Chapter 1

Geographic Information Analysis and Spatial Data

CHAPTER OBJECTIVES

In this first chapter, we:

- Define *geographic information analysis* as it is meant in this book
- Distinguish geographic information analysis from *GIS-based spatial data manipulation* while relating the two
- Review the *entity-attribute model* of spatial data as consisting of *points*, *lines*, *areas,* and *fields*, with associated *nominal*, *ordinal*, *interval,* or *ratio* data
- Note some of the complications in this view, especially *multiple representation* at different scales, *time*, objects with *uncertain boundaries*, objects that are *fuzzy*, and objects that may be *fractal*
- Review *spatial data manipulation operations* and emphasize their importance
- Examine the various *transformations* between representations, noting their utility for geographic information analysis

After reading this chapter, you should be able to:

- List four different approaches to spatial analysis and differentiate between them
- Give reasons why modern methods of spatial analysis are not well represented in the tool kits provided by the typical GIS
- Distinguish between spatial objects and spatial fields and discuss why the vector-versus-raster debate in GIS is really about how we choose to represent these entity types

- Differentiate between point, line, and area objects and give examples of each
- List the fundamental data properties that characterize a field
- Provide examples of real-world entities that do not fit easily into this scheme
- Maintain a clear distinction between a real-world entity, its representation in a digital database, and its display on a map
- Differentiate between nominal, ordinal, interval, and ratio attribute data and give examples of each
- Give examples of at least 12 resulting types of spatial data
- List some of the basic geometrical data manipulations available in the typical GIS
- Outline methods by which the representations of entities can be transformed and explain why this is useful for geographic information analysis

1.1. INTRODUCTION

Geographic information analysis is not an established discipline. In fact, it is a rather new concept. To define what we mean by this term, it is necessary first to define a much older term—*spatial analysis*—and then to describe how we see the relationship between the two. Of course, a succinct definition of spatial analysis is not straightforward either. The term comes up in various contexts. At least four broad areas are identifiable in the literature, each using the term in different ways:

1. *Spatial data manipulation,* usually in a geographic information system (GIS), is often referred to as *spatial analysis*, particularly in GIS companies' promotional material. Your GIS manuals will give you a good sense of the scope of these techniques, as will the texts by Tomlin (1990) and Mitchell (1999).
2. *Spatial data analysis* is descriptive and exploratory. These are important first steps in all spatial analysis, and often are all that can be done with very large and complex data sets. Books by geographers such as Unwin (1982), Bailey and Gatrell (1995), and Fotheringham et al. (1999) are very much in this tradition.
3. *Spatial statistical analysis* employs statistical methods to interrogate spatial data to determine whether or not the data can be represented by a statistical model. The geography texts cited above touch on theses issues, and there are a small number of texts by statisticians interested in the analysis of spatial data, notably those by Ripley (1981, 1988), Diggle (1983), and Cressie (1991).

4. *Spatial modeling* involves constructing models to predict spatial outcomes. In human geography, models are used to predict flows of people and goods between places or to optimize the location of facilities (Wilson, 2000), whereas in environmental science, models may attempt to simulate the dynamics of natural processes (Ford, 1999). Modeling techniques are a natural extension of spatial analysis but are beyond the scope of this book.

In practice, it is often difficult to distinguish between these approaches, and most serious research will involve all four. First, data are collected, visualized, and described. Then exploratory techniques might raise questions and suggest theories about the phenomena of interest. These theories are then subjected to statistical testing using spatial statistical techniques. Theories of what is going on might then be the basis for computer models of the phenomena, and their results, in turn, may be subjected to more statistical investigation and analysis.

It is impossible to consider geographic information without considering the technology that is increasingly its home: geographical information systems (GISs). Although GISs are not ubiquitous in the way that (say) word processors are, they have infiltrated more and more businesses, government agencies, and other decision-making organizations. Even if this is the first time you've read a geography textbook, chances are that you will have already used a GIS without knowing it, perhaps when you used a website to generate a map of a holiday destination or to find driving directions to get you there.

In the above list, current GISs typically include item 1 as standard (since a GIS without these functions would be just a plain old IS!) and have some simple data analysis capabilities, especially exploratory analysis using maps (item 2). GISs have recently begun to incorporate some of the statistical methods of item 3 and only rarely include the capability to build spatial models and determine their likely outcomes (item 4). In fact, it can be hard to extend GIS to perform such analysis, which is why many geographic information analysts use other software environments for work that would be classified as belonging to items 3 and 4. In this book, we focus mostly on items 2 and 3. In practice, you will find that, in spite of rapid advances in the available tools, statistical testing of spatial data remains relatively rare. Statistical methods are well worked out and understood for some types of data but less so for many others. As this book unfolds, you should begin to understand why this is so.

If spatial analysis is so necessary— even worth writing a book about—then why isn't it a standard part of the GIS toolkit? We suggest a number of reasons, among them the following:

- *The GIS view of spatial data and that of spatial analysis are different*. The spatial analysis view of spatial data is more concerned with *processes* and *patterns* than it is with database management and manipulation, whereas the basic requirement for a *spatial database* is far more important to most large GIS buyers (government agencies, utilities) than the ability to perform complex and (sometimes) obscure spatial analysis.
- *Spatial analysis is not widely understood*. Spatial analysis is not obvious or especially easy, although we aim to address that issue in this book. The apparent difficulty means that it is difficult to convince software vendors to include spatial analysis tools as standard products. Spatial analysis tools are a possible addition to GIS that is frequently left out. This rationale has become less significant in recent years as software engineering methods enable GIS vendors to supply "extensions" that can be sold separately to those users who want them. At the same time, third-party vendors can supply add-on components more easily than previously, and open source software has become an increasingly important alternative in some quarters.
- *The spatial analysis perspective can sometimes obscure the advantages of GIS*. By applying spatial analysis techniques, we often raise awkward questions: "It looks like there's a pattern, but is it significant? Maybe not." This is a hard capability to sell!

Despite this focus, don't underestimate the importance of the *spatial data manipulation* functions provided by GIS such as buffering, point-in-polygon queries, and so on. These are essential precursors to generating questions and formulating hypotheses. To reinforce their importance, we review these topics in Section 1.5 and consider how they might benefit from a more statistical approach. More generally, the way spatial data are stored—or *how geographical phenomena are represented in GIS*—is becoming increasingly important for analysis. We therefore spend some time on this issue in Sections 1.2 and 1.3.

