last copy

Photogrammetric Record, 14(79): 113–128 (April 1992)

VARIATION OF THE ACCURACY OF DIGITAL TERRAIN MODELS WITH SAMPLING INTERVAL

By Zhilin Li†
University of Southampton

Abstract

This paper describes some experimental investigations into the variation of digital terrain model (DTM) accuracy with sampling interval, based on data acquired using stereophotogrammetric methods. The main factors affecting DTM accuracy are reviewed and various alternative schemes for carrying out experimental tests on DTM accuracy are outlined. Detailed information is given about the test areas, source data sets and check point data sets, followed by a discussion concerning the special design of this experiment. The experimental results are reported and, finally, a critical analysis of the test results is presented.

INTRODUCTION AND BACKGROUND

For a digital terrain model (DTM) project, the final objective is to produce a DTM with a required or specified accuracy, preferably in an economical and efficient manner. In other words, accuracy, cost and efficiency are the matters of most concern to both the producer and the clients involved in DTM production. Of these attributes, accuracy is probably the most important single concern since it usually lies at the core of a particular DTM project; the accuracy of final DTMs is the main subject of this paper.

The accuracy of a DTM is a result of many individual parameters. However, the thesis by Li (1990) states that the following may be regarded as the main factors:

(i) the three attributes (accuracy, density and distribution) of the source data;
(ii) the characteristics of the terrain itself;
(iii) the method used for the construction of the DTM surface; and
(iv) the characteristics of the DTM surface which is constructed from the source data.

These six factors must therefore be taken into account when the assessment of DTM accuracy is carried out, whether by theoretical analysis, or by experimental investigation or through a combination of both. However, in this study, only an experimental investigation has been performed from which some results are reported.

When attempting to investigate experimentally the variation of DTM accuracy with any one of the six main factors given above, it is necessary to isolate the effects of each of the other (five) factors by keeping them unchanged. There are six possible ways of achieving this objective.

(i) The *accuracy of the source data* could be varied while all the other factors remain unchanged. This can be achieved by using aerial photographs with different scales and different flying heights taken over the same test area,

---

†This paper describes work carried out by the author while engaged in research for a Ph.D. degree at the University of Glasgow.

employing the same sampling strategy and utilising the same modelling method to construct the same type of DTM surface.

(ii) The *density of the source data* could be varied while all the other factors remain unchanged. This can be achieved by using different sampling intervals in the case of regular grid sampling or composite sampling or alternatively by generating new grids from the original (grid-based) source data.

(iii) The *distribution of the source data* could be varied while all the other factors remain unchanged. This can be achieved by using different data patterns such as square grids, rectangular grids, triangular grids, strings obtained from profiling and so on.

(iv) The *test area* could be varied but all other factors remain unchanged. This can be achieved very easily by using several different test areas.

(v) The *method used for the construction of the DTM surface* could be varied while all the other factors remain unchanged. This can be achieved by employing as many methods as possible, for example by constructing the DTM surface either directly from the measured data or from a data set which has undergone a random-to-grid processing.

(vi) The *characteristics of the constructed DTM surface* might be varied (for example the continuity and smoothness could be changed) while all the other factors remain unchanged. This can be achieved by using different modelling approaches to construct different types of DTM surface.

In this study, no attempt has been made to carry out investigations using all of these possible procedures. Only the second alternative is given serious consideration because, as has been recognised by several investigators (for example Makarovič (1972) and Ackermann (1980)), sampling interval is the factor which potentially has the most significant effect on the accuracy of a DTM. Furthermore, only one type of DTM surface (a continuous surface comprising a series of contiguous linear facets) has been considered since it has been found that this type of surface is usually the least misleading (Peucker, 1972). Also, only the method of constructing the DTM surface directly from the measured data points (without any random-to-grid interpolation as pre-processing) is considered. As far as the source data is concerned, only data sets comprising regularly gridded data and composite data have been utilised.

On this basis, a systematic investigation into the variation of DTM accuracy with sampling interval for certain types of terrain using certain data patterns has been carried out and is presented in this paper. Information regarding the source data acquired from aerial photography taken at different scales for three different test areas is given first, followed by a discussion concerning how to generate new grids from the original grids. Subsequently, the accuracy values obtained from these experimental tests are presented and finally an analysis is conducted to try to establish the relationship between the accuracy of the DTMs with sampling interval for the various data patterns used in these tests.

DESCRIPTION OF TEST DATA

The source data sets used in this study are some that were used for the ISPRS DTM test (Torlegård *et al.*, 1986), although all the processing and the tests based on these data sets were carried out independently by the author.

Test Areas

In the DTM tests which were conducted by Working Group 3 of ISPRS Commission III, six areas were used. Data sets for five of these six areas were made available to the author for use in this study. However, it was found that only three of these data sets (those for the Uppland, Sohnstetten and Spitze areas) were suitable for this particular study due to problems related to the data patterns in which the data had been collected and/or limitations of the DTM package used in the tests.

Descriptions of these three test areas have been given by Torlegård *et al.* (1986). Therefore, detailed information is not repeated here but, for convenience, a brief summary of the nature of the three areas is included in the form of Table I.

TABLE I. Locations and descriptions of the test areas.

Test area	*Description*	*Height range (m)*
Uppland (Sweden)	Farmland and forest	7 to 53
Sohnstetten (Germany)	Hills of moderate height	538 to 647
Spitze (Germany)	Smooth terrain	202 to 242

Source Data Sets

The data sets had been measured on a Zeiss (Oberkochen) Planicomp C–100 analytical stereoplotter at the Technical University of Munich and the results were made available to the author through the courtesy of Professor H. Ebner and Dr. W. Reinhardt. The composite sampling method was used. The gridded data sets were stored separately from those of the feature-specific points (including the points measured along break lines and form lines). The grids were arbitrarily oriented. For the Uppland area only, two grids were measured with origins shifted relative to each other by 20 m in the X and Y directions. Information concerning the grid sizes of these data sets, together with the scales of photography, is given in Table II.

