Photogrammetric Record, 14(82): 651–660 (October 1993)

THEORETICAL MODELS OF THE ACCURACY OF DIGITAL TERRAIN MODELS: AN EVALUATION AND SOME OBSERVATIONS

By Zhilin Li
Technische Universität Berlin

Abstract

This paper is an attempt to evaluate existing theoretical models for the accuracy of digital terrain models (DTM) both through a theoretical analysis and by experimental tests. In order to extend the discussion, some comments on the general form in which a DTM accuracy model should be constructed are also made and a possible solution is offered.

Introduction

The subject of optimum sampling interval is of particular interest for digital terrain modelling. In order to produce such an interval for a digital terrain model (DTM) project, a mathematical model for the prediction of DTM accuracy is required. Some work in this respect has been carried out by Balce (1987) using a number of the existing DTM accuracy models. However, it is very disappointing to note that no single accuracy model is capable of producing reliable prediction. The values estimated by these models differed greatly by more than 2·5 times in some cases. Therefore, the question arises "Why and how has this happened"? For such a question to be answered, a systematic evaluation of these models both theoretically and experimentally is required, which is attempted in this paper. In addition, some observations on DTM accuracy models are also made and a solution is offered.

A mathematical model of DTM accuracy could be either a model obtained through theoretical analysis (a theoretical model) or a model based on experience (an empirical model); only a theoretical model is considered in this study. A theoretical evaluation is conducted first, followed by an experimental evaluation. Some observations are then made on the general form of DTM accuracy models and finally a possible solution is presented.

Theoretical Evaluation of Existing Theoretical Models

The task of establishing mathematical models for investigating the height accuracy of digital terrain models through a theoretical analysis started in the early 1970s. A number of attempts have been made by several investigators, such as Makarovič (1972), Kubik and Botman (1976), Frederiksen (1980), Tempfli (1980) and Frederiksen *et al.* (1986). By means of various mathematical tools and making certain assumptions, some mathematical models have already been established for the prediction of DTM accuracy. These existing models are reviewed and evaluated by a theoretical analysis.

Models Based on Fourier Analysis

Makarovič (1972) used Fourier analysis to investigate the fidelity of the DTM surface. He considered sampling and reconstruction from sinusoidal functions. The fidelity of the reconstructed surface is represented by the ratio of the mean value of themagnitude of the linearly constructed sinusoidal waves to the amplitude of the input waves. Transfer functions can also be derived for different interpolation techniques. Makarovič (1974) then tried to convert the fidelity figures into standard deviation values. In this way, the accuracies of different digital terrain models can also be compared for different types of terrain surface. Ackermann (1980) stated "In principle, this theory is complete. . . . The task remains to investigate the frequency distribution of different terrain types and to relate the corresponding theoretical and empirical accuracy results".

Tempfli (1980) considered the digital terrain modelling system as a linear system and then tried to estimate the accuracy of a DTM by a spectral analysis of such a linear system. To the author, this model does not seem to be very different from the approach taken by Makarovič. In both cases, no convenient mathematical expressions have been derived for practical application.

Models Based on (Auto)covariance and Variogram

Kubik and Botman (1976) have made a study of the accuracy of DTMs using different interpolation techniques using the (auto) covariance of data heights as a terrain descriptor. The covariance values are approximated by either exponential or Gaussian function. In this case, simple mathematical expressions have been derived for practical use. The problems remain as to how well covariance functions can describe real terrain and which function should be assigned to different types of terrain. The experience gained by the author seems to show that it is extremely difficult to have a good approximation of covariance values using either of these two functions and the final predictions are not very reliable.

In a manner similar to that used in their work on using covariance, Kubik and his collaborators (Frederiksen *et al.*, 1986) also tried later to use the variogram as a terrain descriptor. They connected this variable to the covariance which was used by Kubik and Botman (1976) to produce yet another model for DTM accuracy prediction. Again, the problems remain concerning how well the variogram functions can describe real terrain and which function should be assigned to different types of terrain.