For all of these reasons, we use the broader term *geographic information analysis* for the material we cover. A working definition of this term is that it is concerned with investigating the *patterns* that arise as a result of *process*es that may be operating in space. Techniques and methods to enable the representation, description, measurement, comparison, and generation of spatial patterns are central to the study of geographic information analysis. Of course, at this point our definition isn't very useful, since it raises the question of what we mean by *pattern* and *process*. For now, we will accept whatever intuitive notion you have about the meaning of the key terms. As we work through the concepts of point pattern analysis in Chapters 4 and 5, it will become clearer what is meant by both terms. For now, we will concentrate on the general spatial data types you can expect to encounter.

1.2. SPATIAL DATA TYPES

Thought Exercise: Representation

Throughout this book, you will find thought exercises to help you follow the text in a more hands-on way. Usually, we ask you to do something and use the results to draw some conclusions. You should find that these exercises help you remember what we've said. This first exercise is concerned with how we represent geography in a digital computer:

1. Assume that you are working for a road maintenance agency. Your responsibilities extend to the roads over a county-sized area. Your GIS is required to support operations such as surface renewal, avoiding clashes with other agencies—utility companies, for example—that also dig holes in the roads and make improvements to the road structure.

   Think about and write down how you would record the geometry of the network of roads in your database. What road attributes would you collect?
2. Imagine that you are working for a bus company in the same area. Now the GIS must support operations such as time-tabling, predicting the demand for existing and potential new bus routes, and optimizing where stops are placed.

   How would the recording of the geometry of the road network and its attributes differ from your suggestions in step 1 above?

What simple conclusion can we draw from this? It should be clear that how we represent the same geographic entities differs according to the purpose of the representation. This is obvious, but it can easily be forgotten.

Quite apart from the technical issues involved, social critiques of geographic information analysis often hinge on the fact that analysis frequently confines itself to those aspects of the world that can be easily represented digitally (see Fisher and Unwin, 2005).

When you think of the world in map form, how do you view it? In the early GIS literature, a distinction was often made between two kinds of system characterized by how the geography is represented digitally:

1. One type of system provides a *vector* view, which records locational (x, y) coordinates of the features that make up a map. In the vector view, we list features and represent each as a point, line, or area *object*. Vector GIS originated in the use of computers to draw maps based on digital data and were particularly valued when computer memory was an expensive commodity. Although the fit is inexact, the vector model is closest to an *object* view of the world, where space is thought of as an empty container occupied by different sorts of objects.
2. Contrasted with vector systems are *raster* systems. Instead of starting with objects on the ground, a grid of small units, called *pixels*, of the Earth's surface is defined. For each pixel, the value, or presence or absence of something of interest, is then recorded. Thus, we divide a map into a set of identical, discrete elements and list the contents of each. Because every location in space has a value (even if it is zero or null), a raster approach generally uses more computer memory than a vector one. Raster GIS originated mostly in image processing, where data from remote sensing platforms are often encountered.

In this section, we hope to convince you that at a higher level of abstraction the vector/raster distinction isn't very useful, and that it obscures a more important division between what we call an *object* and a *field* view of the world.

The Object View

In the object view, we consider the world as a series of *entities* located in space. Entities are (usually) real: you can touch them, stand in them, perhaps even move them around. An *object* is a digital representation of all or part of an entity. Objects may be classified into different object types—for example, *point objects*, *line objects,* and *area objects*—and in specific applications, these types are *instantiated* by specific objects. For example, in an environmental GIS, woods and fields might be instances of area objects. In the object view of the world, places can be occupied by any number of objects. A house can exist in a census tract, which may also contain lampposts, bus stops, road segments, parks, and so on.

Because it is also possible to associate *behavior* with objects, the object view has advantages when well-defined objects change over time—for example, the changing data for a census area object over a series of population censuses. Note that we have said nothing about *object orientation* in the computer science sense. Worboys et al. (1990) give a straightforward description of this concept as it relates to spatial data.

The Field View

In the *field* view, the world consists of properties continuously varying across space. An example is the surface of the Earth itself, where the field variable is elevation above sea level. Similarly, we can code the ground in a grid cell as either having a house on it or not. The result is also a field, in this case of binary numbers where 1 = "house" and 0 = "no house". If a single house is large enough or if its outline crosses a grid cell boundary, it may be recorded as being in more than one grid cell. The key ideas here are spatial *continuity* and *self-definition*. In a field, every location has a value (including "not here" or zero) and sets of values taken together define the field. This is in contrast with the object view, in which it is necessary to attach further attributes to represent an object fully—a rectangle is just a rectangle until we attach descriptive attributes to it.

The raster data model is one way to record a field. In this model, the geographic variation of the field is represented by identical, regularly shaped pixels. Earth's surface is often recorded as a regular grid of height values (a *digital elevation matrix*). An alternative is to use area objects in the form of a mesh of nonoverlapping triangles (called a *triangulated irregular network* or TIN) to represent the same field variable. In a TIN, each triangle vertex is assigned the value of the field at that location. In the early days of GIS, especially in cartographic applications, values of the field given by land height were often recorded using digital representations of the contours familiar from topographic maps. This is a representation of a field using overlapping area objects, the areas being the parts of the landscape enclosed within each contour. The important point is that a field can be coded digitally using either a raster or a vector approach.

Finally, another type of field is one made up of a continuous cover of assignments for a *categorical variable*. Every location has a value, but the "values" are the names given to phenomena. Consider a map of soil type. Every location has a soil, so we have spatial continuity, and we also have self-definition by the soil type involved, so this is a field view. Other examples might be land use maps, even a simple map of areas suitable or unsuitable for some development. In the literature, these types of field variable have gone under a number of different names, among them *k-color maps* and *binary maps*. A term that is gaining ground is *categorical coverage*, indicating a field made up of a categorical variable. The important point is that such categorical coverage can be coded digitally using either a vector or a raster approach.

Choosing the Representation to Be Used

In practice, it is useful to think about the elements of reality modeled in a GIS database as having two types of existence and to keep *both* distinct from the way the same entities might be displayed on a map. First, there is the element in reality, which we call the *entity*. Second, there is the element as it is *represented* in the database. In database theory, this is called the *object* (confusingly, this means that a field is a type of object). Clearly, what we see as entities in the real world depends on the application, but to make much sense, an entity must be

- *Identifiable*. If you can't see it, then you can't record it.
- *Relevant*. It must be of interest.
- *Describable*. It must have attributes or characteristics that we can record.

