O. O. Ayeni
Department of Surveying
University of Lagos
Lagos, Nigeria

Optimum Sampling for Digital Terrain Models: A Trend Towards Automation

Algorithms are presented for determining the optimum number and spacing of terrain elevation points in a DTM.

Abstract: *The relative efficiency of seven different sampling methods-aligned/unaligned systematic sampling, aligned/unaligned stratified sampling, aligned/unaligned random sampling, and unaligned systematic stratified random sampling—were investigated in relation to four major classes of terrain. A "filtering" technique in combination with the Harmonic Vector Magnitude, computed from the Double Fourier Function, was used to determine optimal sample size for any class of terrain through the process of progressive sampling. A computer program* ATOSAP, *developed from this research, can be used interactively by the steroplotter operator or in a fully automated system to determine the optimal size of the sample and also to generate a pattern of sampling for digitizing any given terrain.*

Introduction

Many investigations have been made into techniques of producing a digital terrain model (DTM) using regularly spaced data points. Questions which relate to the problem of the optimum number of data points and their associated distribution in relation to different types of terrain are usually ignored, although the importance of efficient sampling as opposed to interpolation techniques is often acknowledged. This is borne out by Blaschke's (1968) definition of a DTM in which he emphasized the importance of "storing measured coordinates *X*, *Y*, *Z* of characteristic terrain points in sufficient *quantity* and *significance*. . . ." A rather comprehensive definition of a DTM was given by Ayeni (1976a) as "the numerical and mathematical representation of a terrain by making use of adequate elevation and planimetric measurements, compatible in number and distribution with that terrain, so that the elevation of any other point of known planimetric coordinates can be automatically interpolated with specified accuracy for any given application." This definition bears testimony to the importance of evaluating the adequate number of data points as well as the appropriate sampling distribution of such points that constitute a good match for a given terrain. This is what is meant by optimum sampling, which obviously has two components: optimum sample size and optimum sample pattern.

Optimum Sample Size

The Harmonic Vector Magnitude

The problem of how the stereoplotter operator could know when he has measured enough elevations during the process of generating DTM data is not a simple one because it involves a proper assessment of the terrain roughness in relation to the size of the area occupied by the terrain. This implies that, if any terrain roughness parameter is to be used as a criterion to determine optimum sample size, such a parameter should incorporate the size of the entire area. One parameter which seems to satisfy this condition is the Harmonic Vector Magnitude (HVM), which is computed in the following manner. The surface from the stereo model is fitted by least squares to a combination of two mathematical functions—a linear filter, Equation 1, and a double Fourier function, Equation 2, giving rise to Equation 3—that is,

0099-1112/82/4811-1687$02.25/0

$$Z_1 = a + bx + cy \quad (1)$$

$$f(x,y) = \left[\sum_{i=0}^{m} \sum_{j=0}^{n} \lambda_{ij} \left(CC_{ij}\cos \frac{2\Pi x_i}{M} \cos \frac{2\Pi y_j}{N} + CS_{ij}\cos \frac{2\Pi x_i}{M} \sin \frac{2\Pi y_j}{N} + SC_{ij}\sin \frac{2\Pi x_i}{M} \cos \frac{2\Pi y_j}{N} + SS_{ij}\sin \frac{2\Pi x_i}{M} \sin \frac{2\Pi y_j}{N} \right) \right] \quad (2)$$

$$Z = Z_1 + f(x,y) \quad (3)$$

where

M (the fundamental wavelength in x-direction) $= x_m$,
N (the fundamental wavelength in y-direction) $= y_n$, and
$\lambda_{ij} = 1/4$ for $i = j = 0$; $\lambda_{ij} = 1/2$ for $i > 0$, or $i = 0, j > 0$; $\lambda_{ij} = 1$, for $i > 0$ and $j > 0$.

