WFL Publisher
Science and Technology

Meri-Rastilantie 3 B, FI-00980
Helsinki, Finland
e-mail: info@world-food.net

Journal of Food, Agriculture & Environment Vol.11 (1): 709-712. 2013

www.world-food.net

The influence of source data density for generating digital elevation models

Lei Wang [1,2]*, Qinke Yang [1,2], Yongqin Long [2], Weiling Guo [1] and Chunmei Wang [1,2]

[1] *Institute of Soil and Water Conservation, Chinese Academy of Sciences and Ministry of Water Resources, Yangling 712100,China.* [2] *College of Urban and Environment, Northwest University, Xi'an 710127,China.*

**e-mail: montez@nwu.edu.cn*

Received 9 October 2012, accepted 16 January 2013.

Abstract

Digital elevation models have been used in many fields of research related to terrain analysis. The number of source point data in interpolating DEMs is an important factor influencing the quality of DEMs. In order to analyze the influence of source data density for the quality of DEMs, two series of DEMs which have different grid size, ANUDEMs and TIN-DEMs, are generated by the data with different source point densities, which are extracted from aero photos by digital photogrammetric method. The analysis for the relationship between grid size and RMSE of elevation shows that 2.5 m is an appropriate size for building DEM at 1:10,000 scale levels. Another result about relationship among source point density, grid size and RMSE of elevation is that the key point density for building DEMs, with which RMSE of elevation is higher than 2 m and grid size is smaller than 5 m, should be more than 7000 per km^2 for ANUDEMs, and should be more than 8500 per km^2 for TIN-DEMs.

Key words: Digital elevation models, source data density, grid size, RMSE of elevation.

Introduction

At present, as a kind of basic data source, digital elevation models (DEMs) , especially raster-based DEMs have been widely used in many research domains related to terrain, such as hydrographic, climate, soil erosion and so on. When applying DEMs to studies, an agreement has been achieved that the resolution of DEMs influences the results the analysis to a great extent [1-6]. Therefore, how to determine the resolution of DEMs reasonably has been a focus of many researches of DEMs.

Some researchers expressed their views by using the traditional method. Because a large number of DEMs are created by using the contours recorded by a topographic map, the property of contours becomes their main basis to decide the resolution of DEMs. Li [7,8] presented a formula to decide the resolution when setting up DEMs with topographic maps. According to the formula, the grid size of DEMs is decided by the interval of contours and the average slope of ground. Hengl [9] also gave a simple formula to determine the gird size. In contrast to Li's formula, he used the total size of the study area and total cumulative length of all digitized contours instead of the interval of contours and the average slope.

In fact, what the contours expressed is the properties of terrain. Therefore, some researchers point that the topographic properties determine the final choice of grid size of DEMs. For example, Hutchinson [10] suggested another method to determine the optimized resolution of DEMs which is based on the analysis of information content of terrain. This method used a key index, Root Mean Square slope (RMS slope), which changes with grid size, as the basis of resolution optimizing. This method has been applied to decide the resolution when establishing ANUDEM. Liu *et al.* [11] put forward one method to estimate the resolution of DEMs based on independent DEM error analysis which used some hypothetic mathematical faces to create DEMs and to extract the properties of these DEMs . Based on this method, Liu *et al.* [11] gave the formula to calculate the resolution of DEMs according to RMSE of DEM, RMSE of slope and DEM mean slope.

There are other researchers trying to solve this problem in very different ways. They investigated the effects of resolution on the accuracy of their application models and gave their conclusion of determination of resolution. Florinsky and Kuryakova [12] recommended a three-step procedure to define an appropriate DEM grid resolution for the specification of a particular landscape property, such as soil moisture.

Yang *et al.* [13] summarized the methods of determining of DEM's resolution in China. They classified those methods into three categories: 1) using the accuracy of terrain index extracted from DEMs, such as slope, to decide the grid size; 2) using the RMSE of elevation of DEM and the average slope of study area as the foundation to determine the resolution and 3) choosing the cell size according to the interval of contours or the horizontal distance between two adjacent contours.