Model Based on High Frequency Spectral Analysis

Frederiksen (1980) has also designed a mathematical model on the basis of the summation of Fourier spectra of terrain profiles in their high frequency part, in other words those regions higher than $1/(2Dx)$, where Dx is the sampling interval. However, it seems to the author that this model ignores the fact that the magnitude of the spectra will also be reduced if the spectra are computed from data points with a larger interval between these points. As a result, this model may well produce too optimistic a prediction.

Experimental Evaluations of Existing Accuracy Models

Several mathematical models of DTM accuracy have been outlined briefly and an attempt is now made to study how well two models can produce predictions in practice. The evaluation has been limited to models based on variogram analysis, and the summation of Fourier spectra over their high frequency parts, because no convenient expressions have been produced from models based on Fourier analysis for practical use; models based on covariance usually produce very poor results. In this evaluation, only the final expressions of theoretical models are quoted and reference should be made to the original papers cited as references for further details.

Evaluation of the Models based on Variogram Function

Three grid data sets have been used for this study, one for each of the three areas (Uppland, Sohnstetten and Spitze) which have been acquired for the ISPRS DTM test. From each of these three data sets, several new grid data sets with larger intervals have also been generated. The results from both the original and the generated grids have been used for this study. Information concerning these data sets and how the tests were carried out have been described in Li (1990 and 1992).

The so-called variogram is the mean square difference between two data values which are a certain distance, d, apart mathematically, such that

$$2r(d) = \frac{1}{N}\sum_{i=1}^{N}(Z_i - Z_{i+d})^2 \tag{1}$$

where Z_i and Z_{i+d} are the two heights with an interval of d and $r(d)$ is the resulting semi-variogram. It is then assumed that the semi-variogram, $r(d)$, can be approximated by the following expression:

$$2r(d) = Ad^b \tag{2}$$

where A and b are two constants. The mathematical expression of an accuracy model based on variogram is then expressed as follows (Frederiksen *et al.*, 1986):

$$V_{int} = A\left(\frac{Dx}{L}\right)^b\left(-\frac{1}{6} + \frac{2}{(b+1)(b+2)}\right) \tag{3}$$

where V_{int} denotes the DTM accuracy in terms of error variance without taking the errors in the raw data into consideration, Dx is the sampling interval of the raw data and L is the sampling interval of the profiles which were used to compute the two parameters A and b in equation (2). The final expression is

$$V_{DTM} = V_{raw} + V_{int} \tag{4}$$

where V_{DTM} is the accuracy of the resulting DTM, V_{raw} is the accuracy of measured raw data and V_{int} is the accuracy loss due to sampling and reconstruction; all these values are expressed in terms of error variance.

The semi-variograms computed from the three data sets are shown in Fig. 1. For each area, the variogram values produced from the data set with the smallest sampling interval (expressed by solid lines) have been used to compute the coefficients (A and b in equation (2)) using a regression technique. Some examples of the comparison of the accuracy results predicted by this model with the experimental results which were described in Li (1990, 1992) (in terms of standard deviation) are shown in Table I. It can be seen that the predicted accuracy results seem to be in reasonable agreement with the actual results obtained from the experimental test results, except for a comparatively large grid (56·56m) for the Sohnstetten area which has a relatively steep slope. However, when the DTM accuracy figures obtained from the composite data sets were used, the differences were found to be much greater overall (Table I). It can be seen that some results are even unacceptable for the case of a DTM derived from composite data sets.

Table I. Comparison of experimental results with accuracies predicted by the model based on the variogram function.

Test area	*Grid interval (m)*	*Predicted (m)*	*Gridded data* *Tested (m)*	*Gridded data* *Difference (m)*	*Composite data* *Tested (m)*	*Composite data* *Difference (m)*
Uppland	40	1·04	0·76	0·28	0·66	0·38
	56·56	1·18	0·93	0·25	0·70	0·48
	80	1·38	1·18	0·20	0·80	0·58
Sohnstetten	20	0·74	0·56	0·18	0·43	0·31
	28·28	0·98	0·87	0·11	0·56	0·42
	40	1·38	1·45	−0·07	0·78	0·60
	56·56	1·77	2·40	−0·63	1·08	0·69
Spitze	10	0·29	0·21	0·08	0·16	0·13
	14·14	0·37	0·28	0·09	0·17	0·20
	20	0·48	0·35	0·13	0·18	0·30