Formally, an *entity* is defined as a phenomenon of interest in reality that is not further subdivided into phenomena of *the same kind*. For example, a road *network* could be considered an entity and subdivided into component parts called *roads*. These might be further subdivided, but these parts would not be called roads. Instead, they might be considered *road segments* or something similar. Similarly, a *forest* entity could be subdivided into smaller areas called *stands*, which are in turn made up of individual *trees*.

The relationship between the object and field representations is a deep one, which, it can be argued, goes back to philosophical debates in ancient Greece about the nature of reality: a continuously varying field of phenomena or an empty container full of distinct objects? You should now be able to see that the key question, from the present perspective, is not which picture of reality is *correct* but which we choose to adopt for the task at hand. A GIS-equipped corporation concerned with the management of facilities such as individual buildings, roads, or other infrastructure would almost certainly consider an object view most appropriate. In contrast, developers of a system for the analysis of hazards in the environment may adopt a field view. Most theory in environmental science tends to take this approach, using, for example, fields of temperature, wind speed, height, and so on. Similarly, data from remote sensing platforms are collected as continuous rasters, so a field view is the more obvious approach.

It is also a good idea to make a clear distinction not only between the entity and its representation in a GIS database, but also between both of these and the way the same entity is displayed on a map. Representing the content of a map is not the same as representing the world. The objectives of map design are visual—to show map users something about the real world—whereas the objectives of a database are concerned with management, measurement, analysis, and modeling. It pays to keep these objectives distinct when choosing how to represent the world in digital form.

Types of Spatial Object

The digital representation of different entities requires the selection of appropriate spatial object types, and there have been a number of attempts to define general spatial object types. A common approach—reinvented many times—is based on the spatial *dimensionality* of the object concerned. Think about how many *types* of object you can draw. You can mark a *point*, an object with no length, which may be considered to have a spatial dimension or length, L, raised to the power zero, hence L^0. You can draw a line, an object having the same spatial dimension as any simple length, that is, L^1. You can also shade an area, which is an object with spatial dimension length squared, or L^2. Finally, you can use standard cartographic or artistic conventions to represent a volume, which has spatial dimension length cubed, or L^3. The U. S. *National Standard for Digital Cartographic Databases* (DCDSTF, 1988) and Worboy's generic model for planar spatial objects (Worboys, 1992, 1995) both define a comprehensive typology of spatial objects in terms similar to these.

An Exercise: Objects and Fields Decoded

1. Obtain a topographic map at a scale of 1:50,000 or larger of your home area. Study the map, and for at least 10 of the types of entity the map represents—remember that the map is already a representation—list whether they would best be coded as an object or a field. If the entity is to be represented as an object, state whether it is a point, line, or area.
2. If you were asked to produce an initial specification for a data model that would enable a mapping agency to ''play back'' this map from a digital version held in a database, how many specific instances of objects (of all kinds) and fields would you need to record?

Hint: Use the map key. There is, of course, no single correct answer to this question.

1.3. SOME COMPLICATIONS

The view of the world we have presented so far is deceptively simple, and deliberately so. There are a number of complications, which we now examine. In each case, our perspective is that of a geographic information analyst, and the key question to be asked is the extent to which the complication impacts on any analytical results obtained.

Objects Are Not Always What They Appear to Be

Students often confuse the various cartographic conventional representations with the fundamental nature of objects and fields. For example, on a map, a cartographic line may be used to mark the edge of an area, but the entity is still an area object. Real line objects represent linear entities such as railways, roads, and rivers. On topographic maps, it is common to represent the continuous field variable of height above sea level using the lines we call *contours*; yet, as we have discussed, fields can be represented on maps in many different ways.

Objects Are Usually Multidimensional

Very frequently, spatial objects have more than the single dimension of variability that defines them. We might, for example, locate a point object by its (x, y) coordinates in two spatial dimensions, but in many applications it would be much better to record it in three spatial dimensions (x, y, z), with depth or height as a third dimension. A volume of rock studied by a geologist exists at some depth at a location but also has attributes such as it porosity or color; the interest will be in how this attribute varies in (X, Y, Z) space. Many GISs do not cope easily with such data, so frequently it is necessary to record the additional coordinate as another attribute of an object fixed at a location in (x, y). This can make perfectly natural queries and analyses that require the full three spatial dimensions awkward or even impossible. Raper (2000) provides numerous illustrations of the multidimensional nature of geographic objects.

Objects Don't Move or Change

The view of the world presented so far is a static one, with no concept of time except possibly as an attribute of objects. This is fine for some problems, but in many applications our major interest is in how things change over time. Standard GISs do not easily handle location in time as well as location in space. A moment's thought will reveal the problems that incorporating an object's location in time might generate. The idea of change over time, what we call *process*, is of particular importance to most sciences, yet handling it in a digital environment that does not readily incorporate it in any object's description will always be difficult. The problem has been discussed for many years (see Langran, 1992, and O'Sullivan, 2005, for a review of progress), but as far as we are aware, no commercial *temporal GIS* has yet been produced. In research there have been many attempts, such as *PC-Raster*™ (see Wesseling et al., 1996), to create one, almost all of which involve the definition of some generic language for describing change in both space and time.

Objects Don't Have Simple Geometries

Some aspects of geographic reality that we might want to capture are not well represented in either the raster/vector or object/field views. The obvious example here is a transport or river *network*. Often, a network is modeled as a set of line objects (routes) and a set of point objects (transport nodes), but this data structure is very awkward in many applications, and it is difficult to get the representation just right (think of one-way streets, restricted turns, lanes, and so on). Another example is becoming increasingly important: *image* data. An image in a GIS might be a scanned map used as a backdrop or it might be a photograph encoded in a standard format. At the nuts and bolts level, images are coded using a raster approach, but the key to understanding them is that, other than being able to locate a cursor on an image, the values of the attributes themselves are not readily extracted, nor, for that matter, are the individual pixel values important: it is the image *as a whole* that matters. In the most recent revision of the *ArcGIS,* some of these complexities are recognized by having *five* representations of geography, called *locations*, *features* (made up of locations), *surfaces* (fields), *images,* and *networks* (see Zeiler, 1999). As geographic information analysis becomes increasingly sophisticated and is extended to embrace applications that hadn't even been considered when the basic framework we adopt was developed, this issue is likely to be of greater importance.

