5

Urban Boundaries and Edges

The fascination of boundaries lies in their ambivalent role of dividing and connecting at the same time. They mark the transition between different modes of existence. They transmit and control exchange between territories. They are the playground for discovery and conquest ... They are the result of never ending competition and exhibit structure on many scales. (Richter and Peitgen, 1985, p. 571–572.)

5.1 At the Edge of the City

Boundaries, as Richter and Peitgen (1985) so graphically portray, are places which mark the transition between different regimes, different systems, and this is nowhere more so than between the rural and urban worlds at the edge of the city. In one sense, the boundary of the city marks the transition between different epochs, between an older agricultural society and the newer industrial, although the distinction is becoming weaker as contemporary society is beginning to make its transition to a post-industrial era with all its consequences for how cities will be organized. Nevertheless, such zones of transition do reflect the tension between the old and the new, places where more stable, established structures are being continually tested by a newer, ever-changing dynamic. Even in these terms, such boundaries are not likely to be 'smooth' in any sense and as we shall see, their physical form is both irregular but self-similar in that a precise transition between the old and the new can never be definitively marked out.

In defining the physical form of the city, its edge or boundary is the most obvious visual delimiter of its size and shape. Statistical definitions of cities rely upon the definition of boundaries, although such definitions are never comprehensive; there are so many possible ways of cutting the continuum of development from urban to rural that the general idea of a boundary remains a conceptual notion which is only given physical form through narrow definitions. Urban boundaries, however, are not simply linear constructs which mark off one side of the continuum from the other but they imply area, and thus shape (Batty, 1991). As we have argued in earlier chapters, although cities can be visualized across many dimensions, they are usually best pictured in the plane as two-dimensional phenomena and thus their boundaries immediately imply some measure of area. In this sense, the boundary is clearly something more than a one-dimensional line for whenever we examine such an edge, we conceptualize an area.

There are many notions as to what constitutes the boundary of the city, several of which we will be using in this book. In Chapters 3 and 4, we defined cities in terms of concentric rings of different land uses about their CBDs based on von Thunen's original division of land use along a spectrum from highly urban to rural; in contemporary terms, this continuum begins with high density commercial at the core of the city and evolves to low density agricultural uses such as market gardening at the periphery. We also referred extensively in Chapter 4 to the notion of inner and outer suburban areas, while the idea of the suburban fringe as the zone of transition between urban and rural can be extended to the quasi-urban area at the rural edge of the city sometimes called exurbia. In later chapters, we will have recourse to extend our definition of the city to its wider hinterland or field, that area which contains all the development which in one way or another is associated with the city. Definitional problems abound, too, for in the age of the world city, activities may exist everywhere on the globe which are in some sense dependent upon the city in question. Finally, we will make our definitions of the extent of the city much more precise from Chapter 7 on when we begin to introduce the idea that population density must be the delimiter of form and that density itself rather than the shape of land use or physical development, is the true measure of whether or not cities are fractal.

In this chapter and the next, we will begin to define and measure the form of the city in analogy to the way we discussed the definition of a fractal line in Chapters 2 and 3. There we argued that fractal lines are something more than the one-dimensional Euclidean line but something less than the two-dimensional plane; the coastline is the example *par excellence*. In fact, it is likely that the fractal dimension of our urban boundaries will be closer to 1 than to 2 for we will not consider cities which are entirely composed of boundaries, in which the boundary itself twists and turns to fill the two-dimensional space. This point in itself is somewhat controversial, for it is possible to define cities which are entirely composed of boundaries if the level of spatial resolution is chosen accordingly. Moreover, there are recent theories of the post-industrial city which are predicated on the idea that everything significant in the modern city is at the edge: 'Edge Cities' as they have been called (Garreau, 1991), thus giving some meaning to our own notion that the most interesting aspects of urban phenomena depend upon what is happening on their boundaries.

However, urban boundaries or edges can be very different from coastlines in the following sense. Whereas our interest in coastlines is often only over a fixed stretch of the line, our interest in urban boundaries is likely to be over their entirety in that to define a city by its boundary, there is usually some measure of closure to the line. The boundary thus marks out an envelope. In fact, if we examine coastlines in their entirety too, we must consider the same sorts of closed line. It is nonetheless a comment on the rudimentary development of fractal geometry that there are virtually no discussions so far of the implications for measurement posed by objects with closed, in contrast to open, boundaries. We will in fact extend our discussion into these realms, but to anticipate the outcome, much remains to be done. In this chapter we will define the problem of the closed

boundary away by adopting artificial closure, and only in the next will we broach the matter directly.

In Chapter 2, we identified four ways in which we might define objects of interest and measure their scale-dependence, hence their fractal dimension. We noted that we could derive the fractal dimension of a single object by measuring the same object at different scales or by varying the extent or size of the object over which the dimension might be computed. In essence, we will be adopting the first method here, that is taking a given city and examining its physical properties at different scales in contrast to the second method where we change the size of the given object; this we will develop from Chapter 7 on. We can also derive the fractal dimension of a set of objects by examining the size distribution of the objects in question, and we will do this for individual land use parcels in Chapter 6 which will extend the ideas of this chapter. If we have a set of objects, we could also change each of their scales and simply combine all the scale-dependent results and use the methods of this chapter to compute a fractal dimension for the entire set. But for a set of objects, the most appropriate methods are those which involve examining their size, not scale. In short, what we will introduce in this chapter are methods such as those we presented in Chapter 2 for deterministic fractals such as the Koch curve and we will apply these to a single city, deriving its fractal dimensions from the lines which compose its boundary.

Here we will use the town of Cardiff, which is the capital city of Wales, as our example. Cardiff has a very distinct urban edge, and our problems of defining its boundary are considerably less than in many other possible examples. However, the boundary is a closed line, and without any knowledge of fractal geometry, it seems intuitively obvious that such a line implies an object with a dimension somewhat greater than one. Any layman would probably associate the closed line with an area and argue that the purpose of the line was simply to mark out the area. Common sense would thus imply that resulting object was in the plane rather than the line. However, as we shall see, the measurement of the fractal dimension of these closed lines yields values which are much closer to one than two, and which are quite close to the theoretical Koch coastline where $D \approx 1.262$. What these findings will impress is that the concept of fractal dimension is completely dependent upon what is being measured, or rather what physical properties of an object are being selected for measure, and that there are likely to be many different types and values of fractal dimension. We will, of course, elaborate this important point throughout the rest of this book.

We will first discuss the way geographical boundaries might be represented before we move to outline the formal methods which we use to derive the fractal dimension of a line. These methods are those which we have already presented in Chapter 2 but they will be repeated again here for we will adapt them somewhat differently to this context. We illustrate the basic method for the case of Cardiff's urban edge (in 1949) and this serves to point up some problems of measurement and statistical method. We then go on to outline the way these methods might be used to compute the changing dimension of a growing city, using data from the growth of Cardiff from the late 19th century to the middle of the 20th. To generate the relevant dimensions, we will use four different methods of approximating a

fractal line, namely, the structured walk, equipaced polygon, a hybrid of both these, and the so-called cell-count which is a simplification of the well-known box-counting method (Voss, 1988). Finally, we will draw these results together so that we might extend them to measuring the boundaries of different land uses in Chapter 6.

5.2 Cartographic Representation and Generalization of Geographical Boundaries

As we demonstrated in Chapters 2 and 3, the most celebrated example of a fractal is a coastline. Although the development of fractal geometry only really took off after Mandelbrot's famous paper in *Science* in 1967 where he posed the conundrum of length in terms of 'How long is the coast of Britain?', it was Richardson (1961) who first articulated the problem in these familiar terms. Richardson demonstrated quite unequivocally that the length of a coastline depended upon the yardstick or scale with which its length was measured. As we illustrated for the Koch curve, he showed that as the scale became finer, more and more detail could be picked up by the measuring instrument, thus implying no bounds on its length. Although Richardson did not formalize the concept of fractal dimension, which was left to Mandelbrot (1967), he did derive the familiar log–linear relationship between length and scale, and in estimating this, demonstrated that the fractal dimension of coastlines ranged from around 1.02 for South Africa, 1.13 for Australia to 1.25 for the western shore of Britain.

As we also noted in Chapter 2, this conundrum has been remarked upon for at least a hundred years, and it is likely that it was known in some form to Renaissance geometers and thus probably to the Greeks. In the 1960s with the development of mathematical geography, Nysteun (1966) in a seminal paper, not only identified the problem and suggested a solution through the definition of length contingent upon the scale used, but he also pointed to the work of the Polish mathematician Steinhaus (1954, 1960) and geographer Perkal (1958a, 1958b) who had both reflected upon the paradox. Perkal in fact drew attention to the work of the Viennese geographer Penck (1894) who was familiar with the problem in the late 19th century. However, the problem was simply noted, and apart from some attempts at its resolution with respect to associating length with explicit scale, there were no attempts until Mandelbrot (1967) to pose it in a wider framework. It would, in fact, have been remarkable had not the problem been posed in countless guises throughout history, but it probably had to await the arrival of computer graphics, hence fractals, before its universal import could be appreciated.

What is fascinating is that the problem has never been restricted simply to physical systems. Nysteun (1966) described the conundrum of length in discussing the boundary of the town of Ann Arbor, Michigan; Perkal (1958b) illustrated the same for the boundary of the town of Wroclaw in Poland, while Richardson (1961) himself used political frontiers as examples

of fractal curves. In fact, he derived the fractal dimension of the frontier between Spain and Portugal as 1.14 and of the German land frontier in 1899 as 1.15. Although Mandelbrot (1983) developed his new geometry mainly with natural examples in mind, he is strident in maintaining that the geometry is applicable to artificial systems. He says in discussing the amount of circuitry which can be packed onto a chip: "This and a few other case studies help demonstrate that in the final analysis, fractal methods can serve to analyze any 'system', whether natural or artificial, that decomposes into parts in a self-similar fashion, and such that the properties of the parts are less important than the rules of the articulation".

That urban boundaries are fractal in some sense might already seem self-evident, although we still have to demonstrate the point. There is, however, another issue which dominates the definition of boundaries for geographical systems, and this involves the concept of 'generalization' as it appears in cartography. Generalization is the process of aggregating cartographic features which encompass a map, from one scale to another, and as such, the various methods developed have often alluded to the problem posed by the conundrum of length where cartographic lines are involved. In fact, cartographers have made considerable progress in the search for methods for selectively aggregating and filtering geometric detail as lines are generalized from smaller to larger scales, and have, perhaps unwittingly sometimes, invoked the geometry of fractals (Lam and Quattrochi, 1992).

In exploring the extent to which any boundary or line might be dependent on scale, the process of generalization is likely to detect such variations and thus its development is important to the measurement of fractal dimension in cartographic lines. Buttenfield (1985) has conceptualized it as encompassing four related procedures and processes: first, simplification, such as in the removal of unwanted detail and the smoothing of features; second, symbolization, in which line character is graphically encoded according to geographical and perceptual conventions; third, classification, in which cartographic information is aggregated and/or partitioned into categories; and fourth, induction, in which the creative logical assumptions which are made during generalization are applied. As such, it is clear that the depiction of cartographic information is the end result of a variety of codification conventions mediated by a human judgmental process. This is no less the case in the generation of computer-digitized data bases than in traditional cartographic line-drawing (Jenks, 1981).

There are many types of method for line generalization. Buttenfield (1985) develops a comprehensive classification and critique of such algorithms including various random and systematic point weeding routines to simplify detail, the fitting of various mathematical functions to lines, the epsilon neighborhood concept based on linking line length to scale (Perkal, 1958a, 1958b), and the use of both angular and band-width tolerancing to dispense with successive points which fall outside a prespecified angular and/or band-width threshold (Peucker, 1975). She concludes that the choice of method used can depend upon the often-conflicting emphases that different studies have placed upon geographical and perceptual accuracy.

The measurement of shapes through boundaries has a long history in natural science too and has also been absorbed into the locational analysis tradition of human geography (see Haggett, Cliff and Frey, 1977). Many of

the earliest shape indices were based upon simple length, breadth and area relations. This was primarily because the constraints associated with time-consuming manual measurement restricted the assessment of line structure to simple indices of variation at selected points and to the monotonicity of line segments about a base line anchoring the end points of the line. These efforts were nevertheless well-motivated, since even fairly crude measurements and classifications of form can enhance urban analysis. There are numerous examples of such use. At an extreme, Thomson (1977) develops simple areal density measures in order to classify the functional relationship between transport infrastructure and urban form, while closer to the ideas to be developed here, Benguigui and Daoud (1991) have undertaken a detailed empirical analysis of the relationship between the form of the Paris suburban railway system and the distribution of the urban population. As we demonstrated in Chapter 1, many of our conceptions of the city, ancient and modern, are rooted in the idea that the geometry of the city in terms of simple indices of shape can, in some way, be tied to its functioning, and that to change or control its functioning involves manipulating its geometry. In this sense, form follows function and our ultimate aim in this book is to demonstrate how the new geometry of fractals can inform this quest. In this chapter, we will begin by linking the shape of cities to their boundaries.

