Photogrammetric Record, 12(72): 873–877 (October 1988)

Shorter Contributions

ON THE MEASURE OF DIGITAL TERRAIN MODEL ACCURACY

by ZHILIN LI
University of Glasgow

Abstract

In digital terrain model (DTM) production, one important concern is the accuracy and a comprehensive measure of DTM accuracy is of significance. In this paper, the existing measures are evaluated; the potential measures are also discussed and, finally, a more comprehensive measure is recommended.

INTRODUCTION

THE increasing application of digital terrain models (DTMs) in many disciplines underlines the need for the assessment of DTM accuracy. *Accuracy* in relation to DTMs is a term very frequently and widely used. Some terms, such as *fidelity* and so-called *root mean square error*, have also been used but the comprehensiveness of these measures is rarely discussed. This is, indeed, a most important aspect of both DTM producers and users, since a comprehensive precise definition of this concept would provide a sound basis for the discussion of the relative values of various methods of sampling and interpolation and it might also provide a specification for contracts for DTM production by commercial firms.

The aim of this paper is to evaluate the existing measures used for assessing DTM accuracy and to discuss other possibilities. The paper concludes with a recommendation for a more comprehensive measure of DTM accuracy.

EXISTING MEASURES

From DTM literature, it can be found that two measures, fidelity and so-called root mean square error (r.m.s.e.), have been used to assess the accuracy of DTMs; the latter is more widely used.

The term *fidelity* refers to the amount of information transferred from the data source to the reconstructed data. Makarovič (1972) first used this term to measure the information transfer in the reconstruction of data from sampled points. He assumed that a terrain profile can be represented by a Fourier series, then discussed the sampling from a sine wave. If the sampling process is considered as a system, then the sinusoidal curve is the input and the output is a set of points (A, B, C, D, E and F in Fig. 1). Profile $ABCDEF$, reconstructed by linear interpolation, is an approximation to the sinusoidal input. In the figure, Δx is the step (sampling distance), and δy is the height error at X_i, that is the height difference between the sine wave and the reconstructed profile. Suppose a is the original amplitude of the sine wave and m is the estimated mean error level over a sufficient length of the sine wave. Then F in the expression

$$F=(a-m)/a \tag{1}$$

represents the fidelity of the reconstructed data. This measure is not very satisfactory because the terrain profile is normally not periodic. It is not widely used in practice.

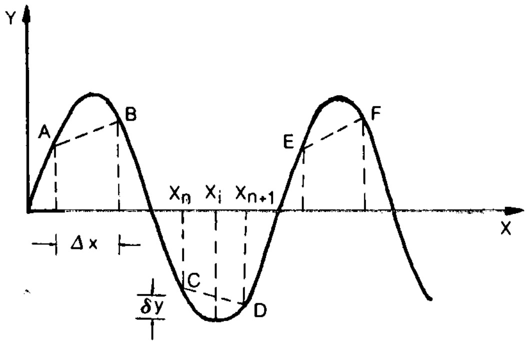

FIG. 1. Sampling from a sine wave and the reconstruction.

The so-called *root mean square error* is the measure most frequently used, both in experimental and in theoretical analysis of DTM accuracy. Fig. 2 shows the principle (in a profile) of the experimental analysis. In this figure, M is the mathematical function constructed using points A, B, C and D; T is the terrain surface. Points 1, 2, . . . , 7 are the check points. The height differences at these points between M and T are DH_1, DH_2, . . . , DH_7. Then the following formula is used to compute the r.m.s.e.

