CHAPTER 6

STATISTICS OF CHANNEL NETWORKS ON FRACTIONAL BROWNIAN SURFACES

Michael F.Goodchild
National Center for Geographic Information and Analysis, University of California, Santa Barbara

Brian Klinkenberg
Department of Geography, University of British Columbia, Vancouver, Canada

ABSTRACT

Many natural forms can be successfully simulated using fractals. We argue that fractals can therefore provide useful models of natural forms in the absence of specific effects and constraints: in effect, null hypotheses for natural forms. The fractional Brownian process is a useful simulation for terrain and offers a method of determining the characteristics of channel networks formed by random processes. We simulate random channel networks and tabulate the statistical distributions of various parameters. The chapter demonstrates that recently observed departures from the random topology model of Shreve are attributable to the geometric constraints imposed by packing networks onto a topographic surface, and not to geologic and geomorphic controls, as previously supposed.

INTRODUCTION

Hypothesis testing is a powerful strategy in science, particularly in areas where data are subject to substantial uncertainty. Suppose that observations suggest the presence of an effect E. If E is not the only effect present, then it is not capable of accounting fully for the observed variation in the data. The general strategy is to consider what would have hap-

pened in the absence of effect E. By comparing this hypothetical outcome with the real one, we may succeed in proving the presence or absence of the effect. We term the hypothetical process in the absence of effect E the null hypothesis or H_0. The strategy has its limitations: it does not evaluate the strength of effect E, merely its existence; if the effect is weak, it may be masked by other sources of variation (termed a Type II statistical error); and other sources of variation may cause us to falsely conclude the presence of the effect (Type I statistical error). In addition, it may be difficult to formulate an appropriate null hypothesis.

Suppose, for example, that we wish to investigate the effect of age on voting behavior. Two samples, one of Republican voters and the other of Democrats, would each show variation in age, but the ages in the two sets would not be the same, and the means would likely be somewhat different. In the absence of the effect of interest (that is, under the null hypothesis), both samples would be drawn from the same population. But if the effect is real, the population of Democrats would be expected to have a different distribution of ages from the population of Republicans.

Many common forms of hypothesis testing involve very simple null hypotheses and standard testing procedures. However, in some cases the null hypothesis is sufficiently complex or mathematically difficult that, while it is easy enough to state, the only way of investigating its outcomes is by simulation. In such cases the general form of the approach is as follows. An effect is suspected, and data are collected. A null hypothesis is stated, and simulation is used to generate a sample of its outcomes. These are compared to the data. If the data are found to be a possible outcome under the null hypothesis, then the effect is declared to be absent. On the other hand, if the data are inconsistent with the outcomes of the simulation, then the effect is declared to be confirmed.

In this chapter, we examine the use of this strategy with a fractal null hypothesis and simulation process to test an important issue concerning stream networks. The complex patterns formed by branching streams fall into the general class of graphs known as trees. Horton (1945) introduced an ordering and numbering scheme for stream networks that was later refined by Strahler (1952), as follows. Each tributary stream (the leafs of the tree network) is order 1; the joining of two order 1 streams forms an order 2; and in general the joining of two order i streams forms an order $i + 1$. Joining of an order i stream with an order j stream ($i \neq j$) produces a stream of order equal to the larger of i and j. The order of the stream forming the root of the tree becomes the order of the basin. There is now a large literature on the analysis of ordered channel networks, much of it prompted by the early work of Horton. Of particular concern here is what is now known as the *Horton law of stream numbers*: if the logarithm of the number of streams of given order is plotted against the order for a network or collection of networks, the points fall close to a straight line.

Although the effect represented by the Horton law is dramatic, we might nevertheless wish to confirm it by a hypothesis test. For bivariate plots, the conventional form of the null hypothesis proposes that the observed values of one variable have no effect on the corresponding values of the other. Although we would almost certainly reject the null hypothesis in this case, confirming the existence of the Horton law, this particular null hypothesis is clearly inappropriate, as it is impossible for the number of streams to be unrelated to order, given the system of ordering: the number of streams of order i can be

no more than one-half the number of order i - 1. We can therefore reject this particular null hypothesis out of hand.

The objective of this chapter is to propose a more appropriate null hypothesis for channel networks and to test its outcomes using fractal simulation. The model is a significant departure from one proposed by Shreve (1966, 1967), described below, which has been the focus of much research in channel networks over the past two decades. Our model accounts for several consistent and puzzling deviations from the Shreve model observed in stream networks.

More generally, the chapter illustrates the value of fractal models as null hypotheses for a whole range of natural systems (Goodchild, 1988). Mandelbrot (1982a) has argued very convincingly that fractals are a useful model of natural systems. But many geomorphologists find the fractal simulations of terrain in Mandelbrot's book too artificial (Mark and Aronson, 1984). Self-affine landscape simulations lack the scale-specific effects of geology and landscape-modifying processes such as wind, rain, and glaciation that provide visual cues as to scale. However, Goodchild and Mark (1987) argue that this lack can be an advantage, in the sense that fractal simulations provide a test-bed for geomorphic processes—an initial or boundary condition landscape on which to simulate the effects of real processes. In general, fractal simulations are a powerful way of visualizing and analyzing what would be expected in the absence of real effects or constraints on natural form.

THE SHREVE MODEL

In the Shreve model of channel network topology (Shreve 1966, 1967), all networks are assumed to be part of an infinite topologically random channel network (TRCN) in which all possible alternative channel network topologies occur with equal probability. There are, for example, five topologically distinct ways in which four first-order streams can combine, four yielding second-order basins and one a third-order basin (Figure 6.1). According to the Shreve model, the five arrangements should be found with equal likelihood in real stream basins. The model has been remarkably successful at explaining the Horton law of stream number, as it predicts not only the observation of straight-line plots, but also the slopes (the ratio of the number of streams of order i - 1 to the number of order i is known as the *bifurcation ratio*, and is constant for a straight-line plot). It has also been extended with success to certain geometrical properties of channel networks, such as link lengths and basin areas (Shreve, 1975).

In the hypothesis testing tradition of statistics, the Shreve model is a null hypothesis, reflecting what would be expected to occur under a simple random process. Thus, if empirical data are found to be consistent with the model, the normal inference would be that the basin or basins from which the data were obtained evolved in a way that was largely free of geological or other constraints on the random combination of streams. On the other hand, deviations from the model's predictions would be interpreted as indicative of the operation of such effects.