In the present context, we suggest that most of these methods involving techniques of generalizing lines or measuring simple geometric properties of shapes are flawed in at least two fundamental ways. First, in a geographical sense, most are heavily reliant upon the *a priori* definition of the scale, starting points and ending points of constituent line features for the synthesis of total line structure. Second, in a perceptual sense, many algorithms fail to preserve the qualitative visual character of a line in terms of its shape when it is generalized (Muller, 1986). Seen in the context of the emergent relationship between measurement and simulation of urban form, this produces two grave shortcomings. First, it is not possible to specify *a priori* those features which we expect to characterize urban boundaries, since their inductive generalization remains one of the primary goals of the measurement exercise. Second, if we are to derive visually acceptable space partitioning rules for land use simulations then our measurement parameters must maintain perceptual accuracy in any of the lines which they are used to generate. In this chapter, we will contend that use of fractal techniques can provide a consistent and feasible route beyond this impasse, since first by using very few parameters, the line can be measured as a total entity rather than as a piecemeal amalgam of constituent features; and second, the line's visual character is preserved by using the concept of self-similarity and by sensitive assessment of the range of scales over which the fractal property holds.

Before we launch into the measurement task, it is important to reflect upon the notion of exact or statistical self-similarity resulting from a single, or small number of processes. Clearly this notion becomes increasingly strained as we make the transition from physical to social systems where, for example, urban edges clearly evolve under a wide range of simultaneous physical and social processes. At a theoretical level, many of the more abstract spatial theories anticipate self-similarity, with central place

theory being perhaps the best example (Arlinghaus, 1985). At a procedural level, we might consider that at worst, the differences in the nature of self-similarity between physical and social systems are of degree rather than of kind. In such circumstances, there would be no rationale why fractal measurement should not proceed in a similar manner to applications in disciplines as diverse as particle science, mineralogy and music (Dearnley, 1985; Kaye, 1978; Mark and Aronson, 1984; Dodge and Bahn, 1986).

All such analyses depend critically upon isolating the most appropriate range of scales over which any statistical property of fractals holds. For example, it is likely to be the case that geographical features will revert to man-made Euclidean dimensions at certain fine scales. Additionally, and in all such instances, the cartographer is the arbiter, and to some extent the architect, of the final depiction of the map feature, and any summary measure must ultimately be viewed in part as the outcome of a human judgmental process. In summary then, the measurement and generalization of cartographic lines using fractals is likely to have a number of advantages over other forms of representation (Muller, 1987) and there are grounds to anticipate that empirical evaluation of the fractal dimension ". . .may be the most important 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" (Goodchild and Mark, 1987). The rest of this chapter will be focussed on demonstrating how such dimensions emerge as a natural consequence of the process of generalization.

5.3 The Basic Scaling Relations for a Fractal Line

The two basic relations for a fractal line associate the number of parts into which the line can be divided, and its length, to some measure of its scale. These relations have already been stated in equations (2.24) and (2.25) respectively and we will proceed in analogy to these. First, consider an irregular line of unspecified length R between two fixed points. Define a scale of resolution r_0, such that when this line is approximated by a sequence of contiguous segments or chords each of length r_0, this yields N_0 such chords. Now determine a new scale of resolution r_1 which is one-half r_0, that is, $r_1 = r_0/2$. Applying this scale r_1 to the line yields N_1 chords. If the line is fractal, then it is clear that ". . .halving the interval always gives more than twice the number of steps, since more and more of the self-similar detail is picked up" (Mark, 1984). Formally this means that

$$\frac{N_1}{N_0} > 2 \text{ and } \frac{r_0}{r_1} = 2. \tag{5.1}$$

This is illustrated for three different scales in Figure 5.1. Using equation (2.25), the lengths of the approximated curves or perimeters, in each case, are given as $L_1 = N_1 r_1$ and $L_0 = N_0 r_0$ and from the assumptions implied in equation (5.1), it is easy to show that $L_1 > L_0$. This provides the formal justification that the length of the line increases without bound, as the chord size (or scale) r converges towards zero.

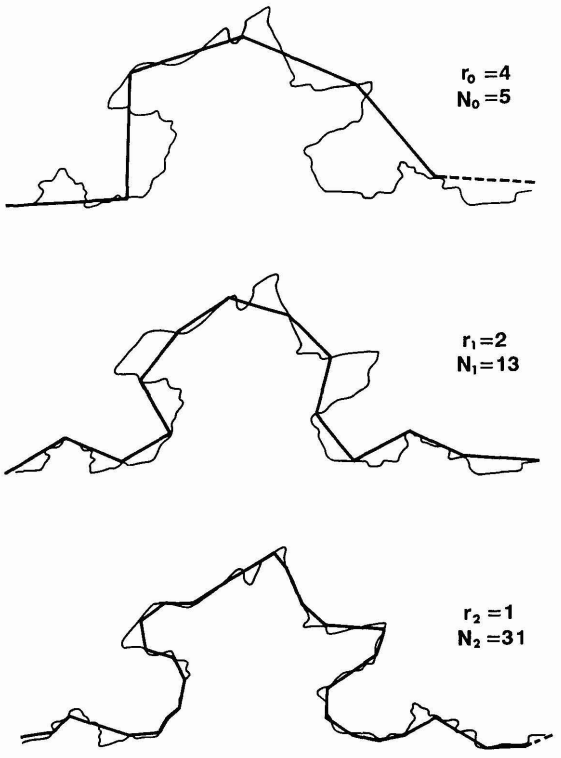

Figure 5.1. Approximating an irregular line and measuring perimeter length at three adjacent scales.

The relationship in (5.1) can be formally equated if it is assumed that the ratio of the number of chord sizes at any two scales is always in constant relation to the ratio of the lengths of the chords. Then

$$\frac{N_1}{N_0} = \left(\frac{r_0}{r_1}\right)^D, \tag{5.2}$$

where D is defined as the fractal dimension. If halving the scale gives exactly twice the number of chords, then equation (5.2) implies that $D = 1$, and that the line would be straight. If halving the scale gives four times the number of chords, the line would enclose the space and the fractal dimension would be 2. Equation (5.2) can be rearranged as

$$N_1 = (N_0 r_0^D)\, r_1^{-D} = \alpha r_1^{-D}. \tag{5.3}$$

where the term in brackets $(N_0 r_0^D)$ acts as the base constant α in predicting the number of chords N_1 from any interval of size r_1 relative to this base.

From equations (5.2) and (5.3), a number of methods for determining D emerge. Equation (5.2) suggests that D can be calculated if only two scales are available (Goodchild, 1980). Rearranging equation (5.2) gives

$$D = \log \frac{N_1}{N_0} \Big/ \log \frac{r_0}{r_1}. \tag{5.4}$$

However, most analyses not only involve a determination of the value of

D but also of whether or not the phenomenon in question is fractal, and thus more than two scales are required. Generalizing equation (5.3) as in equation (2.25) gives

$$N(r) = \alpha r^{-D} \tag{5.5}$$

where $N(r)$ is the number of chords associated with any r. Using logarithms, we can linearize equation (5.5) as

$$\log N(r) = \log \alpha - \mathrm{D} \log r. \tag{5.6}$$

Equation (5.6) can be used as a basis for regression by using estimates of N and r from several scales. The related formula involving the length of the curve or perimeter L analogous to equation (2.25) is derived from equation (5.5) as

$$L = Nr = \alpha r^{(1-D)}. \tag{5.7}$$

Equation (5.7) can also be linearized by taking logarithms,

$$\log L = \log \alpha + \beta \log r \tag{5.8}$$

where $\beta = (1 - D)$. It is clear that the intercepts α in equations (5.6) and (5.8) are identical and the slopes are related to the fractal dimension D in the manner shown. In later sections, we will use equation (5.8) rather than equation (5.6), for equation (5.8) will enable us to check the range of scales used more effectively.

The original method used by Richardson (1961) to measure the length of coastlines and frontiers involved manually walking a pair of dividers along the boundaries at different scales and then determining D from equation (5.8). To enable the entire perimeter to be traversed, the last chord length which always finishes at the last coordinate point is generally a fraction of the step size, and the step sizes used at each scale usually reflect orders of magnitude in geometric relationship; that is $r_n = a^{-n} r_0$, $a > 1$, which enables each step size to be equally weighted and spaced in the log–log regression. In Richardson's (1961) research, about six orders of magnitude or scale were used which is regarded as sufficient to determine a least-squares regression line. Computer simulations of Richardson's manual method are now well established. Kaye (1978, 1989a) refers to the method as a 'structured walk' around the perimeter of an object, and he calls the log–log scatter plot of perimeter lengths versus scale intervals a 'Richardson plot'; this provides a useful visual test of whether or not the phenomenon is fractal. The structured walk method is easy to implement on a computer, and here we have used the algorithm developed by Shelberg, Moellering and Lam (1982) which involves approximating the boundary of an object consisting of line segments between digitized coordinates, with different sized chord lengths.

There are two variants involving this method. First, the number of chords and perimeter lengths will depend upon the starting-point along the curve. To reduce the arbitrariness of this variation, several workers have suggested the structured walk be started at several different points, and averages of the results then formed (Kent and Wong, 1982). For example, Kaye, Leblanc and Abbot (1985) start the walk at five different points along the curves, but there is no reason why, in principle, the walk should not be

started at each of the digitized points which define the base curve. In the illustrative example which we will develop in the next section using the digitized points of Cardiff's urban boundary in 1949, we will initiate successive walks from each of the 1558 digitized points which define the urban edge, the walks proceeding in both directions towards the endpoints of the boundary. This variation on the basic method is inevitably time-consuming in computational terms.

The second variant involves starting the structured walk at different divider lengths and generating sequences of predictions from these different lengths. The range of scales over which the perimeter lengths were computed varied from about half the average chord length associated with the digitized data, to over the maximum distance between any two coordinate points on the perimeter. The average chord length is computed as follows. First, the distances between each adjacent pair of (x,y) coordinates, i and $i + 1$, are computed using the standard triangle equality

$$d_{i,i+1} = [(x_i - x_{i+1})^2 + (y_i - y_{i+1})^2]^{1/2}, i = 1, \ldots, n - 1, \tag{5.9}$$

and then the perimeter L of the base-level curve which has been digitized at resolution r can be summed as

$$L(r) = \sum_{i=1}^{n-1} d_{i,i+1}. \tag{5.10}$$

The average chord length $\bar{d}$ of the original curve is therefore

$$\bar{d} = \frac{L(r)}{n-1}, \tag{5.11}$$

and a lower bound for the chord length used to start the approximation, as suggested by Shelberg, Moellering and Lam (1982), is taken as approximately $\bar{d}$. The maximum distance between any pair of coordinates, which in fine particle science is referred to as Feret's diameter by Kaye (1978, 1989a), is given as

$$F = \max_{i,i+1} (d_{i,i+1}), \tag{5.12}$$

and Kaye (1978), amongst others, suggests that an appropriate upper bound for chord length approximation is $\approx F/2$. The intermediate chord lengths between these lower and upper limits should be ordered a geometrical sequence so as to ensure more equal weighting in the regressions.

5.4 Estimating the Fractal Dimension: the Urban Boundary of Cardiff

We have now presented sufficient method to derive our first example of a fractal dimension for a city. In order to illustrate the procedure, we have digitized the 1158 points composing the boundary of the town of Cardiff

in 1949 which we show in Figure 5.2. This figure provides an excellent example of the problems of measurement which we confront. As such it represents a visual trace akin to the sorts of photographs physicists use to search for the existence of elementary particles, but with an important difference. The points which are defined in Figure 5.2 represent a series of subjective judgements as to the level of detail needed to represent the boundary at this elemental level. As the irregularity of the boundary varies over its length, then more points in general are defined where more detail is observed. It could be argued that we should represent this elemental level at the same level of detail everywhere and this of course would occur if, for example, a grid or other regular tessellation of the plane were used to detect the boundary. However, such a grid would have much redundancy if it were to detect the finest level of detail and therefore we have proceeded on the assumption that it is important to present as much detail as possible at the elemental level. As we shall see, some of the methods we use to derive fractal dimension will be based on regular tessellations of the plane but the problem of measuring 'objective' statistics such as fractal dimensions in subjectively specified data sets will continue to concern us throughout this book and we will return to it again in the sequel.