TABLE II. Details of source (raw) data.

	Absolute orientation accuracy (m)					
	Plan	*Height*	*Height data accuracy (m)*	*Scale of photography*	*Flying height (m)*	*Grid interval (m)*
Uppland	±0·474	±0·497	±0·67	1:30 000	4500	40
Sohnstetten	±0·110	±0·057	±0·16	1:10 000	1500	20
Spitze	±0·072	±0·048	±0·08	1: 4000	600	10

The standard errors (standard deviations) of unit weight for absolute orientation at ground scale (in which the co-ordinate observations are given unit weight) are quoted in Table II. The height accuracy of the raw data, which has been estimated by the author according to the common practice of photogrammetric measurement, is also given in Table II.

Check Points

The check points used in this test were measured from much larger scale photography at the Royal Institute of Technology (the ISPRS DTM test centre) in Stockholm and were made available to the author through the courtesy of Professor K. Torlegård and Dr. M. Li. These check points were used in the ISPRS tests and are also referred to as "ground truth" in the paper by Torlegård *et al.* (1986). Aerial photographs at scales much larger than those used for the acquisition of the DTM source data were used in order to ensure that check points did indeed have a much higher accuracy. Table III shows some of this information.

TABLE III. Details of check points.

Test area	*Scale of photography*	*Flying height (m)*	*Number of check points*	*R.m.s.e. (m)*	*Emax (m)*
Uppland	1:6000	900	2314	±0·090	0·20
Sohnstetten	1:5000	750	1892	±0·054	0·07
Spitze	1:1500	230	2115	±0·025	0·05

The check points are located in a grid pattern which is oriented in an arbitrary direction. This means that the grid directions of the check points are not parallel to those of the source data nor to the X and Y co-ordinate directions. Not every grid node was measured; points which were located on ground where no reliable measurements could be made, such as areas covered by bushes or trees and those with extremely steep slopes were not measured and their height values were recorded as zero. Overall, for each test area, more than 1700 points were measured with analytical plotters in static mode.

The accuracy of the check points was checked at the test centre and the results are given in Table III under the headings of r.m.s.e. (root mean square error) and Emax (maximum error). The check points are not distributed over the whole of each test area, but only over part of each area.

GENERATION OF NEW GRIDS FROM THE ORIGINAL GRIDDED DATA SETS

As has been stated earlier, the main purpose of this experimental investigation is to find some relationship between DTM accuracy and the density of the data set for certain types of terrain using certain data patterns. In practice, as described in the previous section, only three types of terrain and two data patterns are available to check this specific case.

The *type of terrain* is specified by its slope as suggested by Li (1990). Therefore, representative slope values have been derived from the contours for each of these three areas, with values of 6°, 15° and 7° for the Uppland, Sohnstetten and Spitze areas respectively. The two types of *data pattern* are, of course, gridded data and composite data. The data density can be defined in terms of sampling interval since the source data comprises either gridded data or gridded data plus strings.

In order to achieve the aim of this investigation, it would have been best if a few sets of gridded data (which could also be converted to composite data by adding string information) had been measured for each of three areas. This had not been done, which was not in any way the fault of the participants in the test since it was not a requirement of the ISPRS test. So, for each area, only a single specific sampling interval was used. A solution therefore had to be found to overcome this problem.

There are two possible solutions. The first is to measure *more data sets* for each test area using the same diapositives on a photogrammetric instrument. The second is to generate gridded data sets with *different grid intervals* by changing the sampling pattern of the original measured data using wider intervals or other selection criteria. Of course, the latter is not only a quicker and more economic solution but also more appropriate since, in this way, the accuracies of the (new) generated gridded data will be the same as those of the original gridded data. Therefore, the latter procedure has been selected for this study.

The method which has been implemented to generate gridded data sets with different grid intervals is illustrated in Fig. 1. The first four figures, Figs. 1(a) to (d), show the generation of new grids with twice the interval of the original grid. In these diagrams, the points marked with the symbol "o" are those which have been retained by selecting every other point along both the row and column directions. They also show that, by using a different starting point, four such alternative data sets can be generated from the original gridded data.

Figs. 1(e) and (f) illustrate the generation of grids with a grid interval of 1·414 (the square root of 2) times the original one. Again, the symbol "o" denotes those points which have been selected. The directions of the new grids have been rotated and oriented along the diagonal directions of the original grids. It can be seen from these diagrams that two new grid data sets can be generated from the original gridded data. This method can also be viewed as producing a regular triangular network comprising isosceles right-angled triangles with each hypotenuse lying along the row direction. It can also be considered that a regular grid cell is constructed by two such triangles with their common hypotenuse as a diagonal of the grid cell.

In practice, the measurement process has been carried out twice, in a forward

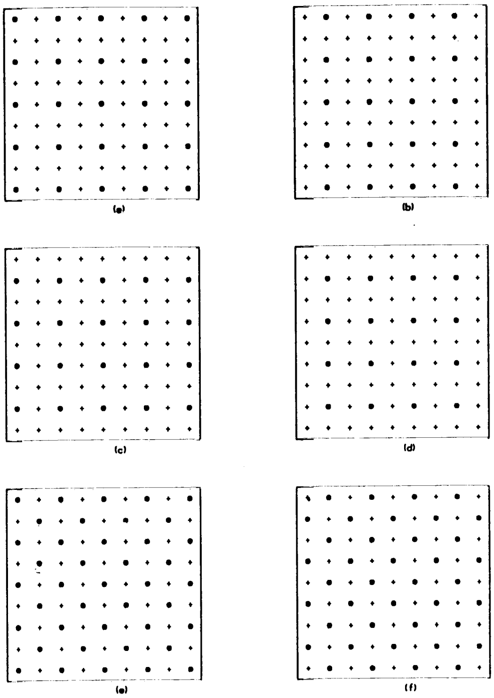

FIG. 1. Generation of new grids with different intervals.

and in a backward direction. Hence, in the generation of new grids, a corresponding rearrangement also had to be made.

