MEASURING THE DIMENSION OF SURFACES: A REVIEW AND APPRAISAL OF DIFFERENT METHODS

André G. Roy
Ginette Gravel
and Céline Gauthier
Département de Géographie
Université de Montréal
C.P. 6128 Succ. "A"
Montréal Qc
Canada H3C 3J7

ABSTRACT

The concept of fractals is being widely used in several cartographic procedures such as line enhancement, surface generation, generalisation, interpolation and error estimation. Such applications rely on the estimation of the fractional dimension (D) of lines and surfaces. Measuring D for surfaces can be achieved from contours and profiles extracted from the surface or from the variability of the surface taken as a whole. In a fractal and self-similar terrain, the values of D should be in agreement regardless of the method used. Mark and Aronson (1984) applied the variogram technique to DEM and observed sharp changes in D with scale suggesting that terrains are composed of nested structures with a highly disorganised and complex component (D=2.6) in the long range and a smooth component (D=2.2) in the short range. The high dimensions may not reflect the terrain itself but be the result of combining residual anisotropic effects at long distances. Tests performed on DEM (or portions of DEM) show that the short range dimensions of the surface variogram are consistent with those extracted from profiles and contours ($2.0 < D < 2.3$). Systematic variations of D with altitude and location were also observed indicating a lack of self-similarity in spite of the apparent self-similarity of the surface variogram.

INTRODUCTION

Several applications of fractals to cartographic lines and surfaces are now well entrenched in the literature. The fractional dimension which characterizes the geometry of a line or surface is a powerful tool for analysis, description and generation of cartographic data. Cartographic lines appear to be more easily reduced to fractal analysis than surfaces. One problem that arises when dealing with surfaces is the difficulty of estimating their fractional dimension. This difficulty is partly explained by the availability of several methods for the computation of the dimension and by the lack of self-similarity in natural terrains. This paper addresses these problems and discusses some implications for cartography.

FRACTALS AND THEIR APPLICATION IN CARTOGRAPHY

This brief review of fractals will emphasize the properties of fractal sets and their applications in the field of cartography. Fractals were introduced by Mandelbrot (1977) to describe, among other things, irregular lines and surfaces. Strictly speaking, fractal applies to entities which have an Hausdorff-Besicovitch dimension (D) greater than the topological dimension. The value of D characterizes the intricacy or the jaggedness of the entity. Lines will have dimensions varying from 1 to 2 while surfaces are described by values of D ranging from 2 to 3. As D increases towards the upper value of the range, the entity becomes highly complex and intricate and the process associated with the line (or surface) is space-filling.

Fractal models of lines and surfaces may be created through fractional Brownian processes (Mandelbrot 1975; Goodchild 1982; Burrough 1983). In practice, fractional Brownian functions may be generated from fractional Gaussian noise. Several properties of such processes are noteworthy. First, the variogram of fractional Brownian functions is described by

$$E[(z_i - z_{i+h})^2] = h^{2H} \tag{1}$$

where h is the distance (or lag) between two points and z_i, z_{i+h} are the values observed at point i and i+h respectively. The variogram takes on the form of a power function in which H should vary between 0 and 1. In the case of profiles

$$D = 2 - H \tag{2}$$

while for fractional Brownian surfaces

$$D = 3 - H. \tag{3}$$

Secondly, the covariance function of such processes also display a relationship with H and consequently with D. As H increases toward its upper limit, the positive auto-correlation between neighbouring values is very strong and the realization of the process is very smooth. The process is Brownian for H=0.5. When H gets below 0.5, then the process tends to become anti-persistent and negatively auto-correlated. Profiles and surfaces generated with values of H lower than 0.5 are very jagged and erratic (Burrough 1983; Goodchild 1982; Culling 1986) and will display a rapid succession of peaks and throughs. Similar interpretations of D in geostatistical terms may be found through the applications of the power spectrum (Mandelbrot 1982; Pentland 1983).