The boundary marking the extent of the urban area of Cardiff was defined from the 1:25,000 Ordnance Survey map published in 1949. The usual problems of definition were encountered in determining the edge of the urban area, and several rules of thumb were invoked. Typically, allotments and other urban fringe land-uses were excluded, villages linked to the urban area by ribbon development were included, man-made alter-

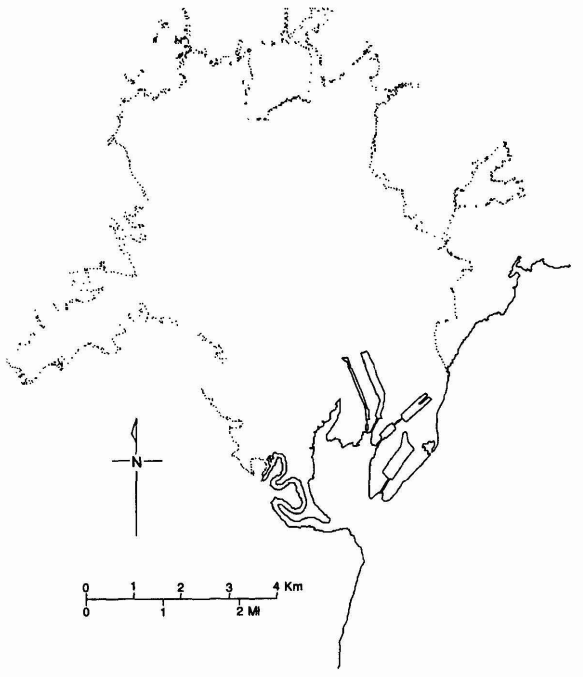

Figure 5.2. The density of point digitization of the 1949 Cardiff urban edge.

ations to rivers and coast were included, but large landed estates which subsequently become part of the urban fabric were included only if development had surrounded them. The entire definition process emphasized the obvious problems that urban processes and constraints operate at different scales, and this casts some doubt on the fractal concept of self-similarity in this context; but perhaps no more doubt than exists in other areas of the physical sciences where fractal concepts have been shown to apply only over restricted scales. Once the boundary had been defined, it was digitized to within 1 mm resolution; the coastline contained some 900 points, whereas the urban boundary was based on 1558 points. Figure 5.2 is thus a fair representation of the land which by 1949 had become 'irreversibly urban' in character, and is consistent with other official standards for defining 'urbanity' (OPCS, 1984).

Figure 5.3 shows the digitized outlines as well as a coarse approximation to the boundary produced by the structured walk method, which is about 30 times the scale of the original data. The approximating polygon touches the original boundary at those points on the base curve which are retained for the approximation, and all of the chords are of equal length except for the end (residual) chord distance(s). The perimeter of the digitized boundary determined from equations (5.9) and (5.10) gives $L = 3104.456$ units, with the average chord length $\bar{d} = 1.993$ from the equation (5.11), and the Feret diameter $F = 432.935$ from equation (5.12). These measures are useful to keep in mind when we discuss the relative merits of different fractal measurement methods. We will deal first with the structured walk method. For a given chord length used to start the sequence of predictions of perimeter lengths, a complete series of 10 chord lengths are used in the

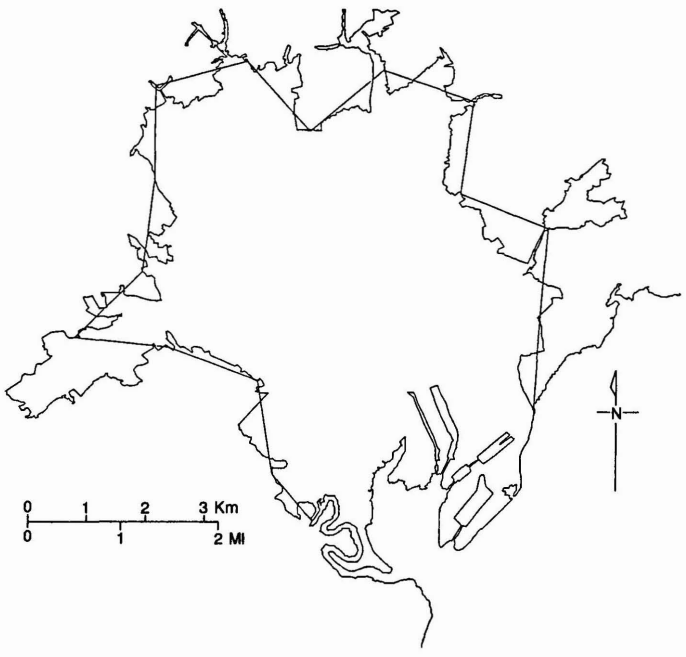

Figure 5.3. A typical scale approximation to the digitized urban edge of Cardiff.

approximations, starting from the finest level of scale now given by r_0 and moving to coarser scales r_n. The sequence of chord lengths is computed from $r_n = 2^n \nabla$, $n = 0, 1, \ldots, 9$, where ∇ is the start length which is always a function of $\bar{d}$, the average chord length. Thus, for example, where $\nabla = \bar{d}/2$ which is the lower bound recommended by Shelberg, Moellering and Lam (1982), the sequence of chord lengths used are in the following ratios: $\frac{1}{2}$, 1, 2, 4, 8, 16, 32, 64, 128, 256. In this case, $r_0 \approx 1$ and $r_9 \approx 510$ which is much larger than Kaye's (1978) upper bound of $F/2$. To provide some feel for this range of approximations, we have plotted the approximated boundaries of Cardiff for r_n, $n = 0, 1, \ldots, 8$, in Figure 5.4. With r_9, the boundary is approximated by only one chord which is clearly inappropriate. Indeed, even with r_7 and r_8, the approximations are too coarse to be of much use. This is clear from Figure 5.4 which shows that this kind of visual test is essential in selecting an appropriate range of measurements for use in the subsequent regressions.

We will illustrate the issue of ascertaining the most appropriate scale

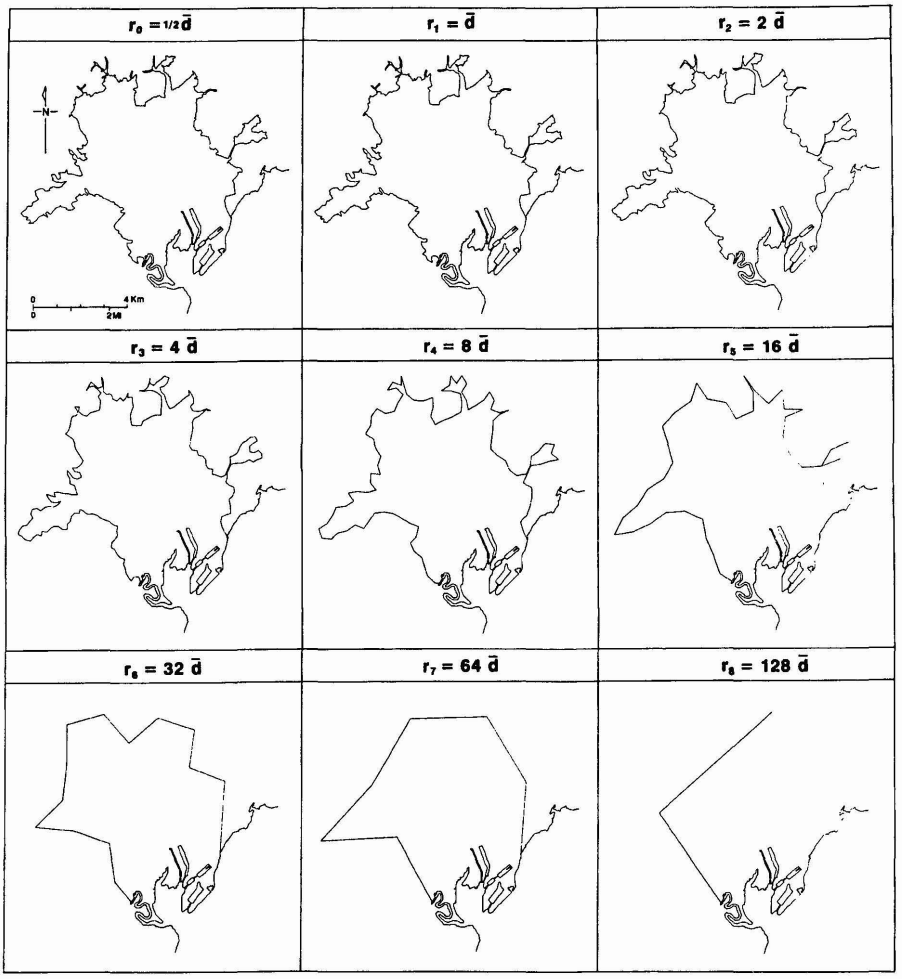

Figure 5.4. A sequence of scale approximations to the urban edge.

range for measuring fractal dimension by selecting 10 different starting values of the chord length ∇, and generating ten sets of measurements for each of these starting values. The values of ∇ chosen are $\nabla = 0.4\bar{d}$, $0.5\bar{d}$, $0.6\bar{d}$, $0.8\bar{d}$, $\bar{d}$, $1.5\bar{d}$, $2\bar{d}$, $3\bar{d}$, $4\bar{d}$ and $5\bar{d}$. From the sequences generated, it is clear that several of the chosen measurements are the same between series, but each of the regressions developed below involves different sets of measures. A visual comparison of each of the ten sequences generated is also contained in the Richardson plots in Figure 5.5 which show the ten measures of log L versus log r for each of the ten starting values of ∇. These plots are all on the same scale for comparative purposes and also show the values of $\bar{d}/2$, $\bar{d}$, $F/2$ and F.

Before we present the results of the regressions, we need to consider how we can systematically narrow the range of results we are able to generate, and to this end, we have devised five criteria. First, we have used the range $0.4\bar{d} \leq r \leq F/2$ to select those observations which are appropriate. Second, we have used the Richardson plots to identify outliers for exclusion. In particular, when $r > F$, then the algorithm always gives the same perimeter length because it always closes the single chord on the last coordinate point. Such points show up horizontally on the Richardson plots and must be excluded. Third, the scale approximation must be acceptable visually. An examination of Figure 5.4 suggests that approximations with ten chords or less are not satisfactory in representing the overall shape, and thus must be excluded. Fourth, we suggest that the r^2 measure of fit (coefficient of determination) should always be better than 0.95, and fifth, the standardized variation in average perimeter length for each chord r should not be greater than 10% of the mean value. This also enables poor approximations to be excluded.

For each of the 10 starting values of ∇ in the structured walk, we have performed regressions on all 10 points shown in the Richardson plots in Figure 5.5, on the first nine, the first eight, seven, six, then five, below which it is not appropriate to carry out such least-squares fitting. The absolute values of the slopes of the regression lines $|\beta| = |(1 - D)|$ are shown in Table 5.1 along with the r^2 values, but as Shelberg, Moellering and Lam (1982) indicate, such r^2 values should be used in a descriptive rather than an inferential sense. In Table 5.1, the figures which are in bold type involve regressions in which the observations meet all the five criteria mentioned above, and this narrows the range considerably. Note that the fractal dimension is given by adding 1 to the absolute slopes in Table 5.1, that is, $D = 1 + |\beta|$.