It is interesting to note that accuracy predictions from this model are closer to the actual results obtained from the grid data sets, whereas they are in very poor agreement with the results obtained from the composite data sets. This might be due to the fact that the values of the variogram used in this model were computed from gridded data sets

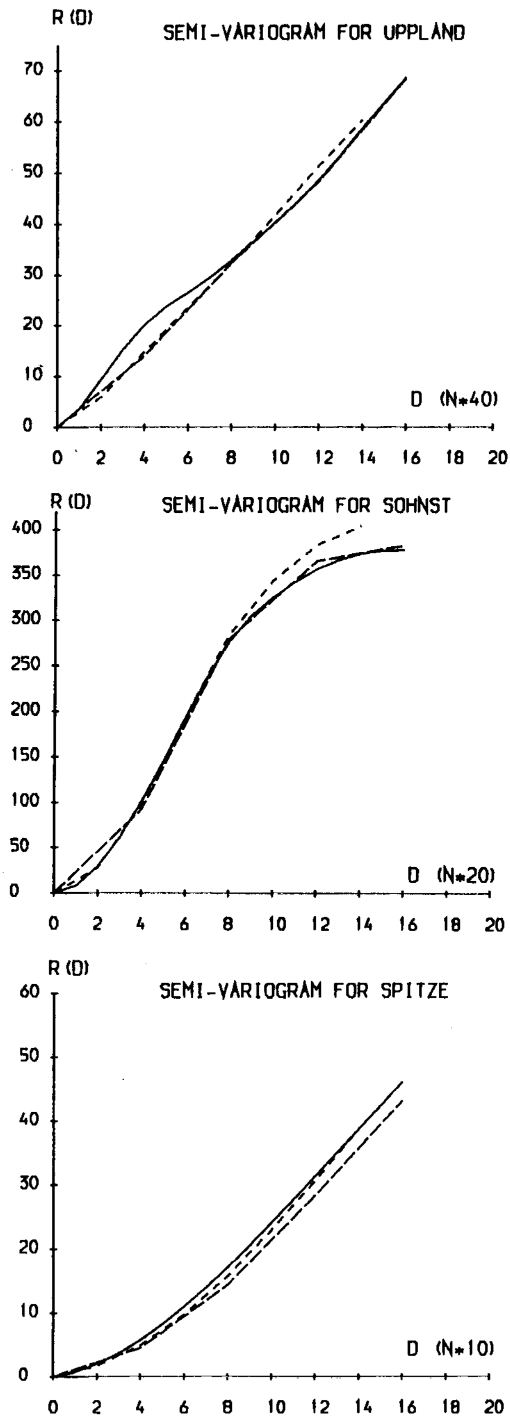

FIG. 1. Semi-variograms for the three test areas, computed from data sets with various grid intervals.

only and not from the composite data sets, since it is complicated and difficult to compute variograms from non-gridded data sets. If the variogram rules had been computed from grids with very small grid intervals, then the prediction produced by this model might be closer to the results obtained from composite data sets.

Evaluation of the Model based on High Frequency Spectral Analysis

The mathematical expression of this model is as follows:

$$V_{DTM} = V_{raw} + \sum_{\lambda = 2Dx}^{0} P_{\lambda} \tag{5}$$

where P_{λ} is the spectral value corresponding to the wavelength λ, Dx is the sampling interval, V_{raw} is the error variance of raw data and V_{DTM} is the error variance of the final DTM.

Some high density profiles were required to evaluate this model. However, no such profiles were available for the data sets used for the evaluation of the previous model and other data sets were therefore employed. These data sets were also used for the ISPRS DTM test. Two areas were the same as for the data sets used in the previous section. However, the grid intervals and measuring accuracies were different. Details of these tests have been described by Li (1990) and are not repeated here. A comparison of the predicted values with the results obtained from experimental tests is given in Table II. In general, it can be seen that the results are not too good but, in the case of the composite data, might be considered as relatively good.