Although many researches about the determination of DEMs' resolution have been published, there are still many problems which are not clear. One of them is whether the density of original data which is used to establish DEMs influences the choice of resolution. In other words, is there any relationship between resolution determination and density of original data? The researches mentioned above did not take this factor into account. In fact, it is well known that the more the source data is used to build DEMs, the better the quality of DEMs. However, the number of source data, which is needed to set up a good enough DEM is still unclear. This paper aims to discuss the relationship among density of source data, resolution and the quality of DEMs, and to find the critical number of source data.

Material and Methods

Study area and original data: The study area is located at the Loess Plateau Gully and Hill Region in Northern Shaanxi. It is a zone with complicated terrain and great soil loss. The area of the study zone is about 22.2 km^2, the elevation ranges from about 1000 to1300 m, and the average slope is about 16 degrees.

The original data in this study are the elevations of terrain feature points and lines (Fig. 1). Those data were obtained from digital aerial images of study area whose ground resolution is 0.36 m by the digital photogrammetric method. The whole process of data collection and the error control are according to the national standard of digital photogrammertry of China. The result of the process mentioned above is that we got 210,225 points with elevation (including the points on the terrain feature lines). The software platform used to edit data is Micro-Station and the data format is *.dgn.

Figure 1. The original data (local region).

Methods: In order to build the DEM, a series of data processing has been carried out. First, we change the format of original data from *.dgn to *.dxf. The mainly reason we do this step is that the dxf file is text format which could be dealt with easily. Then, we designed a program to read the elevations of 210,225 points of original data, and wrote a text file, which recorded the location (x, y) and the elevation (Z) of the points. Next, based on this text file, we used a program to change the density of the point. The program was designed according to the following arithmetic:

- To divide the whole region with a 10×10 grid;
- Calculate the locations of every grid's center;
- To search the nearest points to every grid's center and to pick them up as validate points according to the coordinates of every point;
- Base on the rest points, to use the n×n grid to divide them (in this study, we specified n as 50,100,150,200,250,300,350,400) , to delete the points nearest to grid's center, and to save the rest points.

After the process mentioned above, we obtained a series of point data with different densities and 94 validate points as the basis of our further work. With these point data, we used two kind of software to establish DEMs. One is ANUDEM, which is used to build hydrologically correct DEMs, and the other is ArcGIS, which could establish DEMs with Triangular Irregular Networks. In order to check the relationship between point density and resolution, we chose 10 grid size, 160, 80, 40, 20, 10, 5, 2.5, 1.25, 0.625 and 0.3125 m, when generated two kinds of DEMs mentioned above (because of limit of hardware, we did not establish ANUDEM with cell size 0.3125 m). After these steps, eventually, we got 171 DEMs with different grid sizes and different source data densities, and based on two kind of establishing way. Table 1 gives the details of the data.

Using these DEMs and validate points, we obtained the Root Mean Square Error (RMSE) of elevation of DEMs based on the following formula:

$$RMSE = \sqrt{\frac{1}{n}\sum_{k=1}^{n}(Z_k - z_k)^2} \qquad (1)$$

where the Z_k is the elevation of the location of the validate point extracted from DEMs, z_k is the elevation of the validate point corresponding to the spot. We calculated RMSE of all DEMs, and connected the point density, grid size of DEMs and RMSE of elevation as a whole to analyze the relationship between them.

Results and Discussion

Fig. 2 shows the trend of RMSE when cell size decreases. Although the point density and the methods in generating DEMs are different, the RMSE of elevation of DEMs decreases with the grid size. In the left part of the chart, the RMSE decreases rapidly as cell size decreases. Then, the rate of decrease slows down. When grid size is lower than 2.5 m, the RMSE of elevation achieves a steady state. Namely, the change of RMSE of elevation seems not to be influenced by the cell size even when some RMSE of DEMs increases slightly. According to this trend, we assume that 2.5 m is a key point for the resolution of DEM.

In Fig. 2, we also observed the impact of point density on RMSE of elevation. When the point density of source data is over 7000 per km^2, the RMSE of elevation decreases sharply, especially in the small cell size part of the chart, almost to about 1 m. For attaining the better observation of the impact of point density, we changed the chart to the form showin this chart, we could see clearly that

Table 1. Nine datasets with different point density and the DEMs derived from them.