Unfortunately, standard hypothesis-testing procedures provide only a weak test of the model, since confirmation of a null hypothesis amounts to no more than a failure to

reject. The likelihood of a Type II statistical error, or acceptance of the null hypothesis when in fact it is false, depends both on sample size and on the strength of the effect. In other words, large amounts of data are needed to demonstrate small deviations from the model: the smaller the amount of data available, the more likely is the confirmation of the null hypothesis. Understandably, then, early small-scale empirical tests largely confirmed the predictions of the infinite TRCN model.

A recent paper by Abrahams (1984) undertook a comprehensive review of the field and identified a growing consensus on systematic empirical departures from the model. For large basins (magnitude $M > 50$, where magnitude is defined as the number of first-order streams), there is a tendency for bifurcation ratios (and thus slopes of Horton law plots) to exceed predictions, particularly in areas of high relief. First-order streams can be classified as *S* (source) or *TS* (tributary source) depending on whether they join to other streams of order 1 or to streams of order greater than 1, respectively (in Figure 6.1, the four second-order basins have two *S* and two *TS* links, while the third-order basin has four *S* links and no *TS* links). In theory, the proportion of *TS* links should tend to 50% in very large basins, but in real networks the proportion appears to be higher than expected (networks with a high proportion of *TS* links are said to have *fishbone* topologies). Abrahams (1977) found a strong correlation between the amount of deviation from the model and the relative relief of the topography.

Abrahams (1984, p.164) finds a pattern emerging in empirical results, that "small channel networks fit the model better than large ones." Under the hypothesis-testing framework in which a failure to reject the null hypothesis is often equated with confirmation of the stochastic model, we are tempted to look for explanations of this pattern in erosion processes. But a simpler explanation, suggested by the comments above, is that the null hypothesis is not precisely correct and has been accepted for small data sets because of the greater probability of Type II errors.

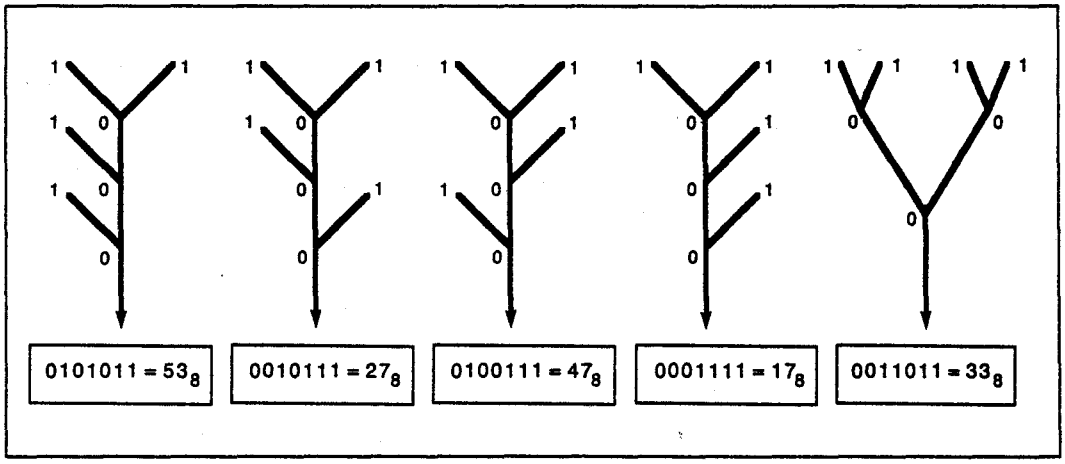

Figure 6.1 The five alternative (topologically distinct) channel networks generated by four first-order streams. The notation denotes the coding schemes for networks and ambilateral classes.

Although much early research supported the random topology model, it is quite possible, given the remarks above, that a reexamination of earlier data would reveal much more widespread disagreement. For example, Smart (1969, pp.1766–1767) analyzed a total of 1157 basins of magnitude 4 (four first-order streams, five possible topologies, see Figure 6.1) and 629 of magnitude 6 (six first-order streams, 42 possible topologies) in a study of the topologies of U.S. networks, separated into eastern and western samples. Although the Shreve random topology model was accepted as a null hypothesis at the 0.05 level in three out of the four analyses, in all cases the data show a bias consistent with those discussed by Abrahams and analyzed later in this paper, which if real would presumably have led to rejection of the null hypothesis given a sufficiently large sample. Smart regarded the results of his analysis as inconsistent, because the null hypothesis was accepted in only three cases, and looked for an explanation in terms of structural controls. Given the work summarized by Abrahams, the alternative explanation of a real bias and three Type II errors seems much more acceptable.

The more straightforward and conventional implication of these results is that there are processes operating on the landscape, in the form perhaps of geological controls, that prevent or bias the random joining of streams, assigning unequal probabilities to certain alternative topologies and resulting in deviations from the model. Clearly, this is an appropriate inference if the model is itself an appropriate null hypothesis and describes correctly the result of random, uninhibited channel network development. However, this is not necessarily the case. Streams developing on a real terrain are subject to constraints of a purely geometrical origin, which will undoubtedly affect the relative likelihoods of alternative topologies. First, basins must pack onto the surface, leading to dependencies between the geometries of adjacent basins, which will in turn affect their topologies. A second class of geometrical constraints is introduced by the requirement of continuity in the surface, which forces dependencies between flow directions of neighboring streams. In principle, then, since all topologies cannot be equally likely if networks are forced to pack onto a surface, the Shreve null hypothesis cannot be true and we can reject it out of hand.

We term a null hypothesis appropriate if it differs from reality only in the absence of the effect of interest, in this case geologic or other control on channel network development, so the implications of rejection are clear and unambiguous. An inappropriate null hypothesis is one that differs from reality in several additional ways, so that rejection may be due either to the effect of interest or to one or more of the additional, irrelevant effects. In this sense, the Shreve model is an inappropriate null hypothesis if indeed the development of an infinite TRCN on real terrain would be inhibited by constraints of a purely geometrical nature. A null hypothesis that failed to include such constraints would be inappropriate, as it would not allow us to separate out the effects of geometrical constraints, which are uninteresting from a geological or geomorphological perspective.

