Review Article

The Fractal Nature of Geographic Phenomena

Michael F. Goodchild* and David M. Mark†

*Department of Geography, University of Western Ontario, London, Ontario N6A 5C2
†Department of Geography, State University of New York, Buffalo, NY 14260

Abstract. Fractal concepts have attracted substantial popular attention in the past few years. The key ideas originated in studies of map data, and many of the applications continue to be concerned with spatial phenomena. We review the relevance of fractals to geography under three headings; the response of measure to scale, self-similarity, and the recursive subdivision of space. A fractional dimension provides a means of characterizing the effects of cartographic generalization and of predicting the behavior of estimates derived from data that are subject to spatial sampling. The self-similarity property of fractal surfaces makes them useful as initial or null hypothesis landscapes in the study of geomorphic processes. A wide variety of spatial phenomena have been shown to be statistically self-similar over many scales, suggesting the importance of scale-independence as a geographic norm. In the third area, recursive subdivision is shown to lead to novel and efficient ways of representing spatial data in digital form and to be a property of familiar models of spatial organization. We conclude that fractals should be regarded as a significant change in conventional ways of thinking about spatial forms and as providing new and important norms and standards of spatial phenomena rather than empirically verifiable models.

Key Words: fractals, fractional dimension, self-similarity, scale, spatial sampling, topography, simulation, spatial data structure, recursive subdivision.

IT IS clear from recent general-interest articles in a variety of journals and magazines that the concepts loosely associated with the term "fractals" have broad general appeal (see for example the *Economist* of 8 September 1984; the *New Scientist* of 4 April 1985; the *New York Times* of 22 January 1985; *Science* of 3 August 1984). It is evident from the text and illustrations of Mandelbrot's two definitive books on the subject that "fractals" are concerned, at least superficially, with spatial phenomena (Mandelbrot 1977, 1982b); those books deal in part with simulation of terrain and in part with measures of length and area of geographic features, expanding on work dating back to the late 1960s (Mandelbrot 1967, 1975a). It is appropriate, then, to undertake a more thorough and dispassionate review of the significance of these ideas to the spatial analytic tradition of geography. Though we find much of the hyperbole of the popular pieces cited above to be extravagant, we do believe that, based on breadth of impact alone, these are among the more significant mathematical ideas of recent times and that they are of direct relevance to a number of areas of spatial analysis.

If a line is measured at two different scales, the second larger than the first, its length should increase by the ratio of the two scales; areas should change by the square of the ratio. Yet because of cartographic generalization, the length of a geographical line will in almost all cases increase by more than the ratio of the two scales because new detail will be apparent at the larger scale. In effect the line will behave as if it had the properties of something between a line and an area. A fractal is defined, nontechnically, as a geometric set — whether of points, lines, areas or volumes — whose measure behaves in this anomalous manner.

The emphasis of this paper is on particular classes of fractals and on their applications to spatial phenomena. We note, however, that other fractal functions should be of interest to geographers and that other aspects of spatial phenomena can be modeled by fractals. For example, Mandelbrot and Wallis (1968, 1969) used fractal functions to model

Annals of the Association of American Geographers, 77(2), 1987, pp. 265–278

hydrologic time series; Mandelbrot (1975b) also presented a fractal analysis of turbulence in fluids, work that should be of interest for both hydrologic and atmospheric flows. Also, Lovejoy and his coauthors have applied a fractals model to link the spatial distributions of rainfall, clouds, and other atmospheric phenomena (Lovejoy 1982; Lovejoy and Schertzer 1983; Lovejoy and Mandelbrot 1985; Lovejoy, Schertzer, and Ladoy 1986). Burrough (1983a, 1983b) applied a fractal model to the spatial distribution of soil properties. Only space constraints prevent us from reviewing these applications of fractals to spatially related phenomena.

This paper is organized in three sections, which consider in turn the relevance of three basic concepts: response of measure to scale, self-similarity, and recursive subdivision of space. Each of these will be defined in the appropriate section. The organization is largely for convenience of presentation and should not be taken to imply any degree of mutual exclusivity between the three concepts or the corresponding sections.

Response of Measure to Scale

A circle of unit radius about the origin, defined by the mathematical function $x^2 + y^2 = 1$, has a well-defined area and circumference length that can be verified empirically to the accuracy of available instruments. The cartographic representation of the shore of an island, on the other hand, is allowed to depend on scale through the process of cartographic generalization, and scale will therefore affect both length and area measures. As scale increases, more and more detailed irregularity will become apparent, raising the question of whether any finite limit exists for length. The fact that measured length increases with increasing accuracy of measurement was noted by Steinhaus (1960) and is sometimes known as "The Steinhaus Paradox" (Coffey 1981, 95; Bibby 1972). Mandelbrot (1967) and Maling (1968) also discussed this point and examined whether the amount of change in measure can be predicted; the same issue can be raised for cartographic lines in general (Buttenfield 1985).

The practical importance of this issue is not immediately apparent but has been growing rapidly in recent years. Richardson's (1961) reasons for carrying out one of the earliest studies of the relationship between cartographic line length and scale were to test hypotheses concerning the propensity to conflict between pairs of nations and the length of common boundary. He observed that any objective method used to measure the length of an irregular line has an implied sampling length, in the form, for example, of the diameter of a wheel rolled along the line or the step size of a pair of dividers. For many years, the forest products industry and its regulatory agencies have relied on dot planimetry to measure the area of irregular stands, with an accuracy clearly dependent on the density of dots (see, e.g., Barrett and Philbrook 1970). More recently the development of remote sensing and geographic information systems has made more and more cellular spatial information available for decision making, much of it in the form of estimated geometrical measures such as areas, lengths, and point counts. In such systems, the role of cartographic scale, and associated generalization, is played by such parameters as pixel size or the precision with which a digitizer operator follows a line.

Richardson was able to show that over a wide range of scales there is a tendency for the length of a variety of cartographic lines to behave in a predictable fashion such that when length is plotted against sampling interval on logarithmic scales the points tend to follow a straight line. Not only does more detail become apparent at larger scales, but it tends to do so at a predictable rate.

The more general implications of this observation were pointed out by Mandelbrot (1967, 1975a; and see Scheidegger 1970, 8–9), and the behavior of cartographic lines, particularly representations of coastlines, has become one of the best-known aspects of the fractal literature. One may define the *Hausdorff-Besicovich* dimension D of an irregular line as follows. Suppose that the length of the line is estimated by stepping dividers along the line with step size s_1, and that n_1 such steps are required to span the line, giving a length estimate of $n_1 s_1$. We now repeat the process with a smaller step size, s_2, and obtain another length estimate, $n_2 s_2$, which is greater than or equal to the previous one. The more irregular the line, the greater the increase in length between the two estimates. Then:

$$D = \log (n_2/n_1) \,/\, \log (s_1/s_2). \qquad (1)$$

If the line is smooth, then a halving of step size or sampling interval will require precisely twice as many steps, and D will equal 1; but if it is irregular, then D will be greater than 1.

A similar argument using square pixels to measure the area of an irregular patch or a rough surface can be used to show that if the length of each pixel's side is halved, at least four times as many pixels will be required to cover the same patch or surface, leading to a value of D of 2 or greater. In the limit of an irregular line that completely fills the plane, the value of D reaches an upper limit of 2. Similarly, an infinitely rugged surface that fills the third dimension has a limiting fractional dimension of 3, whereas a smooth surface has a D of 2. D can also be defined for point sets in a similar fashion.