Fractal models display the property of self-similarity which may be viewed strictly as a cascading mechanism of a fractal generator (Mandelbrot 1977). Self-similarity also implies that H and therefore D, the fractional dimension of an entity, is constant with changes in scale. Thus, small portions of the process are replicates of the global structure. By looking at the realization of a fractional Brownian process, one cannot infer its scale. For fractal surfaces, the same process has operated across the whole entity and self-similarity is associated with isotropy, that is the lack of directional bias in the geostatistical properties of the surface. Thus, profiles extracted from the fractal surface will have the same dimension than that of the surface itself less one. Contours and coastlines will also display the same dimension than the profiles. Self-similarity implies dimensional consistency among the lines and the surface.

Because it deals with the effect of scale on the metric of lines and surfaces, the concept of fractals has proven to be very useful to cartographers. The addition (or elimination) of details into a cartographic entity say a line is a process that may be consistent with the fractal geometry (Buttenfield 1985). Line degeneralization pionnered by Dutton (1981) used the fractal dimension to introduce details into a generalized line. This fractalization process enhances the line. Algorithms for fractal interpolation are well known (Fournier et al. 1982) and are also used to generate terrain profiles (Frederiksen et al. 1985). Muller (in press a) proposed to rely on the property of self-similarity as a standard to assess the quality of line generalization. A generalized line should be self-similar to the original. Furthermore, efficient line generalization may be achieved through the application of walking-step algorithm which is a straightforward application of fractals (Dubuc 1985; Muller in press b).

Errors in sampling and measuring from cartographic data are also related to fractals as it was shown by Goodchild (1980). Because errors increase with the complexity of the entity, they will increase with D. Moreover, fractal dimensions may be used to determine the sampling density required to capture the variability of the phenomenon. Blais et al. (1986) and Dubuc (1985) provided methods of specifying the optimal resolution which are based upon the fractal behavior of the lines or profiles.

Surfaces and more specifically terrain generation has relied heavily on fractals (Mandelbrot 1975; Fournier et al. 1982; Goodchild 1982). Although fractional Brownian landscapes with dimensions of 2.2-2.3 achieve realistic representations of the surface of the earth, little is known about the dimensionality of natural terrains. The estimation of the dimension of a natural surface may be problematic, however. Different methods are available for computing D, all of which should yield similar results if the assumption of self-similarity holds. Given that the processes acting upon the landscape vary with scale, self-similarity may not exist at all scales and for all natural terrains (Mark and Aronson 1984). Goodchild (1982) reported systematic variation in D as we climb from the shorelines to the summit of Random Island. Mark and Aronson (1984) suggested that many landscapes are generated by nesting structures of varying complexity. The problem of estimating D for a surface is compounded by the fact that the different methods may be applied to data coming from different sources which may or may not incorporate a high degree of cartographic generalization.

MEASURE OF THE FRACTAL DIMENSION OF SURFACES

The properties associated with fractional Brownian processes are used to estimate the values of D which may be computed from the variograms of the surface (eq. 1 and 3) or of the profiles (eq. 1 and 2). These dimensions should be consistent with those extracted from contours or coastlines. Mark and Aronson (1984) presented a method to construct the surface variogram of a Digital Elevation Model (DEM) recorded along a regular grid. They proceeded as follows. They randomly selected 32000 pairs of points. Each point had to be within the largest circle drawn within the map. For each pair of points, distance and the squared difference in elevation were computed. The set of measurements was then divided into 100 distance classes of equal size and then the variance of each class was computed. Classes with less than 64 observations were omitted.

They applied their method to seventeen $7\frac{1}{2}$ quadrangles obtained from the USGS. In 15 out 17 cases, they reported that the surface variogram could be described by at least two markedly different slopes (H) and thus two dimensions. At short ranges, for distances smaller than 0.6 - 1 km, they observed relatively low values of D ($D < 2.48$ and close to 2.1) while for longer distances (1 to 4 km) they noted a sharp increase in D sometimes up to 2.8. The average D for this range is 2.6 thus suggesting a very irregular terrain. The low values of D in the short range are more consistent with what has been previously reported in the literature (see Culling 1986) and identifies the strong positive auto-correlation at the hillslope scale. The higher D values are more problematic, however, and although the authors suggest a structural interpretation of the high irregularity, this result is unexpected.