Table IV lists all the test data sets generated from the original data sets (including the original data sets themselves) which are used in this experiment.

TABLE IV. Number of data sets generated from the original data sets.

Grid interval (m)	*Uppland*	*Sohnstetten*	*Spitze*
10·00	—	—	1*
14·14	—	—	2
20·00	—	1*	4
28·28	1	2	—
40·00	2*	4	—
56·56	4	8	—
80·00	8	—	—

*Original set.

RESULTS OF THE EXPERIMENTAL TESTS

In this study, a triangulation-based DTM package (Panacea), which was developed by McCullagh (1983), has been used to produce DTM data points. The processing of the data has been carried out on the ICL 3980 mainframe computer at the University of Glasgow. The test results are reported in the following sections.

Accuracy Parameters Used in this Study

In order to present the results of the experimental tests informatively, some comprehensive measures of DTM accuracy should be used. The accuracy measures which it is possible to use in practice and their value have been discussed by Li (1988). For this study, the extreme values, the mean and the standard deviation of the test results have been used. In addition, the values of root mean square error (r.m.s.e.) have also been computed since this is a traditional measure which is still widely used at present. The definitions of the terms used are as follows:

r.m.s.e. = root mean square error;
mean = average value of residuals (*DH* values);
SD = standard deviation of the residuals from the mean;
Emax = one of the two extreme values, the maximum *DH* value (or residual) of the distribution; and
−Emax = the other extreme value, the minimum DH value of the same distribution.

The values of all these parameters given in the corresponding tables are expressed in metres. The symbol ± before the values for SD and r.m.s.e. has been omitted simply for convenience in this study.

Accuracy of the DTM derived from Regularly Gridded Data Sets

The accuracy results of the digital terrain models formed from the regularly gridded data sets from the ISPRS experiment are given in Tables V, VI and VII. Diagrammatic representations of the test results are presented in Figs. 2(A), (C) and (E).

TABLE V. Accuracy results from the gridded data sets for the Uppland area.

Grid number	*Data set*		*R.m.s.e. (m)*	*SD (m)*	*Mean (m)*	*+Emax (m)*	*−Emax (m)*
40 m Grid1 +Grid 2	28·28 m Grid		0·69	0·63	0·26	3·26	−5·78
40 m Grid 1	40 m Grid		0·77	0·76	0·10	3·49	−6·24
	56·56 m Grid	1	0·91	0·91	0·06	3·41	−6·42
		2	0·94	0·94	0·08	3·87	−6·51
	80 m Grid	F1	1·24	1·24	0·07	6·49	−6·83
		F2	1·15	1·15	0·05	4·68	−7·18
		B1	1·14	1·14	0·06	4·17	−6·82
		B2	1·17	1·17	0·12	4·40	−6·57
40 m Grid 2	40 m Grid		0·84	0·71	0·44	3·84	−5·33
	56·56 m Grid	1	1·01	0·91	0·44	5·04	−5·75
		2	0·99	0·88	0·45	4·74	−5·33
	80 m Grid	F1	1·29	1·23	0·42	6·81	−6·52
		F2	1·28	1·21	0·42	4·92	−5·87
		B1	1·25	1·16	0·46	6·07	−5·36
		B2	1·21	1·14	0·41	5·34	−5·87

F = Forward, B = Backward sets of measurements.

Accuracy of the DTM derived from Composite Data Sets

It is possible to distinguish two types of composite data set: (i) those comprising regularly gridded data with a fixed grid interval plus selected feature-specific points; and (ii) those comprising gridded data with varying grid intervals plus selected feature-specific points. The data sets used in this test belong to the former type and the form lines and/or break lines have been sampled with a moderately high density.

For example, the number of points contained in such lines for the Uppland area is about 1750, while the total number of points produced by the two 40 m grids is only 3638. To give another example, for the data set for the area of Sohnstetten, the total number of points provided by the break lines and form lines is about 1550, while that provided by the 20 m grid is only 1716. Therefore, during the process of triangulation, about 150 points from the form and break lines for each area filtered out automatically by the triangulation program itself because they were considered to be duplicated.

TABLE VI. Accuracy results from the gridded data sets for the Sohnstetten area.

Data sets			*R.m.s.e. (m)*	*SD (m)*	*Mean (m)*	*+Emax (m)*	*−Emax (m)*
20 m	Grid		0·57	0·56	−0·11	3·03	−3·33
28·28 m Grid		1	0·86	0·86	−0·02	4·03	−3·66
		2	0·88	0·88	−0·02	4·95	−4·06
40 m Grid		F1	1·57	1·56	0·19	5·95	−5·42
		F2	1·54	1·53	0·12	7·61	−9·14
		B1	1·34	1·34	0·08	5·71	−7·82
		B2	1·35	1·35	0·10	6·29	−6·37
56·56 m Grid	F1	1	2·42	2·37	0·46	11·16	−5·69
		2	2·34	2·33	0·29	10·68	−10·62
	F2	1	2·59	2·57	0·37	11·16	−11·11
		2	2·27	2·26	0·27	10·52	−6·54
	B1	1	2·34	2·31	0·33	11·97	−9·09
		2	2·49	2·43	0·54	11·35	−8·05
	B2	1	2·46	2·42	0·42	12·18	−6·57
		2	2·54	2·47	0·59	15·08	−6·06

The composite data sets should be providing a higher fidelity in terms of representing the topography of the terrain surface than the regularly gridded data sets. In other words, the accuracy of the DTM formed from a composite data set should be higher than that resulting from the use of regular gridded data only. The results from each of the composite data sets are given in Tables VIII, IX and X. A diagrammatic representation of these test results is shown in Figs. 2(B), (D) and (F).