For the structured walk method, there is still a large variation in fractal dimension from $1.155 \leq D \leq 1.289$, and from Table 5.1 it is quite clear that as finer and finer scales come to dominate the regression, so the value of D decreases. This implies that there is greater irregularity at coarser scales, but it also indicates that where the scale is below the level of resolution of the digitized boundary, that is, where $r < \bar{d}$, then no further detail is picked up and the boundary must be considered Euclidean. This is the case for the first four starting values (first four rows) in Table 5.1, and if these are excluded from consideration, the range of D is from 1.234 to 1.289. In fact, the rule of thumb suggested by Shelberg, Moellering and Lam (1982) that

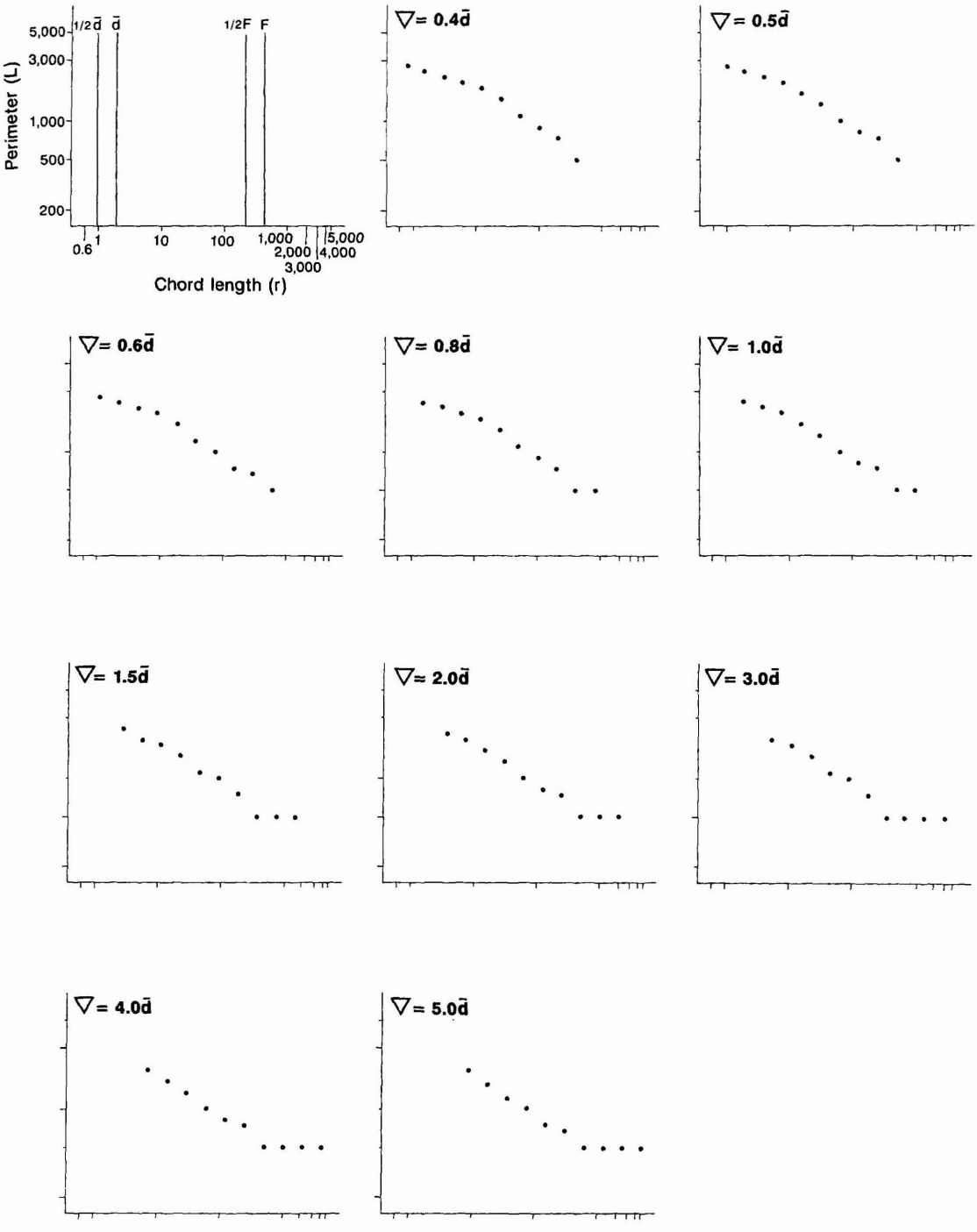

Figure 5.5. Richardson plots based on ten structured walks.

Table 5.1. Logarithmic regression of perimeter on scale associated with the Richardson plots in Figure 5.5

Starting values ∇[1]	Number of observations[2]					
	10	9	8	7	6	5
$0.4\bar{d}$	0.269	0.244	0.231	0.207	**0.177**	**0.155**
	0.953	0.959	0.947	0.944	**0.961**	**0.969**
$0.5\bar{d}$	0.278	0.258	**0.255**	**0.236**	**0.211**	**0.180**
	0.969	0.975	**0.963**	**0.956**	**0.956**	**0.975**
$0.6\bar{d}$	0.279	0.263	0.254	**0.236**	**0.216**	**0.185**
	0.975	0.975	0.964	**0.963**	**0.953**	**0.966**
$0.8\bar{d}$	0.292	0.291	0.266	0.254	**0.231**	**0.198**
	0.976	0.966	0.973	0.962	**0.957**	**0.974**
$\bar{d}$	0.291	0.297	0.276	0.278	**0.261**	**0.234**
	0.982	0.977	0.983	0.975	**0.967**	**0.963**
$1.5\bar{d}$	0.293	0.309	0.308	0.280	**0.261**	**0.254**
	0.975	0.980	0.971	0.980	**0.980**	**0.963**
$2.0\bar{d}$	0.282	0.304	0.315	0.293	0.303	**0.289**
	0.969	0.984	0.982	0.989	0.987	**0.979**
$3.0\bar{d}$	0.274	0.303	0.327	0.331	0.301	**0.282**
	0.945	0.972	0.984	0.977	0.986	**0.985**
$4.0\bar{d}$	0.254	0.284	0.313	0.331	0.306	0.328
	0.924	0.958	0.981	0.983	0.989	0.996
$5.0\bar{d}$	0.245	0.276	0.308	0.331	0.321	0.329
	0.915	0.953	0.984	0.996	0.997	0.996

[1]Starting values in each sequence of the structured walks.
[2]Number of observations of perimeter–chord lengths used in regressions. The first value in each row–column is slope $|\beta|$; the second value in parentheses is r^2.

∇ should begin at about $\bar{d}/2$ should be reevaluated in future work so that the variation around $\bar{d}$ can be considered.

5.5 Form and Process: Cardiff's Changing Urban Edge

Boundaries which partition complex systems from their environment and from one another reflect properties and processes which can be inferred from their morphology, as Richter and Peitgen (1985) imply in the quote introducing this chapter. For example, transport and building technologies, social controls over development as well as physical constraints determine the boundary investigated in the previous section, just as the shape and form of the coastlines referred to in Chapters 2 and 3 reflect the action of a variety of geophysical processes. If we were able to observe the change in

boundaries through time, then this should give us some clue to the various processes at work, and in this section, we will explore this issue with respect to what the changing boundary of Cardiff over a 50 year period implies for the urban growth of that city.

Thus far, we have seen how perimeter–scale relations may be displayed as a Richardson plots as in Figure 5.5, and we have used such plots in order to detect the range of scales over which it is appropriate to extract information from a data base digitized to a given level of resolution. If a fractal dimension is stable over many scales and the scatter of points about a simple regression line is well-behaved (that is, close fitting), we can infer that the morphology is consistent with a single set of processes operating at every scale. In the case of an urban boundary, which evolves as a concatenation of a variety of processes, it is more plausible to anticipate that a multitude of processes leads to the emergence of a particular fractal dimension. In fact, measurements of the fractal dimensions of boundaries, particularly coastlines (Kent and Wong, 1982; Mandelbrot, 1967; Nakano, 1983, 1984; Richardson, 1961) and fine particles (Flook, 1978; Kaye, Leblanc and Abbot, 1985; Orford and Whalley, 1983), have suggested that such phenomena may be 'multifractal', that is, with different (in this case usually lower) fractal dimensions at smaller scales. This is intuitively plausible in that we might anticipate that different processes operate at different scales, especially where man-made and natural processes combine (Kaye, 1984; 1989a). The importance of the fractal dimension thus lies in identification of the range of scales over which processes operate and the different scales at which such properties manifest themselves over time. It also enables changes in the morphological effects of self-similarity to be explored.

Here we will develop the example used above in order to examine how the irregularity of the boundary of an urban area changes at different scales and through time. We will seek to use these measurements to infer changes in the processes which condition urban growth in time and space. Reexamining some of the graphs in Figure 5.5 reveals evidence of a slight curvilinear trend about the points in the Richardson plots, suggesting that a multifractal (rather than conventional straight line, or log–linear) formulation of the perimeter–step length/scale relationship may have been appropriate in this instance. In focussing upon temporal changes in detailed morphology, we have therefore increased the precision with which the boundaries for this substantive analysis were digitally measured. We will consider the urban boundaries of Cardiff in 1886, 1901 and 1922, but in order to avoid ambiguity of interpretation we will not compare our results with the 1949 digitized perimeter, since this data set was not digitized at a directly comparable level of resolution. We do, however, show the four boundaries overlying one another in Plate 5.1 (see color section) which shows the town's urban expansion, its extent and the changing irregularity of the urban edge over time.

Just as the form of a coastline evolves as the outcome of a range of simultaneous physical circumstances, so the morphology of the city is the outcome of a multitude of physical and social processes as we have already implied (Batty, 1992). These include the technology of building, patterns of land tenure, the size of building plots, the demand for residential space, the mobility of the population, and the efficiency and availability of transport

technology. These processes manifest themselves at different scales, for example, building technologies at smaller, transport at larger scales. It is a reasonable assumption that these processes are reflected in the boundary of the city, hence in its degree of irregularity and fractal dimension (Perkal, 1958a, 1958b).

Accordingly, we will advance three hypotheses concerning changes in the fractal dimension of these urban boundaries. First, we consider that the boundary is multifractal across a range of scales; second, that as there is greater control over physical development at smaller scales, the fractal dimension is likely to decrease with scale; and third, that the fractal dimension at smaller scales should decrease over time as greater controls over building technology and land development have been instituted. At larger scales, it is less clear how the fractal dimension changes although increasing mobility and accessibility imply it too will decrease through time. We will test these hypotheses by determining the fractal dimensions of the urban boundary of Cardiff in 1886, 1901 and 1922. These times have been chosen because of the rapid urban growth of the city from a population of 80,000 to 230,000 during this period. This period also marked the development of the tramway system which began in 1872 and was complete by 1914, and it was the period when the predominant style of late Victorian worker housing gave way to more spacious suburban housing. The landed estates which dominated the form of development in Cardiff in the mid-19th century were no longer significant and the period represented the pinnacle of industrial prosperity in Cardiff which was ended by World War I (Daunton, 1977).

The urban boundaries defined from 1:10,560 scale Ordnance Survey maps in 1886, 1901 and 1922, which were digitized to 1 mm accuracy, are displayed Figure 5.6, and overlayed in Plate 5.1. Considerable control was exercised in digitizing to ensure the same level of detail was picked up

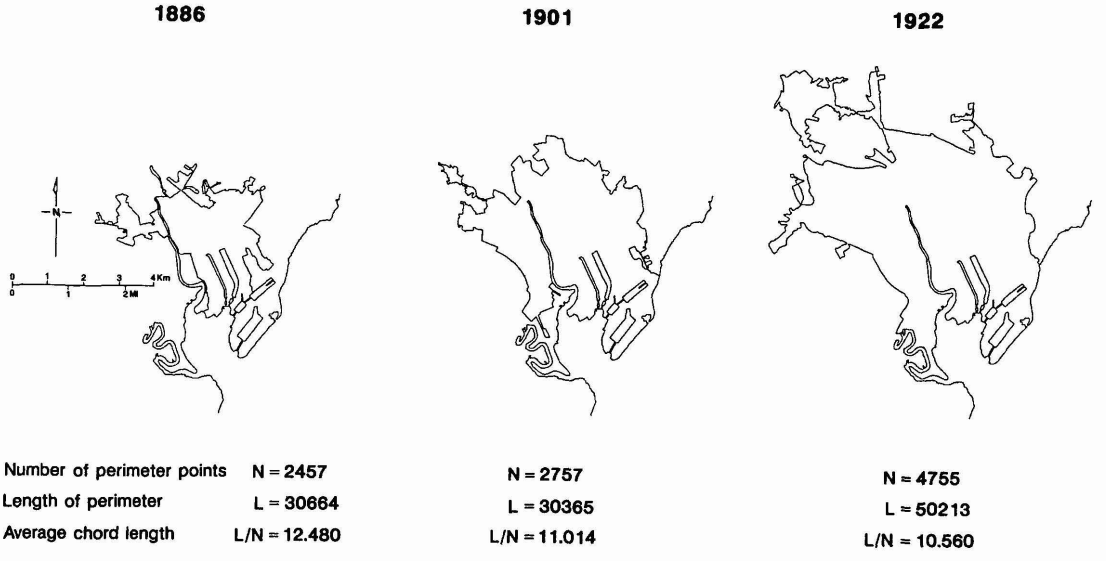

Figure 5.6. The urban edge of Cardiff in 1886, 1901, 1992.

from each map, thus minimizing the possibility that the fractal dimension becomes an artifact of the mapping process. Computing these dimensions involves the same procedure as was outlined above in the previous section: first the length of the perimeter of each boundary is calculated by simulating a traverse of the curve at different scales, and second, these perimeter estimates are related to their associated scales using a curve fitting procedure which yields the fractal dimension. The perimeter L is measured using a simulation of Richardson's method of walking a pair of dividers around the curve, the step length of the dividers r being a measure of scale (Richardson, 1961; Shelberg, Moellering and Lam, 1982). Details of the way in which the algorithm operates will be given below in the next section.