	Data 1	Data2	Data 3	Data4	Data 5	Data 6	Data 7	Data 8	Data 9
NO. of points	133925	144641	156113	168064	180095	191678	201489	207925	210131
Point density (per km^2)	6030	6512	7029	7567	8109	8630	9072	9362	9461
Name of ANUDEM	ADEM1_r	ADEM2_r	ADEM3_r	ADEM4_r	ADEM5_r	ADEM6_r	ADEM7_r	ADEM8_r	ADEM9_r
Name of TINDEM	TDEM1_r	TDEM2_r	TDEM3_r	TDEM4_r	TDEM5_r	TDEM6_r	TDEM7_r	TDEM8_r	TDEM9_r

"r" in the name of DEMs specified the resolution of DEMs. r = 0.3125*2^n (n = 0,1,2,…,9) for TINDEM, r = 0.625*2^n (n = 0,1,2,…,8).

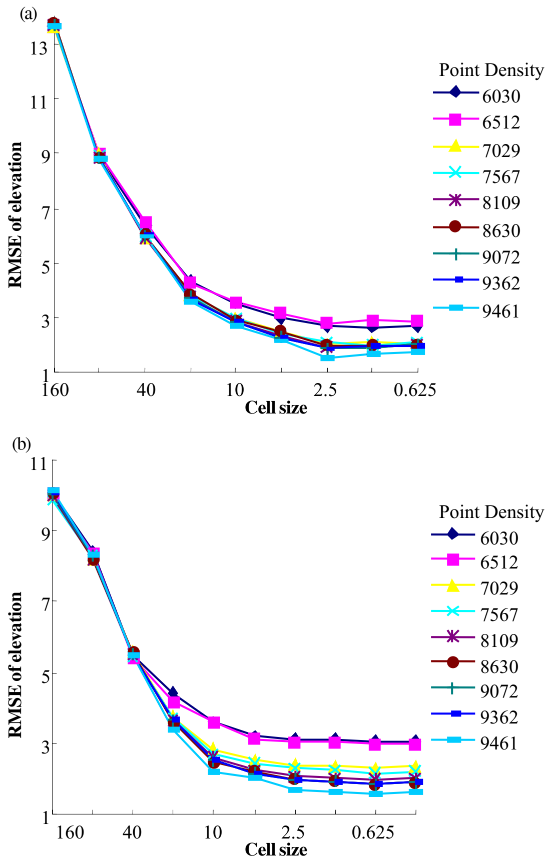

Figure 2. The variation of RMSE of elevation with cell size. Graph (a) is for ANUDEM, (b) is for TINDEM.

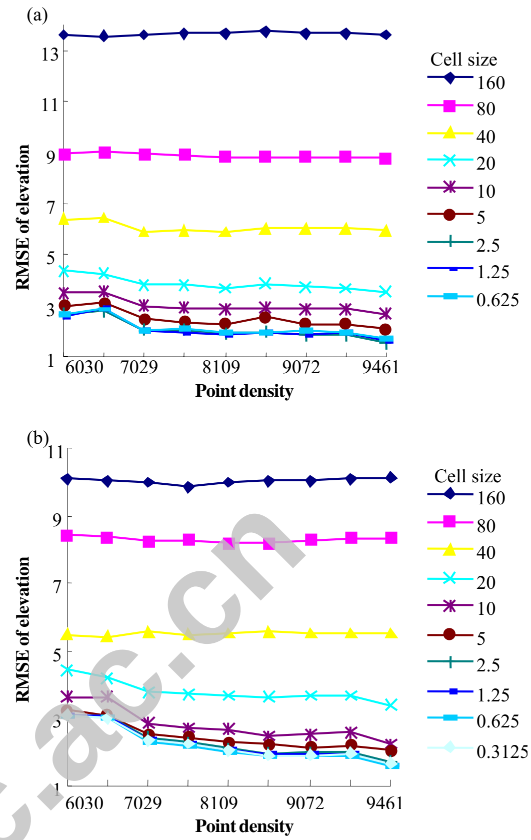

Figure 3. The variation of RMSE of elevation with the point density of: (a) is for ANUDEM, (b) is for TINDEM.