TABLE VII. Accuracy results from gridded data sets for the Spitze area.

Data sets		*R.m.s.e. (m)*	*SD (m)*	*Mean (m)*	*+Emax (m)*	*−Emax (m)*
10 m Grid		0·22	0·21	0·07	1·44	−2·26
14·14 m Grid	1	0·29	0·28	0·07	1·70	−3·06
	2	0·28	0·28	0·06	1·93	−1·77
20 m Grid	F1	0·39	0·39	0·06	2·79	−3·49
	F2	0·34	0·34	0·05	3·17	−1·91
	B1	0·37	0·36	0·06	3·05	−3·02
	B2	0·36	0·36	0·07	2·92	−1·93

ANALYSIS OF TEST RESULTS

An analysis of the test results is given in this section, including a descriptive analysis of the accuracy results obtained from both the regularly gridded data sets and the composite data sets, a regression analysis of the obtained accuracy figures and a descriptive analysis of the occurrence frequencies of large residuals for these different data sets.

TABLE VIII. Accuracy results from the composite data sets for the Uppland area.

Grid number	*Data set*		*R.m.s.e. (m)*	*SD (m)*	*Mean (m)*	*+Emax (m)*	*−Emax (m)*
Grid 1+2	28·28 m Grid+Points		0·64	0·59	0·24	2·41	−5·78
40 m Grid 1	40 m Grid 1+Points		0·67	0·66	0·10	3·33	−6·42
	56·56 m Grid + Points	1	0·71	0·70	0·10	2·69	−6·42
		2	0·71	0·71	0·09	3·29	−6·51
	80 m Grid + Points	F1	0·84	0·83	0·10	3·15	−5·91
		F2	0·82	0·81	0·11	4·68	−5·88
		B1	0·79	0·78	0·13	3·19	−6·47
		B2	0·79	0·78	0·12	3·74	−6·43
40 m Grid 2	40 m Grid 2+Points		0·72	0·63	0·35	2·41	−5·33
	56·56 m Grid + Points	1	0·77	0·70	0·32	3·48	−5·75
		2	0·74	0·67	0·31	2·70	−5·33
	80 m Grid + Points	F1	0·86	0·82	0·24	4·11	−6·52
		F2	0·83	0·77	0·31	3·46	−5·42
		B1	0·83	0·77	0·30	3·76	−5·36
		B2	0·83	0·77	0·32	4·19	−5·16

The term "Points" refers to feature-specific points and those located along form lines and break lines.

TABLE IX. Accuracy results from the composite data sets for the Sohnstetten area.

Data set			*R.m.s.e. (m)*	*SD (m)*	*Mean (m)*	*+Emax (m)*	*−Emax (m)*
20 m Grid+Points			0·43	0·40	−0·15	1·68	−2·55
28·28 m Grid + Points		1	0·55	0·53	−0·14	2·00	−3·63
		2	0·58	0·56	−0·13	2·40	−3·30
40 m Grid + Points		F1	0·79	0·78	−0·15	2·82	−3·00
		F2	0·78	0·77	−0·15	2·90	−4·34
		B1	0·78	0·77	−0·14	3·23	−3·75
		B2	0·78	0·76	−0·15	2·67	−4·63
55-56 m Grid+P	F1	1	1·08	1·08	−0·12	5·12	−5·19
		2	1·09	1·07	−0·22	3·37	−5·19
	F2	1	1·09	1·07	−0·19	3·79	−6·27
		2	1·09	1·08	−0·18	4·45	−5·51
	B1	1	1·07	1·06	−0·16	4·59	−5·59
		2	1·10	1·08	−0·20	3·73	−4·98
	B2	1	1·09	1·07	−0·23	4·38	−5·22
		2	1·12	1·12	−0·10	5·43	−3·73

Descriptive Analysis of the Accuracy Results for the Uppland Area

From Table V, giving the accuracy results obtained from the *regular gridded data sets* for the Uppland area, it can be seen that the mean values and the r.m.s.e. values of the DTM residuals obtained from the first 40 m grid for the Uppland area are quite different to those from the second grid, while the SD values are quite similar. This might suggest that there is a systematic shift between these two sets of measured data. The value of this shift is about 0·30 m. Evidence showing such a trend is also given by the mean resulting from the grid data set with the 28·28 m interval which was generated by adding the two 40 m grids together. This gives a value of 0·26 m which is almost equal to the average of the two means (0·10 m and 0·44 m) resulting from the two 40 m grids.

TABLE X. Accuracy results from the composite data sets for the Spitze area.

Data set		*R.m.s.e. (m)*	*SD (m)*	*Mean (m)*	*+Emax (m)*	*−Emax (m)*
10 m Grid+Points		0·16	0·14	0·07	0·87	−0·79
14·14 m Grid + Points	1	0·17	0·16	0·07	0·88	−2·71
	2	0·17	0·15	0·07	0·88	−2·66
20 m Grid + Points	F1	0·175	0·16	0·05	0·88	−2·38
	F2	0·18	0·17	0·06	0·88	−2·66
	B1	0·174	0·16	0·06	0·88	−2·27
	B2	0·174	0·16	0·06	0·88	−2·72

For the *composite data set* for the Uppland area, the accuracy results are given in Table VIII. From this table, again, a constant shift between the two 40 m grids can be observed. However, the value of the shift becomes smaller (about 0·2 m) in this case, when the feature-specific points are added to the gridded data. The amount of reduction is about 0·1 m. Another interesting point arising from this change is that the value of the mean resulting from the second 40 m grid is reduced by 0·12 m. By contrast, the mean resulting from the first 40 m grid is increased very little by 0·03 m, which is insignificant.