The Fourier coefficients can be used as a measure of the magnitude of deviations of the surface from the plane defined by Equation 1 because the regional fit assumes that the surface oscillates harmonically in two mutually perpendicular directions. The Harmonic Vector Magnitude (HVM) is defined as the square root of the sum of squares of coefficients for terms of specified m and n harmonics. Ayeni (1976a) has demonstrated the use of HVM as a parameter of terrain roughness. Because the fundamental wavelengths, M and N, define the limits of the area concerned, the *HVM*, therefore can be used in determining a criterion for determining optimum sample size. The following steps illustrate the use of HVM, in collaboration with the concept of progressive sampling of terrain models first proposed by Makarovic (1973-75) for the determination of optimum sample size.

Step 1: Start with a suitable sample interval for taking SN^2 elevation measurements ($SN = 4$, say) in a regular grid pattern on the stereo model. This permits a least squares solution of Equation 3 which has ten unknown parameters for $m = n = 1$.

Step 2: Compute HVM for SN^2 sample size.

Step 3: Increase the number of elevation points to $(SN + 1)^2$ using a new interval for the regular grid pattern.

Step 4: Repeat Steps 2 and 3 until HVM satisfies the following criteria:

$$\text{HVM}_i - \text{HVM}_{(i+1)} < CC \quad (4)$$

where the quantity CC can be determined by Equation 5 for the $(i + 1)^{th}$ sample; i represents the i^{th} sample.

$$CC = \text{HVM}_{i+1} \times f_{\text{HVM}} \quad (5)$$

where f_{HVM} may be taken as the measure of accuracy for photogrammetric representation of relief given by Markarovic (1973-75) as

$0.2 - 0.5\%$ of elevation for minimum σ_H
$0.4 - 1.0\%$ of elevation for maximum σ_H

where σ_H = pointing accuracy of elevation measurement for a given stereoplotter.

The steps outlined above were used to determine the optimum sample size of nine simulated terrains. The results are shown in Table 1. By optimum sample size we mean the number of elevation points on a terrain which will be ideal for adequately describing the roughness of the terrain. It will be observed from Table 1 that the HVM remains fairly constant after the optimum sample size is attained according to the above criterion in Equation 4.

MULTIPLE LINEAR REGRESSION EQUATIONS FOR OPTIMUM SAMPLE SIZE

An attempt was made to develop empirical linear regression equations for determining optimum sample size which may require less computation than the HVM method. It was hypothesised that the optimum sample size, SN, is functionally related to nine terrain roughness parameters (see Table 2). Using statistical selection procedures—maximum and minimum r^2 improvement, backward and forward elimination, and stepwise regression—for choosing the "best" equations (see Ayeni (1976a) for details), the following regression equations (Equations 6 to 15) have been found to give results comparable to those of the HVM method:

$$SN = 29.387 + 541.907X_9 \quad (6)$$
$$SN = 32.008 + 0.0000083X_3 + 525.653X_9 \quad (7)$$
$$SN = 32.228 + 0.205X_2 + 482.495X_9 \quad (8)$$
$$SN = 35.024 + 0.0073X_1 + 0.174X_2 + 480.008X_9 \quad (9)$$
$$SN = 32.008 + 0.00013X_1 + 0.00000X_3 + 525.618X_9 \quad (10)$$
$$SN = 34.880 + 0.173X_2 + 0.0000045X_3 + 483.066X_9 \quad (11)$$
$$SN = 1{,}277 + 0.722X_4 + 572.632X_9 \quad (12)$$
$$SN = 8.192 + 0.097X_4 + 539.553X_9 \quad (13)$$
$$SN = 10.917 + 0.103X_2 + 0.551X_4 - 29.488X_8 + 537.647X_9 \quad (14)$$
$$SN = 7.511 + 0.199X_2 + 0.000094X_3 + 0.540X_4 + 18.29X_6 - 1.81X_7 + 510.72X_9 \quad (15)$$

where SN is the optimum sample size (number of elevation points on the terrain). A useful criterion for stopping the progressive sampling procedure when using any of the Equations 6 through 15 may be stated as

$$SN_k - SN_{k+1} < TR$$

where

$TR = SN_{k+1} f$, $f = 0.5\%$, and
k represents the k^{th} sample.