TABLE II. Comparison of test results with the values predicted by the model based on the summation of high frequency Fourier spectra.

			Gridded data		*Composite data*	
Test area	*Grid (m)*	*Predicted (m)*	*Tested (m)*	*Difference (m)*	*Tested (m)*	*Difference (m)*
Sohnstetten	15	0·26	0·46	− 0·20	0·35	− 0·09
Spitze	15	0·10	0·31	− 0·21	0·20	− 0·10
Drivdalen	20	1·25	1·57	− 0·32	1·47	− 0·22

It should be noted that the results predicted by this model are closer to the accuracy figures obtained from composite data sets. This may be due to the fact that the spectra here were computed from high density profile data. In such a case, it is likely that information about feature specific points and lines was to some extent included in the profile data. Thus the results produced by this model are very close to those obtained from the composite data sets. It is also interesting to note that, in the case of these limited results, this model, as expected, always produced too optimistic a prediction.

Discussion

Two existing mathematical model of DTM accuracy have been evaluated experimentally. From the comparisons of the values predicted by these models with the test results, it can be found that these models behave very differently. The results predicted by the model based on the variogram function seem to be quite close to those obtained from gridded data sets, but not to those resulting from the composite data sets. By contrast, the results predicted by the model based on spectral analysis are closer to those obtained from the composite data sets.

It must be emphasised that the parameters of the model based on variogram analysis were estimated from the whole set of data points. In practice, it is impossible to do this with confidence since the DTM accuracy for a given sampling interval needs to be predicted before the actual measurement of the data points can be carried out. Thus it may not be easy to obtain reliable estimates for the parameters in this model, which may therefore be impractical to implement, even though it may produce reasonable results in some cases.

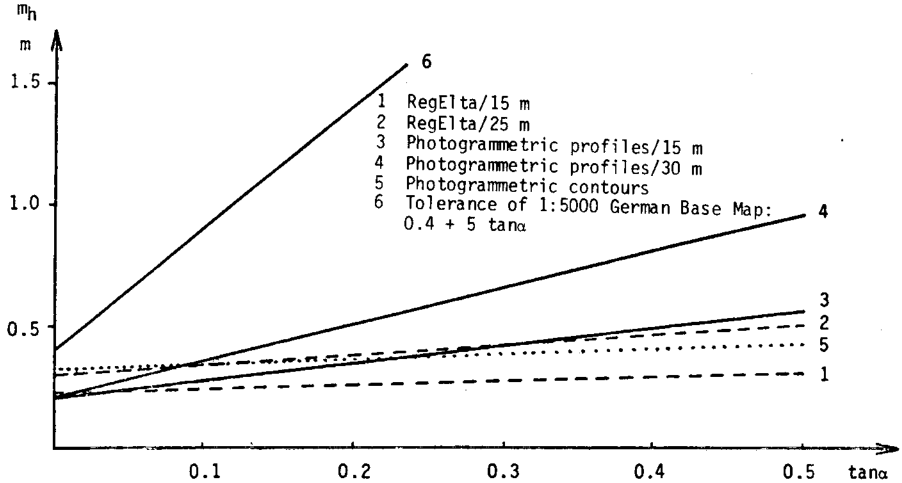

FIG. 2. DTM accuracy (r.m.s.e.) varying with slope angle for DTMs derived from various types of data. (After Ackermann, 1980.)

SOME OBSERVATIONS ON THEORETICAL MODELS OF DTM ACCURACY

It has been shown that among existing accuracy models, some are not convenient for practical application, some are relatively weak in theory and some are not capable of producing reliable prediction. Therefore, these models do not seem to be suitable for practical use. It can be argued that new mathematical models should be considered and the question then arises as to what kind of models should be sought.

A General Form of an Accuracy Model

Firstly, the general form in which DTM accuracy models should be expressed needs to be considered. From experience, a formula such as the following expression can be suggested:

$$V_{DTM} = KV_{raw} + V_{loss} \tag{6}$$

where K is a constant between 0 and 1, V_{DTM} is the accuracy of the final DTM, V_{raw} is the accuracy of raw data and V_{loss} is the accuracy loss due to sampling and reconstruction, all in terms of error variance.