It can also be determined from inspection of Figs. 2(A) and (B) that the r.m.s.e. and SD values increase with an increase in the sampling interval. In both cases, the trend appears to be quite linear. The only difference between these two trends is that the speed of decrease in accuracy is faster in the case of regular gridded data than when composite data are used. It is additionally of interest to note that, in this test, the accuracy figures from the data sets measured in the forward direction are almost the same as those of the corresponding data sets measured in the backward direction.

Descriptive Analysis of the Accuracy Results for the Sohnstetten and Spitze areas

The accuracy results from the gridded data for the Sohnstetten and Spitze areas are listed in Table VI, VII, IX and X. From Table VI, it can be seen that, in general, the mean error values increase with the increase in the grid interval. The magnitudes of the extreme errors (the positive and negative maximum errors) also show a similar trend.

The reason for these phenomena may be found from the characteristics of the test area itself. It can easily be seen that the area covered by the check grid lies, in the most part, along and on either side of a steep sided valley. It is not difficult to imagine that, with the increase in the grid interval, the DTM surface which has been constructed linearly from the gridded data has been lifted up over the valley area. If this is the case, then the magnitude of the errors (positive in this case) in the points lying on the DTM surface will increase. The consequence is an increase in the magnitude of the positive maximum errors. Also, a small part of the test area covers a ridge line. As in the case of the steep sided valley, the DTM surface covering this small area will be lowered when a larger grid interval is used. This would result in an increase in the magnitude of the negative errors and thus of the maximum value. However, since most of the area lies on either side of the ravine line, the increase in the sum of the positive errors will almost certainly have been far larger than that of the negative errors. Thus the resulting mean value increases in a positive sense with the increase in the grid interval.

The plot of the DTM errors for the 20 m grid data for the Sohnstetten test is reproduced in Fig. 3. The size of a circle indicates the magnitude of the DTM error. The solid circles indicate the positive errors and the pecked circles denote the negative errors. From this diagram, it can be seen clearly that the positive errors are almost all located along the bottom of the steep valley, while the largest negative errors are located along the ridge lines (for example, top left on the diagram). This

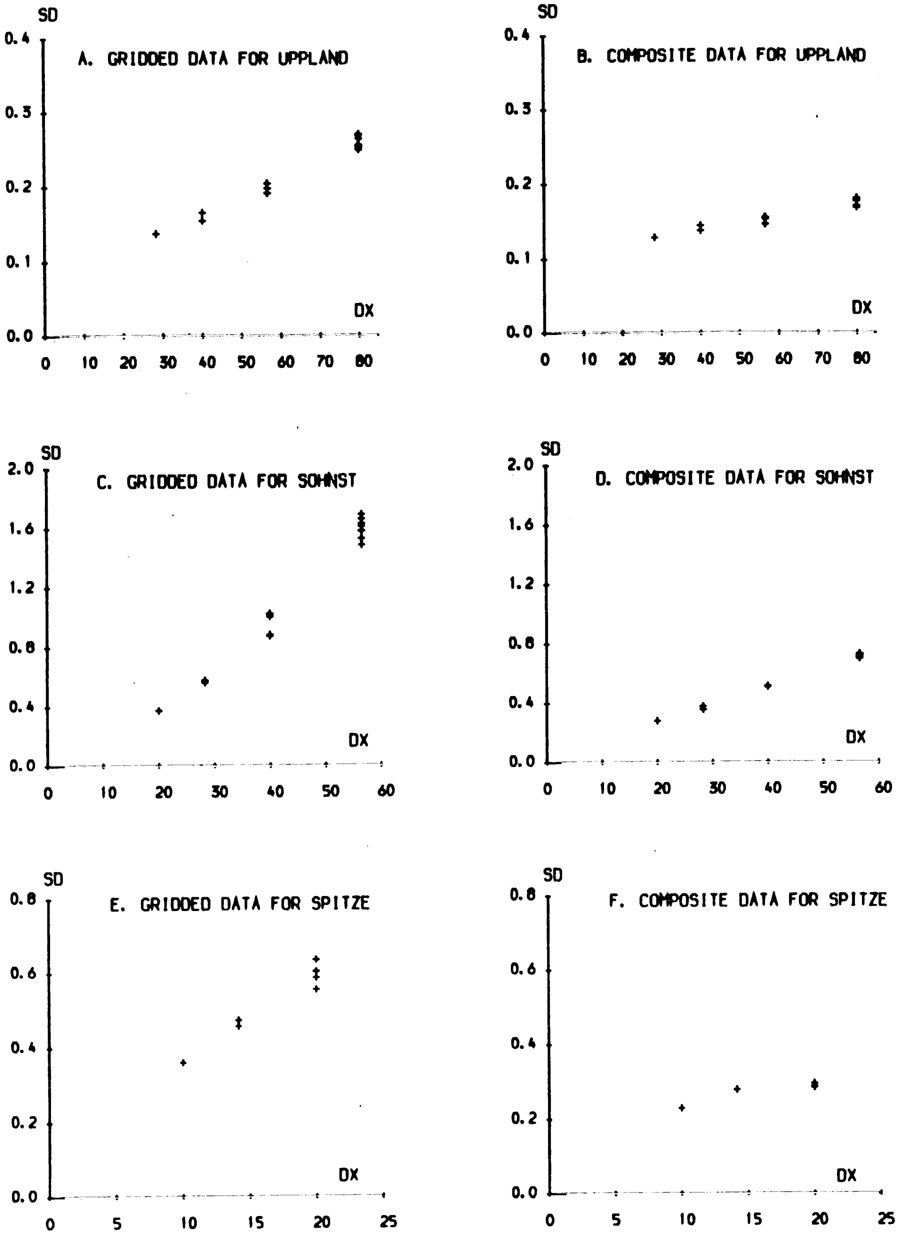

FIG. 2. Variation of DTM accuracy with sampling interval.

diagram provides strong confirmation of the analysis given above. Further powerful evidence to substantiate this reasoning is provided by the fact (see Table IX) that, after adding the points measured along the valley break lines and ridge lines, the magnitude of the extreme errors is significantly reduced and the mean value is kept almost constant.