Objects Depend on the Scale of Analysis

Different object *types* may represent the same real-world phenomenon at different scales. For example, on his daily journey to work, one of us used to arrive in London by rail at an entity called Euston Station. At one scale this is best represented by a dot on a map, which in turn is an instance of a point object that can be represented digitally by its (x, y) co-ordinates. Zoom in a little, and Euston Station becomes an area object, best represented digitally as a closed string of (x, y) coordinates defining a polygon. Zooming in closer still, we see a network of railway lines (a set of line objects) together with some buildings (area objects), all of which would be represented by an even more complex data description. Clearly, the same entity may be represented in several ways. This is an example of the *multiple representation* problem in geographic information analysis. Its main consequence is to make it imperative that in designing a geographic information database and populating it with objects of interest, it is vital that the type of representation chosen will allow the intended analyses to be carried out. As the next exercise illustrates, this is also true for any maps produced from the same database.

Scale and Object Type

We can illustrate this idea using a convenient example from Great Britain accessed via your Web browser. The same exercise can easily be done using paper maps or Web-delivered map extracts from another national mapping agency.

1. Go to www.ordnancesurvey.co.uk. This will bring you to a screen with an option to "Get-a-map." At the window, enter "Maidwell" (without quotation marks) which is the name of a small village in the English Midlands and hit GO.
2. You arrive at a screen with an extract from the 1:50,000 topographic map of the area around this village.
3. To the left of the map are some balloons labeled "+" and "−". If you run the mouse over them, you will see that each balloon corresponds to a map of the area at scales of 1:25,000 (zoom level 5), 1:50,000 (zoom level 4), and 1:250,000 (zoom level 3 in two versions, a "Miniscale" and a "Simplified Miniscale").
4. The exercise is simple. Make a table in which columns represent each of the five mapping scales and the rows are entities of interest—we suggest "roads", "houses", "public house" (there is a rather good one), "rivers," and "land-height"; enter a code into each cell of this table to indicate how the feature is represented. It will help if you use codes such as P (point feature), L (line feature), A (area object), F (field), and X for features that are absent at that scale.
5. What does this tell you about multiple representation?

CEOs of national mapping agencies know to their cost that it is almost impossible to produce maps from a single digital database at scales other than that for which the original database design was intended.

Objects Might Have Fractal Dimension

A further complication is that some entities are *fractals*, having the same or similar level of detail no matter how closely we zoom in. Fractals are difficult to represent digitally unless we accept that the representation is only a snapshot at a particular resolution. The classic example is a linear feature such as a coastline whose "crinkliness" remains the same no matter how closely we examine it. No matter how accurately we record the spatial coordinates, it is impossible to capture all the detail. A rather unexpected consequence is that when dealing with irregular lines, their length appears to increase the more "accurately" we measure it!

Imagine measuring the coastline of (say) a part of the North Island of New Zealand using a pair of dividers set to 10 km, as in the top left panel of Figure 1.1. We will call the dividers' separation the *yardstick*. With a yardstick

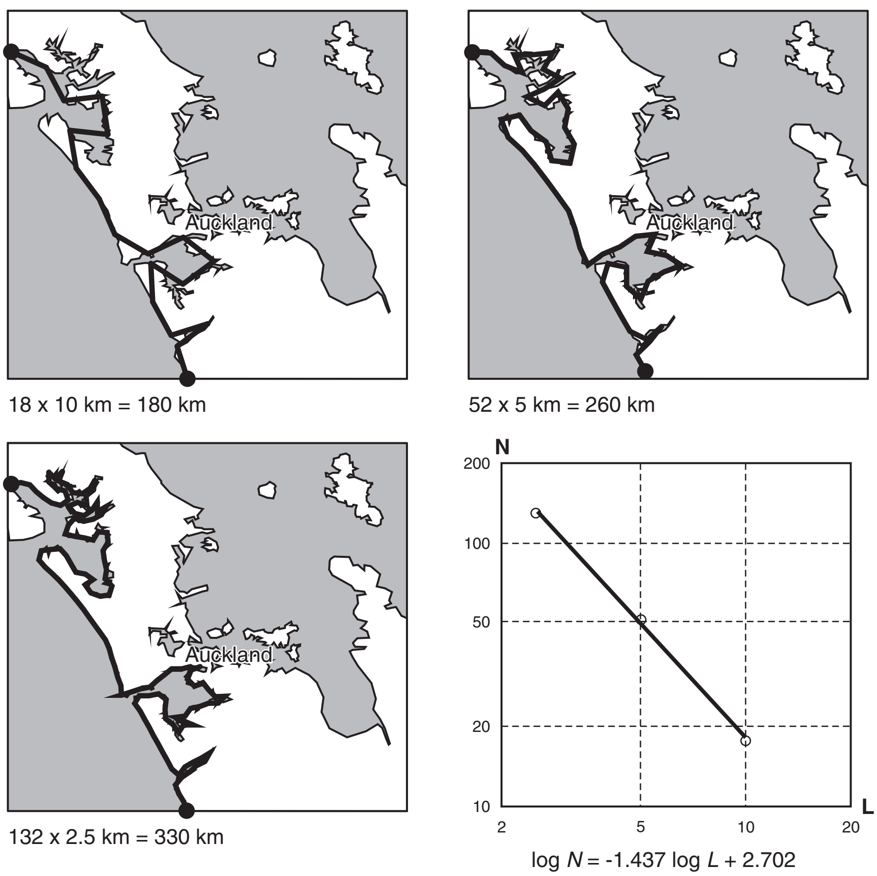

Figure 1.1 Determining the fractal dimension of part of the New Zealand coastline.

of 10 km, we count 18 segments and get a coastline length of 18×10 km = 180 km. Repeating the process with a 5 km yardstick, as in the top right panel, we count 52 segments and get a length of $52 \times 5 = 260$ km. Finally, using a yardstick of 2.5 km, we obtain a total length of $132 \times 2.5 = 330$ km. The coastline appears longer the closer we look. What would happen if we had a 1-km or a 100-m yardstick? What about 100 mm? What is the "real" length of this coastline?