Several problems seem embedded in the method presented by Mark and Aronson (1984). First, the sampling plan is biased towards the long range of the variogram and the random selection of pairs of points within a circle will always generate many more middle and long distances than short ones. Thus, by allowing the random selection of pairs of points, the emphasis is put on the part of the variogram which is farther away from the origin. In view of the fact that the analysis of the variogram tends to rely on the proximal part of the plot, this sampling bias may be important and yield unreliable D values in the short range. Furthermore, normal use of variograms tends to exclude the variances computed for the range of distances farther away than $\frac{1}{4}$ of the maximum distance on the map. Finally, the surface variogram may be viewed as a composite of profile variograms which may display different characteristics according to the direction. Such directional biases will represent anisotropies of the terrain. In order to avoid a sampling bias, variances should be computed for selected distance classes. A point could be randomly chosen within the largest circle contained in the map. The selected distance could be walked from that point in a direction which would be determined randomly. If the end point of the walk is outside the circle, the pair of points would be rejected from the analysis. This scheme is advantageous because we control the distance classes which could be specified in a geometric progression and also because the number of pairs in each class could be determined a priori.

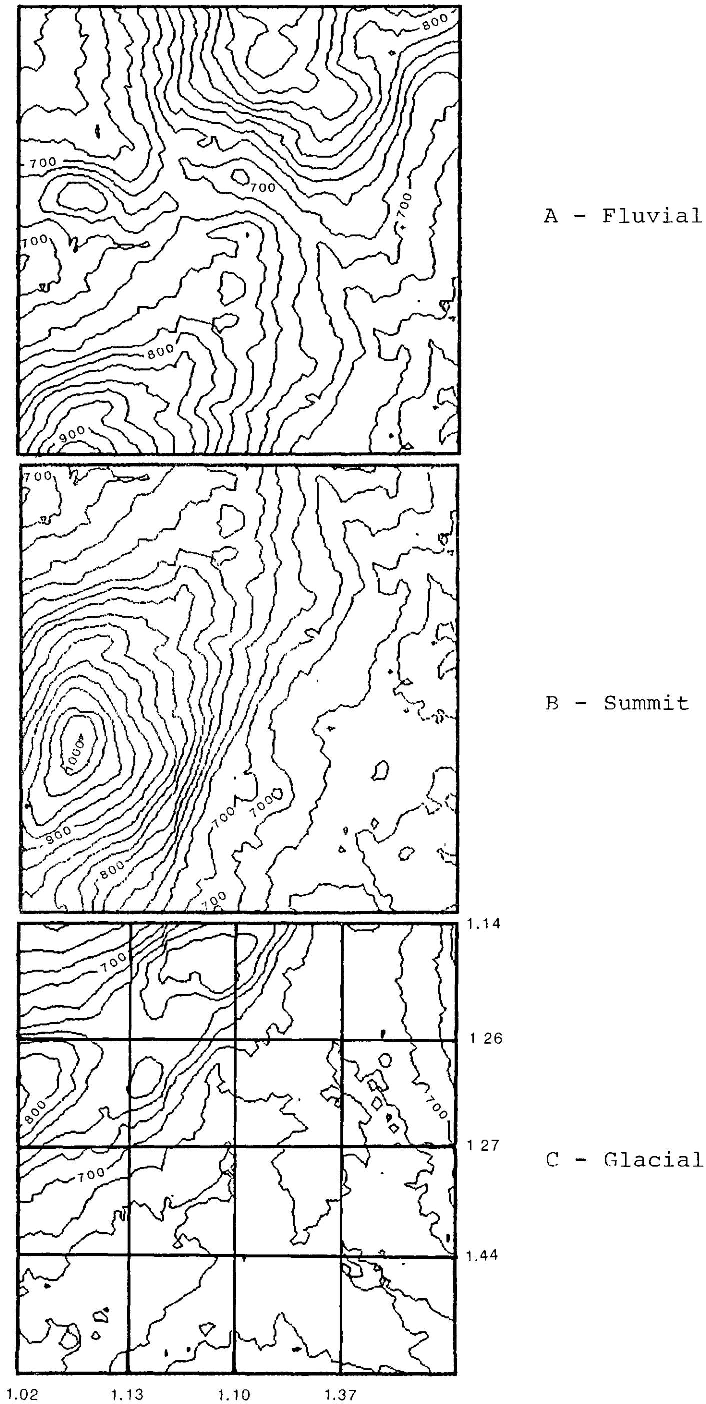
Nonetheless, the surface variogram should be preceded by an analysis of profile variograms which would allow to detect the presence of anisotropies. The search for directional bias can only be done in the NW, NE-SW, NW-SE, NS directions, however. All other directions would involve interpolated values and the variogram would not reflect the variability of the raw data. Profile variograms are simply build from the systematic sampling of all possible pairs of points separated by a distance h. Thus, the estimates of the variance at a longer range are derived from fewer pairs of points and the variogram should be reliable for distances shorter than one fourth of the maximal distance.