The two variants in the method described in Section 5.3 above have both been invoked. First, successive measurements are started at every digitized point on the curve and the perimeter taken as an average of each walk to remove any dependence on starting values. The method is extremely time-consuming, each pass of the method taking 65 minutes of CPU time for a curve involving 4755 digitized points (the 1922 boundary), running on a computer operating at 2 MIPS. Second, the scales used in each walk varied from a step length r_0 computed as the average of the chords linking the digitized points, to a scale which gave not less than eight chords, below which any approximation to the boundary was deemed unacceptable in accordance with the criteria developed at the end of Section 5.4. Thirty changes in scale were used and each scale was related to the lowest step length r_0 by $r_n = 2^n \nabla$, $n = 0, 1, \ldots, 30$, where ∇ is the parameter controlling the geometric scaling related to $\bar{d}$ in equation (5.11). As in Section 5.4, these scales ensure equal weighting of values in the log–log regressions based on the log–log plots of perimeter against scale. These are shown in Figure 5.7 for each boundary.

As we have seen in Chapter 3, geophysical boundaries are characterized by a wide variation in the value of D (Burrough, 1981), but for coastlines, the value of D is likely to be less than 1.3 as first shown by Richardson (1961) and confirmed many times since (Kent and Wong, 1982; Mandelbrot, 1984; Shelberg, Moellering and Lam, 1982). The slope α and intercept β are once again determined by a linear regression of $\log L$ on $\log r$ as in equation (5.8). Richardson plots describing the scaled perimeter measurements for the three time periods in Figure 5.7 form the basic data for the regressions and the results of fitting straight lines through these scatters are given in Table 5.2. The fractal dimensions D decrease as hypothesized with the largest falling in the period 1886–1901. However, both Figure 5.7 and Table 5.2 reveal that the phenomena are multifractal. It is impossible to identify clear breaks in the slopes of the plots and thus approximating the plots by several linear functions would be arbitrary. It would appear that the fractal dimension itself is a function of scale, and thus we have postulated that the scaling coefficient β is determined as

$$\beta = \lambda + \phi r. \tag{5.13}$$

Substituting (5.13) into (5.8) gives

$$\log L = \log \alpha + \lambda \log r + \phi r \log r, \tag{5.14}$$

from which it is clear that

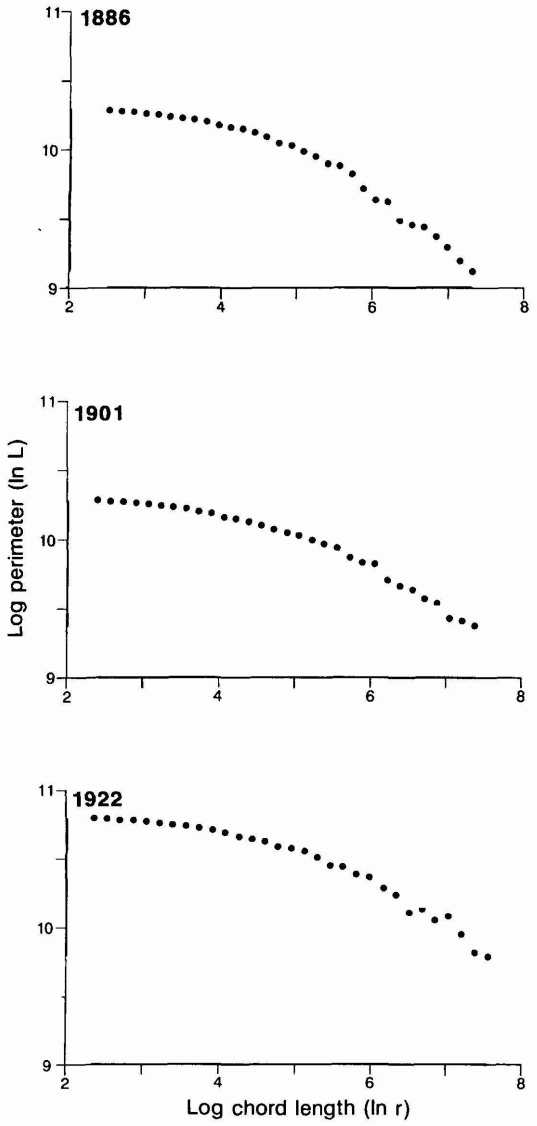

Figure 5.7. Richardson plots over many scales for Cardiff in 1886, 1901 and 1992.

Table 5.2. Scaling constants and fractal dimensions from equation (5.8)

Data set	Log α	$D = 1 - \beta$	Goodness-of-fit (r^2)
1886	11.080	1.239	0.914
1901	10.866	1.184	0.927
1922	11.393	1.185	0.907

Table 5.3. Scaling constants and fractal dimensions from equation (5.14)

Data set	Log α	$D = 1 - \lambda$ (when $r = 0$)	$\phi \times 10^{-5}$	Goodness-of-fit (r^2)
1886	10.719	1.141	5.865	0.983
1901	10.622	1.117	3.947	0.985
1922	11.114	1.109	3.901	0.984

$$D = 1 - \lambda - \phi r. \tag{5.15}$$

As the scale $r \rightarrow 0$, $D \rightarrow 1 - \lambda$. Thus the term ϕr log r in equation (5.14) acts as a dispersion factor which increases the fractal dimension as the scale increases. If $\phi = 0$, then this factor which introduces the non-linearity into the plots is redundant and equation (5.14) collapses back to equation (5.8). The model is thus consistent with increasing fractal dimension with scale.

Regressions based on equation (5.14) are shown in Table 5.3 and the performance of each model measured by r^2 dramatically improves in comparison with equation (5.8) and Table 5.2. Changes in the fractal dimensions based on equation (5.15) are plotted in Figure 5.8 from which it is quite clear that the smallest scale dimension where $r = 0$, declines over time in the manner hypothesized. The effect of scale given by ϕ also decreases over time, and in both cases, the greatest decreases in λ and ϕ occur between 1886 and 1901 when the greatest changes in transport technology – new docks and tramways – were developed. These results are consistent with

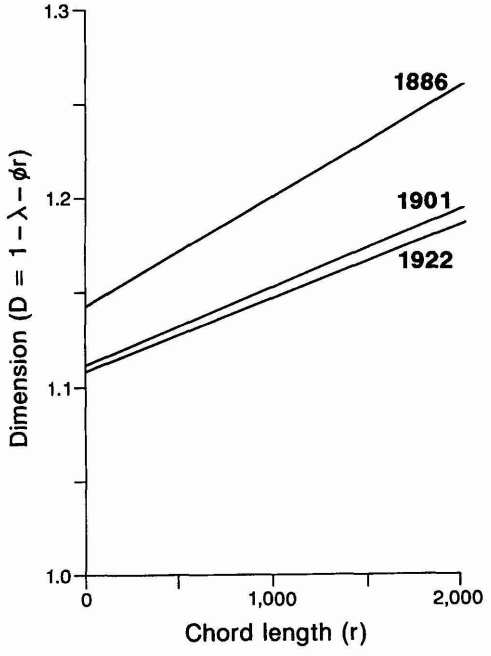

Figure 5.8. Predicted variations in fractal dimension over scale for the 1886, 1901 and 1992 data sets.

the three hypotheses originally stated, although the decrease in the irregularity of Cardiff's urban boundary between 1886 and 1922 cannot be specifically attributed to changes in any single process of development. However, the traditional image of urban growth becoming more irregular as tentacles of development occur around transport lines is not borne out by this analysis. It would appear that greater social and physical controls over development in the late 19th and early 20th century city, together with increased accessibility due to improvements in transport, have combined to gradually reduce the irregularity of urban areas such as Cardiff. These results will only apply to West European cities and similar analyses of North American and other world cities are required. It is also tempting to speculate that these results reflect the general notion of man's increasing control over environment, but such a conclusion should be avoided because there is greater variation in the dimensions produced by different methods than by different temporal data sets on the same city (Batty and Longley, 1987).

These empirical findings suggest that it is necessary to postulate fractal models based on processes which operate at different scales and which thus generate multifractal geometries. Nakano (1983) has indicated how this is possible for a coastline, and Suzuki (1984) has demonstrated how such geometries can emerge theoretically over time. These ideas involve the notion of transient self-similarity and transfer the analysis to models of varying self-similarity with respect to morphology and scale. In fact, since the mid-1980s, there has been increasing concern for the concept of multifractals and the notion that all physical objects are likely to imply a multitude of fractal dimensions has become accepted as the basic notions of what constitutes a fractal have been relaxed and broadened through empirical examples (Feder, 1988; Stanley and Ostrowsky, 1986). In this context, it may now be possible to examine detailed changes in the form of a city, developing an incremental model of urban change in which changes in shape through the boundary are associated with different processes, different degrees of irregularity, different fractal dimensions, all persisting through time, a theme we will return to in Chapters 7 and 8. For the moment, however, it is sufficient to note that fractal dimensions of urban boundaries are a function of scale. Other published data such as that pertaining to coastlines, fine particle morphologies, indeed a host of other related examples throughout the physical and natural sciences (Kaye, 1989a), should be reexamined in the light of this argument.

5.6 Fractal Measurement Methods Compared I: the Structured Walk

Thus far, we have described how fractal dimensions are calculated for 'real world' or 'empirical' irregular curves, and this has been developed using the analogy between automated computation and the manual process of obtaining scaled measurements through dividers. The process of changing the divider span with which a base curve is measured is, in fact, just one

way of adjusting the scale or resolution at which that curve is measured. In this and the following sections, we will first review the process by which the divider-based measurement algorithm works, and will assess the accuracy and computational burden associated with its use. We will then describe three other methods of measuring fractal irregularity and will discuss the relative merits of each with reference to the basic structured walk method. We will evaluate each of these methods using the examples introduced in Sections 5.4 and 5.5 above.

In his original application, Richardson (1961) manually 'walked' a pair of dividers along a mapped boundary, and obtained scale-dependent measurements by systematically increasing the divider span. Shelberg, Moellering and Lam (1982) were among the first to automate this procedure with an algorithm designed to approximate a digitized curve using a prespecified range of chord lengths. In Section 5.4, we described how the initial (base) scale length for each curve was computed by first calculating the distances between each adjacent pair of (x,y) coordinates i and $i + 1$ using equation (5.9). Successive scale changes were then incremented using a geometrical progression of chord lengths. In our later examples, we specified criteria to define the maximum and minimum chord lengths to be used in the measurement process, and interpolated 30 scale changes across the scale range bounded by these two extremes.

The walk at any given scale begins by calculating the distance $d_{s,i}$ from a starting point (x_s, y_s) to the second coordinate pair (x_i, y_i) using equation (5.9). If this distance is less than the chord length r, the next coordinate pair (x_{i+1}, y_{i+1}) is selected, the distance $d_{s,i+1}$ is computed and the test against chord length r is made again. This process continues until the distance $d_{s,i+k} > r$ and when this is achieved, a new point (x_{s+1}, y_{s+1}) is interpolated onto the line segment which joins points (x_{i+k-1}, y_{i+k-1}) and (x_{i+k}, y_{i+k}). The walk then recommences from this interpolated point and proceeds through painstaking use of trigonometry to span the curve with chords of exactly length r. As the end of the curve is approached, the distance between the last interpolated point and the end point will invariably be less than r; in this instance, a fraction of the chord length r is computed in order to close the interpolated curve. Measured perimeter lengths (L) at any scale (r) can be obtained from any starting point on the digitized base curve; if the walk begins other than at either the end points, the interpolation proceeds along the curve in both directions and the final recorded length is the sum of these computed values. As stated previously, the empirical measurements recorded in the Cardiff example comprise the average of the lengths measured from every possible starting point on the curve, repeated, of course, for every scale change.