Up to this stage, the problem arising concerns the kind of form that is suitable for V_{loss}. This question may be answered either based on experimental tests or theoretical analysis or a combination of both methods. Based on experimental tests, Ackermann (1980) found that the relationship between final DTM accuracy and sampling interval is linear; he therefore considered the following as the general form:

$$V_{DTM} = V_{raw} + (C \times Dx)^2 \tag{7}$$

where C is a parameter depending on the characteristics of terrain surface, Dx denotes the sampling interval, V_{DTM} is the error variance of the final DTM and V_{raw} is the error variance of raw data. Based on the arguments presented earlier, the following model seems more appropriate:

$$V_{DTM} = K_1 V_{raw} + K_2 (C \times Dx)^2 \tag{8}$$

where K_1 and K_2 are constants depending on the interpolation method used.

Equation (8) should represent the general form of a DTM accuracy model. The problem now occurs concerning what kind of function should be used to represent C in equation (8). From experimental tests, Ackermann (1980) has demonstrated (Fig. 2) that there is a linear relationship between the accuracy of a DTM and the value of $\tan\alpha$, where α is the mean slope of the area. This relationship has been confirmed by Li (1990) as existing in the case of DTMs derived from composite data (regularly gridded data with additional feature specific data). However, in the case of DTMs derived from grid data only, C is a function of not only mean slope, but also the mean wavelength (horizontal variation) and the grid interval. Therefore, in general, it can be written that

$$C = f(\tan\alpha, W, Dx) \tag{9}$$

where α is the mean slope value, W is the mean wavelength value and Dx is the sampling interval. The remaining problem is related to the adequacy of slope and wavelength as terrain descriptors.

The Adequacy of the General Model

It has been realized that the roughness, or complexity, of a topographic surface cannot be completely defined by any single parameter, but must be represented by a roughness vector or set of parameters (Mark, 1975). In this set of parameters, relief is used to describe the vertical dimension (or amplitude of the topography), while the term wavelength is used to describe the horizontal variation. The parameters for these two dimensions are connected by slope. Since these three variables are not independent, only two are required. For many reasons, a combination of slope and wavelength can be recommended as the main terrain descriptors for DTM purposes (Li, 1990), although some researchers (Makarovič, 1990) may feel that they are still insufficient.

In practical terms, the relief parameter for a given area will often be known. Therefore the value of the corresponding wavelength will also be known if the slope value is known. Hence the adequacy of this descriptor depends primarily on the availability of slope values for acceptable terrain description.

The efficacy of using slope as the main terrain descriptor for DTM purposes can be justified for the following reasons.

(i) In geomorphology, it has long been recognised that slope is a very powerful and very effective terrain descriptor. As quoted by Evans (1972), Strahler (1956) pointed out that "slope is perhaps the most important aspect of surface form, since surfaces may be formed completely from slope angles . . .".

(ii) Slope is the first derivative of altitude on the terrain surface. It shows the rate of change of height of the terrain surface with distance.

(iii) Traditionally, slope has also been recognised by the mapping community to be an adequate terrain descriptor and is well utilized in mapping practice. For example, map accuracy specifications for contours are given worldwide in terms of slope angle.

(iv) Most importantly, in DTM practice, Ackermann (1980) and Ley (1986) found that a high correlation exists between the vertical error of a regional DTM and the mean slope of the region. Ley (1986) concluded that it is possible to predict the vertical accuracy of DTMs purely by analysing the mean slope of the model.

The Advantages of the General Accuracy Model over Others

The advantages of the general accuracy model over others rests on the advantages of slope as a terrain descriptor over other descriptors. It seems to the author that slope has the following advantages.

(i) Cartographers have considerable experience in dealing with slopes when making topographic maps. Thus prior knowledge of slope is available that can be put to practical use.