Fractal dimension is a mathematical idea for dealing with this difficulty. Although it was popularized by Mandelbrot (1977), the idea that the length of lines varies with the scale of their representation was spotted long before the fractal concept was widespread in mathematics (Richardson, 1961). *Fractal* is a compression of *fraction* and *dimensional* and expresses the idea that a line may be somewhere *between* one- and two- dimensional, with a fractal dimension of, say, 1.2 or 1.5. One understanding of an object's dimension is that it expresses how its apparent size (in this case length), measured by counting smaller elements (in this case yardsticks), changes as the linear size (or *resolution*) of the smaller elements changes. A simple nonfractal one- or two- dimensional entity's size, as measured by counting subelement yardsticks, increases with the power of its dimension, so that the number of shorter segments in a simple straight line doubles if we halve the yardstick size. The count of small cube yardsticks in a large volume increases eightfold if we halve the dimension of the small cubes. If the linear dimensions of the yardsticks for two measurements are L_1 and L_2, and the respective counts of yardsticks are N_1 and N_2, then the fractal dimension D is given by

$$D = \frac{\log(N_1/N_2)}{\log(L_1/L_2)} \tag{1.1}$$

This definition does not restrict D to whole number values, and it can be used to estimate the fractal dimension of irregularly shaped entities. The lower right panel of Figure 1.1 shows each yardstick length and yardstick–count combination on a log-log *Richardson plot*. The three points lie roughly on a straight line, whose negative slope, fitted by simple linear regression, gives the fractal dimension. In the example, we arrive at an estimate for the fractal dimension of 1.44. In fact, we can properly make this measurement only on the coastline itself, because any stored representation, such as the map we started with, has a limiting resolution at which the length of the line will become fixed as we set our dividers smaller. However, the yardstick length–count relationship is often stable over several orders of magnitude, and we can estimate the fractal dimension of an entity from a large-scale object representation.

Some More Work: Do It for Yourself

As so often is the case, the best way to understand this is to do it yourself.

1. Find a reasonably detailed topographic map at a scale of about 1:25,000 or 1:50,000 and select a length of river as your object of study. Obviously, since we are interested in the sinuosity of linear objects, it makes sense to choose a river that shows meandering behavior. The equivalent of a 20-km length is about right and will not involve too much work.
2. Now set a pair of dividers at a large equivalent distance on the ground, say 1 km, and ''walk'' them along the river counting the number of steps. Record the yardstick length and the number of segments.
3. Repeat using a halved yardstick, equivalent to 500 m on the ground.
4. Repeat the process again and again until the yardstick becomes so short that the experiment is impractical. You should now have a table of values.
5. Convert both the number of steps and the yardstick length to their logarithms and plot the resulting numbers with log(number of steps) on the vertical axis and log(yardstick length) on the horizontal axis. If you have access to a spreadsheet, you should be able to do this easily. The points should fall roughly along a straight line, although success isn't guaranteed.
6. Finally, use the spreadsheet, (or a straight edge and a good eye) to fit a best-fit line to your data and estimate the fractal dimension of your river.

So, what does this all *mean*? The simplest interpretation of the fractal dimension of a line is as a measure of its “wiggliness” or “crinkliness.” The fractal dimension of an entity expresses the extent to which it “fills” up the next dimension from its *topological dimension*. A line has a single topological dimension of length L^1. A line with fractal dimension 1.0 (note the decimal point) is an idealized line and takes up no space in two dimensions. However, a line with fractal dimension 1.1 or 1.15 begins to “fill up” the two dimensions of the plane in which it is drawn. Many linear features in geography have a fractal dimension somewhere between about 1.1 and 1.5.

Variants of the Richardson plot can be used to estimate the fractal dimension of area and volume objects (or surfaces) by counting the number of elements they contain at different linear resolutions. A considerable amount of work has been done on measuring the fractal dimension of the developed area of cities (Batty and Longley, 1994). While a perfectly smooth surface has fractal dimension 2.0, a rough one might have fractal dimension 2.3. Often, the fractal dimension of surface topography can be related to other characteristics. Other examples of fractal dimension in geographic phenomena are provided by soil pH profiles and river networks (see Burrough, 1981; Goodchild and Mark, 1987; Lam and De Cola, 1993; Turcotte, 1997).

The fractal concept is strongly related to the notion of scale. Some researchers have tried to make use of this in cartographic generalization, with mixed results. It should be clear that fractal dimension is indicative of how measures of an object will change with scale and generalization, and, in general, scale-varying properties of phenomena can be related to their fractal dimension (Goodchild, 1980). More recently, it has been suggested that measuring the fractal dimension of digitized data can help to determine the scale of the source map line work from which it is derived (Duckham and Drummond, 2000).

Objects Can Be Fuzzy and/or Have Indeterminate Boundaries

The preceding discussion has assumed that the objects we deal with are what is technically called *crisp*, and that if they have a spatial extent, their boundaries can in principle be recognized exactly. Many spatial entities that we might want to describe and analyze aren't crisp, and some may also have uncertain boundaries. The archetypal example is soil. On a map, the soil type will be represented by nonoverlapping area polygons, with hard and fast lines separating the various soil types that the surveyor recognizes. This is an example of a k-color map discussed in Section 3.7, but as any soil surveyor knows, it is really a fiction—for two possible reasons.

First, although the soil can change very abruptly, such that a line can be drawn on a map to separate different soil types, soil types may also grade almost imperceptibly from one to another, such that there is no certain boundary between them. Honest soil surveyors often recognize the uncertain nature of such a boundary by marking the transition with a dotted line. In a GIS and in subsequent spatial analysis, this uncertainty is often simply erased by assuming that such lines are in fact certain. The same issue arises, for example, in geology, where rock types can change imperceptibly; in marketing, where the boundaries of some trade area might be uncertain; and when describing mental maps. Every Londoner knows that a part of the city is called "Soho," but this has no legislative basis and most people would have difficulty deciding where it begins and ends. In the same city, "Westminster" has a legislative basis and so, in principle, has a certain boundary, but we doubt that many Londoners would know this and instead think of it in much the same uncertain way that they would think of Soho. To add to the complexity, and as with Soho and some soil types, some parts of an object's boundary might be uncertain but other parts of the same boundary might be certain. One possible way to handle this boundary uncertainty is to assign some probability of membership of the defined type to each location, so that instead of saying "Here we are on soil type such and such," our maps and data would say "Here there is a probability of (say) 0.7 that we are on soil type such and such."