TABLE 1. OPTIMUM SAMPLE SIZE USING MEAN HARMONIC VECTOR MAGNITUDE CRITERION (Units in Metres)

Surface	Sample Size	Mean Harmonic Vector Magnitude
Surface with Line Trend (1)	16	0.00260
	25*	0.00260
	36	0.00263
	49	0.00256
Exponential Surface (2)	16	0.449
	25	0.137
	36*	0.135
	49	0.131
	441	0.154
Logarithmic Surface Surface (3)	16	2.060
	25	1.906
	36	1.736
	49*	1.730
	64	1.764
	441	1.978
Double Fourier Surface with Random Coeffs (4)	100	6.978
	121	4.848
	144	4.672
	169*	4.679
	225	4.674
	256	4.680
	289	4.687
	400	4.882
Polynomial Surface (5)	144	39.592
	109	38.214
	196	37.004
	225	36.092
	256*	35.257
	289	34.533
	324	33.897
D. Fourier Surface with Synthetic Coeffs (6)	81	35.119
	256	266.603
	289	23.241
	324	270.834
	361	272.760
	400	26.5841
	441	276.741
	484*	277.811
	529	278.275
Polynomial Surface No. 2 (7)	289	4847.969
	361	4683.835
	400	4511.977
	441	4553.312
	484	4497.551
	529	4447.608
	576	4400.941
	625	4358.812
	676*	4349.053
	729	4340.406
D. Fourier Test Surface (8)	36	14.726
	47	10.136
	64	13.901
	81	9.717
	100	10.569
	121	8.834
	144*	8.820
	169	8.811
	196	8.801
	225	8.802

TABLE 1.—Continued

Surface	Sample Size	Mean Harmonic Vector Magnitude
Test Polynomial Surface (9)	121	50.157
	144	46.782
	169	44.034
	196	41.759
	225	39.8654
	256	38.336
	289	37.439
	324*	37.431
	361	37.677
	100	37.939

* Optimum Sample Size

OPTIMUM SAMPLE PATTERN

In the previous section we have used a regular grid pattern (systemic pattern) to generate data for determining the optimum sample size. The problem is that the systematic pattern may not be optimum for a given terrain, and this may lead to sampling error. Two other sampling patterns well known in statistics—stratified and random patterns—are illustrated in Figure 1 with their minor variations. Morrison (1970) has shown that the first six sampling patterns can theoretically represent nearly all the possible sample point scatter in a two-dimensional plane. This author feels intuitively that a seventh sampling plan—unaligned stratified systematic random—has some merits which deserve investigation. See the apendix for the definitions of the seven sample patterns used in this research.

CORRELATION FOR OPTIMUM SAMPLING PATTERN

The correlation characteristics of topography may be considered from a two-dimensional viewpoint. For example, the correlation between elevation points separated by distances u and v along the x- and y-directions may be expressed as

$$\rho(u,v) = \mathrm{E}\left[\left(Z_{i+u,\, j+v} - \mu\right)\left(Z_{ij} - \mu\right)\right] \quad (17)$$

where μ is the mean of the elevations Z_{ij}.

The same correlation function may be approximated by

$$r(u,v) = \frac{1}{(n-u)(m-v)} \sum_{i=1}^{m-q} \sum_{j=1}^{n-p} (Z_{i+q,\, j+p} - \bar{Z})(Z_{ij} - \bar{Z}) \quad (18)$$

where $q = 0, 1, 2, \ldots, T$ (Lags in x-direction);
$p = 0, 1, 2, \ldots, T$ (Lags in y-direction);
m, n represent the number of data points; along x- and y-directions, respectively; and
$\bar{Z}$ = sample mean of Z_{ij}.