when the cell size is equal to or coarser than 40 m, no matter which kind of DEMs, ANUDEM or TIN-DEM, the impact of point density on RMSE could almost be ignored. While the cell size is smaller than 40 m, we noted the change of RMSE could be classified into three stages. When the points density is lower than 6512 per km^2, the RMSE keeps steady, while the points density changes from 6512 per km^2 to 7029 per km^2, the RMSE decreases about 1 m suddenly, and then comes to stable until the points density is larger than 9362 per km^2.At that point, the RMSE plummets again. Because the limit of original data, we have no more dense data. We don't know the trend of RMSE's change when the point density is larger than 9362 per km^2. We presumed that another stable stage should appear again. In despite of this limit, we still could infer that these two point densities, about 7000 per km^2 and 9400 per km^2, may be the significance densities which could influence the quality of DEMs remarkably in generating DEMs.

In order to see the change of RMSE of DEMs with small cell size with point density clearly and to obtain the difference of two kind of DEMs, we amplified the bottom part of the charts (the cell size is equal and lower than 5 m) in Fig. 3 and merged the two chart (Fig. 4). From this figure, we could see the three stages clearly and noticed the difference between ANUDEMs and TINDEMs. According to Figs 3 and 4, when the cell size is equal to or lager than 10 m, the RMSE of ANUDEMs is bigger than those of TINDEMs in spite of point density. However, the trend begins to change when the cell size becomes finer. If the point density is lower than 8000 per km^2, the RMSE of ANUDEMs is lower than those of TINDEMs. While the point density becomes denser, the RMSE of these two kinds of DEMs come to the same level.

Another fact we noticed from Fig. 4 is that the DEMs with the smallest grid size don't have the lowest RMSE. For TINDEM, the lowest RMSE attaches to the cell size 0.625 m but not to 0.3125 m, while the lowest RMSE appears at 2.5 m level instead of 0.625 m. This fact seems to prove our speculation that the DEMs with high resolution are not always better than those with coarse grid size.

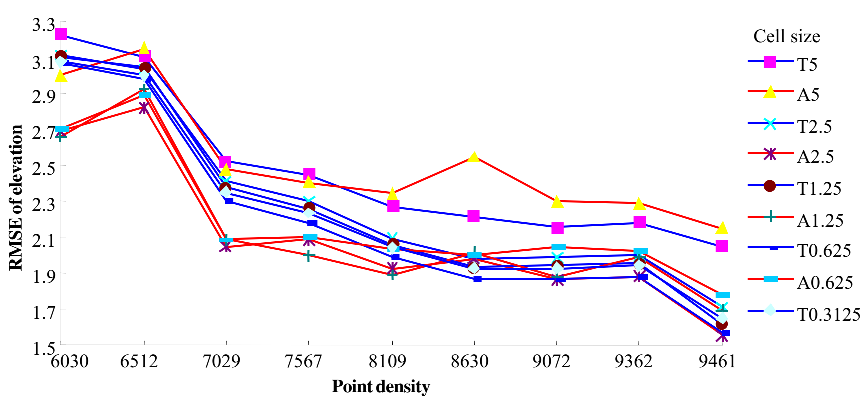

Figure 4. The comparison between the RMSE of elevation of ANUDEM and those of TINDEM. In legend, “T” specified TINDEM, “A” specified ANUDEM.

Conclusions

The research reported in this paper discussed how the RMSE of elevation of DEMs varies with the density of original data and grid size. On the basis of this study, the following conclusions could be reached:

1. To get the DEMs whose RMSE of elevation is about 2 m or better, for establishing ANUDEM, it is necessary that the density of original point data be higher than 7000 per km^2; while for TINDEM, the density should be more than 8500 per km^2.
2. No matter which kinds of DEMs are used, it is not that the smaller grid size of DEMs is, the lower RMSE of elevation. For ANUDEM, about 2.5m gird size is small enough, while the size for TINDEM is 0.625 m.
3. When the grid size is coarse (lower than 10 m), establishing ANUDEM or TINDEM has no significant difference. However, to obtain the DEMs with small gird size, ANUDEM is a better choice than TINDEM according to RMSE of elevation. Because generating an ANUDEM with the same level or better RMSE of elevation and with same grid size, fewer points are needed than generating a TINDEM.