Mandelbrot (1977) coined the term *fractal* to describe any function for which the Hausdorff-Besicovich dimension exceeds the topological dimension (e.g., 0 for points; 1 for lines; 2 for areas); for such functions, D is commonly termed the *fractal dimension*.

The problem of measuring D from a digitized version of a cartographic line has aroused a certain amount of interest, as it is desirable that the procedure be as simple as possible. Suitable algorithms have been described by Shelberg, Moellering, and Lam (1982) and Shelberg and Moellering (1983), and more recently Eastman (1985) has described a measure based on angles between adjacent segments that is closely related to D.

It is easy to show that the slopes of Richardson's (1961) plots are equal to $1\text{-}D$ and that if the points fall on a straight line, this is diagnostic of a feature with a constant fractional dimension over a wide range of sampling intervals. The implications of this observation are quite profound, and they led ultimately to the choice of title for Mandelbrot's 1982 book, *The Fractal Geometry of Nature,* and to many of the more extravagant claims being made of the field. First, the concepts of a fractional dimension and dependence of measure on scale are quite foreign to much of mathematics and may therefore offer the first effective tools for understanding the irregularity widely observed in the geometry of real phenomena. Second, the numerical value of D may be the most important single parameter of an irregular cartographic feature, just as the arithmetic mean and other measures of central tendency are often used as the most characteristic parameters of a sample. And finally it appears that such seemingly chaotic features as coastlines may behave in certain ways with substantial regularity.

The response of measure to scale has some other interesting implications. Woronow (1981) characterized the shapes of ejecta blankets on Mars by a measure of shape relating perimeter to area. As the shapes are irregular, they can be treated as fractal curves. To compensate for the effect of scale on the estimated length of perimeter, Woronow created a dimensionless shape measure by dividing perimeter by area to the $D/2$ power, rather than the square root of area. Church and Mark (1980) discussed the relation of the fractals model to a similar situation involving mainstream length and basin area.

Several recent studies have explored the generality of Richardson's observations, in effect evaluating the goodness of fit between real phenomena and a fractal of constant D. Goodchild (1980) showed that although Richardson's data for the west coast of Britain had produced a straight line, a similar analysis of the east coast showed two distinct domains, one being relatively smooth (low D) at large scales and the other relatively rugged (high D) at small scales. In the same paper, a reanalysis of data on Swedish lakes collected by Håkanson (1978) showed substantial departure from straight lines. In a more systematic study of the topography of Random Island, Newfoundland, Goodchild (1982) found that the coastline, the 250-ft. and 500-ft. contours, and the lake shorelines all displayed straight-line plots of number of steps against step size, and thus constant D, but that D varied systematically from low values near the shore to higher values on the plateau of the island. In an analysis of digital elevation models of the topography of several areas of the U.S., Mark and Aronson (1984) found distinct horizontal scale domains within which D remained substantially constant but had sharp breaks of slope at certain sampling intervals. D was usually much higher at longer horizontal scales (intervals greater than about 600 m) than over shorter distances. Both studies concluded that the domains could be interpreted in terms of the different geomorphological processes and geological constraints operating either at different elevations or at different scales. Bradbury, Reichelt, and Green (1984; see also Bradbury and Reichelt 1983; Mark 1984) found differences in D with scale for coral reef contours, and related these differences to ecological processes. Similar comparisons of real phenomena with the fractal model have been reviewed by Burrough (1981). Brown and Scholz (1985) examined the fractal dimension of topography of natural rock surfaces at scales between 20 microns and 1 m.

It is quite clear from this work that most real coastlines and other spatial entities are not fractals in the pure sense of having a constant D but in a looser sense of exhibiting the behavior associated with noninteger dimensions. D thus provides a characteristic parameter whose variation can be usefully interpreted in terms of the processes that have influenced the entity's development.

Fractional dimension is clearly unique in its ability to predict the effects of generalization and spatial sampling. Goodchild (1980) showed that D can be used to estimate errors in dot planimetry and by extension to optimize dot density, and a similar approach can be applied to pixel sizes in remote sensing. Consider for example a LANDSAT scene for which all pixels have been correctly classified to show presence or absence of woodland. The total area of woodland is computed by counting pixels, but the size of the pixels will clearly affect the accuracy of the estimate, which will be much lower if the woodland is scattered over the scene in small parcels than if it is concentrated in one, singly bounded circular patch. The standard error as a percentage of the area estimate can be shown to be proportional to $a(1\text{-}D/4)$ where a is the area of a pixel and D is a characteristic parameter of the phenomenon. Standard error will thus depend on $a^{1/2}$ for highly scattered woodland and $a^{3/4}$ for single, circular patches with smooth boundaries (see Goodchild 1980 for the derivation of these results and their relationship to the literature in this area). The former result is readily obtainable from the standard deviation of the binomial distribution.

Self-Similarity

A feature is said to be self-similar if any part of the feature, appropriately enlarged, is indistinguishable from the feature as a whole. The term "indistinguishable" can be taken in its precise sense for regular features such as the constructions discussed in the next section, but clearly requires further interpretation for irregular features. We define an irregular feature such as a coastline as statistically self-similar if both the feature as a whole and any parts of it, suitably enlarged, were generated by the same stochastic process. Empirically, this means that we will be unable to reject a null hypothesis to that effect: in more general terms, all differences can be ascribed to chance.

It follows from this definition that irregular self-similar objects must have constant fractional dimension. Furthermore, as such objects are unaffected by enlargement, they must appear indistinguishable at all scales. An image or model of a self-similar object would thus possess no visual cues as to its scale, and an observer would be unable to estimate its size.

The probability distributions governing the sizes of self-similar phenomena must also lack characteristic scales, and it follows that they must therefore be hyperbolic or Pareto distributions of the form:

$$\Pr(X > x) = k\, x^{a} \tag{2}$$

where $\Pr(X > x)$ denotes the probability that the size X of an object exceeds a value x, k is a constant of proportionality, and a is a constant power. Hyperbolic distributions have the rank-size property that if the size of the object is plotted against its rank in the sample, with the largest object assigned the rank 1, and using logarithmic scales, then the points will fall on a straight line. Curl (1960, 1966, 1986) has shown that cave lengths have this property and has explored the possibilities of a fractal model for caves. The well-known rank-size property (for review see Richardson 1973) thus establishes self-similarity in city size distributions, although not in their spatial arrangements. Hyperbolic distributions are observed for an enormous variety of phenomena, many of them geographic, ranging from the areas of the world's lakes to the heights of tall buildings in major American SMSAs. Korčak (1940) and Fréchet (1941) were among the first to recognize this empirical regularity of the areas of lakes, areas of islands, and lengths of rivers, and it has come to be known as "Korčak's Law." The heights of the tallest mountains are not hyperbolic, however, because of obvious limits to growth.