Since the correlation properties are related to the spatial distribution of surface irregularities, correlation characteristics will be applied to the problem of optimum sampling pattern.

The problem of finding the relative efficiency of

TABLE 2. REGRESSION OF OPTIMUM SAMPLE SIZE (DEPENDENT VARIABLE) ON THE PARAMETERS OF TERRAIN ROUGHNESS (INDEPENDENT VARIABLES)

Surface	Optimum Sample Size (SN)	Mean Gradient (X_1)	Mean Curvature (X_2)	Comparea (X_3)	Mean Dip (X_4)	Mean Bump Freq. (X_5)	Var. of Direction Cosines (X_6)	Resultant Kurtosis of Direction Cosines (X_7)	No. of Breakline Per Unit Area (X_8)	No. of Points Per Unit Area (X_9)
1	25	1.141	0.0	1.998	27.506	0.010	0.276	1.3374	0.0	0.00025
2	36	0.6646	1.101	1.3196	42.453	0.399	0.4072	1.1682	0.0	0.000036
3	49	0.099	0.094	1.0003	69.973	7.723	0.4246	2.2615	0.0	0.000049
4	169	23.897	23.922	213.447	1.794	68.576	0.9905	6.750	0.111	0.2934
5	289	227.247	31.316	1087.07	10.915	606.361	0.8409	8.303	0.029	0.5017
6	484	207.129	247.358	19730.746	0.6341	945.158	0.9975	12.2968	0.151	0.8403
7	676	2499.869	344.405	32732.96	−0.175	133505.873	0.9997	180.317	0.049	1.1736
8	196	25.680	14.56	132.068	4.880	48.48	0.983	4.169	0.0243	0.391

DEFINITION OF TERMINOLOGIES FOR SURFACE ROUGHNESS (*See* Ayeni (1976a) for details)
Mean gradient = mean of the first spatial derivatives (i.e., magnitude of mean gradient) at various points on the surface computed from

$$S = \sqrt{\left(\frac{\partial Z}{\partial X}\right)^2 + \left(\frac{\partial Z}{\partial Y}\right)^2}$$

Mean Curvature = mean of second spatial derivatives (i.e., magnitude of mean curvature) computed from

$$C = \sqrt{\left(\frac{\partial^2 Z}{\partial X^2}\right)^2 + \left(\frac{\partial^2 Z}{\partial Y^2}\right)^2}$$

Comparea = ratio between the surface area and the plane area computed from the product of length and breadth of the area occupied by the terrain.
Mean Dip = Mean of the slopes of a set of triangular intersecting planes fitted to adjacent groups of three elevations.
Mean Bump Frequency = Mean of the distances from elevation points or the terrain to the best-fit planar surface in a direction normal to the latter.
Variance of Direction Cosines = Variance of all direction cosines computed between any two points in the X, Y, Z directions on the terrain. Statistically, variance is the second movements of these direction cosines.
Kurtosis of Direction cosines = Third movements of all direction cosines computed between any two points in the X, Y, Z directions on a terrain.
Breakline = Sudden or abrupt change in slope. The objective method of detecting the number of breaklines on a terrain from sample data is described in Ayeni (1976a).