The above conclusions are just based on the analysis of RMSE of elevation of DEMs. Although the RMSE of elevation of DEMs could not evaluate the quality of DEMs comprehensively and even some researchers questioned its reliability [14, 15], it is still an important index to measure the quality of DEMs. Further works are needed to use other indices of topography to study the relationship among the method of establishing DEMs, the resolution and the quality of DEMs.

Acknowledgements

This research would not have been possible without the support of National Natural Science Foundation of China (Grant No. 41101264, 40971173) and the Science Research Project of Shaanxi Provincial Department of Education (Grant No. 11JK0740). The Science Research Foundation of Northwest University (Grant No. 10NW07) funded us partially.

References

[1]Horritt, M. S. and Bates, P. D. 2001. Effects of spatial resolution on a raster based model of flood flow. Journal of Hydrology **253**(1-4):239-249.

[2]Thompson, J. A., Bell, J. C. and Butler, C. A. 2001. Digital elevation model resolution: Effects on terrain attribute calculation and quantitative soil-landscape modeling. Geoderma **100**(1-2):67-89.

[3]Wu W., Fan, Y., Wang, Z. and Liu, H. 2008. Assessing effects of digital elevation model resolutions on soil-landscape correlations in a hilly area. Agriculture, Ecosystems & Environment **126**(3-4):209-216.

[4]Liu, X. and Zhang, P. 2008. Effective scale of slope and aspect derived from grid-based digital elevation mode. Geomatics and Information Science of Wuhan University **33**(12):1254-1258.

[5]Wu, X. F., Liu, C. M. and Wang, Z. G. 2003. Effect of horizontal resolution of raster DEM on drainage basin characteristics. Journal of Natural Resources **18**(2):148-154.

[6]Sørensen, R. and Seibert, J. 2007. Effects of DEM resolution on the calculation of topographical indices: TWI and its components. Journal of Hydrology **347**(1-2):79-89.

[7]Li, Z. 1990. Sampling Strategy and Accuracy Assessment for Digital Terrain Modelling. Ph.D. thesis, University of Glasgow, Glasgow, 299 p.

[8]Li, Z. 1994. A comparative study of the accuracy of digital terrain models (DTMs) based on various data models. ISPRS Journal of Photogrammetry and Remote Sensing **49**(1):2-11.

[9]Hengl, T. 2006. Finding the right pixel size. Computers & Geosciences **32**(9):1283-1298.

[10]Hutchinson, M. F. 1996. A locally adaptive approach to the interpolation of digital elevation models. In Proceedings Third International Conference/Workshop on Integrating GIS and Environmental Modelling. http://www.ncgia.ucsb.edu/conf/SANTA_FE_CD-ROM/sf_papers/hutchinson_michael_dem/local.html.

[11]Liu, X. J., Gong, J.Y., Zhou, Q.M. and Tang, G. A. 2004. Research on error of derived slope and aspect related to DEM data properties. Geography and Geo-Information Science **20**(6):15-39.

[12]Florinsky, I. V. and Kuryakova, G. A. 2000. Determination of grid size for digital terrain modelling in landscape investigations - exemplifed by soil moisture distribution at a micro-scale. International Journal of Geographical Information Science **14**(8):815-832.

[13]Yang, Q. K., Zhang, C.X., Li, L.T., McVicar, T. R. and Van Nie, T. G. 2006. Optimizing DEM resolution with information content analysis. Journal of Yangtze River Scientific Research Institute **23**(5):21-23.

[14]Hu, P., Wu, Y. and Hu, H. 2003. Research on fundamental theory of assessing the accuracy of DEMs. Geo-information Science **5**(3):65-70.

[15]Hu, P., Wu, Y. and Hu, H. 2005. A new research on fundamental theory of assessing the accuracy of DEMs. Geo-information Science **7**(3):28-33.