The effect of geometrical constraints on the random topology model is unknown, and it is unlikely that a more appropriate null hypothesis would be amenable to mathematical analysis. The approach used in this paper has been to simulate the random development of stream networks on surfaces under a broad range of conditions. If agreement is found with the infinite TRCN model, then the implication will be that real deviations are due to additional, geological or other forms of control over network development. On the

other hand, deviations observed in the simulations will tend to confirm the inappropriateness of the Shreve model as a channel network null hypothesis. Finally, if the form of the simulated deviation agrees with observed deviations in real networks, the implication will be that the origin of such deviations lies in geometrical constraints rather than in physical processes. Earlier results from this study appeared in Goodchild et al. (1985).

FRACTIONAL BROWNIAN MOTION

There have been several previous simulations of random channel development on surfaces. Leopold and Langbein (1962), Schenck (1963), and others allowed streams to wander randomly over a square lattice, generating a move at each lattice cell from a random integer corresponding to one of four move directions (we refer to four move directions as the rook's case, to distinguish it from the queen's case of eight move directions, including four diagonal moves). Special means had to be devised to avoid topological inconsistencies such as closed loops and spirals. The packing constraint is imposed in these models in the form of a uniform density of one stream move per cell: the elimination of loops and spirals has the effect of a weak surface continuity constraint. Seginer (1969) imposed a stronger continuity constraint by first generating a random surface and then extracting the stream network from it, thus ensuring topological consistency. The initial random surface consisted of a square grid of independent elevations, characterized by frequent pits and no general trend. A study by Craig (1980) of simulated erosion processes on surfaces extracted stream networks and compared them to the random topology model as a means of validation of the simulation process, a dubious procedure given the known departures of real networks from the model.

A much more appropriate means of generating suitable random surfaces is available in a class of stochastic processes known as fractional Brownian motion (fBm). Their use for terrain simulation was first explored by Mandelbrot (1977, 1982a) and has attracted considerable attention, due at least partly to striking computer-generated illustrations. More significantly, fBm surfaces have several attractive features that make them suitable for the purposes of this study. The variogram of an fBm surface has the form

$$E[z(x) - z(x+d)]^2 = k|d|^{2H} \tag{6.1}$$

where x denotes a location, d is a displacement from this location, z is the elevation of the surface at location x, E is the statistical expectation, $|d|$ is the length of the displacement d, k is a constant, and H is a parameter in the range 0 to 1. $H = 0$ corresponds to independent elevations and thus an infinitely rugged surface with no trend, while as H tends to 1 surfaces become locally smoother with stronger general trend. $H = 0.5$ corresponds to Brownian motion, in which the difference in elevation between adjacent points has a probability distribution that is symmetrical about zero.

Because of the power law form of the variogram, fBm surfaces are scale free and termed self-affine: a small piece of the surface if sufficiently enlarged is statistically indistinguishable from the surface as a whole. Since most if not all geomorphological processes are scale dependent, having different amounts of influence on the landscape at different

scales, fBm surfaces appear to lack evidence of erosion. So, by extracting drainage networks from fBm surfaces of varying H, we can simulate network development under a broad range of surface conditions and subject to packing and continuity constraints. Kirkby (1985) used a fractal surface as the random, erosion-free starting point for simulations of slope and stream evolution.

There have been several comparisons between the fBm model and real terrain. Goodchild (1982) examined the topography of Random Island, Newfoundland, and found deviations from self-similarity that could be interpreted in terms of the varying influence of geomorphological processes. Mark and Aronson (1984) reached similar conclusions in their analyses of areas in the eastern United States. fBm surfaces have equal abundances of peaks and pits; rugged surfaces of low H, with high peak and pit frequencies, are consequently more suggestive of karst or dead ice topographies than real, eroded terrain.

The surfaces used in the study (Figure 6.2) consisted of square arrays of 256 by 256 elevations, generated by the method described by Mandelbrot (1977) in which an initially flat surface is faulted along randomly located, straight lines. For H=0.5, the surface on either side of each fault is displaced upward or downward by a constant amount; for other values of H a nonuniform displacement is used. This method was adopted in preference to the more commonly used algorithm of Fournier et al. (1982) because of Mandelbrot's theoretical objections to the latter (Mandelbrot, 1982b). A total of 1000 faults was applied to each surface to ensure a high probability that at least one fault occurred between every pair of adjacent elevations. No scaling was applied to the elevations, since both the self-similarity properties and the process of network extraction are invariant under linear transformation. Surfaces were generated with values of H ranging from 0.3 to 0.7 in steps of 0.1. Three replications of the 0.7 surface were generated because this value of H was thought by Mandelbrot to give the closest resemblance to certain real terrains. Figure 6.2 shows examples of fBm surfaces for values of H ranging from 0.3 to 0.9.

DRAINAGE NETWORK EXTRACTION

fBm surfaces have no eroded stream channels, so the derivation of channel networks must be based on local terrain slopes. O'Callaghan and Mark (1984) have described the extraction of networks under similar conditions from real digital elevation models, and see also Jensen (1985). In their algorithm, each cell may flow to one of eight queen's case neighbors; in this study moves were restricted to the four rook's case neighbors because of the problems that would be introduced by allowing up to seven streams to join at once.

Each cell was assigned an integer between 0 and 4 according to the following rules. If any of the four neighboring cells were strictly lower, flow was to the lowest neighbor, designated by an integer between 1 and 4. Otherwise, the cell was designated a sink and assigned 0. This is termed the deterministic or D assignment.

