j=1}^{n} CF_{(i,j)}}{mn} \]  

(3)

It is clear from Eq. (3), \( E_{CF} \) can be used to represent the overall terrain complexity degree of DEM. DEM vibration is more significant when the value of \( E_{CF} \) is bigger.

1.2 Terrain complexity factor for triangulated irregular network DEM

For a DEM established by triangulate irregular network, terrain complexity factor is also determined by convergence angle between the positive normal vector of each triangulate unit and its adjacent triangulate unit as well. \( cf \) is used here to express the terrain complexity of each triangulate unit. In Fig. 3, for the non-boundary triangulate plane \( S \), \( S_1, S_2, S_3 \) are the adjacent triangulate planes, \( a, b, c, d \) are the positive normal vectors for each triangulate plane, and \( \alpha, \beta, \gamma \) are the convergence angles between \( a \) and \( b, c, d \). Therefore, the terrain complexity factor of the triangulate unit \( S \) can be calculated as:

\[ CF = (cf_a + cf_\beta + cf_\gamma)/3 \]  

(4)

Fig.3 Terrain complexity value of triangle \( S \) in a TIN DEM

If the triangulate plane \( S \) is a boundary triangulate plane, the adjacent terrain triangulate plane has a terrain complexity factor value as 0, so the terrain triangulate has the terrain complexity factor as the mean value of all non-zero adjacent triangulates. If the DEM is composed of TIN triangulate with size of \( N \), overall terrain complexity factor \( E_{CF} \) is calculated as:
\[ E_{CF} = \frac{\sum_{i=1}^{N} CF_{(i)}}{N} \]  \hspace{1cm} (5)

Where \( CF_{(i)} \) is the complexity factor of the \( i \)th triangular unit.

2 Simulation analysis

Terrain complexity factor can quantify the terrain complexity degree of overall DEM. Emphasis is laid on the relationship between terrain complexity factor and terrain representation error of DEM when this simulation testing is performed. Latent rule of the relationship is then discussed when LiDAR data is used to establish the DEM. \( E_{CF} \) is calculated as the terrain complexity factor to represent the real terrain complexity degree. However, the simulation method is also adopted when considering the lack of prior statistic information of the used diverse terrain relief.

2.1 Calculation of terrain complexity factor based on LiDAR data

Airborne LiDAR data is used here to calculate terrain complexity factor \( E_{CF} \). The data was captured on Oct 20, 2006, and was located in the littoral region at Yantai, Shandong Province. 4000 points (Fig.4) are selected in the simulation, whose elevation ranged from 26 m to 29 m. Grid DEM is obtained by second conicoid fitting, and Eq. (3) is used to calculate the terrain complexity factor \( E_{CF} \). The result is:

\[ X_{\text{min}} = 597233.410 \text{ m}, \quad X_{\text{max}} = 597400.190 \text{ m}, \]
\[ Y_{\text{min}} = 244375.000 \text{ m}, \quad Y_{\text{max}} = 244435.130 \text{ m}, \]
\[ Z_{\text{min}} = 26.000 \text{ m}, \quad Z_{\text{max}} = 28.960 \text{ m}, \]

spatial resolution of DEM is 1 m, and terrain complexity factor \( E_{CF} \) is 1.0159

Fig.4 Sketch of LiDAR data in study area

From the result of the experiment, terrain complexity factor \( E_{CF} \) is feasible in representing the real DEM, and can quantify real terrain complexity degree. Considering the lack of prior statistic information of real irregular terrain, this experiment simulates real terrain by Gauss synthetical surface to seek the relationship between terrain complexity factor and terrain representation accuracy.

2.2 Relationships between terrain complexity degree and terrain representation accuracy

2.2.1 Design of experiment

It is difficult to represent the variety of a real irregular terrain surface using a simple mathematic surface model at present. Although a high-order polynomial can approach any complex surface theoretically,[16] computation error and unstable solution still emerge easily. By considering restrictions caused from variety condition, Gauss synthetical surface model is used to simulate the practical grid DEM, which is also used to explain the relationship between terrain complexity factor and terrain representation accuracy. The formula of the Gauss synthetic surface is:

\[ z = f(x, y) = A \times \left( 1 - \left( \frac{x}{m} \right)^2 \right) \times e^{-\left( \frac{x^2}{m^2} + \frac{y^2}{n^2} \right)} - B \times \left( 0.2 \times \left( \frac{x}{m} \right)^3 - \left( \frac{y}{n} \right)^3 \right) \times e^{-\left( \frac{x^2}{m^2} + \frac{y^2}{n^2} \right)} - C \times e^{-\left( \frac{x^2}{m^2} + \frac{y^2}{n^2} \right)} \]  \hspace{1cm} (6)

Where, \( A, B, \) and \( C \) are the parameters of topographic relief. \( m \) and \( n \) are the region control parameters. Actually, diverse types of terrain DEM can be simulated by adjusting the parameters of \( A, B, \) and \( C \). The terrain complexity factor is also used to represent the terrain complexity degree in quantity.