Fractional Brownian motion (fBm) provides a relatively straightforward method of generating irregular, self-similar surfaces that resemble topography and that have known fractional dimension (Mandelbrot 1975a). Since the publication of Mandelbrot's first book (Mandelbrot 1977), this class of functions has attracted considerable attention as a method of simulating topographic surfaces. The functions are characterized by variograms of the form:

$$E[z(\mathbf{x}) - z(\mathbf{x}+\mathbf{d})]^2 = k\,(|\mathbf{d}|)^{2H} \tag{3}$$

where $E[\ \ ]$ denotes the statistical expectation, $z(\mathbf{x})$ is the height of the surface at coordinates denoted by the vector $\mathbf{x}$, $\mathbf{d}$ is a displacement vector, k is a constant, $|\mathbf{d}|$ is the magnitude (length) of the displacement vector, and H is a parameter in the range 0 to 1. In other words, two points a given distance apart are expected to have elevations whose squared difference is proportional to that distance to the power $2H$. Figure 1 shows example realizations of fBm for various values of H.

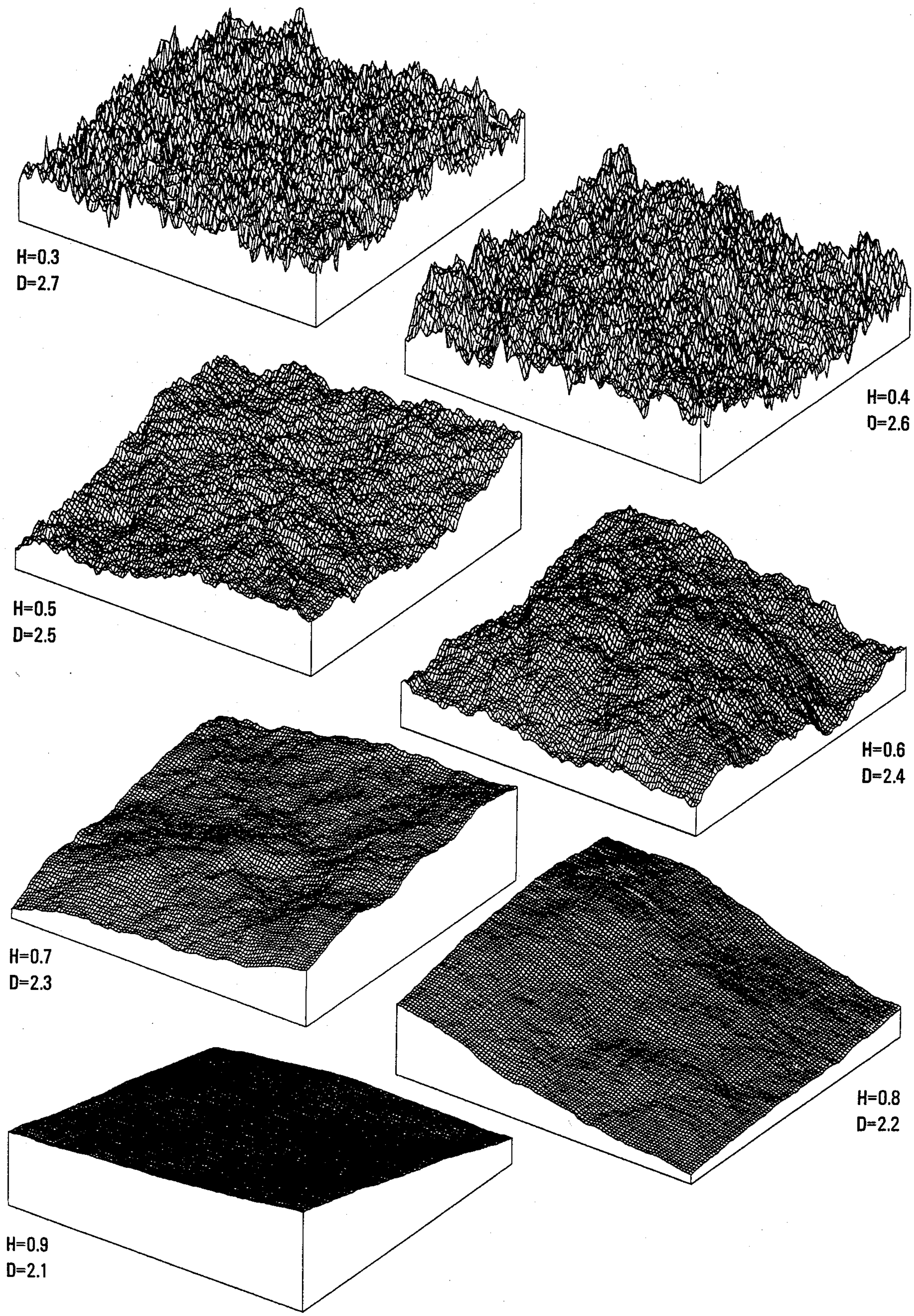

Figure 1. Self-similar surfaces generated with a fractional Brownian process using a range of values for the H parameter.

Orey (1970) has shown that the parameter H of an fBm surface is linearly related to fractional dimension D. Contours of the surface are lines with dimension 2-H, and the surface itself has dimension 3-H. We thus have a convenient way of generating self-similar objects of known D. Of greater significance, however, is Mandelbrot's discovery that certain values of H in the range 0.6 to 0.9 generate surfaces with striking similarity to real terrain. With sophisticated graphics display devices, snow patches, lakes, and tree cover can be simulated very convincingly (see for example Mandelbrot 1982b; Greenberg et al. 1982; Smith 1982). Such simulations have now been carried to a fine art by computer graphics specialists working in the motion picture, video game, and flight simulator industries.

The most successful value of H for simulating topography is approximately 0.7 (contour dimension 1.3), which results in surfaces suggestive of lunar topography or the very rugged terrain with strong local relief typical of dead ice topography. By flooding the surface to a certain level it is possible to generate reasonable replicas of complex island archipelagoes such as those of Southeast Asia. Few physical geographers would be willing, however, to extend the range of successfully simulated topographies very far. The characteristic of fBm realizations that peaks and pits occur randomly and with equal density is clearly a limiting factor, as pits are very rare in most terrestrial landscapes. In addition many real landscapes possess strong trends in elevation because of underlying geological structures, whereas any trends that may be apparent on fBm surfaces, particularly when H approaches 1, must be due strictly to chance. Finally, at global scales there are strong upper limits to topographic variance, implying an H of 0.

Informal experiments by Mark and Aronson (1948) using matched pairs of fBm and real terrain have shown that geomorphologists and cartographers seem to have little difficulty in successfully discriminating between them. The set of topographies acceptably simulated by fBm appear to be those with the greatest degree of self-similarity, whereas most real landscapes possess strong cues to scale attributable to the geomorphological or geological processes that have influenced them. For example, such features as glacial cirques and drumlins occur over relatively narrow ranges of scale, and their appearance in a photograph of landscape is sufficient to establish scale within narrow limits. In general, most geomorphological processes appear to have differential effects across scales, so that self-similar simulations appear to the eye as raw and unmodified. Nevertheless fBm provides a sufficient fit to real terrain for some purposes and undoubtedly a better fit than any other available random process (Mark 1978).