The dimension of a surface may be found from the dimensions of the contours and coastlines. In doing so, we are concerned with three problems. First, several techniques are currently used to estimate D for such lines. Most of these techniques involve the estimation of the rate of change in length with an increase in the sampling interval. The slope (b) of the log-linear relationship between the length of the line and the length of the divider used to measure it is given by

$$b = 1 - D . \tag{4}$$

Other methods rely on cell counting algorithms (Goodchild 1982; Shelberg _et al_. 1983). Goodchild (1982) compared several techniques and obtained higher estimates of D when length was measured from a cell counting method. Reliability of each method is difficult to assess, however. Secondly, the selection of the contours that we submit to fractal analysis may be critical. In a self-similar terrain, this would not be of concern since all contours display similar complexity. Natural terrains may exhibit systematic variations in complexity as was pointed out by Goodchild (1982). Shelberg _et al_. (1983) suggested that a set of contours should be used in the analysis. Finally, should the contours be taken from the maps (Goodchild 1982) or derived from the DEM itself (Shelberg _et al_. 1983) ? If cartographic generalization preserves self-similarity, then the source of the data would not affect the estimation of D. Such a postulate remains to be shown.

FIGURE 1: Three 80 x 80 windows extracted from the DEM



APPLICATIONS TO DIGITAL ELEVATION MODELS

The different methods of calculating D were applied to a USGS DEM of an area located in the White Mountains at the border of Quebec, Maine and New Hamshire (Moose Bog 7½ Quadrangle). Maximum relief in the quadrangle is 700 m. Three 80 x 80 windows illustrating different landscapes within the area - a fluvial landscape at the headwater of a stream (Fig. 1 A), a summit area (Fig. 1 B) and a valley filled with glacial sediments (Fig. 1 C) - were also submitted to fractal analysis. For the whole quadrangle and the three windows, D was estimated using four techniques:

- the surface variogram sampled using the fixed length technique;
- the variograms of profiles taken across the DEM in the EW, NS directions and along the diagonals;
- the contours digitized from the topographic map;
- the contours threaded into the altitude matrix.

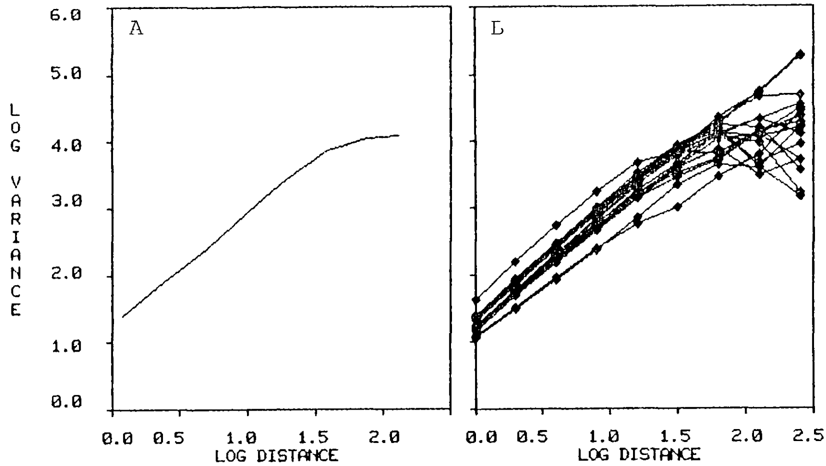
The dimensions of contours were evaluated using the dividers technique.