According to the definition of terrain representation accuracy,[2] the terrain representation error is defined as the elevation difference between the center point and average of the four corner point for each grid, and is described as:

\[ E_t = H_o - H'_o = H_o - (H_A + H_B + H_C + H_D) / 4 \]  \hspace{1cm} (7)

Where, \( H_A, H_B, H_C, H_D \) are the elevation values of the grid corner points, and \( H'_o \) is the mean elevation of the four points \( A, B, C, D \). \( H_o \) is the elevation of the grid center point.

For a DEM with the size of \( m \times n \), the terrain representation accuracy is:

\[ RMSE_{E_t} = \sqrt{\frac{\sum_{i=1}^{m} \sum_{j=1}^{n} (E_{t(i,j)} - E(ET))^2}{(m-1)(n-1) - 1}} \]  \hspace{1cm} (8)
2.2.2 Experiments analysis

According to Eq. (6), suppose \( x \in [-500, 500], y \in [-500, 500] \), \( m=500, n=500 \). Parameters of \( A, B, \) and \( C \) are supposed respectively, \( h \) is the mean elevation. Therefore, the DEMs in diverse complexity degree are simulated based on the above parameters (in Fig.5). Fig.5(a) is regarded as a plain because of a simple surface shape with minimum fluctuate, whose elevation range is from \(-3\) m to \(4\) m. Figs.5(b) and 5(c) are a bit more fluctuate than Fig.5(a), and are regarded as a hill. Figs.5(d), 5(e) and 5(f) have more complex surface shapes. Fig.5(f) is regarded as a mountain with maximum fluctuates among the three, whose elevation range is from \(-600\) m to \(600\) m. There is a variation of terrain complexity from simple to complex according to the parameters in Figs.5(a) to 5(f).

From Figs.5(a) to 5(f), the coordinates of four corner points \((x_i, y_i, z_i), i = 1, 2, 3, 4\) for each grid are fitted to obtain the optimal plane and positive unit normal vector.

Fig.5 Simulated DEM terrains with Gauss synthetical surface (DEM resolution is 10 m)
Besides, the terrain complex factor of each grid \((i,j)\) can be calculated with Eq.(2), and the terrain complex factor of the DEM is obtained subsequently with Eq.(3).

An analysis window with the size of 3×3 is designed. Terrain representation accuracy of DEM is obtained with Eq.(8) after representation error of each grid is calculated with Eq.(7).

After setting diverse resolution as 5 m, 10 m, 20 m, 25 m, 50 m and 100 m, operations are performed respectively to obtain terrain representation accuracy under different resolutions (Table 1).

| Resolution (m) | Fig.5(a) | Fig.5(b) | Fig.5(c) | Fig.5(d) | Fig.5(e) | Fig.5(f) |
|---------------|----------|----------|----------|----------|----------|----------|
| 5             | 0.9954   | 0.9960   | 0.9971   | 1.0032   | 1.0093   | 1.0568   |
| 10            | 0.9811   | 0.9821   | 0.9844   | 0.9951   | 0.9992   | 1.0347   |
| 20            | 0.9601   | 0.9615   | 0.9639   | 0.9764   | 0.9790   | 1.0267   |
| 25            | 0.9495   | 0.9515   | 0.9540   | 0.9628   | 0.9689   | 1.0179   |
| 50            | 0.8976   | 0.8985   | 0.9023   | 0.9116   | 0.9173   | 0.9742   |
| 100           | 0.7901   | 0.7910   | 0.7933   | 0.7942   | 0.8161   | 0.9100   |

### 3 Result analysis

#### 3.1 Relationship between spatial resolution and terrain complexity factor \(E_{CF}\)

It can be seen from Table 1 that \(E_{CF}\) becomes larger when the terrain complexity degree becomes greater under the same spatial resolution condition. But under the condition with the same terrain complexity degree, \(E_{CF}\) becomes smaller when the spatial resolution is reduced. \(E_{CF}\) becomes larger when terrain becomes more complex and spatial resolution improves. Besides, Section 2 used LiDAR data to build the DEM with a resolution of 1 m. The terrain complex factor is then calculated as 1.0159. Meanwhile, a simulated terrain plane in Fig.6(a) is also obtained. The terrain complex factor is 0.9954 under the resolution of 5 m. By comparing the LiDAR DEM with the simulated DEM, the terrain complex result is coincident logically.