$$\text{r.m.s.e.} = \sqrt{[DH_i^2]/N}, \tag{2}$$

where DH_i is the ith height difference; N is the number of DH; and [ ] denotes summation. In this example,

$$\text{r.m.s.e.} = \sqrt{[DH_1^2 + DH_2^2 + \ldots + DH_7^2]/7}.$$

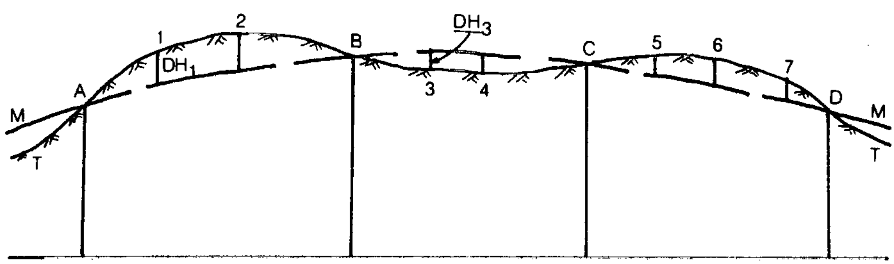

FIG. 2. The terrain surface, DTM surface and check points.

In a theoretical investigation, Tempfli (1980) applies spectral analysis to this profile and he regards the terrain profile as a univariate continuous space signal $f(x)$. Suppose the reconstructed curve is $f'(x)$, then he treats the difference $e(x)$,

$$\text{where } e(x) = f(x) - f'(x) \tag{3}$$

as the error of reconstruction. Finally, he uses the mean square error (m.s.e.) as the accuracy measure.

$$\text{m.s.e.} = \int_0^L e^2(x)\, dx/L \tag{4}$$

where [0, L] is the interval on which $f(x)$ is given. The concept of this measure is identical to the r.m.s.e. used previously. From the discussion in the next section, it will be seen that such a measure is not always appropriate in a statistical sense.

Statistical Concepts applied to DTM Accuracy

Statistics is a science dealing with random variables, so it seems pertinent to first introduce the concept of a random variable. A *random variable* is a variable X which may take many possible values ($X_1, X_2, \ldots, X_n$), each being associated with a probability ($P(X_1), P(X_2), \ldots, P(X_n)$). The set of possible values of the variable is referred to as the sample space (or population). The characteristics of this space are studied by analysing a subset sampled from it. In particular, an attempt is made to find:

(a) the magnitude of the random variable; and

(b) the spread or dispersion of the random variable.

To measure the former, some parameters can be used such as the extreme values (X_{min} and X_{max}), the mode (the most likely value), the median (the frequency centre) and the mathematical expectation (weighted average). The last is adopted in most cases.

To measure dispersion, some parameters such as range ($X_{max} - X_{min}$), the expected absolute deviation and the standard deviation can be used. Among them, the last is the best in most cases.

In the DTM case, in order to measure the accuracy of the DTM surface, some check points are used. However, the problem that arises is which should be considered as the random variable. Should it be the height on the terrain surface, the height on the DTM surface, the height difference, all three, or something else? Some possibilities will be discussed in the next paragraphs.

Case 1

The height on the terrain surface is considered as the true value and the height on the DTM surface is considered as a random variable. In this case, for every check point, there is a corresponding true value, so the height of every DTM point is a random variable. So this is a multivariable problem. The joint distribution of these variables needs to be taken into consideration for such a problem. This complicates the matter. This will also happen when the DTM height is considered as the true value and the terrain height is a variable.

With regard to the use of r.m.s.e., each DH is considered as the deviation of the DTM point from its true value. In this case, it is a multivariable problem but it is treated as a univariable problem by the assumption that the DH are normally distributed with zero mean. It has been found that this assumption is not always valid (Torlegård *et al.*, 1986). Therefore this measure is not always appropriate in a statistical sense, although it does give some information about the accuracy of DTMs.

Case 2

Another line of thought is that the height difference between the terrain surface and the DTM surface is the random variable. Thus the sample of height differences (DH) may be used to estimate the DH on the whole surface. Based on this approach, some new measures of DTM accuracy will be introduced in the next section.

Potential Measures of DTM Accuracy

Considering the height difference (DH) between the DTM surface and the terrain surface as the random variable, the following statistical measures may be utilised.

The two most *extreme values* of a value set of the variable DH indicate the general location of all the other values. The *range R*

$$\text{where } R = DH_{max} - DH_{min} \tag{5}$$

might also be taken as a measure of the dispersion of the random variable. This is used in the sense that DH_{max} and DH_{min} are the values of DH most distant from each other, the measure of spread of DH.