The form of a real topographic surface at any point in time is the result of the action of process on the forms that existed at previous times, and models of physical landscape development must therefore postulate some initial form on which modeled processes can operate. The Davisian cycle proposed an uplifted block as the initial form at the beginning of each cycle (Davis 1899), while Sprunt (1972) and Hugus and Mark (1984, 1985) used a tilted plane (interpretable as fBm with $H = 1$) and Craig (1980) adopted a surface of random, independent elevations, which is equivalent to fBm with $H = 0$. Both the block and tilted plane are regular surfaces, so that in order to simulate an irregular outcome it is necessary for an element of randomness to be included in the simulated process. That Davis failed to do this is evident from the regular nature of his diagrams. Sprunt (1972) introduced a random precipitation input, whereas Hugus and Mark (1984, 1985) included a random factor in the process-response part of the model. The problem is avoided if the initial surface form is irregular, and this, together with the raw, unmodified appearance of fBm surfaces noted above makes them particularly attractive as neutral, initial forms on which to simulate geomorphic processes. Kirkby (1986) adopted this approach, using an initial surface consisting of the sum of an inclined plane and a fractal. In summary, we argue that fBm provides a terrain that is unreal only in its lack of modification and is therefore of significant value for process simulation. By contrast, scale-dependent alternatives such as periodic (wave-like) surfaces are more readily interpreted as the outcomes of certain types of processes.

The hypothesis-testing tradition of statistics is concerned with establishing the presence or absence of an effect by comparing an observed sample with what would have been expected had the effect been absent. To be appropriate, then, a null hypothesis should propose the absence of the effect but should reflect reality in every other respect. If it differs in other respects as well, as is frequently the case in practice, and the null hypothesis is rejected, then it will be impossible to know with certainty whether it has been rejected because of the presence of the effect of interest or because of some other reason. In this conceptualization, the other effects are equivalent to the assumptions that often accompany a test. For example, the null hypothesis in a t-test of means may be rejected not only because of the presence of a real effect, but also because the parent populations are not normal or the variances are not similar.

It follows that the lack of an appropriate null hypothesis can lead to serious difficulties of interpretation. The significance of Horton's (1945) "laws" of channel networks became much clearer after Shreve (1966, 1967) provided a random model, or null hypothesis, of network topology and showed that a good fit to the Horton law of stream numbers probably implies an acceptance of the null hypothesis. More recent studies (James and Krumbein 1969; Abrahams and Flint 1983; Abrahams 1984b; see Abrahams 1984a for a review of other similar results) have shown systematic deviations from the Shreve model, or rejection of H_0, and these have often been interpreted in geomorphological terms. However the random model is concerned with topology alone and does not include geometric constraints, which may be added when stream networks are packed onto a surface in an area of uniform precipitation and which may influence the relative abundances of different topologies. Specifically, some networks are found (Abrahams 1984a) to show higher-than-expected bifurcation ratios and to favor significantly "fishbone" topologies over other types.

Abrahams (1984a), in a recent review of this literature, found that such tendencies appeared more common in high-relief basins, suggesting a geomorphological interpretation. Many studies have concluded by accepting the Shreve model, or H_0, but this outcome could be ascribed to insufficient data, in other words Type 2 statistical errors (see Abrahams and Mark 1986). The possibility still exists, therefore, that these effects are due to the constraints imposed by packing or to the inappropriateness of the null hypothesis of topological randomness. Simulation would appear to offer the only reasonable way of investigating this alternative.

Goodchild et al. (1985) described a series of experiments in which tree networks were obtained from fBm surfaces by a process intended to simulate the development of drainage. Each cell of a 258 by 258 array of elevations was compared to each of its rook's-case neighbors and coded according to the following rules: 0, if no neighboring elevation is strictly lower; 1,2,3,4, if at least one neighboring elevation is lower and the lowest is up, right, down or left respectively. The bordering cells were then discarded. Each code was taken to indicate direction of drainage, with 0's representing sinks. All flows directly into sinks were interpreted as the discharging links of individual basins, and the entire population of basins was extracted by search.

The simulated network populations obtained from fBm surfaces differed from the predictions of the topologically random model in several ways, and in many of these there was agreement with previously observed deviations between the random model and real networks. The relative abundance of "fishbone" basins exceeded the model's predictions, leading to differences in several aggregate indices such as bifurcation ratios. This suggests that the packing of basins onto surfaces does indeed bias populations to the point where deviations are observed from the random model and that therefore such deviations do not necessarily relate to the processes of basin formation.

It would be inappropriate to argue at this stage that the random model be replaced by fBm simulations as a null hypothesis for drainage networks, as we have already discussed several arguments on which fBm could be rejected a priori. The results indicate, however, the usefulness of fBm surfaces in providing norms for the interpretation of geomorphological observations. The Shreve model provided such a norm for the Horton laws and led to a radical change of interpretation. But it seems unlikely that other areas of quantitative geomorphology will be similarly amenable to statistical analysis, and, as we have seen, even the Shreve model may be an inappropriate norm in certain respects.

In addition, we would argue that the irregularity and unmodified appearance of fBm surfaces makes them much more suitable for the simulation of geomorphic process than are the alternatives that have appeared in the literature, such as uplifted and tilted blocks and independent, random elevations.

Recursive Subdivision of Space

Lines with the property of self-similarity can readily be generated by one of a number of suitable recursive procedures, one of which is illustrated in Figure 2. A single square cell is first divided into four equal cells, and the four central points are then connected to form the letter N. The