| Resolution (m) | Fig.5(a) | Fig.5(b) | Fig.5(c) | Fig.5(d) | Fig.5(e) | Fig.5(f) |
|---------------|----------|----------|----------|----------|----------|----------|
| 5             | 0.0007867| 0.0048   | 0.0096   | 0.0138   | 0.0383   | 0.1437   |
| 10            | 0.0032   | 0.0192   | 0.0385   | 0.0511   | 0.1535   | 0.5755   |
| 20            | 0.0126   | 0.0770   | 0.1541   | 0.4244   | 0.6152   | 2.3061   |
| 25            | 0.0197   | 0.1204   | 0.2408   | 0.6442   | 0.9616   | 3.6049   |
| 50            | 0.0788   | 0.4801   | 0.9601   | 1.6769   | 3.8630   | 14.3846  |
| 100           | 0.3069   | 1.8650   | 3.7300   | 5.2772   | 14.9227  | 55.9669  |

It can be seen from Table 2, under the same spatial resolution, that the terrain representation accuracy becomes lower (\(RMSE_{E_t}\) is larger) when the terrain becomes complex. However, under the sample condition of terrain complex, terrain representation accuracy becomes higher (\(RMSE_{E_t}\) is smaller) when the spatial resolution is improved. Actually, real terrain becomes complex when the \(RMSE_{E_t}\) is larger; meanwhile real terrain becomes simple when the \(RMSE_{E_t}\) is smaller.

#### 3.2 Influence analysis with terrain complex factor \(E_{CF}\) on terrain representation accuracy \(RMSE_{E_t}\)

Under different types of terrain relief, the statistical data in Table 1 and Table 2 are used to perform the regression between \(E_{CF}\) and \(RMSE_{E_t}\), and its relationship is obtained in Table 3.

Linear regression fitting is performed with the principle of least squares under different terrains described in Fig.5. A group of parameters \(\beta_0\) and \(\beta_1\) is obtained
in Table 3. The linear regression relationship is described as:

\[ \text{RMSE}_{E_t} = \beta_0 + \beta_1 \times E_{CF} \]  

(9)

Table 3 Linear regression coefficients of \( E_{CF} \) and \( \text{RMSE}_{E_t} \) under different DEM terrains

| \( \beta_0 \) | \( \beta_1 \) | Related coefficient \( \gamma \) |
|---|---|---|
| Fig.5(a) | 1.4785 | -1.5145 | -0.9742 |
| Fig.5(b) | 9.0806 | -0.9031 | -0.9738 |
| Fig.5(c) | 18.3157 | -18.7241 | -0.9760 |
| Fig.5(d) | 25.2942 | -25.4599 | -0.9925 |
| Fig.5(e) | 77.5979 | -78.2207 | -0.9738 |
| Fig.5(f) | 403.5642 | -389.4164 | -0.9518 |

Fitting data for each group has a distinct relevance, so \( E_{CF} \) influences the terrain representation accuracy significantly. However, for different simulated terrain reliefs, \( \beta_0 \) and \( \beta_1 \) have no distinct rule or relevance. Therefore, it is difficult to describe the influence that \( E_{CF} \) acts on \( \text{RMSE}_{E_t} \) using an accordant formula.

Meantime, \( \beta_0, \beta_1, \) and mean elevation \( h \) are found have a good linear relation from Fig.6 and Fig.7. Therefore, linear regression is performed to obtain the relationship between \( h \) and \( \beta_0 \), and between \( h \) and \( \beta_1 \). The relationships are described as:

\[ \beta_0 = -8.0125 + 2.0246h, \quad r = 0.9959 \]  

(10)

\[ \beta_1 = 7.1363 - 1.9247h, \quad r = 0.9950 \]  

(11)

Eqs. (11) and (12) are substituted in Eq.(10), and fitting relation between \( \text{RMSE}_{E_t} \) and \( E_{CF} \) is:

\[ \text{RMSE}_{E_t} = \left(7.1363 - 1.9247h \right) 	imes E_{CF} + 2.0246h - 8.0125 \]  

(12)

From the Eq. (12), \( \text{RMSE}_{E_t} \) has the influence both from \( E_{CF} \) and \( h \). Actually, the quality of DEM can be determined by supposing a mean elevation and terrain complex factor.