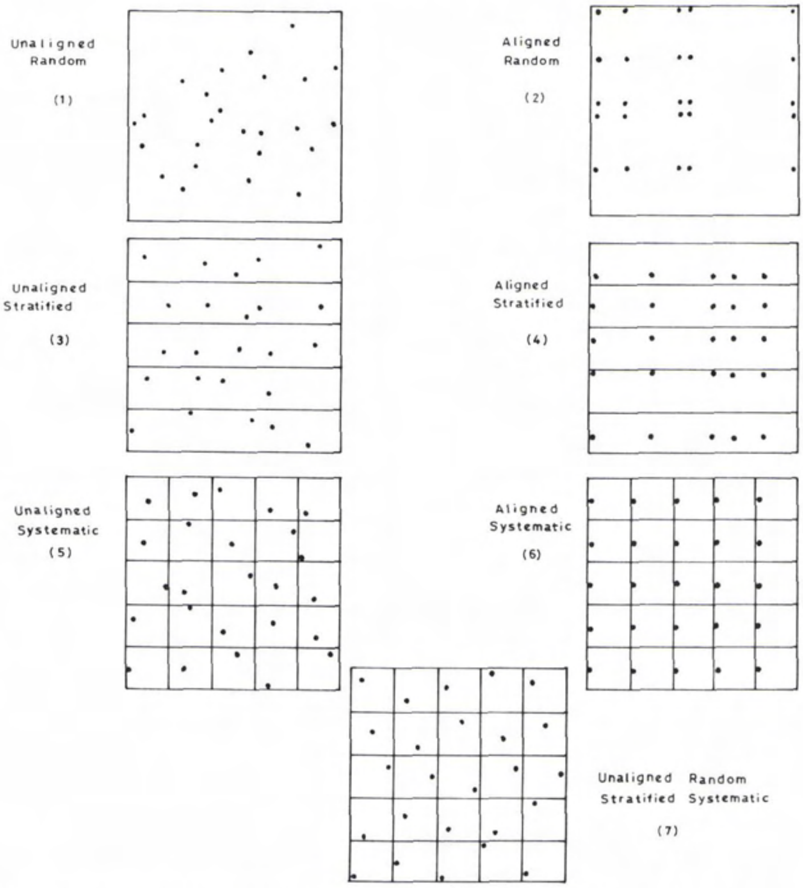

FIG. 1. Seven sampling patterns (see Appendix for the description of these sampling patterns).

various sample patterns for certain categories of populations has been the subject of many published papers in statistics, namely, Das (1950), Quenoulle (1949), Cochran (1946), and others. Cochran (1946) working with one-dimensional autocorrelation has developed some theorems to show when a particular sample type (Random, Stratified, or Systematic) is more efficient than the other. Das (1950) has demonstrated a two-dimensional extension of Cochran's theorems, to establish the relative efficiency of these three major sampling patterns. Das theorems are now stated without proofs. (see Das (1950) for proofs).

DAS THEOREMS FOR RELATIVE EFFICIENCY OF SAMPLING PATTERNS

Theorem 1: For all infinite populations in which

(i) $\Delta_1\Psi \leqslant 0$ where $\Delta_1 = \rho(u+1,v) - \rho(u,v)$; $\Delta_2 = \rho(u,v+1) - \rho(u,v)$
(ii) $\Delta_2\Psi \leqslant 0$ where $\Psi = \rho(u,v) + \rho(-u,v)$
(iii) $2\Psi(u,v+1) \geqslant \Psi(u+1,v) + \Psi(u+1, v+1)$ and
(iv) $2\Psi(u+1,v) > \Psi(u,v+1) + \Psi(u+1, v+1)$

$\sigma_{st}^2 \leqslant \sigma_r^2$ (i.e., stratified sampling is more efficient than random sampling) for any size of the sample and $\sigma_{st}{}^2 < \sigma_r{}^2$ unless equality holds in each case. Note that $\rho(u,v)$ is the correlation defined in Equation 17.

Corollary: For all infinite populations in which $\Delta_1\psi = \Delta_2\psi < 0$, $\sigma_{st}^2 \leqslant \sigma_r^2$ for any size of the sample and $\sigma_{st}{}^2 < \sigma_r{}^2$ unless equality holds in each of the above cases.