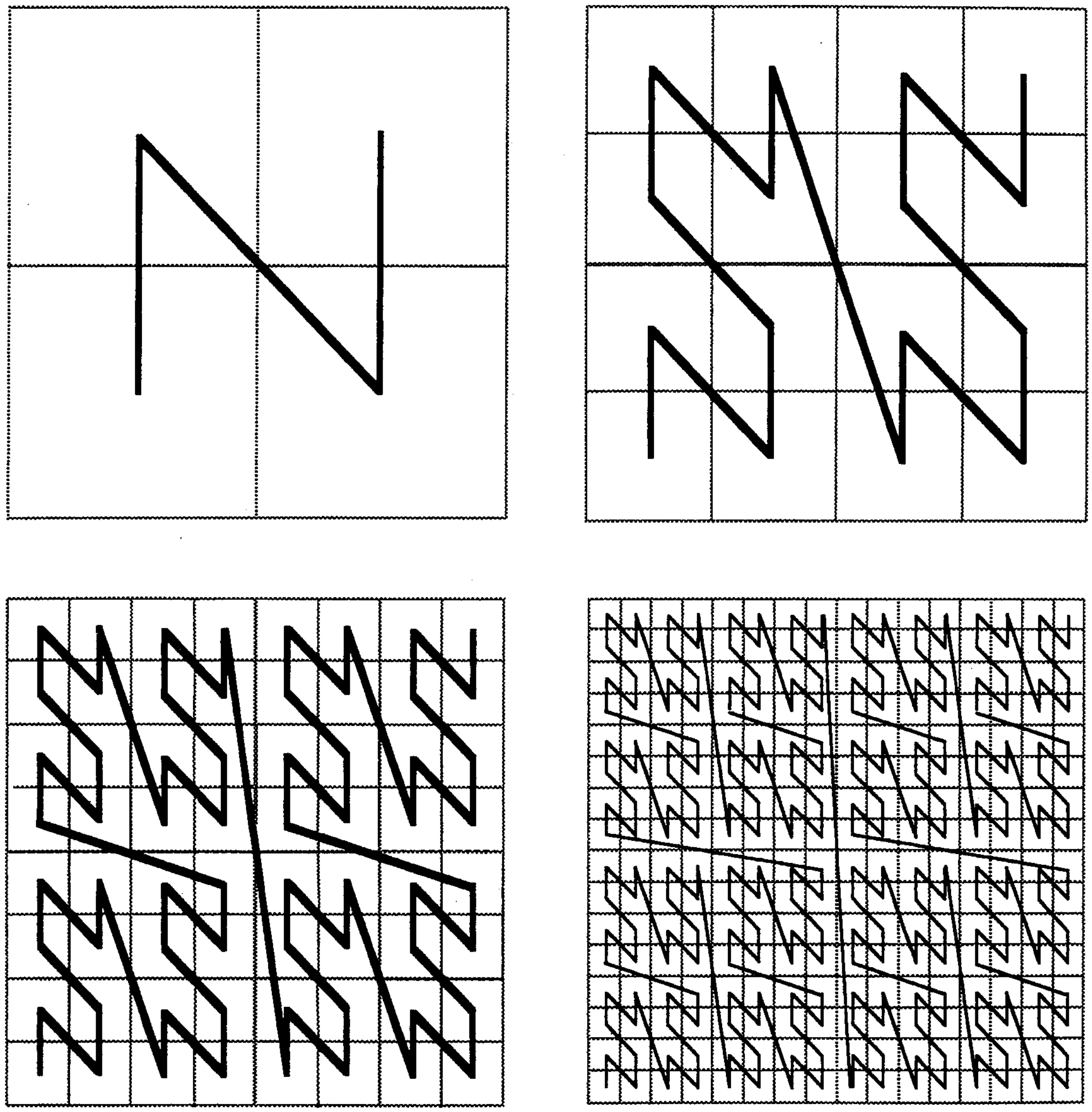

Figure 2. Recursive, self-similar generation of a space-filling (Peano) curve, the Morton sequence or N-tree.

process is then repeated in each of the four cells to give a line connecting 16 central points. The recursion, which is shown in Figure 2 up to the next step of 64 cells, can clearly be repeated indefinitely. This particular sequence was first described by Morton (1966), and many other similar recursive constructions are illustrated in the fractal literature (see for example Mandelbrot 1977, 1982b). The method of generation is sufficient to ensure that these regular curves are self-similar with a constant D, in this case equal to 2.

Fournier, Fussell, and Carpenter (1982a, 1982b; also Carpenter 1980; Fournier and Fussell 1980) applied the same concept of recursive subdivision to develop algorithms for the generation of irregular fractal curves and surfaces. An irregular line can be generated by beginning with a straight line and offsetting the position of its midpoint perpendicularly by an amount determined by a random number. The same process is then applied to the two straight halves, and recursively ad infinitum. In the case of surfaces, four random elevations are first generated in the form of a square (Fig. 3). The elevations of the five points at the next level of subdivision are obtained from the means of the existing neighbors, adjusted up or down by a random number. The standard deviation of this random disturbance controls the ruggedness of the resulting surface. Similar processes can be used to generate surfaces based on recursive subdivision of triangles.

Mandelbrot (1982a) has objected to these meth-

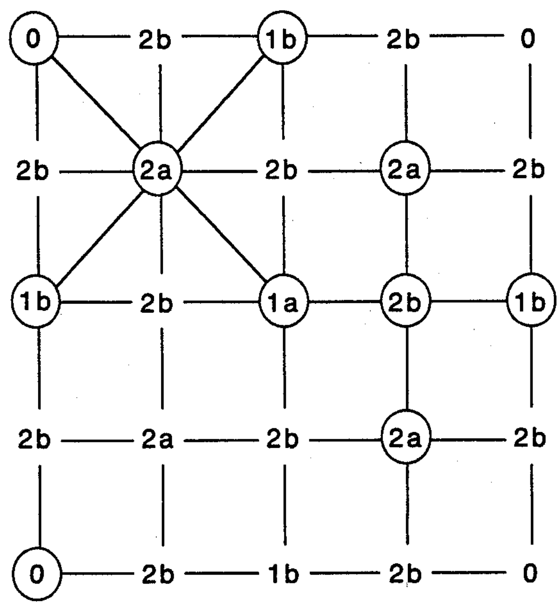

Figure 3. Order of computation for the recursive filling of a grid in two dimensions. Given values for the cells marked 0, the algorithm fills the others in the order: 1a, 1b, 2a, 2b, ... This is the sequence used by Fournier, Fussell, and Carpenter's (1982a) algorithm for the generation of surfaces having certain fractal properties.

ods on several related grounds: the resulting forms are not self-similar; they lack the power-law variograms characteristic of fBm; and their visual appearance is not as convincing. Significantly, he identified the final objection as the most critical, arguing that "the basic proof of a stochastic model of nature is in the seeing" (Mandelbrot 1982a, 581). Self-similarity is a general property of real landscape, he implies, and therefore any simulations that lack it will be visually unacceptable and furthermore will be widely perceived as such. The fBm surfaces discussed in the previous section of this paper were all generated by a different and theoretically sound method involving the fracturing of an initially flat surface by randomly located, straight cliffs.

Despite these objections, both the regular curves generated in Figure 2 and the recursive algorithms of Fournier, Fussell, and Carpenter illustrate an important implication of the self-similarity property — that spatial form is the same at all scales. The notion of scale independence of form has a long history in the study of spatial organization, appearing strongly in the hierarchies of central place theory. Some of the implications of the fractal concept of self-similarity for spatial organization have been explored by Batty (1985), who notes that recursion has also been applied to branching processes in the simulation of networks and trees. A recent paper by Arlinghaus (1985) has shown that the central place hierarchy has the properties of a fractal set, and Batty and Longley (1985) have used fractal structures in an experimental simulation of the spatial structure of London. Lovejoy, Schertzer, and Ladoy (1986) showed that the spatial distribution of world meteorological stations is a fractal set, which has implications for its ability to detect certain types of weather phenomena.

The field of spatial data handling (which underlies automated cartography, geographic information systems, and much of remote sensing) has traditionally seen the question of the digital representation of space in terms of two alternatives, raster and vector (for reviews of this issue and compromise alternatives see Peuquet 1979, 1981a, 1981b). Both are intuitively familiar, raster as an analog of the written page and vector corresponding to the free movement of the eye in surveying a scene or the pen in drawing a map. The fixed cell size of a raster representation implies a constant scale of resolution.

Recently, the quadtree has emerged as an alternative form of spatial representation based on recursive subdivision of space (for a review see Samet 1984); this representation has no obvious analog in everyday human experience. Subdivision is continued independently in each area of the map to a level that depends on the local complexity of data: in the case of a map of soil types, for example, subdivision might stop when a cell contained only one soil type or when some minimum size was reached. Quadtree structures can be searched more efficiently than rasters can be and generally require much less storage space.

Lauzon and Mark (Mark and Lauzon 1984; Lauzon et al. 1985) described a variant of quadtree spatial representation known as two-dimensional run encoding (2DRE). Let the four quadrants of a rule-of-four subdivision be ordered as lower left, upper left, lower right, upper right, in other words in the form of an N. The process of recursive subdivision of quadrants is then precisely the process of construction of the curve in Figure 2. Suppose all cells are subdivided to the same lowest level. Then in essence the structure consists of a reordering of a row-by-row raster in a regular, fractal curve. White (1983) has called a structure based on this ordering an N-tree for this reason, although the order itself was first implemented in a spatial data handling system by Morton (1966) in the Canada Geographic Information System (see Tomlinson, Calkins, and Marble 1976).