The rudiments of this procedure are given visual expression in Figure 5.9(a) and (b) and in Plate 5.2. The displays in Plate 5.2 are of the 1949 urban edge and were produced using an interactive version of the structured walk algorithm. In this algorithm, the user specifies the initial and final chord lengths (the former as a percentage of mean chord length on the base curve, the latter as an absolute value), the starting point on the base curve, and the number of generalizations (levels) that are to be produced using the walk algorithm. Screen annotation for each level records (from the bottom line to the top) the chord length to be interpolated onto the base curve, the

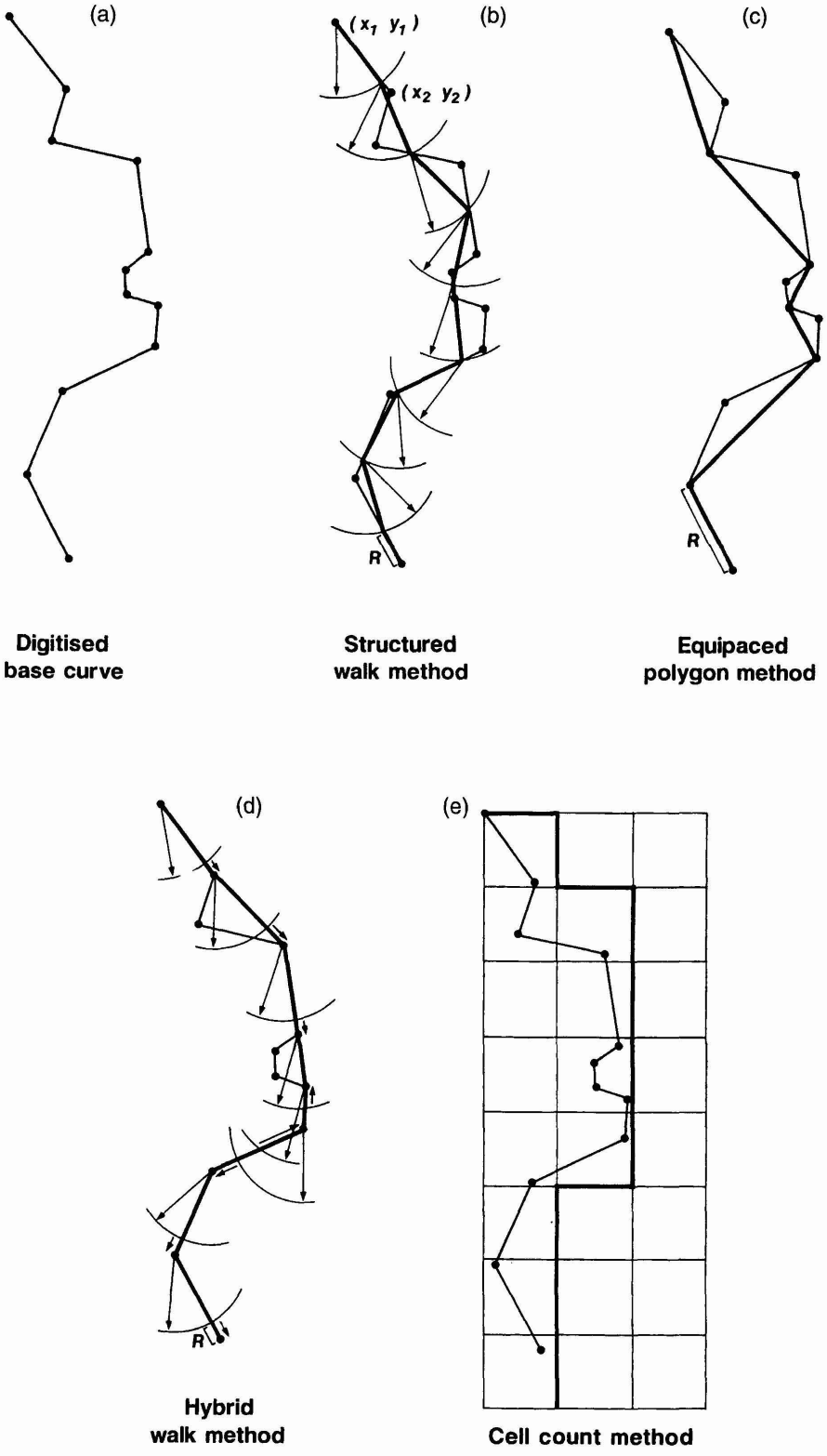

Figure 5.9. The mechanisms underlying the four measurement methods.

measured perimeter of the curve at this scale, and the number of (complete or partial) traverses that are necessary to close the curve on its end points. The sequence of displays allows the user to gain a visual appreciation of the manner in which measured perimeter lengths (and number of chords) decrease as scale increases, and could also assist a decision on the level of fractal detail most appropriate to the storage and display of a given digitized data set.

We have recorded the computer time required to make repeated measurements of the 1949 urban boundary information used in Section 5.4 as well as the 1886, 1901 and 1922 data used in Section 5.5. Scaled measurements for each of the four data series are shown together in Figure 5.10, and the range of scales common to all four analyses are highlighted here. The rate of increase in CPU time in relation to increased numbers of digitized points is shown in Figure 5.11. In order to compare the methods, we have fitted both of the functional forms (5.8) and (5.14), and the results for this basic structured walk are shown alongside the CPU times in Table 5.4. One of our earlier substantive findings was that a 'transient dimension' model (in which fractal dimension is itself a function of scale) was appropriate to the measurement of urban boundaries and these results are presented in summary form in Table 5.4, together with the computer processing times

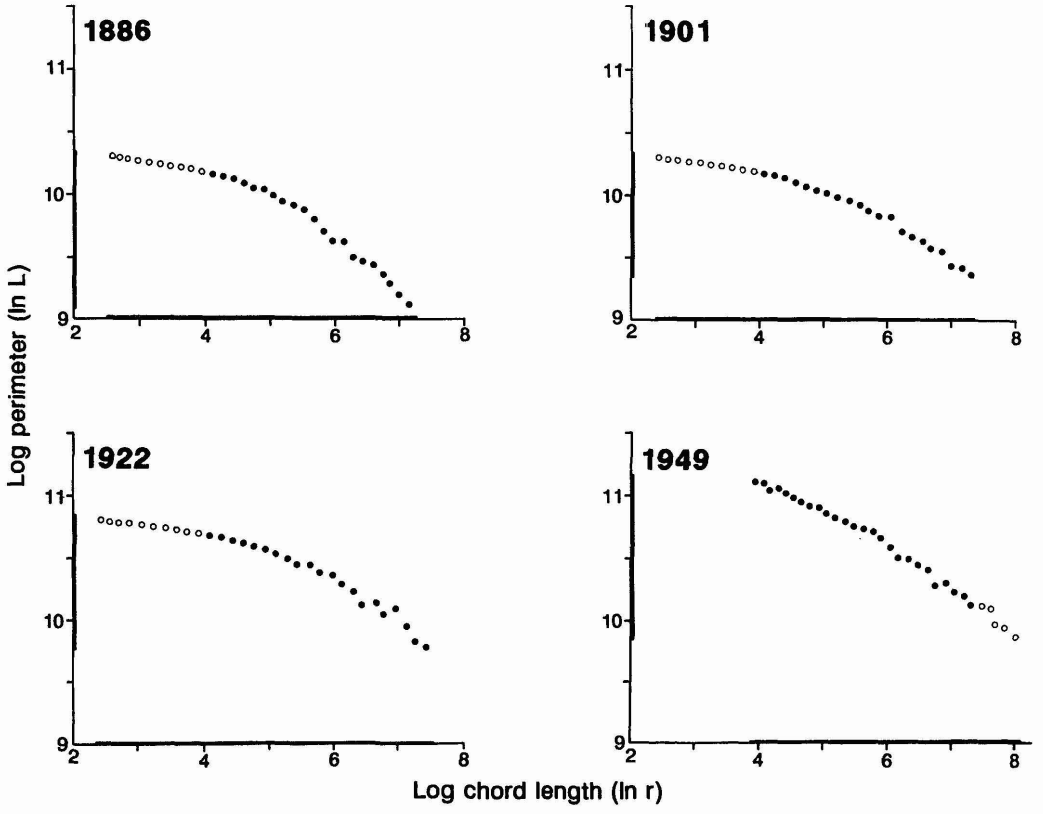

Figure 5.10. Richardson plots of perimeter–scale relations from the structured walk method. •: observation falling within scale range common to all four temporal data bases, 1886, 1901, 1922 and 1949.

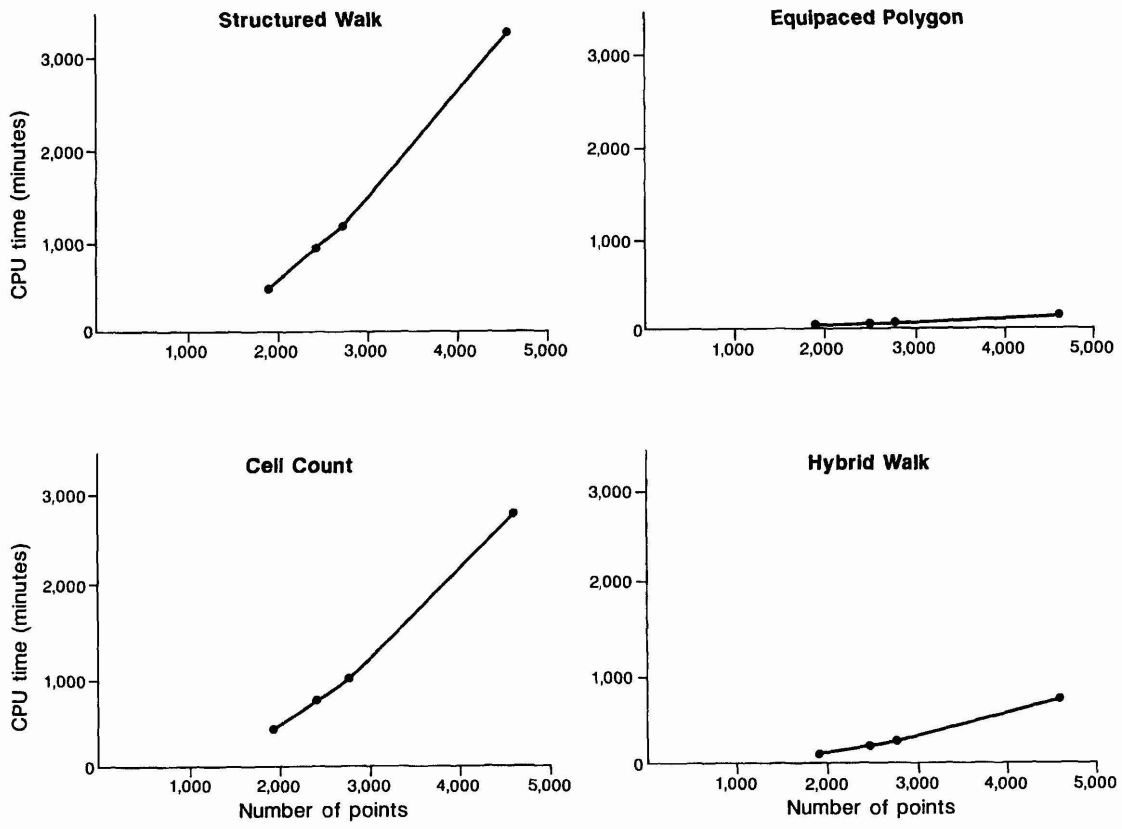

Figure 5.11. CPU usage associated with the four measurement methods.

Table 5.4. The structured walk method: computational costs and statistical performances

Data set	CPU usage Day:h:min	Log–linear form equation (5.8)			Transient dimension model equation (5.14)			
		Log α	D	r^2	Log α	$1-\lambda$	$\phi \times 10^{-5}$	r^2
1886	0:15:23	11.080	1.239	0.914	10.719	1.141	5.865	0.983
1901	0:19:11	10.886	1.184	0.927	10.622	1.117	3.947	0.985
1922	2:07:10	11.393	1.186	0.907	11.114	1.109	3.901	0.984
1949	0:07:49	12.150	1.267	0.975	11.883	1.211	1.202	0.991

associated with each of the analyses. The r^2 values show that the transient dimension model produces a consistently better statistical fit than the standard log–linear form for every one of the four time slices under analysis.