Second, and again best illustrated using soils, objects might be *fuzzy*. In saying that a given soil belongs to a specific soil type, we are asserting that this type (or *set*) is itself *crisp*, by which we mean that we can unambiguously state whether or not the soil really is of that type. Yet, some sets might defy such assignment, so all we can say is that the given soil is more or less of a given sort, replacing our certainty with a value that expresses the extent to which it might belong to the given type. This isn't the same as the boundary uncertainty discussed above, where we are certain about the types but uncertain about the given soil. Here we are uncertain about the type of soil itself but, in a sense, certain about the soil belonging to that uncertain type. Practically, the extent to which any given entity is a member of such a fuzzy set is often recorded using a "membership" value ranging from 0 to 1. This can give rise to confusion with the probabilities associated with uncertain boundaries. Thinking through the exercise below might help you distinguish the two sorts of uncertainty.

Thought Exercise: Certainty and Uncertainty

Consider the following hierarchy of statements:

John is over 1.8 m tall.

This is a certain statement about John being a member of the crisp set of all people over 1.8 m tall.

I think John is over 1.8 m tall.

This set is still crisp, but we aren't sure about John's membership of it and might assign a probability to our uncertainty. This is analogous to the uncertain boundary issue.

(*continues*)

(*box continued*)

John is tall

''Tall'' is a fuzzy set, but we are certain that John belongs to it. It is the fuzzy category ''tall'' that now encapsulates the uncertainty. It is also possible that John is in the fuzzy set ''really tall'' or ''of average height'' (although probably not in all three of these fuzzy sets, since they now cover quite a range of circumstances). For any particular height, we might assign a membership value for these sets that records how ''tall,'' ''really tall,'' or ''of average height'' it is. If John's actual height is 1.9 m, then his membership in the set ''tall'' might be 1.0, ''really tall'' 0.6, and for ''of average height'' 0.05. This is the fuzziness issue.

I think John is really tall.

This combines the two types of uncertainty into one statement about John

Now think about how these same statements can be translated into spatial examples. What are the implications of both types of uncertainty for allegedly simple measures such as the total area of a given soil that could be extracted from a GIS in a matter of seconds?

1.4. SCALES FOR ATTRIBUTE DESCRIPTION

In addition to point, line, and area object types, we need a means of assigning attributes to spatially located objects. The range of possible attributes is huge, since the number of possible ways we can describe things is limited only by our imagination. For example, we might describe buildings by their height, color, age, use, rental value, number of windows, architectural style, ownership, and so on. Formally, an *attribute is any characteristic of an entity selected for representation*. In this section, we explore a simple way of classifying attributes into types based on their *level of measurement*. The level of measurement is often a constraint on the choice of method of analysis and, ultimately, on the inferences that can be drawn from a study of that attribute's spatial structure.

It is important to clarify what is meant by measurement. When information is collected, *measurement* is the process of assigning a class or value to an observed phenomenon according to some set rules. It is not always made clear that this definition does not restrict us to assignments involving numbers. The definition also includes the classification of phenomena into types or their ranking relative to one another on an assumed scale. You are reading a work that you assign to the general class of objects called books.

You could rank it relative to other books on some scale of merit as good, indifferent, or bad. It is apparent that this general view of measurement describes a process that goes on in our minds virtually all our waking lives as we sense, evaluate, and store information about our environment.

If this everyday process is to yield useful measurements, it is necessary to insist that measurements are made using a *definable process*, giving *reproducible* outcomes that are as *valid* as possible. The first requirement implies that the measurer knows what he or she is measuring and is able to perform the necessary operations; the second is that repetition of the process yields the same results and gives similar results when different data are used; the third implies that the measurements are true or accurate. If any of these requirements are not met, the resulting measurements will be of limited use in any GIS, or at any rate, we will need good information about the ways in which the measurements fail to comply with these requirements in order to make effective use of them. In short, we need to know what we are measuring, there must be a predefined scale on which we can place phenomena, and we must use a consistent set of rules to control this placement.

Sometimes what we need to measure to produce attribute data is obvious, but at other times, we are interested in analyzing concepts that are not readily measured and for which no agreed-upon measurement rules exist. This is most common when the concept of interest is itself vague or has a variety of possible interpretations. For example, it is easy to use a GIS to map the population density over a region, but because it involves people's reactions, standard of living, and available resources, the concept of *overpopulation* cannot be measured simply by the population density. Note that these ideas do not prevent us from creating measures based on opinions, perceptions, and so on, and therefore admit the development of GIS dealing with qualitative data, provided that attention is paid to the difficulties.

The rules defining the assignment of a name, rank, or number to phenomena determine what is called the *level of measurement*, different levels being associated with different rules. Stevens (1946) devised a useful classification of measurement levels that recognizes four levels: *nominal*, *ordinal*, *interval,* and *ratio*.

Nominal Measures

Because no assumptions are made about relative values being assigned to attributes, *nominal* measures are the lowest level in Stevens's scheme. Each value is a distinct *category*, serving only to label or name the phenomenon. We call certain buildings "shops," and there is no loss of information if these are called "category 2" instead. The only requirement is that categories are *inclusive* and *mutually exclusive*. By *inclusive*, we mean that it must be possible to assign all objects to some category or other ("shop" or "not a shop"). By *mutually exclusive*, we mean that no object should be capable of being placed in more than one class. No assumption of ordering or of distance between categories is made. In nominal data, any numbers used serve merely as symbols and cannot be manipulated mathematically in a meaningful way. This limits the operations that can be performed on them. Even so, we can count category members to form frequency distributions. If entities are spatially located, we may map them and perform operations on their (x, y) locational coordinates.

Ordinal Measures

For nominal measures, there are no implied relationships between classes other than their mutual exclusivity. If it is possible to rank classes consistently according to some criterion, then we have an *ordinal* level of measurement. An example is the classification of land into capability classes according to its agricultural potential. We know the order, but not the differences, along an assumed scale. Thus, the difference between the first and second classes may be very different from that between the ninth and tenth classes. Like nominal data, not all mathematical operations are clearly meaningful for ordinal data, but some statistical manipulations that do not assume regular differences are possible.

Attributes measured on the nominal and ordinal scales are often collectively referred to as *categorical data*.