In a 2DRE structure, runs of cells of the same type are compressed. Goodchild and Grandfield (1983) compared 2DRE structures in terms of storage efficiency using several different orderings (see Fig. 4): conventional row-by-row rasters (row order), row-by-row rasters with alternate rows reversed (row-prime order), N (Morton) order, and another fractal curve known as the Hilbert-Peano curve (pi order, so-named for its pi-shaped primitive elements). In terms of storage efficiency, all orderings should produce the same volume of data if no spatial autocorrelation exists. If there is a tendency for local homogeneity, however, the most efficient ordering would be the one that best preserves local spatial relationships. Using simulated spatial data obtained from fBm surfaces, Goodchild and Grandfield (1983) were able to find some limited storage advantage in orderings in which every move was to a rook's-case neighbor. This is a property of the pi and row-prime orders, but not of the Morton order or the conventional row-by-row raster (see Fig. 4). Effective one-dimensional ordering of two-dimensional space is a problem of considerable theoretical and practical interest.

Discussion and Conclusions

In dividing this paper into three sections, we do not intend to imply that the three concepts represented are independent or unrelated. Indeed, self-similarity follows directly from the notion of a constant fractional dimension, and recursive subdivision from self-similarity. Rather, the intention has been to emphasize a successive broadening of

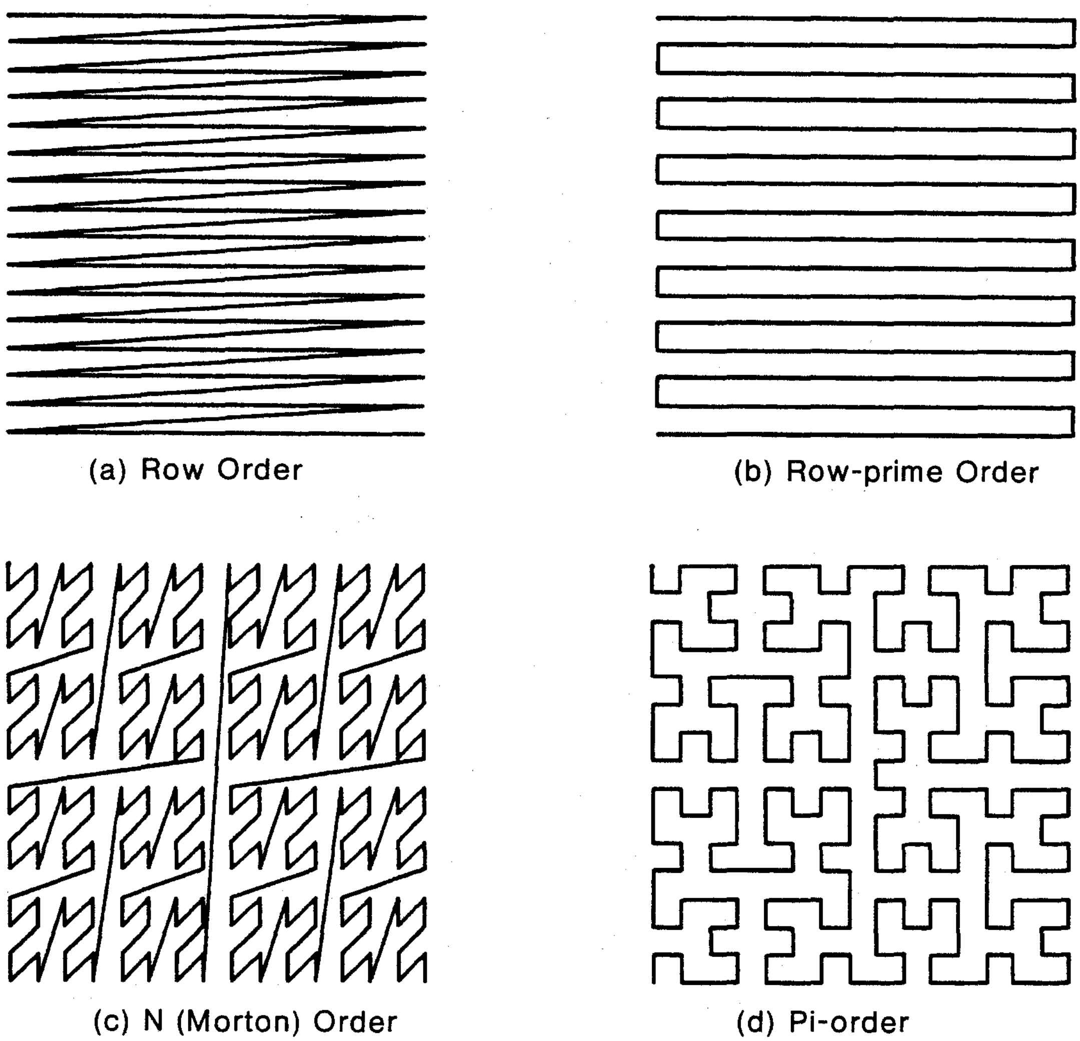

Figure 4. Four space-filling curves that can be used to order the cells in a square image (after Goodchild and Grandfield 1983). Note that the N-tree presented is a mirror image of the one constructed in Figure 2.

scope from technical problems of measurement and the explanation of Richardson's empirical results to the issue of a null hypothesis for terrain and finally to more conceptual developments in the recursive representation of space.

Also, our discussion has stressed mathematical and theoretical aspects of fractals; fractals are not, however, without applications in cartography and spatial data handling. Often, generalized versions of digital cartographic lines look angular and unrealistic. Dutton (1980, 1981) has described the use of fractal simulations to add realistic (yet spurious) detail to generalized cartographic lines in a reversal of the usual process of generalization; there have been several other similar applications (Hill and Walker 1982; Shelberg, Moellering, and Lam 1982; Armstrong and Hopkins 1983; Dell'Orco and Ghiron 1983; Shelberg and Moellering 1983; Shelberg, Lam, and Moellering 1983). In another important application of fractals, several authors (Goodchild and Grandfield 1983; Lam 1982; Lauzon et al. 1985; Mark and Lauzon 1985) have used fBm surfaces as test data sets to investigate issues of spatial data handling. They have argued that by varying H, it is possible to produce a range of lines and surfaces similar to many real cartographic lines but with controlled statistical properties. Other researchers can conduct comparative studies by generating similar simulated test data sets.

The significance of fractals clearly does not lie in the explanatory power of fBm as a model of terrain, despite Mandelbrot's strong suggestions to this effect (Mandelbrot 1975a, 1977, 1982b). The issue of whether a stochastic process can ever be regarded as an explanation has been discussed at length in the literature and need not concern us here; rather, it is evident that self-similarity is exhibited only in limited regions and over limited ranges of scale in real spatial phenomena. Nevertheless, both self-similarity and the concept of fractional dimension provide useful reference standards against which real phenomena can be compared and measured. For some purposes, such as the generation of test data for spatial data structures and the prediction of accuracy of spatial measures, the assumption of a constant D is frequently acceptable. Finally, there are no obvious alternatives to D as a single characteristic parameter of cartographic features with application in the prediction of the effects of generalization and scale change.

The fractional Brownian process has been used as a convenient way of generating self-similar surfaces, and certainly such surfaces more closely resemble some types of real topography than do the results of any other available method of simulation. We have argued that in addition to its popular acceptance in video games and science fiction movies, fBm offers a unique tool to geomorphology as a null hypothesis terrain. Its self-similarity gives it the appearance of rawness or lack of geomorphic modification, suggesting further application as an initial form for simulation of process.