Theorem 2: When stratification is made by parallel strips along u-direction if

(i) $\Delta_i\Psi \leqslant 0$
(ii) $\Delta_2\Psi \leqslant 0$ and
(iii) $2\ \Psi(u+1,v) \geqslant (u,v+1) + \Psi(u+1, v+1)$, then, $\sigma_{st}{}^2 \leqslant \sigma_r^2$

Corollary: When stratification is made by parallel u-strips if $\Delta_2\psi \leqslant \Delta_1\psi \leqslant 0$ then $\sigma_{st}{}^2 < \sigma_r^2$ (i.e., stratified is better than random).

Theorem 3: For all patterns of stratification if

(i) $\Delta_i\Psi \leqslant 0$
(ii) $\Delta_2\Psi \leqslant 0$ and
(iii) $\Delta_1\Delta_2 \leqslant 0$ then $\sigma_{st}{}^2 \leqslant \sigma_r^2$

Theorem 4: For all infinite population which

(i) $\delta_1^2\rho(u,v) \leqslant 0$ where $\delta_1^2\rho(u,v) = \rho(u + 1,v) + \rho(u - 1,v)$
(ii) $\delta_2^2\rho(u,v) \geqslant 0$ where $\delta_2^2\rho(u,v) = \delta_1^2[\delta_1^2\rho(u,v)]$
Then $\sigma_{sy}{}^2 \leqslant \sigma_{st}{}^2$ for any sample size and $\sigma_{sy}{}^2 < \sigma_{st}{}^2$ unless equality holds;
i.e., systematic is better than stratified.

Theorem 5: For all infinite population in which

(i) $\Delta_1\Psi(u,v) \leqslant 0$
(ii) $\Delta_2\Psi(u,v) \leqslant 0$
(iii) $\delta_1^2\rho(u,v) \leqslant 0$
(iv) $\delta_2^2\rho(u,v) \geqslant 0$
Then $\sigma_r^2 \geqslant \sigma_{st}{}^2 \geqslant \sigma_{sy}{}^2$
i.e., *systematic* is better than *stratified* and *stratified* is better than *random*.

It should be noted that the physical interpretation of these conditions is very difficult to perceive except in one-dimension. For example, in Theorem 5, conditions (i) and (ii) are nothing but strictly monotonic correlation properties and (iii) and (iv) are properties of second differences of correlation in the x and y directions, for elevation points separated by distances u and v.

THEORETICAL EXPECTATION AND EMPIRICAL RESULT OF THE EFFICIENCY OF SAMPLING PATTERNS

The objective of this Section is to investigate the validity of these theorems when they are applied to a topographic surface as the infinite population of interest, and when efficiency of sampling pattern relates to interpolation accuracy. An experiment set up to determine the practical efficiency of the various sampling patterns shown in Figure 1 is briefly described below.

Step A: Determine the theoretical relative efficiency of the six standard sampling patterns according to Das' theorems. A computer program called AUTOCO computes the correlation properties described in Theorems 1 to 5. From this, the theoretical efficiency of a sampling pattern can be determined. For example, if all the conditions in Theorem 5 are satisfied, then systematic sampling is "better" than stratified sampling, which in turn is "better" than random sampling.

Step B: Determine the practical (empirical) relative efficiency of the six standard sampling patterns in the following manner:

- A surface is generated (or simulated) by a mathematical function using the program TERRAIN. Let the parent function be $Z = a_0 + bX + cX^2 + dY + eY^2$, for example.
- A program called GENSAP was written to generate Z values for the seven sampling patterns discussed above, at a given or optimum sample size using this parent function.
- A "deficient" function is established by dropping one or two terms of the parent function, e.g., $Z_1 = BX + CX^2 + EY^2$. Then, by least-squares technique, the Z values so generated for the seven sample patterns are used in fitting the deficient function to determine the parameters B, C, and E which are, in turn, used to compute the Z_1 values at some points whose X, Y, Z coordinates are known on the parent surface.

Step C: The RMS (root mean squares) of differences between the Z's from parent surface and the Z_1's from "deficient" surface are then computed. The sampling plan with the smallest RMS is then chosen as the most efficient for that terrain. This result is then compared with the most efficient, theoretically speaking, sampling pattern obtained from Step A.