Several objections can be raised to this simple approach. First, it assigns precisely one channel flow direction to each cell, whereas real channels are observed to develop only after the accumulation of a sufficient amount of nonchannel flow. We allow for this by the truncated or T assignment, which modifies the networks obtained under assignment

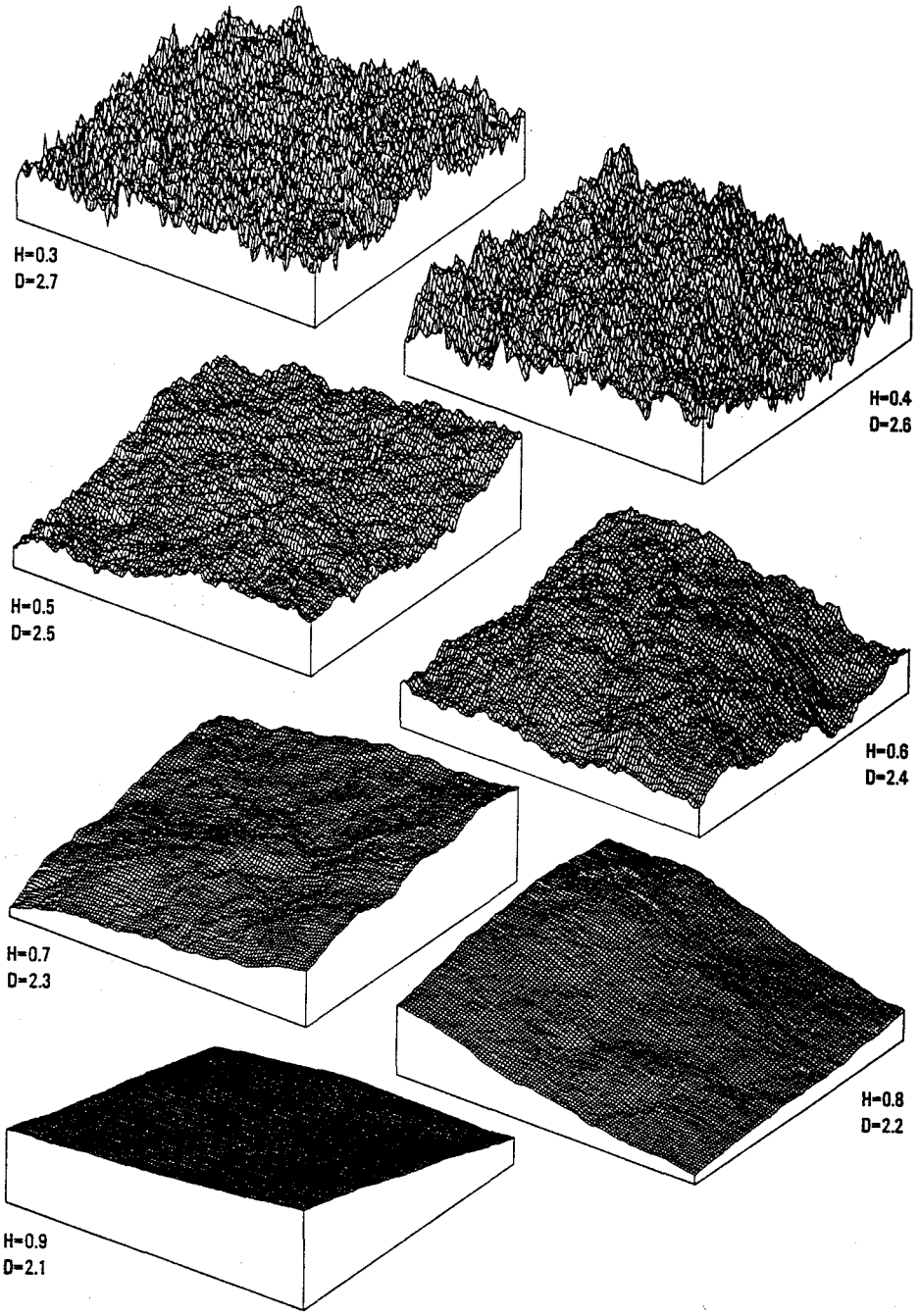

Figure 6.2 Self-affine surfaces generated by the fractional Brownian process with H ranging from 0.3 to 0.9 (D=3-H).

D above by deleting downstream from each first-order stream head until a certain minimum flow level is reached, measured by the number of upstream cells (see also Mark, 1983; Hugus and Mark, 1984; O'Callaghan and Mark, 1984). Thus, the T assignment assumes a certain threshold-contributing upstream area before a channel can develop. This process can result in the deletion of junctions and redesignation of stream orders.

Second, the use of a limited set of move directions introduces a potential bias in flow direction. Consider an area of uniform slope. In reality all streams would flow in the direction of maximum gradient without joining. Under the D assignment, all streams would flow without joining in the permitted move direction closest to the direction of maximum gradient, introducing a consistent bias of up to 45°. The random or R assignment allows for this through the following rules:

> If no neighbor is strictly lower, the cell is a pit.
>
> Else if exactly one is lower, flow is in that direction,
>
> or if exactly two are lower and in opposite directions, flow is to the lowest,
>
> or if exactly two are lower and in neighboring directions, or more than two are lower, assign direction based on probabilities proportional to the differences in elevation between the cell and the lower neighbors.

The R assignment also deals to some extent with a third problem. Because of the self-similarity property, fBm surfaces vary within-cells as well as between them, although the use of a discrete array has the effect of deleting within-cell variation. The consequent straightening of derived channels within cells reduces the probability of joining. The R rules compensate for this by allowing streams to join even in areas of uniform slope.

The problem of the edge was dealt with in all cases by adding a border of cells with very high elevations. This ensured that flow either sank at the edge of the 256 by 256 array or flowed inward from it, but never outward.

Channel networks were obtained from the integer arrays in the form of binary strings (Shreve, 1967; Scheidegger, 1967). In our implementation, the binary string representation of each basin is computed by traversing the tree beginning and ending at the root, or outlet stream, in a clockwise or left to right direction, turning left at every intersection and reversing direction at every leaf or source stream. A 0 is encoded every time a junction is encountered from downstream and a 1 every time a leaf or source stream is reached. This binary notation provides a unique coding for every possible network topology, in which the number of 1s is one more than the number of 0s and that always terminates in a 1. Furthermore, at any point before the end of the string the number of 1s is less than or equal to the number of 0s. Figure 6.1 demonstrates the coding scheme with the five topologies of magnitude $M = 4$.