The surface variogram (Fig 2 A) obtained from the whole DEM shows an initial straight segment up to a lag of 2.0 km (64 pixels) with a constant slope (H = .84). D is therefore equal to 2.16 and it indicates a strong positive auto-correlation of elevations. The distal part of the variogram for longer lags also has a trend (H = .18). The break in slope is sharp as was the case of the examples presented by Mark and Aronson (1984) but it occurs close to the limit of reliability of the variogram (one fourth of the maximum distance is 79 pixels). The variances computed for the surface result from the composite effects of the profiles. This is shown in Figure 2 B where all 20 profile variograms are plotted. We note that the slopes of the initial segment are relatively constant while the distal parts of the variograms are highly variable. The residual trend observed in the surface variogram is clearly the amalgam of highly variable behaviors at longer distances and cannot be meaningfully interpreted. Thus, we conclude from the surface variogram that it describes an apparently self-similar terrain. This conclusion was also confirmed by plotting the profiles to scale and sampling them at various intervals. Smoothness of the terrain was always evident.

TABLE 1: Dimensions computed from different methods for the DEM as a whole

METHOD	D	DMIN	DMAX
Surface Variogram	2.16	----	----
EW Profile Variograms(9)	1.13	1.06	1.19
NS Profile Variograms(9)	1.17	1.09	1.28
Diagonal Profile Variograms(2)	1.21	1.17	1.25
Digitized Contours(13)	1.17	1.06	1.33
Threaded Contours(47)	1.09	1.01	1.28

FIGURE 2: Surface (A) and profile (B) variograms for the whole DEM



The average values of D computed from all methods are strickingly consistent (Table 1). The low D confirm the smoothness of the landscape despite a great amount of vertical relief in the area. Minimum and maximum D values show some variability in the estimation of D. The variability is greater for the digitized contours. This is explained by the sampling plan which attempted to capture the whole range of contour complexity.

TABLE 2: Dimensions computed from different methods for three 80 x 80 windows

METHOD	WINDOW 1	WINDOW 2	WINDOW 3
Surface Variograms	2.13	2.10	2.21
EW Profile Variograms	1.11	1.10	1.28
NS Profile Variograms	1.17	1.13	1.15
Threaded Contours	1.07	1.08	1.10

The comparison of D values obtained for the three windows shows differences among the terrain complexities especially when we look at the values obtained from the variograms. The valley filled with glacial deposits (Fig. 1 C) has a higher complexity than the summit area (Fig. 1 B) or the fluvial landscape (Fig. 1 A). This difference had to be expected and becomes even more important when the dimensions of individual profiles are compared (Fig. 1 C). All profiles that entirely cut through the valley bottom have a high dimension ($D \simeq 1.37 - 1.44$) while those on the hill side are very smooth ($D < 1.10$). Some profiles combine the attributes of both types of terrain. Thus within a relatively small terrain we assist to rapid changes in complexity depending on the nature of the sediments. At this scale, the lack of self-similarity is shown through a juxtaposition of terrain rather than by nesting smooth within complex structures as is evident from an examination of the contours. Contours become less intricate with altitude (Fig. 3). This is due to the erratic nature of the

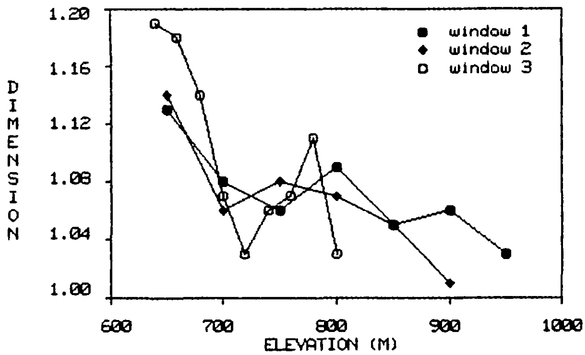


FIGURE 3: Variation of D with elevation

glacial deposits which were laid upon the valley bottom and to the gradual disappearance, as we climb towards the summits, of the crenulations associated with fluvial erosion.