Interval and Ratio Measures

In addition to ordering, the *interval* level of measurement has the property that differences or distances between categories are defined using fixed equal units. Thermometers typically measure on an interval scale, ensuring that the difference between, say, 25°C and 35°C is the same as that between 75.5°C and 85.5°C. However, interval scales lack an inherent zero and so can be used only to measure *differences*, not absolute or relative magnitudes. *Ratio* scales have an inherent zero. A distance of 0 m really does mean no distance, unlike the interval scale 0°C, which does not indicate no temperature. By the same token, 6 m is twice as far as 3 m, whereas 100°C is not twice as hot as 50°C.

The distinction is clarified by examining what happens if we calculate the ratio of two measurements. If place A is 10 km (6.2137 miles) from B and 20 km (12.4274 miles) from C, then the ratio of the distances is

$$\frac{\text{distance}\,AB}{\text{distance}\,AC} = \frac{10}{20} \equiv \frac{6.2137}{12.4274} \equiv \frac{1}{2} \tag{1.2}$$

whatever units of distance are used. Distance is fundamentally a ratio-scaled measurement. Interval scales do not preserve ratios in the same way. If place B has a mean annual temperature of 10°C (50°F) and place C is 20°C (68°F), we cannot claim that C is twice as hot as B because the ratio depends upon our units of measurement. In Celsius it is $20/10 = 2$, but in Fahrenheit it is $68/50 = 1.36$. In spite of this difference, interval and ratio data can usually be manipulated arithmetically and statistically in similar ways, so it is usual to treat them together. Together, they are called *numerical measures*.

Although data may have been collected at one measurement level, it is often possible and convenient to convert them into a *lower* level for mapping and analysis. Interval and ratio data can be converted into an ordinal scale, such as high/low or hot/tepid/cold. What is *generally* not possible is to collect data at one level and attempt to map and analyze them as if they were at a higher level, as, for example, by trying to add ordinal scores.

It is important to note that not everybody is convinced by Stevens's scheme for classifying levels of measurement. Velleman and Wilkinson (1993) have pointed out that it may be unnecessarily restrictive to rule out various types of analysis because the level of the attribute measurement seems not to support it (they also point out that this was not Stevens's intention). A good example is where a nominal attribute—say, a county ID number—seems to have some relationship with another variable of interest. Often in spatial numbering schemes there is a spatial pattern to the numbering—perhaps from east to west or north to south, or from an urban center outward. In such cases, relationships might very well be found between a theoretically nominal attribute (the ID number) and some other variable. Of course, in this case it would be important to determine what is responsible for the relationship and not simply to announce that zip codes are correlated with crime rates!

Later, Stevens himself added a *log interval* scale to cover measures such as earthquake intensity and pH in which the interval between measures rises according to a power rule. Later still, Chrisman (1998) pointed out that there are many types of attribute data in GISs that don't fit this scheme. For example, many types of line objects are best represented by both their magnitude and direction as *vector* quantities, and we often refer measures to cyclical scales such as angles that repeat every 360°. Such criticism of the measurement level approach emphasizes the important principle that it is always good to pursue investigations with an open mind. Nevertheless, the nominal, ordinal, interval, ratio scheme remains useful in considering the possibilities for analysis.

Dimensions and Units

Apart from their level of measurement, attributes have the property of *dimensionality* and are related to some underlying *scale of units*. If we describe a stream as a line object, variables we might consider important include its velocity, cross-sectional area, discharge, water temperature, and so on. These measurable variables are some of its so-called *dimensions* of variability. The choice of dimensions depends on the interests of the researcher, but in many problems in science it can often be reduced to combinations of the three fundamental dimensions of mass, length, and time, indicated by the letters M, L, and T. For example, a velocity dimension is distance L divided by time T, or L/T. This is true regardless of whether velocity is recorded in miles per hour or meters per second. LT^{-1} is another way of writing length divided by time.

Similarly, cross-sectional areas can be reduced to the product of two length dimensions, or L^2, discharge is a volume L^3 per unit of time T with dimensions L^3T^{-1}, and so on. Nondimensional variables are an important class whose values are independent of the units involved. For example, an angle measured in *radians* is the ratio of two lengths—arc length and circle radius—whose dimensions cancel out ($LL^{-1} = L^0$) to give no dimension. An important source of nondimensional values is observations recorded as proportions of some fixed total. For example, the proportion of the population that is white in some census district is a nondimensional ratio.

Dimensional analysis is an extremely valuable method in any applied work. Because equations must balance dimensionally as well as numerically, the method can be used to check for the existence of variables that have not been taken into account and even to help in suggesting the correct form of functional relationships. Surprisingly, geographers have shown little interest in dimensional analysis, perhaps because in a great deal of human geographic work no obvious fundamental dimensions have been recognized. Yet, as Haynes (1975, 1978) has shown, there is nothing to stop the use of standard dimensions such as P (= number of people) or $ (= money), and this usage may often suggest possible forms of equations.

Finally, interval and ratio attributes are related to a fixed scale of *units,* the standard scales used to give numerical values to each dimension. Throughout history, many systems of units have been used to describe the same dimensions. For example, in distance measurement, use has been made of "British" or Imperial units (inches, feet, miles), metric units (meters, kilometers), and other traditional systems (hands, rods, chains, nautical miles), giving a bewildering and confusing variety of fundamental and derived units. Although many systems were used because of their relevance to everyday life and are often convenient, in science they are unsatisfactory and can become confusing. This is something that NASA found out to enormous cost in 1998 when confusion over the system of units used to measure the gravitational acceleration of Mars spelled disaster for the Mars Climate Orbiter mission.

Thought Exercise: Spatial Data Types in Everyday Life

Look at Figure 1.2, which attempts to cross-tabulate measurement level with the geometric object types we have discussed to arrive at 12 possible spatial data types.

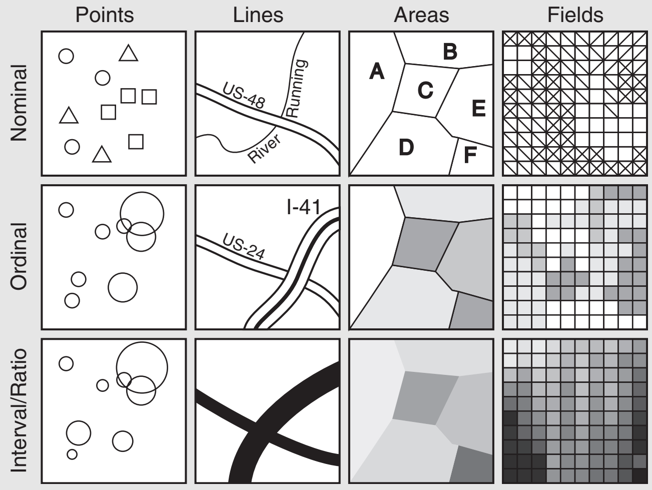

Figure 1.2 A schematic representation of entity-attribute spatial data types.