It is of interest to note that the two extreme values of the DTM errors for the grid data for the Sohnstetten test area appear very large when the grid interval reaches a value of 56·56 m. The values for the maximum errors range from 10·52 m to 15·08 m. These values are larger than 0·67 per mille of the flying height and

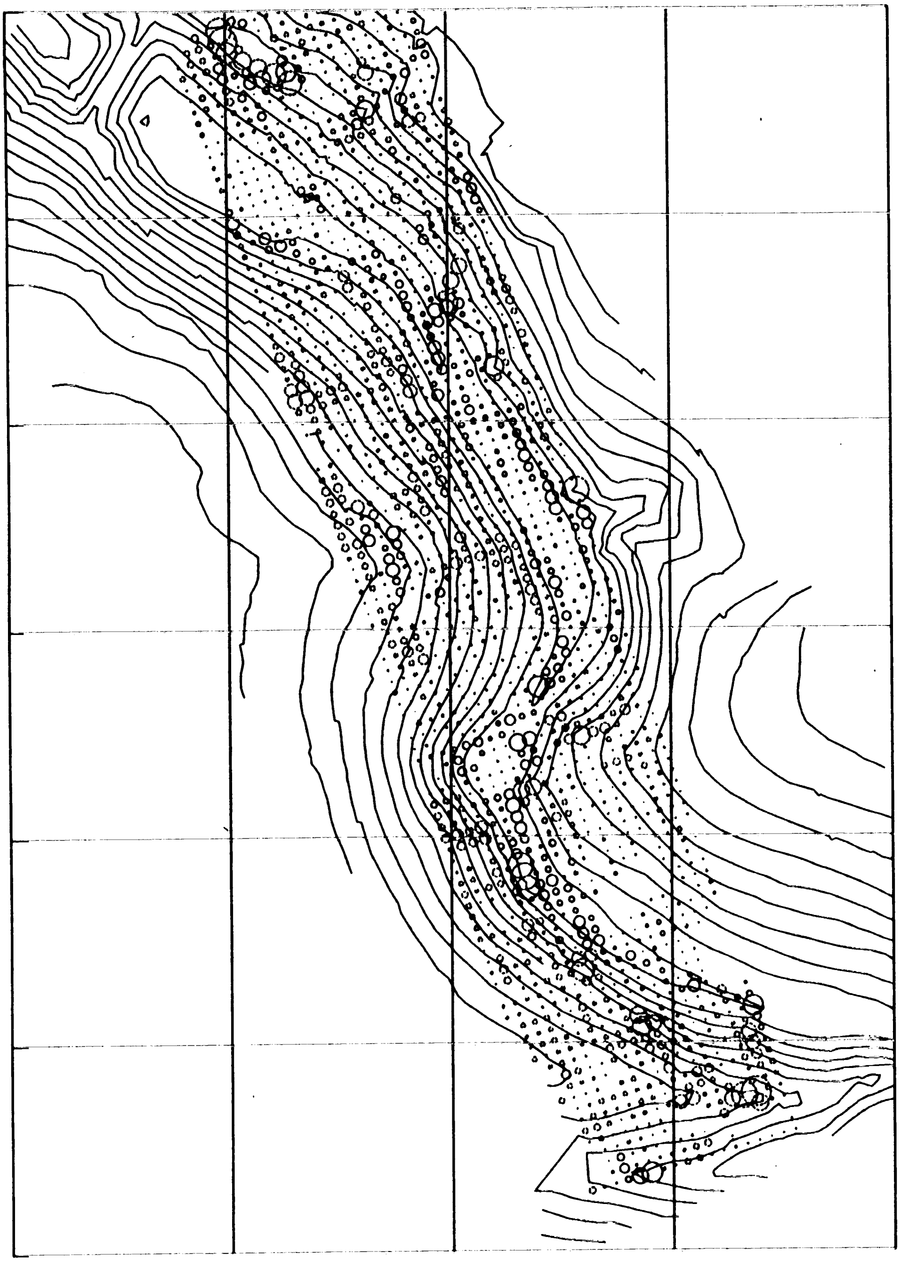

FIG. 3. Distribution of DTM errors for the 20 m gridded data for the Sohnstetten area.

amount to about 10 per cent of the height range. At first sight they seem enormous. However, they can by no means be considered as gross errors and indeed they are due to the inherent nature of the terrain surface in the test area. Strong evidence to support this point is that, after the addition of the form lines and break lines for this area, the extreme values are significantly reduced to a level of about 5 m. The theoretical analysis which has been described by Li (1990) also shows that it is still possible for such large extreme values to occur. Similarly, the large errors for the grid data for the Spitze area (amounting to 0·4 per mille of flying height and 5 per cent of the height range) can also be expected to occur.

From the analysis conducted above, it can be seen that, in the estimation of the accuracy of a DTM which is derived from data acquired by regular grid sampling (in the case when feature-specific lines are not included), the proportion of the grid cells across which form lines and break lines may pass should be taken into consideration. A more detailed discussion of this matter has been carried out by Li (1990) in which a family of mathematical models for DTM accuracy prediction has been described.

As expected, the trend that the r.m.s.e. and SD values will increase with an increase in the grid interval can be seen clearly from the results for both the Sohnstetten and the Spitze areas. The speed of increase in the r.m.s.e. and SD values is again faster in the case of gridded data sets than for composite data sets. It is also faster for the Sohnstetten area than for the Spitze area. The trend is approximately linear.