DISCUSSION

Despite the apparent self-similarity of the whole quadrangle, spatial variations in D occur within the DEM. These effects are not detected from the surface variogram because they are not scale-related and the averaging process cancel their individual effect. Goodchild (1982) has also reported similar changes in the fractional dimensions of contours with altitude. In fact, one should anticipate that the dimensionality of most natural terrains should vary spatially. Variations in processes and/or structures may be responsible for these changes in dimensions. For example, creep will produce smoother surfaces than rill erosion. Systematic variations in the dimension within the surface bear important cartographic consequences. For instance, the degeneralization of contours should not be carried out using a unique fractalization process. Elevation and physiographic location must be used to guide the interpolation and enhancement procedures. A similar rationale also applies to terrain sampling as the optimal density should be a function of terrain complexity. Hilltops (in this case study) should be represented with fewer points than the valley floors. The fractal description of a surface should provide useful information for terrain generation and reconstruction and more attention should be given to the fractal signature of characteristic terrains (e.g. fluvial, morainic, eolian landscapes). This conclusion is not unlike that of Mark and Aronson (1984) who viewed the nested structure of terrains as a key component of surface generation. We suggest, however, that the nesting effect is not terrain-related at least in the range of distances where it was observed but rather that the detection of self-similar patches of terrains could be used advantageously by cartographers.

ACKNOWLEDGEMENTS

We acknowledge the financial support of EMR, Canada.

REFERENCES

BLAIS, J.A.R., CHAPMAN, M.A. and LAM, W.K., 1986, Optimal interval sampling in theory and practice: Proc. 2nd Conf. Spatial Data Handling, 185-192.

BURROUGH, P.A., 1983, Multiscale sources of spatial variation in soil: I. The application of fractal concepts to nested levels of soil variation: Jour. of Soil Sci., 34, 577-598.

BUTTENFIELD, B., 1985, Treatment of the cartographic line: Cartographica, 22, 2, 1-26.

CULLING, W.H.E., 1986, Highly erratic spatial variability of soil-pH on Iping Common, West Sussex: Catena, 13, 81-98.

DUBUC, O., 1985, La résolution optimale dans la généralisation de lignes, Unpublished M.A. thesis, Dep. de Géographie, Université de Montréal.

DUTTON, G.H., 1981, Fractal enhancement of cartographic line detail: The American Cartographer, 8, 23-40.

FOURNIER, A., FUSSELL, D. and CARPENTER, L., 1982, Computer rendering of stochastic models: Graphics and Image Processing, communication of the ACM, 25, 371-384.

FREDERIKSEN, P., JACOBI, O. and KUBIK, K., 1985, A review of current trends in terrain modelling: ITC Jour., 101-106.

GOODCHILD, M.F., 1982, The fractional Brownian process as a terrain simulation model: Modelling and Simulation, 13, 1133-1137.

GOODCHILD, M.F., 1980, Fractals and the accuracy of geographical measures: Math. Geo., 12, 2, 85-98.

MANDELBROT, B., 1975, Stochastic models for the earth's relief, the shape and the fractal dimension of the coastlines, and the number-area rule for islands: Proc. Nat. Acad. Sci. USA, 72, 3825-3828.

MANDELBROT, B., 1977, Fractals, Form, Chance and Dimension, San Francisco, Freeman, 365 p.

MANDELBROT, B., 1982, The Fractal Geometry of Nature, San Francisco, Freeman, 468 p.

MARK, D.M. and ARONSON, P.B., 1984, Scale-dependent fractal dimensions of topographic surfaces: An empirical investigation, with applications in geomorphology and computer mapping: Math. Geo., 16, 671-683.

MULLER, J.-C., in press a, Fractal dimension and inconsistencies in cartographic line representations: The Cartographic Journal.

MULLER, J.-C., in press b, Fractal and automated line generalization: The Cartographic Journal.

PENTLAND, A.P., 1983, Fractal-based description: Proc. of I.J.C.A.L., 973-981.

SHELBERG, M.C., MOELLERING, H. and LAM, N., 1983, Measuring the fractal dimensions of surfaces: Proceedings, AUTO-CARTO 6, 319-328.