(ii) There is evidence of a high correlation between slope and relief (Evans, 1972). Therefore relief information may be used as a rough guide for the value of slope. For example, in deciding the contour intervals to be used in small scale topographic mapping, Imhof (1965) takes 45° as the average slope value used for high mountainous areas of rugged relief, 26° for lower mountainous areas with less rugged relief and 9° for relatively flatter areas.

(iii) Due to the high correlation between slope and relief, the concept appears very intuitive, although obviously there are exceptions such as high plateaux with large elevations but possibly a relatively gentle terrain in terms of slope.

A Realization of the General Model

Based on theoretical analysis and some experimental results, Li (1990) has realized a general accuracy model. The formulae are as follows.

For the DTM surface linearly constructed from composite data

$$V_{DTM} = K_1 V_{raw} + K_2(Dx \tan\alpha)^2 \quad (10)$$

and for the DTM surfaces linearly constructed from grid data only

$$V_{DTM} = K_1 V_{raw} + K_2\left(1 + \frac{4Dx}{W}\right)(Dx \tan\alpha)^2. \quad (11)$$

Where K_1 is a constant between 0 and 1, it is suggested that a value of 4/9 might be appropriate and K_2 is also a constant, depending on the characteristics of the terrain surface and the resulting DTM surfaces. It is suggested that approximate values for the DTM surfaces comprising contiguous bilinear surfaces and contiguous triangular facets should be 5/768 and 17/3072, respectively. In equation (11) α is the mean slope angle and W is the average wavelength of the terrain variation. For relatively large areas, W can be roughly determined as follows:

$$W = (H_{max} - H_{min})\cot\alpha \quad (12)$$

where H_{max} and H_{min} are the maximum and minimum height values in the area. The definitions of the other parameters are the same as used before.

The first term on the right hand side of equations (10) and (11) is included to represent the errors propagated from the raw data, with an error variance of V_{raw}. The coefficient K_1 has been derived by taking into consideration all the possible point positions on a bilinear surface. The second term on the right hand side of equations (10) and (11) represents the errors generated by sampling and reconstruction. The coefficient K_2 has been derived from the analysis of the distribution of such errors, both theoretically and experimentally.

These models have also been experimentally evaluated by Li (1990), using the results which have been used to evaluate previously existing models. In this evaluation, $K_1 = 4/9$ and $K_2 = 17/3072$. The results of these models are given in Tables III and IV and they seem to be very satisfactory.

TABLE III. Comparison of accuracy predicted by new models with test results obtained from the first data sets.

		Gridded data			*Composite data*		
Test area	*Grid interval (m)*	*Predicted (m)*	*Tested (m)*	*Difference (m)*	*Predicted (m)*	*Tested (m)*	*Difference (m)*
Uppland	28·28	0·53	0·63	−0·10	0·50	0·59	−0·09
	40	0·62	0·76	−0·14	0·55	0·66	−0·11
	56·56	0·80	0·93	−0·13	0·63	0·70	−0·07
	80	1·16	1·18	−0·02	0·78	0·80	−0·02
Sohnstetten	20	0·58	0·56	0·02	0·42	0·43	−0·01
	28·28	0·90	0·87	0·03	0·58	0·56	0·02
	40	1·44	1·45	−0·01	0·81	0·78	0·03
	56·56	2·38	2·40	−0·02	1·15	1·08	0·07
Spitze	10	0·16	0·21	−0·05	0·11	0·16	−0·05
	14·14	0·23	0·28	−0·05	0·14	0·17	−0·03
	20	0·35	0·35	0·00	0·19	0·18	0·01

It should be noted that it is not the purpose of this paper to say why the new models are so good or how they could work well. However, it has been demonstrated that complicated models are not always necessary and, in many cases, simple models may even work better.

TABLE IV. Comparison of test results from the second data sets with the values predicted by new accuracy models.