The literature on channel networks has consistently assumed all junctions to be three valent, consisting of one outlet and two inlet streams, whereas the method of network extraction described above will frequently produce four-valent junctions where three inflowing streams merge to form one outflowing stream. The extraction algorithm allowed for this by breaking all four-valent junctions into two three-valent junctions plus a dummy link. To avoid favoring certain topologies, breaks were made randomly, the two alternatives being given equal probabilities. This is illustrated in Figure 6.3. Suppose the eleva-

tions predict that flow is into the central cell from the left, from above and from the right, and that outflow is downward, thus creating a four-valent junction (this could be the result if the cells to the left, above, and to the right had higher elevations, and the cell below had a lower elevation). The letters *a*, *b*, and *c* in Figure 6.3 denote the binary strings describing the three subbasins upstream of the junction. Figure 6.3 shows the two equally likely networks created by randomly breaking the junction of four streams into two junctions of three streams each. The results reported earlier (Goodchild et al., 1985) did not randomize the breaking of four-valent junctions and were limited to the *D* method of flow assignment.

Channel-length information was encoded in the form of a second string of the same length, giving the length of the link downstream of each junction for each 0 in the binary string and the length of the source link for each 1.

All analyses of the basins were carried out by manipulations of these pairs of strings. Stream order was defined using the Strahler (1952) system and obtained with the following algorithm (see also Liao and Scheidegger, 1968). Replace all instances of 0*xx* in the string by $x+1$ (x not equal to 0), and increment a counter for the number of streams of order $x+1$ by 1. Replace all instances of 0*xy*, x not equal to y and neither x nor y equal to 0, by the greater of x and y. This process must terminate with a single digit equal to the order of the basin.

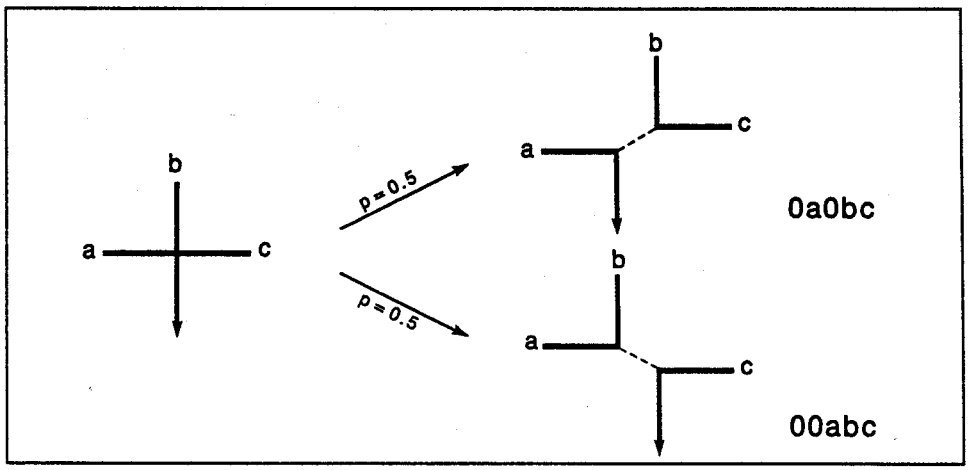

Figure 6.3 A four-valent junction and the two equiprobable pairs of three-valent junctions derived from it.

ANALYSIS: *D* ASSIGNMENT

Under the extraction rules described above, a basin is defined as a tree network rooted or terminating in a sink. This has no direct equivalent in the infinite TRCN model, since all links are part of one infinitely large, singly rooted tree. A randomly chosen basin in the infinite TRCN model is best defined as the subbasin rooted in and upstream of a randomly chosen link. However, we might assume that the basins in the simulation have arisen by a process in which any link in the model is equally likely to become a sink. This would allow us to compare the abundances of basins by order in the simulation with the known probabilities that a randomly chosen basin or subbasin in the model is of order *w*. The relative abundances of basins by order under the *D* assignment rules are shown in Table 6.1.

TABLE 6.1 Abundances of Basins by Order

H	1	2	3	4	5	Total
0.3	14,363	4,949	642	1	0	19,955
0.4	10,646	4,502	806	5	0	15,959
0.5	7,524	3,589	827	16	0	11,956
0.6	3,900	2,130	751	40	1	6,822
0.7a	621	380	178	50	9	1,238
0.7b	1,543	808	383	56	4	2,794
0.7c	1,801	1,009	458	68	0	3,336

The infinite TRCN model predicts that the number of basins of order i be half the number of order i -1. This is clearly more nearly true of the smoother surfaces of high H than of the more rugged, low H surfaces with much more frequent pits. Although basins on the smoother surfaces reach higher orders, there is clearly severe truncation, which may be due to limitations on the abundances of high-magnitude basins or to the relative infrequency of basins of high order for a given magnitude. The latter interpretation is more likely for the high H surfaces where pits are much less common.

The analysis of relative abundances of streams and links on the simulated surfaces is similarly affected by the presence of pits, since the infinite TRCN model is again concerned with a randomly chosen feature in an infinite network. The simulations show similar overabundances of low-order streams and links, particularly on the more rugged surfaces, and truncation at higher orders.

In the infinite TRCN model, all possible topologies of basins of a given magnitude are equally likely. Table 6.2 shows the relative abundances of the five possible topologies of magnitude 4 basins (basins with four first-order streams; see Figure 6.1). Each basin topology is identified by the octal representation of its binary string; thus, the string 0001111 becomes basin 17. Note that in this and subsequent analyses in this paper, a basin

is defined as terminating in a sink; subbasins of a given magnitude that might exist embedded in basins of larger magnitude are not included in the sample.

TABLE 6.2 Abundances of Topologies for Magnitude 4 Basins

H	17	27	33	47	53	Total
0.3	110	150	178	126	163	727
0.4	128	157	170	158	137	750
0.5	118	131	115	119	113	596
0.6	66	87	63	77	64	357
0.7a	10	17	7	7	14	55
0.7b	37	30	20	26	29	142
0.7c	31	33	26	30	43	163

The clearest trend in the table is that shown by basin 33, or 0011011, which is overabundant for the rugged $H = 0.3$ and $H = 0.4$ surfaces and underabundant for the remainder, although only the $H = 0.3$ results are significantly different from the model predictions at the 0.05 level on a chi-squared test. Basin 33 is unique in the magnitude 4 basins as the only one of third-order. As noted earlier, the second-order basins, 17, 27, 47, and 53, have been loosely referred to as fishbone topologies in the literature, as a high proportion of first-order streams, in these cases 50%, join streams of a higher order (*TS* or tributary source links). In the case of basin 33, all first-order streams join other first-order streams (*S* or source links). We will informally refer to the degree to which a basin approximates a fishbone form as its *linearity* and use the proportion of *TS* links as a more objective measure.