It should also be mentioned that the accuracy results from the two 40 m grid data sets measured in a forward mode are a little different from those resulting from the corresponding two data sets measured in a backward mode. However, after the addition of the form lines, these differences disappeared.

Analysis of the Occurrence Frequency of Large Residuals

In this experiment, the occurrence frequencies of large residuals have also been computed, although the results are not presented in this paper due to limitation in space. The purpose of recording these values is to provide information about the distribution of DTM errors since it has been found that this distribution may vary with the type of terrain and the pattern of data points. Some display an approximately normal distribution while others do not (Torlegård *et al.*, 1986). The simple histograms included in Figs. 4 (a) and 4 (b) are two examples which prove these two characteristics since they not only show different distributions, but also show differences in the frequencies of error occurrence. The occurrence frequencies of large residuals may also vary with terrain type and with the pattern of the data points. It would be expected that the occurrence frequencies of large residuals would be much higher for broken terrain than for smooth terrain, if the data for the break lines and form lines are not used for surface reconstruction. In view of this discussion it is, therefore, necessary to undertake a more detailed examination of the occurrence frequencies of large residuals.

These large residuals have been divided into three classes: those larger or equal to two times SD (standard deviation), three times SD and four times SD. From the test results obtained from this study, it was found that the frequency of residuals larger than two times SD is smaller than 7·0 per cent. For those residuals larger than three times SD, it is 2·0 per cent. Finally, the frequency is much smaller than 1·0 per cent for those errors larger than four times SD in all cases, except for the gridded data set of Spitze area where terrain discontinuities exist.

Therefore, from this series of tests, it can be stated that the vast majority of DTM errors are smaller than four times SD (the standard deviation). Comparing this with the normal distribution, it can also be found that the probability with which DTM errors fall within the range from $-4 \times \text{SD}$ to $4 \times \text{SD}$ from the mean is very approximately equal to the probability with which random errors of normal distribution fall within the range of three times the corresponding SD.

Regression Analysis of the Accuracy Results

A regression analysis is carried out in order to obtain some quantitative results. The procedure used in this study is as follows.

(i) Firstly, a mathematical model is selected based on experience. Thus, in practical terms, the selected model is usually an empirical model.

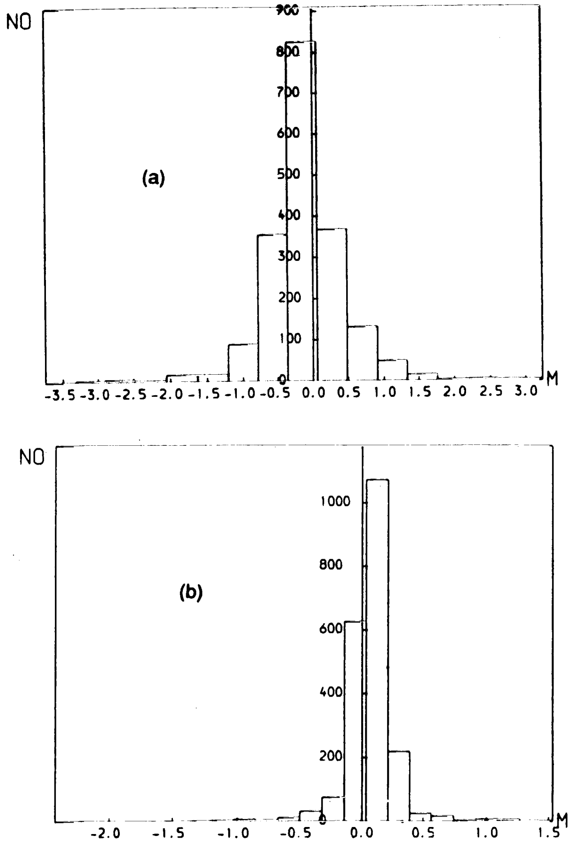

FIG. 4. Histograms for the distribution of DTM errors: (a) from composite data (20 m grid) for the Sohnstetten area; (b) from gridded data (10 m grid) for the Spitze area.

(ii) The experimental test results are then used to compute the parameters (coefficients) of this model.

(iii) Finally, the adequacy of this model is examined by the value of its correlation coefficient.

The mathematical model used in this analysis is

$$\mathrm{VAR(DTM)} = a^2\ \mathrm{VAR(PMD)} + b\ (Dx \tan \alpha)^2 \qquad (1)$$

where a and b are two coefficients; VAR(PMD) denotes the variance of PMD (the photogrammetrically measured data); α refers to the typical slope angle in the area; Dx is the sampling interval (the grid interval); and VAR(DTM) denotes the variance of the resulting DTM. This is a model which is similar to that suggested by Ackerman (1980).

It can be seen that, for the same data set, the first term of the right hand side of (1) is a constant; $b \tan^2\alpha$ will also be a constant for the same area. Therefore, for the case of this study, this model can be simplified as follows:

$$\mathrm{VAR(DTM)} = c + d(Dx)^2 \qquad (2)$$

where c and d are the coefficients. The regression results are shown in Table XI.

For the Uppland area, due to the constant shift between two 40 m grids, the SD values were used instead of the r.m.s.e. values. For the accuracy results for the

TABLE XI. Regression results for the coefficients.

Test area	*Data set*	*c*	*d*	*r*	*a*
Uppland	Gridded	0·2575	0·0001775	0·983	0·7574
	Composite	0·3335	0·0000459	0·953	0·8620
Sohnstetten	Gridded	−0·8269	0·0019505	0·990	None
	Composite	0·0253	0·0003719	0·999	0·9990
Spitze	Gridded	0·0240	0·0002749	0·955	1·9365
	Composite	0·0251	0·0000151	0·890	2·8228

a = coefficient in equation (1).
c, d = coefficients in equation (2).
r = correlation coefficient.
None = coefficient does not exist as a real number.