The use of range may lead to a specification of DTM accuracy something like the American National Map Accuracy Standard. But some characteristics of this measure are particularly objectionable:

(a) the value R depends on only two values of the random variable; others are ignored and hence it may be a poor measure; and

(b) the probability of the values in DH is ignored.

Another powerful measure is the *mathematical expectation and standard deviation*. Considering DH as the random variable, then DH has a value set DH_1, $DH_2, \ldots DH_N$, each occurring once in this set. In other words, each is with a probability of $1/N$ if N check points are used. Thus, by definition, the mathematical expectation (u) and the standard deviation ($SD(DH)$) of this variable may be computed as follows:

$$u=[DH_i]/N \tag{6}$$

$$\text{and } SD(DH)=\sqrt{[DH_i-u]^2/N}. \tag{7}$$

According to Chebyshev's theorem, most of the probability distribution is massed within the $4.SD$ distance from u. Thus $SD(DH)$ gives strong limits to the range or the dispersion of DH with respect to the probability of DH. Chebyshev's theorem states that the probability is at least as large as $1-1/k^2$ that an observation of a random variable X will be within the range from $u-k.SD(X)$ to $u+k.SD(X)$; or

$$P(|X-u|>k.SD(X))<1/k^2$$

where k is any constant larger than or equal to 1. If the normal distribution is used to approximate the distribution of this variable DH, the standard deviation computed by equation (7) has the special meaning which is familiar to us.

Consequently, a combination of the mathematical expectation and standard deviation can be used as a measure of DTM accuracy as follows:

$$Ac=u\pm SD(DH). \tag{8}$$

Discussion and Conclusion

From the discussion in the preceding sections, it can be concluded that the existing measure, the r.m.s.e., is not always appropriate in a statistical sense. The range of extreme values can also be used as a measure, but is objectionable in some respects. The expression (8) is the most comprehensive measure and thus the best measure of DTM accuracy. This measure has the following characteristics:

(a) the mean represents a shift of the surface. If $u=0$, then the $SD(DH)$ is equal to the r.m.s.e. of the existing measure. The shift may be due to improper position of the control points (as in Fig. 2) as well as systematic errors;

(b) $SD(DH)$ is a measure of dispersion. If the sampling is adequate then this might be a measure of the *goodness of fit* of the interpolation (mathematical) function to the terrain surface.

In DTM production, the user may specify the requirements for u and $SD(DH)$.

Acknowledgement

The author would like to express his thanks to Mr. B. D. F. Methley for his assistance in writing this paper and to Mr. A. Azizi for his help. Thanks are also due to Professor F. Ackermann of the University of Stuttgart for his comments.

References

Makarovič, B., 1972. Information transfer in reconstruction of data from sampled points. *Photogrammetria*, 28(4): 111–130.

Tempfli, K., 1980. Spectral analysis of terrain relief for the accuracy estimation of digital terrain models. *I.T.C. Journal*, 1980-3: 478–510.

TORLEGÅRD, K., ÖSTMAN, A. and LINDGREN, R., 1986. A comparative test of photogrammetrically sampled digital elevation models. *Photogrammetria*, 41(1): 1–16.

Résumé

La connaissance de la précision des modèles numériques du terrain (MNT) est une préoccupation importante lorsque l'on produit un MNT. Aussi est-il fondamental de savoir mesurer globalement la précision d'un tel MNT.

On examine dans cet article les diverses méthodes de mesure existant actuellement ainsi que celles que l'on pourrait envisager. Il en résulte une recommandation sur une méthode de mesure plus complète.

Zusammenfassung

Bei der Erzeugung digitaler Geländemodelle (DTM) ist ein wichtiger Gesichtspunkt die Genauigkeit, und ein umfassendes Maß der Genauigkeit von DTM ist von Bedeutung. In dem Artikel werden die vorhandenen Meßmöglichkeiten bewertet und diskutiert. Eine mehr umfassende Bewertungsmöglichkeit wird empfohlen.