Two topologies are said to be of the same ambilateral class (Smart, 1969) if one can be obtained from the other by exchanging the two subbasins incident at one or more junctions, in effect removing the distinction between left and right. In Figure 6.1, basins 17, 27, 47, and 53 are all of the same ambilateral class, but cannot be transformed into basin 33 by mirroring at junctions. Smart first suggested aggregation to ambilateral classes as a means of reducing the number of alternative topologies in high-magnitude basins, while retaining the significant hydrological differences between classes (Smart, 1969, p.1761). In addition, all members of one ambilateral class clearly share the same proportions of *S* and *TS* links, and therefore the same degree of linearity, or similarity to a fishbone form.

To identify a basin's ambilateral class, each 0, or junction, in the basin's binary string is examined. The two subbasins incident at the junction are described by two consecutive blocks of binary code immediately following the 0 and can be identified from the rule that any subbasin must have exactly one more 1 than it has 0s. The binary representations of the two subbasins are compared numerically and, if necessary, switched so that the first has a numerical value less than or equal to the second. When all junctions have been

examined in this way, the binary string will have been reduced to that of a common ambilateral class. For example, consider the basin 00111 (basin 7, magnitude 3). The first junction has incident subbasins 011 and 1. Since the first is greater in value than the second, they are reversed, giving the string 01011. The second 0 has symmetrical incident subbasins, 1 and 1, so no reversal is necessary, and the ambilateral class is identified as 13, the octal representation of 01011.

The magnitude 5 basins, for which there are 14 possible topologies and 3 ambilateral classes, are shown tabulated by ambilateral class in Table 6.3. The most linear topology, in this case class 253, which has 60% *TS* links, is overabundant for all surfaces and significantly so in all but one.

TABLE 6.3 Abundances of Basins by Ambilateral Class, M = 5

H	233	153	253	Chi squared
0.3	26	129[a]	252[a]	20.8[b]
0.4	33	124[a]	221[a]	10.7[b]
0.5	28	93	246[a]	19.0[b]
0.6	19	64	194[a]	21.6[b]
0.7a	2	7	32[a]	7.6
0.7b	3	21	63[a]	11.3[b]
0.7c	5	30	83[a]	12.3[b]
Model probability	2/14	4/14	8/14	
Proportion *TS* links	20%	20%	60%	

a. Denotes observed abundance exceeds model predictions.
b. Denotes significant at the 0.05 level

The pattern for magnitude 6 basins, shown in Table 6.4, is similar, as is the analysis of magnitude 7 basins. In both cases there is increasing bias toward the most linear topologies, 1253 and 5253, respectively, with increasing *H*. In Table 6.4, it is evident that 553 is also overabundant on all surfaces but one. Since 553 has the same proportion of *TS* links as three other ambilateral classes, it is clear that the proportion of *TS* links is not a completely satisfactory predictor. Basin 553 is symmetrical, with two linear subbasins (01011) meeting at the top of the root link.

In general, basin order and proportion of *TS* links are inversely related; as Table 6.4 shows, the most linear basins are also those of lowest order. The overabundance of linear basins thus has the effect of truncating the distribution of basin orders so that fewer high-order basins are found than expected under the infinite TRCN model.

TABLE 6.4 Abundances of Basins by Ambilateral Class, $M = 6$

H	633	1233	553	1153	653	1253	Chi squared
0.3	4	15	18[a]	26	36[a]	82[a]	7.4
0.4	6	20	33[a]	37	54[a]	102[a]	10.7
0.5	2	14	25[a]	28	36	123[a]	32.3[b]
0.6	0	7	18[a]	26	20	105[a]	42.8[b]
0.7a	1	2	3	4	2	20[a]	c
0.7b	0	6	9[a]	11	8	53[a]	23.0[b]
0.7c	2	5	8[a]	7	12	49[a]	17.4[b]
Model probability	2/42	4/42	4/42	8/42	8/42	16/42	
Proportion *TS* links	0/6	2/6	2/6	2/6	2/6	4/6	
Basin order	3	3	3	3	3	2	

a. Denotes overabundant.
b. Denotes significant at the 0.05 level.
c. Denotes chi squared not computable due to small numbers.

Table 6.5 shows the proportions of *TS* links observed in basins of magnitude 4 through 20, compared to the proportions expected under the model, for three of the surfaces. It is easy to identify *TS* links in a basin's binary string since the two 1s in any sequence 011 are *S* links and all other 1s are *TS*. The expected proportions in the model are given by $(M\text{-}2)/(2M\text{-}3)$, which is asymptotic to 0.5 as M tends to infinity (Mock, 1971). In the limit, the probability that a randomly chosen bit is a 1 in the infinitely long basin string is 0.5, and each bit is independent of its neighbors. There are eight equiprobable sequences of the two bits preceding and the one bit following the 1, four of which dictate that the 1 denote an *S* link and four a *TS* link:

0010, *TS* 0011, *S* 0110, *S* 0111, *S*
1010, *TS* 1011, *S* 1110, *TS* 1111, *TS*

Table 6.5 shows that the observed proportion of *TS* links is generally higher than expected, and increasingly so with increasing basin magnitude and increasing H.

Finally, we tabulate the bifurcation ratios for these three surfaces and for basin magnitudes 4 through 20 in Table 6.6. The ratios are calculated in two ways. $B0$ is the $(i - 1)$th root of the number of first-order streams, where i denotes the basin order, while $B1$ is the ratio of first-order to second-order streams. For given magnitude, the model predicts bifurcation ratios close to 4.0. Instead, observed ratios tend to be higher than 4.0, particularly on surfaces of high H, $B1$ is consistently higher than $B0$ and both tend to increase with increasing magnitude.