Spitze and Sohnstetten areas, the mean values of the residuals are comparatively small, therefore the SD values are almost equal to the r.m.s.e. values. Thus either of these two values will do; in practice, the r.m.s.e. values have been used in this analysis.

By substituting the value of VAR(PMD) into (1), the value of the coefficient *a* can be obtained for each of the different test areas. The values of *a* should lie between the values 0·0 and 1·0. The reason is obvious because, if any parameter in the second term on the right side of (1) is zero, then this term will be zero. In that case, the errors in the DTM points are simply propagated directly from the source data points. In this case, the DTM height is the mean (possibly weighted) of the reference points which have been used for interpolation. The accuracy of the mean value is, of course, higher than the accuracy of the reference points. Thus the value of *a* should lie within the range 0·0 to 1·0.

It was found from Table XI that this mathematical model is approximately suitable for the results from the data sets of the Uppland area and the composite data set of the Sohnstetten area. However, the results for the gridded data set of the Sohnstetten area show that the value of the variance of the DTM errors from that data set is proportional to the sampling interval (grid cell in this case) to a power greater than two. Of course, the results from the regression process may not be absolutely reliable. However, in any case, the results obtained from the gridded data set for the Sohnstetten area indicate strongly that quite different mathematical models should be used for gridded data sets and composite data sets.

The results from the data sets for the Spitze area indicate that this mathematical model does not fit the experimental data. One of the reasons could be that too few samples of grid intervals were used for the analysis; therefore the results which were obtained are very unreliable.

Of course, it would be possible to try to discuss and employ different mathematical models to see how well they fit these test results. Realistically, even if a mathematical model does fit the practical data very well, it does not necessarily mean that this model expresses the fundamental or inherent relationship between the variables under investigation. On the other hand, a model which does not fit a specific set of experimental data is not necessarily inadequate. Thus, for a particular mathematical model, the important assessment that needs to be made is to estimate how wrong it is, instead of just rejecting it. Therefore, for the time being, the only conclusion which can be reached is that the model expressed by (1) does apply in some cases but not in all. Indeed, the theoretical analysis carried out by Li (1990) also shows that this model expresses only the DTM accuracy in the case of composite data sets.

CONCLUSIONS

(i) The accuracy of a DTM formed from photogrammetrically measured data is highly correlated with the sampling interval (grid interval) if only gridded data is used.

(ii) When *feature-specific points* (including points measured along break lines and form lines) are added to the data set, the accuracy of the DTM can be improved. This improvement is greater if the sampling interval is large. With a small sampling interval, this effect may not be significant.

(iii) Large residual errors do occur but the *occurrence frequencies* of these residual errors being greater than 4 × SD are usually very small. With the inclusion of feature-specific points the magnitude of large residuals can be reduced. If the area has many break lines and terrain discontinuities, this reduction could be significant.

(iv) It was also found that the accuracy of DTMs is correlated with the *slope angle* of the terrain surface. In those areas with steeper slopes, the r.m.s.e. and SD values increase with a faster speed.

(v) The accuracy results obtained from *two data sets of the same area* may be quite different, even if the sampling interval is the same for both of them. Therefore, it is impossible to be definite about the accuracy value which can be obtained from a data set employing a given sampling interval. However, an approximate value or a range of values can be given for the accuracy.

ACKNOWLEDGEMENTS

This paper is based on work carried out for the author's Ph.D. thesis. The author would like to express his sincere gratitude to the following: the University of Glasgow for financial support of his study; Professor H. Ebner and Dr. W. Reinhardt of the Technical University of Munich for providing the photogrammetric data sets; Professor K. Torlegård and Dr. M. Li of The Royal Institute of Technology, Stockholm for providing the check point data sets; and Professor G. Petrie for his assistance during this research work and the writing up of this paper.

REFERENCES

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

LI, Z., 1988. On the measure of digital terrain model accuracy. *Photogrammetric Record*, 12(72): 873–877.

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

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

MCCULLAGH, M., 1983. "If you're sitting comfortably we'll begin". *Workshop Notes on Terrain Modelling*, Australian Computing Society, Siren Systems. 158 pages.

PEUCKER, T., 1972. *Computer cartography*. Association of American Geographers, Commission on College Geography, Washington, D.C. 75 pages.

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

Résumé

On presente dans cet article quelques études expérimentales sur la variation de l'exactitude d'un modèle numérique du terrain (MNT) en fonction de l'intervalle d'échantillonnage de données obtenues à partir de méthodes stéréophotogrammétriques. On examine les facteurs principaux qui interviennent sur l'exactitude d'un MNT et l'on esquisse divers projets permettant de mener au choix des essais expérimentaux sur l'exactitude d'un MNT. On fournit des informations détaillées sur les polygones d'essai utilisés, les jeux de données concernant les points de vérification et ceux de l'altimétrie initiale, suivies d'une présentation de la conception particulière de ces études. On décrit les résultats expérimentaux obtenus et l'on termine cet article par une analyse critique des résultats de l'essai.

Zusammenfassung

Im Artikel werden einige experimentelle Untersuchungen zur Variation der Genauigkeit von Digitalen Geländemodellen (DTM) beschrieben, wobei ein Rasterintervall verwendet wurde, das auf stereophotogrammetrischer Datenerfassung basierte. Es werden die Haupteinflußfaktoren auf die Genauigkeit der DTM betrachtet und verschiedene alternative Wege zur Testung der DTM-Genauigkeit skizziert. Detaillierte Angaben erfolgen zu den Testgebieten, den Quelldaten-Sätzen und den Datensätzen der Kontrollpunkte mit anschließenden Ausführungen zur speziellen Gestaltung dieses Versuchs. Die Versuchsergebnisse werden angegeben, und abschließend erfolgt ihre kritische Analyse.