The results of the empirical investigations into the relative efficiency of the seven sampling patterns shown in Figure 1 in relation to eight terrains chosen from the four major terrain types objectively classified by the author (Ayeni, 1976b) may be enumerated as follows (see Table 3):

- Das's theorems can be used as a good a priori indicator of which of the six standard sampling patterns is optimum for any given terrain irrespective of the roughness, class, or sample size of the terrain. This may suggest that the paramount characteristics of the terrains that determines the optimum sample pattern is correlation.
- The seventh sample pattern—the unaligned systematic stratified random pattern—was found to be most efficient in nearly all the terrains investigated.
- The relative efficiency of sampling patterns may also be determined by comparing the determinants or traces of the dispersion matrices (D) (i.e., normal coefficient matrices) in a least-squares solution for interpolation. For example, for two sampling patterns, dispersion matrices D_1, D_2 are obtained. If $|D_1| < |D_2|$ or if trace $(D_1) <$ trace (D_2) then sampling pattern 1 is better than sampling pattern 2.

AUTOMATIC OPTIMUM SAMPLING FOR A DTM

As a result of the investigations performed in this research, a FORTRAN IV program called ATOSAP was developed, which could be used interactively by the stereoplotter operator or in a fully automated system to determine the optimum sample size and also to generate the optimum sample pattern for digitizing any given terrain. The essential features of ATOSAP are as follows:

- It uses the HVM program to interactively sample a stereo-model to obtain the optimum sample size using the criterion in Equation 4. As an alternative, the operator could use any of the linear regression equations. Equations 6 through 15 in conjunction with Equation 16.
- ATOSAP then calls AUTOCO to determine apriori

TABLE 3. THEORETICAL EXPECTATION AND EMPERICAL RESULTS OF THE EFFICIENCY OF SAMPLING PATTERNS

Surface	Class	Sample Size	Theoretical Expectation	Emperical Result
1	I	25	rand. < str.	rand. < str.:
		64	sys. < str.	sys. < str.
				sys. str.
2	I	36	rand. < str.	rand. < str.:
				sys. < str.
		100	No decision	No decision
		49	rand. < str.	rand. < str.
				sys. < str.
3	I	100	rand. < str.;	rand. < str.:
			rand. < sys.	rand. < sys.
		100	No decision	No decision
4	II	169	sys. < str.	sys. < str.
5	III	256	sys. < str.	sys. < str.
		324	sys. < str.	sys. < str.
		400	sys. < str.	sys. < str.
		100	sys. < str.	rand. < str.
				sys. < str.
6	IV	196	No decision	No decision
		484	No decision	sys. < str.
		256	sys. < str.	sys. < str.
7	III	394	sys. < str.	sys. < str.
		400	sys. < str.	sys. < str.
		256	sys. < str.	sys. < str.
8	III	324	sys. < str.	sys. < str. < random
		400	sys. < str.	sys. < str.

"<" = better than
sys. = systematic sampling
str. = stratified sampling
rand. = random sampling

which of the sampling patterns is relatively most efficient according to Das's theorems.

- GENSAP is called by ATOSAP to generate the *X,Y* coordinates for the theoretically most efficient sampling pattern. If the operator wishes, he might call any type of sampling pattern of his own choice, especially for an unusual occurrence of "Breaklines" on a terrain—a situation which can also be detected by ATOSAP.

ACKNOWLEDGMENTS

The author wishes to express grateful thanks to Professor S. K. Ghosh for his supervision of the research reported in this paper. Extensive use of the facilities of the Instruction and Research Computer Center of the Ohio State University is gratefully acknowledged.

REFERENCES

Ayeni, O. O., 1976a. *Considerations for Automated Digital Terrain Models with Applications in Differential Photo Mapping*, Ph.D. dissertation, Ohio State University, Columbus, Ohio.