Test area	Gridded data			Composite data		
	Predicted (m)	*Tested (m)*	*Difference (m)*	*Predicted (m)*	*Tested (m)*	*Difference (m)*
Spitze	0·25	0·31	−0·06	0·16	0·20	−0·04
Sohnstetten	0·40	0·46	−0·06	0·32	0·35	−0·03
Drivdalen	1·60	1·57	0·03	1·57	1·47	0·10

CONCLUDING REMARKS

In this paper, existing theoretical models have been evaluated. The author considers that some of these models may not be convenient for practical use, some may have a theoretical weakness and some may not be capable of producing reliable prediction. Thus, it seems to the author that they are not necessarily suitable for practical application.

In addition, some observations on theoretical models have been made and a general model has been suggested. A realization of this general model is presented; the main advantage of this realization is that the models are an analogue of a conventional map accuracy specification, and are thus very convenient for practical application. However, further research is still needed to evaluate these models intensively to see if they are really reliable.

ACKNOWLEDGEMENTS

This paper is partly based on material included in the author's Ph.D. thesis submitted to the University of Glasgow. The author would like to express his thanks to the University of Glasgow and the CVCP for their financial support of his study. He is also grateful to Professor G. Petrie for his assistance. The comments made by the referee concerning this paper were much appreciated.

REFERENCES

ACKERMANN, F., 1980. The accuracy of digital terrain models. *Proceedings of 37th Photogrammetric Week*, University of Stuttgart. 234 pages: 113–143.

BALCE, A. E., 1987. Determination of optimum sampling interval in grid digital elevation models (DEM) data acquisition. *Photogrammetric Engineering and Remote Sensing*, 53(3): 323–330.

EVANS, I., 1972. General geomorphology, derivatives of altitude and the descriptive statistics. *Spatial analysis in geomorphology* (Editor R. Chorley). Methuen & Co. Ltd., London. 393 pages: 17–90.

FREDERIKSEN, P., 1980. Terrain analysis and accuracy prediction by means of the Fourier transformation. *International Archives of Photogrammetry and Remote Sensing*, 23(4): 284–293.

FREDERIKSEN, P., JACOBI, O. and KUBIK, K., 1986. Optimum sampling spacing in digital terrain models. *Ibid.*, 26(3/1): 252–259.

IMHOF, E., 1965. *Kartographische Geländedarstellung*. Walter de Gruyter & Co., Berlin. 425 pages.

KUBIK, K. and BOTMAN, A. G., 1976. Interpolation accuracy for topographic and geological surfaces. *I.T.C. Journal*, 1976–2: 236–274.

LEY, R., 1986. Accuracy assessment of digital terrain models. *Auto Carto London*, 1: 455–464.

LI, ZHILIN, 1990. *Sampling strategy and accuracy assessment for digital terrain modelling*. Ph.D. thesis, University of Glasgow. 299 pages.

LI, ZHILIN, 1992. Variation of the accuracy of digital terrain models with sampling interval. *Photogrammetric Record*, 14(79): 113–128.

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

MAKAROVIČ, B., 1974. Conversion of fidelity into accuracy. *I.T.C. Journal*, 1974–4: 506–517.

MAKAROVIČ, B., 1990. Personal communication.

MARK, D., 1975. Geomorphological parameters: a review and evaluation. *Geografiska Annaler*, 57A: 165–177.

STRAHLER, A., 1956. Quantitative slope analysis. *Bulletin of the Geological Society of America*, 67: 571–596.

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

Résumé

On cherche à faire dans cet article, une évaluation des modèles théoriques actuellement disponibles sur la précision des modèles numériques de terrain (MNT), en procédant à la fois à une analyse théorique et à des essais expérimentaux. Afin de ne pas limiter le sujet, on fournit quelques commentaires sur la façon générale dont il faudrait établir ces modèles de précision des MNT et l'on présente une solution possible.

Zusammenfassung

Im Beitrag wird der Versuch unternommen, existierende theoretische Modelle für die Genauigkeit digitaler Geländemodelle (DTM) sowohl durch theoretische Analyse als auch experimentelle Untersuchungen zu bewerten. Um die Diskussion weiter zu beleben, werden auch einige Bemerkungen zur allgemeinen Form, nach der ein Genauigkeitsmodell für DTM aufgebaut sein sollte, gemacht und mögliche Lösungen angeboten.