TABLE 6.5 Observed and Expected Proportions of *TS* Links

		$H = 0.3$		$H = 0.5$		$H = 0.7a$	
M	Expected	n	Observed	n	Observed	n	Observed
4	0.400	727	37.8	596	40.4	55	43.6
5	0.429	407	44.8	367	46.8	41	51.2
6	0.444	181	47.7	228	51.0	32	53.1
7	0.455	120	51.4	194	51.7	32	61.6
8	0.462	69	46.7	125	53.6	24	57.3
9	0.467	44	54.0	94	54.4	13	62.4
10	0.471	18	55.6	56	55.0	13	66.2
11	0.474	13	58.0	46	56.9	12	63.6
12	0.476	3	55.6	28	53.0	7	73.8
13	0.478	4	50.0	29	53.8	14	62.6
14	0.480	2	35.7	14	56.1	7	75.5
15	0.482	0		17	53.7	7	63.8
16	0.483	1	62.5	11	54.5	11	67.0
17	0.484	0		4	58.8	11	65.8
18	0.485	2	55.6	3	51.9	4	69.4
19	0.486	0		5	60.0	8	68.4
20	0.487	0		2	60.0	4	72.5

ANALYSIS: *T* AND *R* ASSIGNMENTS

The above analyses were repeated using the T assignment, truncating each stream downstream from the head until the flow from a prescribed number of cells F had been accumulated. This minimum number of cells was set to 1, 2, and then 5, and results were compared in each case to those obtained under the D assignment.

Truncation tends to delete large numbers of small first-order basins and to reduce the order of larger basins. For $F = 1$, the model was rejected as a predictor of the relative abundances of the five $M = 4$ topologies (compare Table 6.2) for all surfaces except $H = 0.7c$, in all cases due to the underabundance of basin 33 and the overabundance of the linear topologies. The most linear ambilateral class was overabundant in every case where chi squared

could be calculated for every surface and for each of $M = 4$, 5, 6, and 7. Observed proportions of *TS* links tended to be higher for given magnitudes than those shown in Table 6.5, and bifurcation ratios higher than those in Table 6.6. Although sample sizes were substantially lower, the same trends were observed for increasing truncation at $F = 2$ and $F = 5$. In summary, truncation produces increasing divergence from the model.

TABLE 6.6 Bifurcation Ratios by Basin Magnitude

	$H = 0.3$			$H = 0.5$			$H = 0.7a$		
M	*n*	*B0*	*B1*	*n*	*B0*	*B1*	*n*	*B0*	*B1*
4	727	3.51	3.51	596	3.61	3.61	55	3.75	3.75
5	407	3.95	4.05	367	4.09	4.18	41	4.39	4.45
6	181	4.06	4.34	228	4.36	4.61	32	4.67	4.84
7	120	4.17	4.67	194	4.24	4.71	32	5.50	5.80
8	69	3.43	4.14	125	4.15	4.87	24	4.98	5.50
9	44	4.36	5.13	94	3.89	4.87	13	5.31	6.12
10	18	4.68	5.42	56	4.51	5.27	13	6.32	7.05
11	13	3.91	5.29	46	3.96	5.20	12	4.60	6.11
12	3	3.46	4.67	28	3.42	4.57	7	8.34	9.14
13	4	3.29	4.44	29	4.17	5.17	14	4.28	5.88
14	2	3.74	3.15	14	3.55	4.83	7	6.67	9.00
15	0			17	3.54	4.69	7	3.67	5.89
16	1	4.00	5.33	11	3.73	4.70	11	4.00	6.59
17	0			4	4.12	5.31	11	4.12	6.31
18	2	4.24	4.80	3	4.24	4.50	4	4.24	7.13
19	0			5	4.36	5.19	8	4.36	6.53
20	0			2	4.47	5.83	4	4.47	7.92

By increasing the probability of streams joining, the *R* assignment tends to increase basin order. Larger numbers of order 5 basins were obtained compared to Table 6.1, and the $H = 0.7$b and $H = 0.7$c surfaces produced order 6 basins. For the magnitude 4 ambilateral classes, surfaces $H = 0.6$, $H = 0.7$a, and $H = 0.7$b showed overabundance of the linear class, but the remaining four surfaces showed underabundance. For magnitude 5, the most

linear class, 253, was overabundant for all surfaces except H =0.3, and similarly for magnitude 6, the linear classes were overabundant for all except H =0.3 and H =0.4. Proportions of *TS* links and bifurcations ratios tended to be higher than expected, although again not to the same degree as with the D and T assignments, and the differences again tended to increase with increasing magnitude. As we might expect, then, by increasing the probability of joining, randomization tends to reduce the linearity of the extracted basins, but not sufficiently to produce consistency with the infinite TRCN model.

DISCUSSION

The simulated networks derived from fBm surfaces under a range of assignment rules show significant deviations from the infinite TRCN model. These deviations agree in many ways with those identified by previous empirical research and summarized by Abrahams (1984). Furthermore, there is evidence that a reexamination of data previously interpreted as supporting the random model would also show deviations of the same type, although not sufficiently strong or based on sufficiently large sample sizes to cause rejection of the model.

More specifically, the analysis has identified the following grounds for rejection of the random model:

1. Ambilateral classes for magnitude 4 through 7 basins do not occur with predicted frequencies: instead most surfaces show a bias toward classes with linear or fishbone form, low order, and high proportions of *TS* links. The bias is strongest with smoother surfaces of high H.
2. The proportion of *TS* links is higher than expected for basins of magnitude 4 through 20, and the deviation increases with magnitude and with the parameter H.
3. There is a tendency for both evaluations of the bifurcation ratio to exceed model predictions. *B1* tends to be higher than *B0*, and the bias increases with basin magnitude and with H.

There are other forms of observed deviation from the model, such as asymmetry and the distributions of *cis* and *trans* links, whose interpretation is quite different and that are not therefore relevant to this paper.

In his 1966 paper, Shreve proposed the hypothesis that "in the absence of geologic controls a natural population of channel networks will be topologically random" (Shreve, 1966, p.27). Subsequent research has tended to follow the same paradigm, interpreting deviation from the model as due to the influence of geologic factors and acceptance of the null hypothesis as confirming their absence. Abrahams (1984, p.162) identifies the first of the underlying postulates of the random topology model as "In the absence of environmental controls, channel networks are topologically random."

We have argued in this paper that the packing of channel networks onto a continuous topographical surface imposes additional constraints of a purely geometrical nature, raising the possibility of a second basis for rejection of the model. Given this alternative inter-

pretation, the model becomes an inappropriate null hypothesis for testing for the presence or absence of geologic controls. Many results that have previously been taken as confirming the model can be reinterpreted as failing to reject it simply because of inadequate sample size.