The Richardson plots shown in Figure 5.10 illustrate that the structured walk method produced estimates which correspond closely to this functional form, with the clearest continuous trend being discernible for the smaller step lengths. Although the positioning of the points does become slightly more erratic for the largest step lengths, there is no evidence of any

sudden 'flattening' of the curve, which would have indicated that the scales were too coarse to pick up further fractal detail. This cohesion of the larger scale points about the best fitting functional form is the result of the averaging of each scale observation through measurements from every single possible starting point. Finally, Figure 5.11 shows that the structured walk method is consistently associated with the highest CPU usage of the four methods to be described here. This is a consequence of the precise trigonometric interpolation of points upon the base curve. It might be conjectured that this precision obviates the need to average out the measured perimeter lengths by using every conceivable starting point, although the decay in the trend in the points at larger scale steps suggests this is not necessarily the case.

5.7 Fractal Measurement Methods Compared II: Equipaced Polygon, Hybrid Walk and Cell-Count Methods

The structured walk method provides a precise means of calculating the fractal dimensionality of vectorized boundary data. As we have seen in Chapter 2, fractal measurement and compression provides a general and powerful means of storing coordinate information. It can be used on information stored in both vectorized and rasterized formats, and its use in association with these different data structures can make alternative measurement methods more appropriate. Moreover, data processing requirements for large data sets can make computer processing time an important consideration in devising measurement algorithms. Three such alternative measurement procedures are the equipaced polygon, hybrid walk and cell-count methods. In this section, we will describe their computation and evaluate the comparative performance of each using the Cardiff urban edge data. Repeated averaging of measurements is carried out as earlier, and similar ranges of scale changes are also used.

The equipaced polygon method was first suggested by Kaye (1978, 1989a) and elaborated in Kaye and Clark (1985) as a measurement method in which there is no need to compute new base-level points. The first perimeter length for the sequence of scale changes is computed by summing the distance between adjacent coordinates; the second perimeter length represents the summed distance between every second coordinate; the third sums the distance between every fourth coordinate; and so the progression continues, weeding out all but every 8th, 16th, 32nd, ... point. This geometric point weeding series is contrived so as to give observations a more equal spacing in the Richardson plots, and hence a more equal weighting in the regression analysis. In terms of the Richardson plots and regression analysis, the chord length r which is to be paired with an associated measured perimeter length is given by the average chord length spanning the points at the corresponding level of the point weeding sequence. This is illustrated in Figure 5.9(c).

Formally then, a direction is established from a given starting point on the base-level curve (x_i, y_i) and a chord is constructed to a digitized point (x_{i+k}, y_{i+k}) which is k steps away from (x_i, y_i); k is thus an index of scale. The distance $d_{i,i+k}$ is computed using equation (5.9), and then the next chord involving the point (x_{i+2k}, y_{i+2k}) is constructed from (x_{i+k}, y_{i+k}). Eventually the endpoint of the base-level curve is approached, and the level k curve is closed on this endpoint when the remaining number of base points is less than step length k (this is equivalent to the 'remainder' as described for the structured walk). Computations in both directions from the starting point are added to determine the perimeter and mean chord lengths.

The Richardson plots from the Cardiff data which are associated with this method are shown in Figure 5.12 and the results of regression analysis in Table 5.5. At an intuitive level, one might anticipate that this method yields results of a slightly more arbitrary nature, because exact perimeter lengths will be dependent upon the evenness with which the base curve has been digitized. For example, points are unlikely to have been 'forced' on long straight sections, so these sections are unlikely to contain chord end points; moreover, the entire shape of a measured curve is likely to change if the base curve contains major irregularities or fissures (for exam-

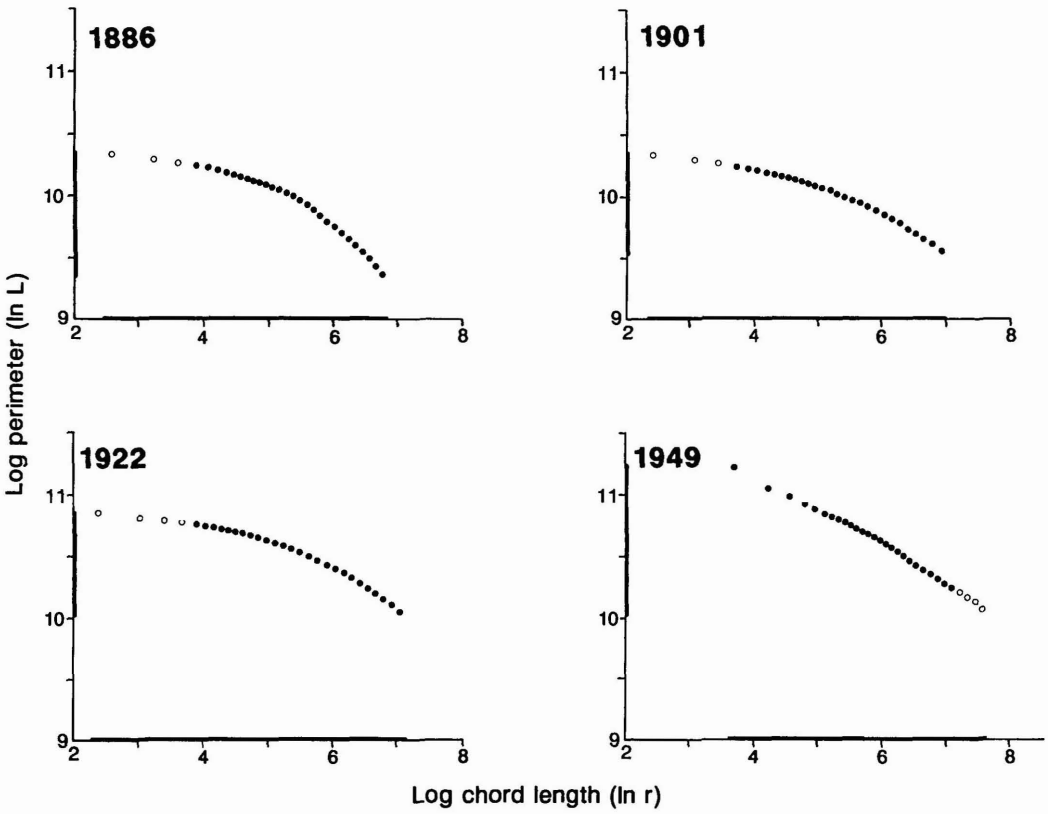

Figure 5.12. Richardson plots of perimeter–scale relations from the equipaced polygon method. •: observation falling within scale range common to all four temporal data bases, 1886, 1901, 1922 and 1949.

Table 5.5. The equipaced polygon method: computational costs and statistical performances

Data set	CPU usage Day:h:min	Log–linear form equation (5.8)			Transient dimension model equation (5.14)			
		Log α	D	r^2	Log α	$1-\lambda$	$\phi \times 10^{-5}$	r^2
1886	0:00:36	11.176	1.236	0.875	10.589	1.086	11.200	0.995
1901	0:00:45	10.923	1.178	0.917	10.594	1.094	5.920	0.994
1922	0:02:09	11.420	1.172	0.902	11.078	1.085	5.187	0.992
1949	0:00:20	12.342	1.293	0.992	12.132	1.250	1.211	0.998

ple, lines bounding suburban communities connected to the main urban area by ribbon development) which will be detected abruptly at a shift between two scale changes. The equipaced polygon method was particularly susceptible to such phenomena, since the measured curve could suddenly dislocate when the point weeding criteria missed some fissured points for the first time. Figure 5.13 shows an example of this susceptibility

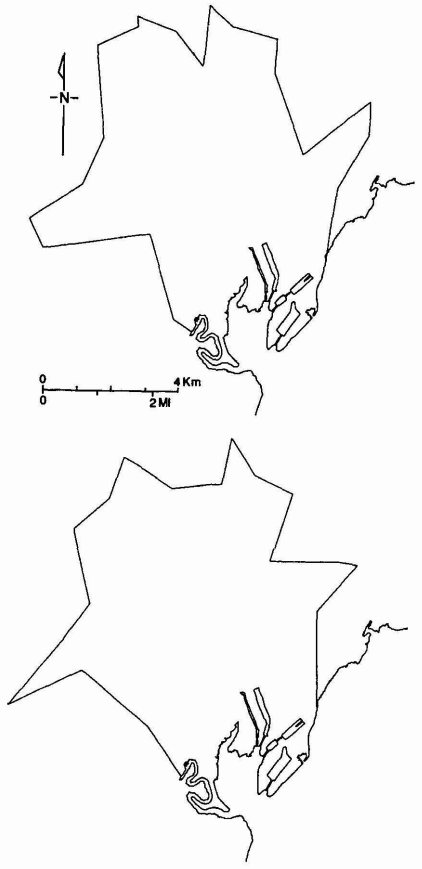

Figure 5.13. Sudden changes in approximation at two adjacent scales using the equipaced polygon method.

at two adjacent scales. In fact, the Richardson plots show that this effect is removed by the averaging process, and the points actually follow a clearer trend than the structured walk plots. The regression results compare directly with the structured walk results, both in terms of measured fractal dimensions and the statistical fits of the two competing functional form specifications. The biggest apparent difference between the two methods seems to be CPU usage as seen in comparing Tables 5.4 and 5.5, in that the equipaced polygon method used less than 5% of the resources required for the structured walk in a fully averaged run. However, intermediate polygon plots are more erratic than those for the structured walk when full averaging does not take place.

The second alternative method is the hybrid walk which was suggested by Clark (1986) as a method which retains some favorable characteristics of both the structured walk and the equipaced polygon methods. It is based directly upon the same prespecified geometric chord length series as the structured walk, which makes it less vulnerable than the equipaced polygon method to the spacing of points on the base curve. However, it is similar to the equipaced polygon method in that no new points are interpolated onto the base curve; rather, each chord is either extended or contracted to coincide with the nearest digitized point, which is then used as the origin from which the next chord is sought. Removal of the time-consuming trigonometric interpolations thus serves to speed up the computations. It is based on the same lowest level of resolution r_0 as the previous two methods and entails similar treatment of the 'remainder' distance as the end of the curve is approached. This is illustrated in Figure 5.9(d).

Formally, the method proceeds in the same way as the structured walk, except that when a point (x_{i+k}, y_{i+k}) is reached where $d_{s,i+k} > r$, no new point is interpolated using the Shelberg–Moellering–Lam algorithm. If $|d_{s,i+k} - r| \leq |d_{s,i+k-1} - r|$, then point (x_{i+k}, y_{i+k}) is selected; if not, then the point (x_{i+k-1}, y_{i+k-1}) is selected, because this point is the closest to the point at which chord length r intersects the base curve. The Richardson plots associated with this method illustrated in Figure 5.14 show a similar pattern to those of the structured walk method in Figure 5.4. The analytical results given in Table 5.6 are also comparable with the first two methods, although the method is unable to discriminate between the log–linear forms of the 1901 and 1922 series. The graphs of CPU usage in Figure 5.11 show that only comparatively modest savings are made compared to the structured walk method, and the equipaced polygon method remains the least demanding by far in this respect.