Data structures intimately related to fractals and based on recursive subdivision of space are still relatively novel in the field of spatial data handling, and geographic information systems that exploit them are still under development. It is already clear, however, that they add substantially to the more traditional, intuitive options of raster and vector.

The publicity that fractals and related mathematical concepts have received is largely due to their strong visual impact, and there is not surprisingly a tendency to evaluate them in these terms. Mandelbrot has argued repeatedly that visual appearance is the most important test of a stochastic model of a natural phenomenon and that on this basis fBm surfaces must be accepted as models of terrain. This argument is of course unacceptable to most geomorphologists, both in practice given the limited number of terrain types with any resemblance to fBm and in principle given the known properties of these simulations. We have argued that it would be inappropriate to evaluate the usefulness of fractal concepts in spatial analysis in this limited and somewhat superficial way. Instead we have identified three interrelated conceptual themes, with associated and substantial applications to certain areas of spatial analysis. Each one represents a significant change in conventional thinking, and together they seem to us to offer the potential for significant advances.

Acknowledgments

Our work on this topic has been supported by Natural Sciences and Engineering Research Council of Canada grants A-6915 and A-7978. Mark's work on the recursive subdivision of space is supported by National Science Foundation Grant SES-8420789; the support of NSERC and NSF is gratefully acknowledged.

References

Abrahams, A. D. 1984a. Channel networks: A geomorphological perspective. *Water Resources Research* 20:161–88.

———. 1984b. The development of tributaries of different sizes along winding streams and valleys. *Water Resources Research* 20:1791–96.

Abrahams, A. D., and Flint, J. J. 1983. Geological controls on the topological properties of some trellis channel networks. *Bulletin of the Geological Society of America* 94:80–91.

Abrahams, A. D., and Mark, D. M. 1986. The random topology model of channel networks: Bias in the statistical tests. *The Professional Geographer* 38:77–81.

Arlinghaus, S. L. 1985. Fractals take a central place. *Geografiska Annaler* 67B:83–88.

Armstrong, M. P., and Hopkins, L. D. 1983. Fractal enhancement for thematic display of topologically stored data. In *Proceedings, Sixth International Symposium on Automated Cartography (Auto-Carto VI),* vol. 2, pp. 309–18. Ottawa.

Barrett, J. P., and Philbrook, J. S. 1970. Dot grid area estimates: Precision by repeated trials. *Journal of Forestry* 68:149–51.

Batty, M. 1985. Fractals—geometry between dimensions. *New Scientist* 105(1450):31–35.

Batty, M., and Longley, P. 1985. The fractal simulation of urban structure. Papers in Planning Research, No. 92, Department of Town Planning, University of Wales Institute of Science and Technology, Cardiff.

Bibby, J. 1972. Infinite rivers and Steinhaus' paradox. *Area* 4:214.

Bradbury, R. H., and Reichelt, R. E. 1983. Fractal dimension of a coral reef at ecological scales. *Marine Ecology Progress Series* 10:169–71.

Bradbury, R. J.; Reichelt, R. E.; and Green, D. G. 1984. Fractals in ecology: Methods and interpretation. *Marine Ecology Progress Series* 14:295–96.

Brown, S. R., and Scholz, C. H. 1985. Broad bandwidth study of the topography of natural rock surfaces. *Journal of Geophysical Research* 90(B14):12575–82.

Burrough, P. A. 1981. Fractal dimensions of landscapes and other environmental data. *Nature* 294:240–42.

———. 1983a. Multiscale sources of spatial variation in soil. I. The application of fractal concepts to nested levels of soil variation. *Journal of Soil Science* 34:577–97.

———. 1983b. Multiscale sources of spatial variation in soil. II. A non-Brownian fractal model and its application in soil survey. *Journal of Soil Science* 34:599–620.

Buttenfield, B. 1985. Treatment of the cartographic line. *Cartographica* 22:1–26.

Carpenter, L. C. 1980. Computer rendering of fractal curves and surfaces. In *SIGGRAPH '80 Conference Proceedings*. Seattle.

Church, M., and Mark, D. M. 1980. On size and scale in geomorphology. *Progress in Physical Geography* 4:342–90.

Coffey, W. J. 1981. *Geography: Towards a general spatial systems approach*. London: Methuen.

Craig, R. G. 1980. A computer program for the simulation of landform erosion. *Computers and Geosciences* 6:111–42.

Curl, R. L. 1960. Stochastic models of cavern development. *Bulletin, National Speleological Society* 43:3–11.

———. 1966. Caves as a measure of karst. *Journal of Geology* 74(5):798–830.

———. 1986. Fractal dimensions and geometries of caves. *Mathematical Geology* 18(8):765–84.

Davis, W. M. 1899. The geographical cycle. *Geographical Journal* 14:481–504.

Dell'Orco, P., and Ghiron, M. 1983. Shape representation by rectangles preserving fractality. In *proceedings, Sixth International Symposium on Automated Cartography (Auto-Carto VI),* vol. 2, pp. 299–308. Ottawa.

Dutton, G. H. 1980. A fractal approach to the control of cartographic detail. In *Proceedings, Computer Graphics '80,* pp. 371–81. Brighton, U.K.

———. 1981. Fractal enhancement of cartographic line detail. *American Cartographer* 8:23–40.

Eastman, J. R. 1985. Single-pass measurement of the fractal dimension of digitized cartographic lines. Paper presented at the annual meeting of the Canadian Cartographic Association, June, at Fredericton, New Brunswick.

Fournier, A., and Fussell, D. 1980. Stochastic modeling in computer graphics. *SIGGRAPH '80 Conference Proceedings*. Seattle.

Fournier, A.; Fussell, D.; and Carpenter, L. 1982a. Computer rendering of stochastic models. *Communications of the ACM* 25:371–84.

———. 1982b. Author's reply to "Comment on computer rendering of fractal stochastic models," by B. B. Mandelbrot. *Communications of the ACM* 25:583–84.

Fréchet, M. 1941. Sur la loi de répartition de certaines grandeurs géographiques. *Journal de la Societé de Statistique de Paris* 82:114–22.

Goodchild, M. F. 1980. Fractals and the accuracy of geographical measures. *Mathematical Geology* 12:85–98.

———. 1982. The fractional Brownian process as a terrain simulation model. *Modeling and Simulation* 13:1133–37. Proceedings, 13th Annual Pittsburgh Conference on Modeling and Simulation.

Goodchild, M. F., and Grandfield, A. W. 1983. Optimizing raster storage: An evaluation of four alternatives. In *Proceedings, Sixth International Symposium on Automated Cartography (Auto-Carto VI),* vol. 2, pp. 400–407. Ottawa.

Goodchild, M. F.; Klinkenberg, B.; Glieca, M.; and Hasan, M. 1985. Statistics of hydrologic networks on fractional Brownian surfaces. *Modeling and Simulation* 16:317–23. Proceedings, 16th Annual Pittsburgh Conference on Modeling and Simulation.

Greenberg, D.; Marcus, A.; Schmidt, A. H.; and Gorter, V. 1982. *The computer image: Applications of computer graphics*. Reading, Mass.: Addison Wesley.