It would be very difficult to devise a more appropriate null hypothesis because of the wide range of possible packing and continuity conditions. The scale-free fBm surfaces used in this study provide continuity conditions ranging from extreme ruggedness to local smoothness and thus possess a reasonable degree of generality. Quite different constraints might be imposed by surfaces with strong scale-dependent or periodic variation, such as glaciated landscapes, or surfaces with regional variation in smoothness. It is evident from the results obtained above that much stronger biases are imposed by packing basins onto relatively smooth surfaces of high H than onto rugged ones.

Another source of difficulty lies in the presence of pits on the fBm surfaces. At low H, with high pit frequencies, we suspect that removing downstream drains from the surface, as in a karst landscape, has the effect of reducing the severity of packing constraints, which may account for the smaller divergences observed from the random model on these surfaces. The assignment rules used in these experiments to simulate channel development may also be unrealistic. Real channels modify their own landscapes, perhaps in ways that affect the operation of packing and continuity constraints. In summary, although they are useful for illustrative purposes, fBm surfaces are not sufficiently general that we would argue for their use in a more appropriate null hypothesis.

We would agree therefore with the conclusion reached by Abrahams (1984, p.185) that, despite commonly observed deviations, "the random model will continue to serve as a standard against which to compare natural channel networks." However, it is not necessarily true that deviations from the model indicate the operation of geologic constraints. Furthermore, it is demonstrable by simulation that channel networks simulated on a wide range of random surfaces are constrained so as to deviate significantly from the random model, irrespective of geologic or environmental constraints. It is therefore difficult to see how acceptance of the model or, more correctly, failure to reject can be interpreted as anything other than a Type II statistical error; equally, it would be difficult to place an unequivocal interpretation on a rejection.

Finally, we hope this chapter has demonstrated the value of fractal simulations, not only in providing interesting visual effects, but also in providing standards and models for investigating a wide range of physical effects. There are numerous instances in science where it is necessary to compare reality to some hypothetical version in which a given effect is absent, but often we have no clear notion of what that hypothetical version might be. Fractal simulations provide precisely that, particularly, as we have demonstrated in this chapter, in the case of topography. fBm is useful both as a starting point for simulating the effects of landform processes and also as a null hypothesis in investigating the effects produced by those processes.

REFERENCES

Abrahams, A. D. 1977. The factor of relief in the evolution of channel networks in mature drainage basins. *American Journal of Science* 277:626–645.

———, 1984. Channel networks: a geomorphological perspective. *Water Resources Research* 20:161–188.

Craig, R. G. 1980. A computer program for the simulation of landform erosion. *Computers and Geosciences* 6:111–142.

Fournier, A., Fussell, D., and Carpenter, L. 1982. Computer rendering of stochastic models. *Communications of the Association for Computing Machinery* 25:371–384.

Goodchild, M. F. 1982. The fractional Brownian process as a terrain simulation model. *Modeling and Simulation* 13:1133–1137. Proceedings of the Thirteenth Annual Pittsburgh Conference on Modeling and Simulation.

———, 1988. Lakes on fractal surfaces: a null hypothesis for lake-rich landscapes. *Mathematical Geology* 20:615–630.

———, and Mark, D. M. 1987. The fractal nature of geographic phenomena. *Annals of the Association of American Geographers* 77:265–278.

———, et al. 1985. Statistics of hydrologic networks on fractional Brownian surfaces. *Modeling and Simulation* 16:317–323. Proceedings of the Sixteenth Pittsburgh Conference on Modeling and Simulation.

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 of the Fall ASP–ACSM Convention*, San Antonio, TX, pp.683–693.

Jensen, S. K. 1985. Automated derivation of hydrologic basin characteristics from digital elevation model data.In *Proceedings, Seventh International Symposium on Automated Cartography (Auto-Carto 7)*, Washington, DC, pp.301–310.

Kirkby, M. J. 1985. A two-dimensional simulation model for slope and stream evolution. Paper presented at the Binghamton Conference, State University of New York at Buffalo.

Leopold, L. B., and Langbein, W. B.1962. The concept of entropy in landscape evolution. *U.S. Geological Survey Professional Paper* 500A. Reston, VA: U. S. Geological Survey.

Liao, K. H., and Scheidegger, A. E. 1968. A computer model for some branching-type phenomena. *Bulletin of the International Association for Scientific Hydrology*13:5–13.

Mandelbrot, B. B. 1977. *Fractals: Form, Chance and Dimension.* San Francisco: W.H. Freeman & Co.

———, 1982a. *The Fractal Geometry of Nature*. San Francisco: W.H. Freeman & Co.

———, 1982b. Comment on computer rendering of stochastic models. *Communications of the Association for Computing Machinery* 25:581–583.

Mark, D. M. 1983. Automated detection of drainage networks from digital elevation models. In *Proceedings of the Sixth International Symposium on Automated Cartography (Auto-Carto 6)*, Vol.2, Ottawa, Canada, pp.288–298.

———,and Aronson, P. B. 1984. Scale-dependent fractal dimensions of topographic surfaces: An empirical investigation, with applications in geomorphology and computer mapping. *Mathematical Geology*16:671–684.

Mock, S. J. 1971. A classification of channel links in stream networks. *Water Resources Research* 7:1558–1566.

O'Callaghan, J. F., and Mark, D. M. 1984. The extraction of drainage networks from digital elevation data. *Computer Vision, Graphics, and Image Processing* 28:323–344.

Scheidegger, A. E. 1967. On the topology of river nets. *Water Resources Research* 3:103–106.

Schenck, H. S. Jr. 1963. Simulation of the evolution of drainage basin networks with a digital computer. *Journal of Geophysical Research* 68:5739–5745.

Seginer, I. 1969. Random walk and random roughness models of drainage networks. *Water Resources Research* 5:591–607.

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

———, 1967. Infinite topologically random channel networks. *Journal of Geology*75:178–186.

———, 1975. The probabilistic-topologic approach to drainage basin geomorphology. *Geology* 3:527–529.

Smart, J. S. 1969. Topological properties of channel networks. *Bulletin of the Geological Society of America* 80:1757–1774.

Strahler, A. N. 1952. Hypsometric analysis of erosional topography. *Bulletin of the Geological Society of America* 63:1117–1142.