The final method that we will consider is the cell-count method. This method is more akin to a rasterized conception of the digitized base curve and has been suggested by a number of authors (Dearnley, 1985; Goodchild 1980; Morse *et al.*, 1985). In effect, the computer algorithm imposes a square lattice for a range of different spacings on the base curve. The spacings of the different lattices introduce the sequence of scale changes over which the irregularity of the base curve is to be measured. At each scale (grid spacing), the cell-count algorithm simply enumerates all of the cells that the base-level curve passes through. Counts are made at each scale change for grids originating at each point on the base curve: these are averaged to produce the final observations for each scale change in the now familiar

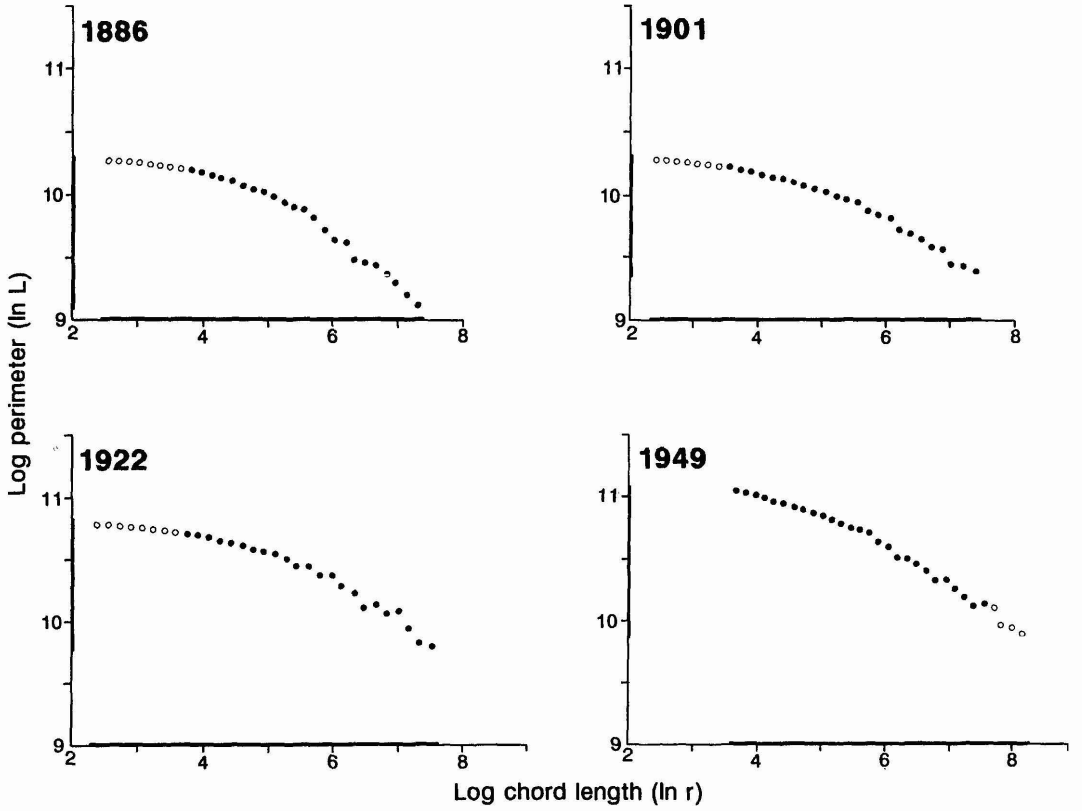

Figure 5.14. Richardson plots of perimeter–scale relations from the hybrid walk method. •: observation falling within scale range common to all four temporal data bases, 1886, 1901, 1922 and 1949.

Table 5.6. The hybrid walk method: computational costs and statistical performances

Data set	CPU usage Day:h:min	Log–linear form equation (5.8)			Transient dimension model equation (5.14)			
		Log α	D	r^2	Log α	$1-\lambda$	$\phi \times 10^{-5}$	r^2
1886	0:12:34	11.119	1.248	0.913	10.715	1.137	7.256	0.987
1901	0:15:52	10.895	1.190	0.929	10.633	1.117	4.560	0.990
1922	1:21:56	11.412	1.190	0.906	11.111	1.106	4.567	0.989
1949	0:06:52	12.416	1.308	0.989	12.197	1.262	1.001	0.996

way. Strictly speaking, each grid scale should be defined with respect to the start and endpoints on the base-level curve, although for reasons of convenience and comparability, the empirical results reported and depicted below are based on the same 31 scales used for the structured and hybrid walk methods. This is illustrated in Figure 5.9(e). This cell-counting procedure is related but not identical to box-counting, and its associated dimension, the box dimension (Voss, 1988). Falconer (1990), however,

includes all four of the techniques introduced here under the broader heading of box-counting to distinguish these from spectral methods for computing dimension which we alluded to in Chapter 3.

Formally, from a given starting point (x_s, y_s) with a selected cell size r and direction of traverse, the next coordinate (x_i, y_i) on the base curve is alighted upon. A test is made to see if this point lies within the same cell by considering whether $|x_s - x_i| \geq r$ or $|y_s - y_i| \geq r$. If either of these conditions hold, a new point is established where the coordinate in question is updated in the direction of greatest increase. Thus if $|x_s - x_i| \leq |y_s - y_i|$, $x_{s+1} = x_s + r$ and $y_{s+1} = y_s$, whilst if the converse holds, $x_{s+1} = x_s$ and $y_{s+1} = y_s + r$. If the increase along both the x and y axes is less than the grid size r, then a new coordinate point (x_{i+1}, y_{i+1}) is selected and the tests are made once again. Each time the direction is updated, a cell has been crossed and is thus counted. Unlike the previous methods, when the end point of the curve is approached, the cell approximation simply finishes when the cell in which the end point exists has been identified.

The way in which this procedure works is illustrated in Plate 5.3 where the aggregation shows that the intricate form of the line is lost at an early stage in the cell-count process. It is for this reason that the method has been advocated as a computationally inexpensive first approximation to measurement. Figure 5.11 shows that the cell-count method is closest to the equipaced polygon method in its meager CPU requirements. Although the Richardson plots shown in Figure 5.15 exhibit generally smooth trends, there is some evidence of the 'bottoming out' of the curves at the coarsest scales. This indicates that the method does not detect fractal detail at these scales, despite the averaging that has taken place. Although the choice of starting point makes little or no difference to the results when the base curve is being traversed in very small increments, Figure 5.16 shows that this is not invariably the case for large step increments which only crudely approximate the curve. Such measurements are highly sensitive to the lengths of the residual steps which are left at a fairly coarse resolution using the cell-count method. This is illustrated in Figure 5.16 where the outline initiated at coordinate 456 (Figure 5.16(a)) is very different from that initiated at coordinate 1234 (Figure 5.16(b)). Largely because of this, the fractal dimensions and statistical fits shown in Table 5.7 bear less direct comparison with the other methods than has been the case in the previous sections.

5.8 Beyond Lines to Areas

The algorithms described in this chapter have been used to investigate a wide range of physical phenomena (Burrough, 1984), but rarely has the irregularity of artificial boundaries been investigated. The preceding sections have illustrated that fractal measurement provides a plausible and flexible means of detecting the structure and character of cartographic boundaries, while our substantive example suggests that the processes which structure urban form and urban edges might be investigated with

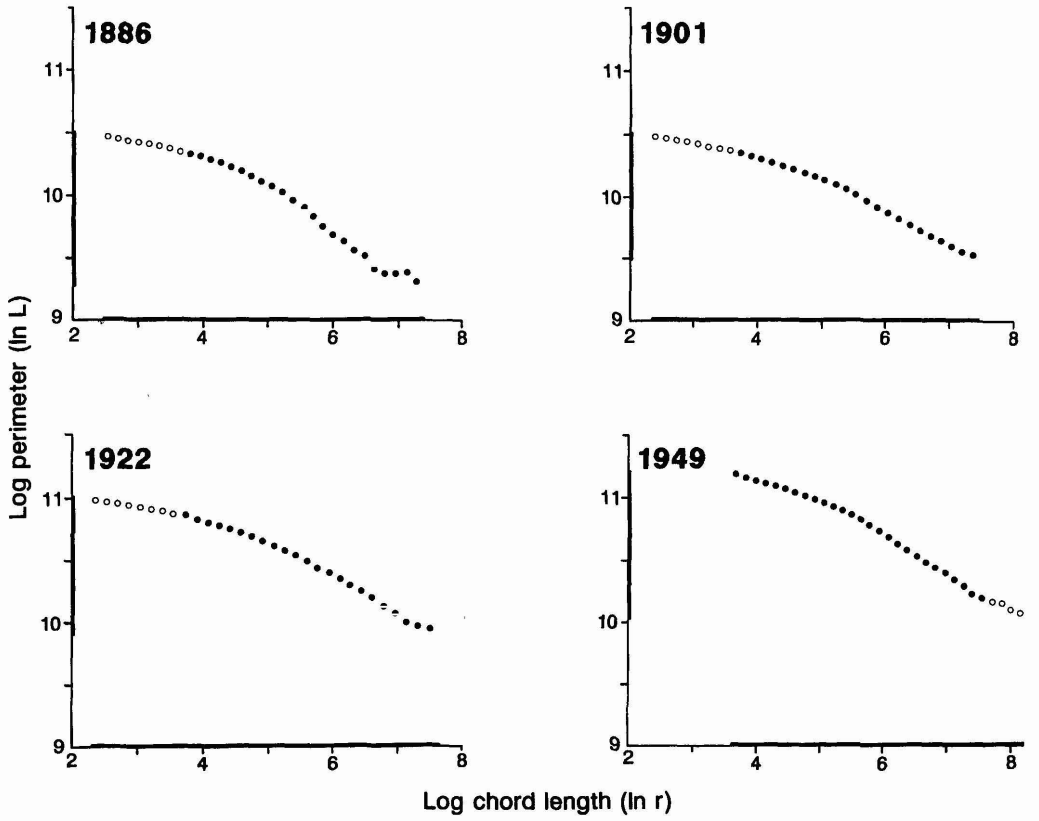

Figure 5.15. Richardson plots of perimeter–scale relations from the cell-count method. •: observation falling within scale range common to all four temporal data bases, 1886, 1901, 1922 and 1949.

respect to the manifest fractal irregularity which characterizes different cities in time and space. This latter objective might be accomplished by having regard to what urban theory suggests about the concatenation of processes, but also by recognizing different types of irregularity at different scales and over different ranges of the same phenomenon. Historical variations in fractal dimension are indeed likely, for the development of cities has been influenced by processes whose form and scale has changed over time. Whilst fractal methods can be used to generate 'semi-realistic' tessellations of the plane in order to facilitate routine spatial forecasting as we demonstrated in Chapter 4, such measurements are likely to be more useful in developing appropriate physical theory for cities. We will return to this issue of reconciling form with function in later chapters.

Describing the fractal form of cities from cartographic lines which mark their edge is perhaps the most simplistic approach we can take to linking form to function. Although it is clear that planned cities are likely to have dimensions which are integer in contrast to the organically growing cities whose irregularity gives fractional dimension, urban boundaries simply provide the envelopes for urban form and as such give little clue as to how much of the two-dimensional space is filled by the city. Envelopes do not

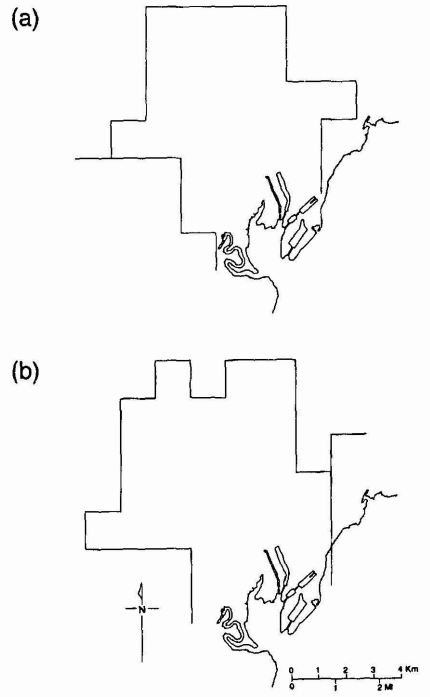

Figure 5.16. Variations in shape approximation using different starting points for the cell-count method.

Table 5.7. The cell-count method: computational costs and statistical performances

Data set	CPU usage Day:h:min	Log–linear form equation (5.8)			Transient dimension model equation (5.14)			
		Log α	D	r^2	Log α	$1-\lambda$	$\phi \times 10^{-5}$	r^2
1886	0:03:04	11.326	1.267	0.953	11.109	1.207	3.525	0.973
1901	0:03:51	11.079	1.200	0.967	10.919	1.156	2.592	0.989
1922	0:11:30	11.617	1.209	0.957	11.426	1.156	2.686	0.988
1949	0:01:37	12.288	1.274	0.985	12.144	1.244	0.646	0.990

pick up the detailed texture and irregular fabric of urban development and thus offer little by way of linking dimension to the density of development. We will in fact explore these notions from Chapter 7 onwards where we will switch our focus away from boundaries to cellular development where the focus will be upon density, occupancy and area rather than upon lines and edges.

In Chapter 10, we will return once again to questions of the urban boundary, but then we will explore the way boundary length and area are related across different sizes of city, in this way seeking to model the relation between area and perimeter and deriving fractal dimensions which pertain to different size classes of city. Here, however, we will continue to explore the fractal form of single cities, and to this end, we will examine the pattern

of land uses which occur within the urban envelope, seeking to generate the fractal dimension and properties of sets of different land use. The notion that boundaries and edges do not exist in their own right but simply serve to define space by marking off different regimes from one another, while closing some from others, is central. In the next chapter we will examine the patterns of several sets of land uses in a small English town, measuring their boundaries and computing their fractal dimensions, but with the explicit intention of exploring the extent to which we can relate boundaries to areas through the area–perimeter relations which enable the dimension of *sets of objects* rather than a *single object* to be computed. However, in providing an unambiguous link between boundaries of entire cities and their areas, we will have to wait until the last chapter before we tie together these ideas formally and empirically. In the meantime, we will disaggregate not aggregate our spatial focus, exploring land use inside the city rather than relations between cities of different sizes.