Håkanson, L. 1978. The length of closed geomorphic lines. *Mathematical Geology* 10:141–67.

Hill, J. S., Jr., and Walker, S. E., Jr. 1982. On the use of fractals for efficient map generation. In *Proceedings, Graphics Interface '82*, pp. 283–89. Toronto.

Horton, R. E. 1945. Erosional development of streams and their drainage basins; hydrophysical approach to quantitative morphology. *Bulletin of the Geological Society of America* 56:275–370.

Hugus, M. K,. and Mark, D. M. 1984. Spatial data processing for digital simulation of erosion. In *Proceedings, Fall Meeting ASP/ACSM*, pp. 683–93. San Antonio.

———. 1985. Digital simulation of three-dimensional landscape evolution. *Modeling and Simulation* 16:305–9. Proceedings, 16th Annual Pittsburgh Conference on Modeling and Simulation.

James, W. R., and Krumbein, W. C. 1969. Frequency distribution of stream link lengths. *Journal of Geology* 77:544–65.

Kirkby, M. 1986. A two-dimensional simulation model for slope and stream evolution. In *Hillslope processes*, ed. A. D. Abrahams, pp. 203–22. Boston: Allen and Unwin.

Korčak, J. 1940. Deux types fondamentaux de distribution statistique. *Bulletin de L'Institut de Statistique* 30:295–99.

Lam, N. S. 1982. An evaluation of areal interpolation methods. In *Proceedings, Fifth International Symposium on Computer-Assisted Cartography (Auto Carto V)*, vol. 2, pp. 471–79. Washington, D.C.

Lauzon, J. P.; Mark, D. M.; Kikuchi, L.; and Guevara, J. A. 1985. Two-dimensional run-encoding fôr quadtree representation. *Computer Vision, Graphics, and Image Processing* 30:56–69.

Lovejoy, S. 1982. Area-perimeter relation for rain and cloud areas. *Science* 216:185–87.

Lovejoy, S., and Mandelbrot, B. B. 1985. Fractal properties of rain, and a fractal model. *Tellus* 37A:209–32.

Lovejoy, S.; Schertzer, D.; and Ladoy, P. 1986. Fractal characterization of inhomogeneous measuring networks. *Nature* 319:43–44.

Lovejoy, S., and Schertzer, D. 1983. Buoyancy, shear, scaling and fractals. In *Sixth Symposium on Atmospheric and Oceanic Waves and Stability*. Boston.

Maling, D. M. 1968. How long is a piece of string? *Cartographic Journal* 5:147–56.

Mandelbrot, B. B. 1967. How long is the coast of Britain? Statistical self-similarity and fractional dimension. *Science* 156:636–38.

———. 1975a. Stochastic models of the Earth's relief, the shape and the fractal dimension of coastlines, and the number-area rule for islands. *Proceedings of the National Academy of Sciences* 72:3825–28.

———. 1975b. On the geometry of turbulence, with stress on the fractal dimension of the iso-surfaces of scalars. *Journal of Fluid Mechanics* 72:401–16.

———. 1977. *Fractals: Form, chance and dimension.* San Francisco: Freeman.

———. 1982a. Comment on computer rendering of fractal stochastic models. *Communications of the ACM* 25:581–83.

———. 1982b. *The fractal geometry of nature.* San Francisco: Freeman.

Mandelbrot, B. B., and Wallis, J. R. 1968. Noah, Joseph and operational hydrology. *Water Resources Research* 4:909–18.

———. 1969. Some long-run properties of geophysical records. *Water Resources Research* 5:321–40.

Mark, D. M. 1978. Comments on Freiberger and Grenander's "Surface patterns in theoretical geography." *Computers and Geosciences* 4:371–72.

———. 1984. Fractal dimension of a coral reef at ecological scales: A discussion. *Marine Ecology Progress Series* 14:293–94.

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. *Mathematical Geology* 11:671–84.

Mark, D. M., and Lauzon, J. P. 1984. Linear quadtrees for geographic information systems. In *Proceedings, International Symposium on Spatial Data Handling*, vol. 2, pp. 412–30. Zurich, Switzerland.

———. 1985. The space efficiency of quadtrees: An empirical examination including the effects of two-dimensional run-encoding. *Geo-Processing* 2:367–83.

Morton, G. 1966. A computer oriented geodetic data base, and a new technique in file sequencing. Report prepared for IBM Canada Ltd., Toronto.

Orey, S. 1970. Gaussian sample functions and the Hausdorff dimension of level crossings. *Zeitschrift fur Wahrscheinlichkeitstheorie und verwandte Gebiete* 15:249–56.

Peuquet, D. J. 1979. Raster processing: An alternative approach to automated cartographic data handling. *American Cartographer* 6:129–39.

———. 1981a. An examination of techniques for reformatting digital cartographic data. Part 1: The raster-to-vector process. *Cartographica* 18(1):34–48.

———. 1981b. An examination of techniques for reformatting digital cartographic data. Part 2: The vector-to-raster process. *Cartographica* 18(3):21–33.

Richardson, H. W. 1973. Theory of the distribution of city sizes: Review and prospects. *Regional Studies* 7:239–51.

Richardson, L.F. 1961. The problem of contiguity. *General Systems Yearbook* 6:139–87.

Samet, H. 1984. The quadtree and related hierarchical data structures. *Computing Surveys* 16:187–260.

Scheidegger, A. E. 1970. *Theoretical geomorphology.* 2nd ed. New York: Springer-Verlag.

Shelberg, M. C.; Moellering, H.; and Lam, N. S. 1982. Measuring the fractal dimensions of empirical cartographic curves. In *Proceedings, Fifth International Symposium on Computer-Assisted Cartography (Auto-Carto V)*, pp. 481–90. Washington, D.C.

Shelberg, M. C., and Moellering, H. 1983. IFAS: A program to measure fractal dimensions of curves and surfaces. *Proceedings 45th Annual Meeting ACSM*, pp. 483–92. Washington, D.C.

Shelberg, M. C.; Lam, N. S.; and Moellering, H. 1983. Measuring the fractal dimensions of surfaces. In *Proceedings, Sixth International Symposium on Automated Cartography (Auto-Carto VI)*, vol. 2, pp. 319–28. Ottawa.

Shreve, R. L. 1966. Statistical law of stream numbers. *Journal of Geology* 74:17–37.

———. 1967. Infinite topologically random channel networks. *Journal of Geology* 75:178–86.

Smith, A. R. 1982. The Genesis Demo: Instant evolution with computer graphics. *American Cinematographer* 63:1038–39, 1048–50.

Sprunt, B. 1972. Digital simulation of drainage basin development. In *Spatial analysis in geomorphology,* ed. R. J. Chorley, pp. 371–89. London: Methuen.

Steinhaus, H. 1960. *Mathematical snapshots*. London: Oxford University Press.

Tomlinson, R. F.; Calkins, H. W.; and Marble, D. F. 1976. *Computer handling of geographical data: An examination of selected geographic information systems*. Paris: UNESCO Press.

White, M. 1983. N-trees: large ordered indexes for multidimensional space. Paper presented at the Lincoln Institute of Land Policy's Colloquium on Spatial Mathematical Algorithms for Microcomputer Land Data Systems, May 9–10, Cambridge, Mass.

Woronow, A. 1981. Morphometric consistency with the Hausdorff-Besicovich dimension. *Mathematical Geology* 13:201–16.