———, 1976b. Objective Terrain Description and Classification for Digital Terrain Models, *XIII ISP Congress Proceedings*, Vol. 3, paper 3-12 com. III, Loftstrom *et al.* (Eds).

Blaschks, W., 1968. "Problems in Obtaining and Processing of Photogrammetric Data for Digital Terrain Models." Invited Paper, XI International Congress of Photogrammetry, Inter-Commission Working Group V/IV Lausanne.

Cochran, W. G., 1946. Relative Accuracy of Systematic and Stratified Random Samples for a Certain Class of Population, *Annals of Mathematical Statistics*, Vol. 17.

Das, A. C., 1950. Two dimensional Systematic Sampling and Associated Stratified and Random Sampling, *Sankhya*, Vol. 10.

Makarovic, B., 1973-75. Progressive Sampling for Digital Terrain Models, *I.T.C. Journal.*

Morrison, J. L., 1970. A Link Between Cartographic Theory and Mapping Practice; The Nearest Neighbor Statistic, *Geographical Review*, Vol. 60.

Quenoville, M. H., 1949. Problems in Plane Sampling, *American Association of Mathematical Statistics*, Vol. 20.

(Received 15 October 1981; accepted 14 January 1982; revised 1 March 1982)

Appendix
Major Sampling Patterns

There are three major sample types well known in statistics. These are

- Random sampling
- Stratified sampling
- Systematic sampling

Quenoville (1949) has rightly pointed out that there are many ways in which we can sample a two-dimensional space because there is flexibility in employing random, stratified, or systematic sampling in either direction. However, Morrison (1970) has demonstrated that six sampling patterns can theoretically represent nearly all the possible sample point scatters in a plane. This author feels intuitively that a seventh sampling pattern which combines all the features of the three major sampling patterns has some merits which deserve investigation. These seven sample types are now briefly described.

UNALIGNED RANDOM SAMPLING

Unaligned random sampling is a type of sampling in which each point is chosen randomly. In a two-dimensional plane this means that the coordinates X and Y of a given point are selected at random either by using a table of random numbers or by generating random numbers through a computer program. This gives rise to an uneven areal coverage. It is believed by statisticians that this type of sampling is efficient if there is periodic variation or any type of trend in the population because it gives rise to an uneven areal coverage.

ALIGNED RANDOM SAMPLING

This type of sampling is similar to the previous one except that the random number in one direction, X or Y, is fixed and in the other is chosen randomly. This type of sampling does not have as good areal coverage as its counterpart without alignment.

UNALIGNED STRATIFIED SAMPLING

This is a type of sampling in which the area concerned is subdivided into strata within which sampling points are chosen randomly in a manner similar to unaligned random sampling. The advantage of stratified sampling is that it tends to increase the precision of the estimate of a population parameter without increasing the number of points because the areal coverage of points seems to be more representative of the population.

ALIGNED STRATIFIED SAMPLING

This is the same as its unaligned counterpart except that one of the coordinates is aligned within each stratum.

UNALIGNED SYSTEMATIC SAMPLING

Unaligned systematic sampling is generated by dividing the area into sections (rectangulars or squares) and points are sampled randomly in each section.

ALIGNED SYSTEMATIC SAMPLING

This is by far the most popular type of sampling used in digital terrain model studies because it is the easiest to generate. The initial point is selected randomly or purposefully and all others are determined by a fixed interval.

UNALIGNED STRATIFIED SYSTEMATIC RANDOM SAMPLING

As the name implies, this sampling type is a combination of all the three major sampling types. The area concerned is covered with squares or rectangular grids and the first point is selected at random in the first square. The X coordinate of this first point is then used with a new random Y coordinate to locate the new point in the second square. A new point is similarly treated in the subsequent squares in the first row. The second and subsequent rows of squares are treated like the first row to generate the points required. Figure 1 shows the examples of sample point scatters resulting from each of the seven sampling plans discussed above.