4 Conclusion

Based on the idea of convergence angle between positive normal vectors of adjacent terrain units, this paper puts forward a terrain complex factor to describe the terrain complex degree of a DEM. The computation formula of terrain complex factor is given on a grid and irregular DEM respectively. In order to discuss the influence of terrain complex factor on mean elevation of DEM, the Gauss synthetical surface model is used to simulate the terrain relief with diverse complex degree. Experiment is performed with simulated terrain relief and results are obtained as follows:

(1) Terrain complex factor can describe the complex degree of a real terrain surface in quantity. The experiment results indicate the feasibility subsequently.

(2) The terrain complex factor of DEM is influenced by spatial resolution, and becomes larger when the spatial resolution improves.

(3) The terrain representation accuracy of DEM is influenced by spatial resolution, and becomes smaller when the spatial resolution improves.

(4) Under the same condition of terrain relief, terrain complex factor has a linear influence on terrain representation accuracy. However, there is no accordant linear model to describe the influence of the
mean elevation on terrain representation accuracy under different conditions of terrain relief.

The statistical relationship among terrain complex factor, mean elevation and terrain representation accuracy can help to evaluate the terrain representation accuracy of a DEM. The relationship can also provide a theoretical accordance after applying the error analysis on the application from DEM. However, the scope of application shall also be further discussed if the relationship between terrain complex degree and terrain representation accuracy is used in the practical data production, and the quality standard of DEM shall be drawn up as well.

References

[1] Zhou Q M, Liu X J (2006) Digital terrain analysis[M]. Beijing: Science Press
[2] Tang G A, Gong J Y, Chen Z J, et al. (2001) A simulation on the accuracy of DEM terrain representation[J]. Acta Geodaetica Et Cartographica Sinica, 30(4): 361-365
[3] Wang G X, Zhu C Q, Shi W Z, et al. (2004) The further study on the accuracy of DEM terrain representation[J]. Acta Geodaetica Et Cartographica Sinica, 33(2): 168-173
[4] Li T W, Liu X J, Tang G A (2004) Influence of terrain complexity on slope and aspect[J]. Journal of Mountain Research, 22(3): 272-277
[5] Liu X J, Zhang P (2008) Effective scale of slope and aspect derived from grid-based digital elevation model[J]. Geomatics and Information Science of Wuhan University, 33(12): 1254-1258
[6] Wang Y G, Wang Z W, Zhu C Q (2008) Analysis of digital terrain error based on spatial autocorrelation[J]. Geomatics and Information Science of Wuhan University, 33(12): 1259-1262
[7] Wang L, Tang G A, Liu X J, et al. (2004) Terrain complexity index and its derivation from DEMs[J]. Bulletin of Soil and Water Conservation, 24(4): 55-58
[8] Zhou T, Long Y, Tang G A, et al. (2006) A fractal method to describe the terrain complexity reflected by the raster DEM[J]. Geography and Geo-Information Science, 22(1): 26-30
[9] Zhang Y, Tang G A (2003) A mathematical simulation of DEM terrain representation error—a case study in the loess hill gully areas of China[J]. Journal of Mountain Research, 21(2): 252-256
[10] Long Y, Zhou T, Tang G A, et al. (2007) Research on terrain complexity of several typical regions of loess landform based on fractal method[J]. Journal of Mountain Research, 25(4): 385-392
[11] Zhu C Q, Wang Z W, Liu H Y (2008) Accuracy evaluation model of DEM based on reconstructed contour[J]. Geomatics and Information Science of Wuhan University, 33(2): 153-156
[12] Zhu C Q, Shi W Z (2005) Estimation of average DEM accuracy under linear interpolation considering random error at the nodes of TIN model[J]. International Journal of Remote Sensing, 26(24): 5509-5523
[13] Smith M P, Zhu A X, Burt J E, et al. (2006) The effects of DEM resolution and neighborhood size on digital soil survey[J]. Geoderma, 137(1-2): 58-69
[14] da Paz, A R, W Collischonn, et al. (2008) Errors in river lengths derived from raster digital elevation models[J]. Computers & Geosciences, 34(11): 1584-1596
[15] Deng Y, Wilson J P (2007) DEM resolution dependencies of terrain attributes across a landscape[J]. International Journal of Geographical Information Science, 21(1-2): 187-213
[16] Liu X J (2003) Error analysis and estimation on grid digital elevation model algorithm[M]. Wuhan:Press of Wuhan University
[17] Cao Z T, Liu X J, Wang Y F, et al. (2005) Contrastive analysis of methods of terrain simplification based on DEM[J]. Modern Surveying and Mapping, 28(1): 9-13