Now we want you to think about the rather abstract ideas we have been discussing.

What types of spatial object do you move among in your day-to-day life? For example:

- Is your house a point, an area, or both?
- Is your route to work, school, or college a line? What attributes might be used to describe it?
- Are you a nominal point data type? Perhaps you are a space–time (hence four-dimensional) line?

(*continues*)

(*box continued*)

- What measurement scales would be suitable for the attributes you would use to describe each of these (and any other) spatial objects you have suggested?

The answers to these questions give a sense of how potentially rich but also how reductive the entity-attribute framework is. As we explore spatial analysis further, remember this point: regardless of the insights that spatial analysis may yield, it is always performed on a representation of reality that may ultimately limit its usefulness.

1.5. GIS AND SPATIAL DATA MANIPULATION

We have noted that it is the ability to perform spatial manipulations on its data that distinguishes a GIS from any standard database management system. In this section, we examine a selection of these spatial data manipulations from our perspective as geographic information analysts. We do not intend to cover all these operations in detail since the number is very large and their implementation varies from system to system. In this section, we develop two spatial analytical perspectives on them. First, we develop the idea that these geometric operations involve some form of *transformation* between spatial data types. Second, we draw attention to the impact of *error* in our coding of the (x, y) coordinates used on the various outcomes.

Sometimes the geometry involved is simple—for example, finding the total length of line objects with a given characteristic (rivers, railways, roads needing repair) or calculating the total area and perimeter of some area objects (woodlands, crops of a certain type). At other times, it is *intersecting* types of spatial units in different ways that is the key. For example, we can easily use a GIS to determine how many cases of a disease occur within various distances of certain kinds of factories or other point objects. We need geo-coded data for cases of the disease and also for the facilities. These are usually available in the form of mailing addresses both for those afflicted by the disease and for the suspect facilities. We can then *buffer* the facilities to some distance (say 1 km) and use *point-in-polygon* operations to determine how many cases of the disease occur in the relevant buffer areas. The end result is a set of numbers recording how many cases of the disease occurred in the vicinity of each factory and how many occurred nowhere near a factory. Having determined these numbers, we could use appropriate statistical methods to determine whether or not the rates exhibit some non-random pattern.

Similarly, map *overlay* is where two or more map layers are combined in various ways to produce new combined layers. The classic example involves combining a number of development suitability classifications into a single composite index. This application was one of the original inspirations for GIS technology (see Ian McHarg's 1969 classic book *Design with Nature*). Input map data might include land slope, woodland density, transport accessibility (which might have been generated from buffer operations on the transport system), environmental sensitivity, and geological suitability for building. Map overlay produces a composite map formed from multiple intersections of all the inputs. Areas in the composite map have multiple attributes, derived from the attributes of their "parents," and can be assigned an overall suitability rating for development. The fundamental operation here is *geometric intersection* of the polygon areas in each map. A related operation *merges* polygons in different maps, depending on the similarity of their attributes. Incidentally, both of these operations are examples of the interchangeability of the raster and vector models, since either can readily be performed in a system based on either model. In fact, the two operations are developments—in geographic space—of the intersection and union operations familiar from set theory and Venn diagrams. Because it is so often a part of geographic information analysis, map overlay and the issues it raises are further discussed in Chapter 11.

Whether we are referring to point, line, area, or field entities, these operations (length, area, perimeter, intersection, buffer, merger, point-in-polygon, overlay, etc.) all involve relatively simple geometric manipulations of locational (x, y) coordinates. A useful way to think of them is as transformations between the various spatial data types that we recognized in Section 1.2. For example, if we have a data set made up of point objects, we might be interested in the areas within, say, 5 km of these objects defined by a series of circular buffers centered on each object. The buffered areas form a set of area objects, and we have transformed from points of length dimension L^0 to areas of length dimension L^2 (L^0 to L^2). In fact, such a buffer can also be considered to be a defined isoline on a continuous surface of distances from the points which is L^0 to L^3. Had the original buffer been along a line object, the transformation would have been from line to area (L^1 to L^2), and a buffer around an area object creates a second area object (L^2 to L^2). Reverse operations are also possible. We could start with area objects, and transform them into lines by computing their skeleton network (L^2 to L^1) or find their centroids, thus generating point objects (L^2 to L^0) (see Chapter 7 for more on these operations). Table 1.1 attempts to summarize this transformational view of GIS operations.

Rows of the table represent the data type from which we transform, and columns represent the resulting data type. Each row and column

Table 1.1 Spatial Geometric Operations as Transformations Between Data Types

		TO			
		Point, L^0	*Line, L^1*	*Area, L^2*	*Field, L^3*
FROM	*Point, L^0*	Mean center	Network graphs	Proximity polygons TIN, point buffer	Interpolation. Kernel density estimation Distance surfaces
	Line, L^1	Intersection	Shortest distance path	Line buffer	Distance to nearest line object surface
	Area, L^2	Centroid	Graph of area skeleton	Area buffer, Polygon overlay	Pycnophylatic interpolation and other surface models
	Field, L^3	Surface specific points VIPs	Surface network	Watershed delineation, Hill masses	Equivalent vector field

intersection defines one or more possible geometric operations. Whether you are a novice or expert user of a GIS, it is worth while spending a little time on the following thought exercise.

Thought Exercise: Geometry and GIS

If you already know enough to navigate your way around a GIS, we invite you to see how many of these operations can be achieved using it. If you are new to geographic information analysis, it is worthwhile to check your favorite GIS textbook for examples of each of these operations. We would be the first to admit that our matrix is certainly incomplete! We also recognize, following Chrisman (1999), that it oversimplifies the transformations involved.

Why is the ability to change the way we represent spatial entities important? First, as our example of cases of a disease around a suspect facility indicates, we might have hypotheses that can only be tested if we make use of appropriate geometric transformation. Second, it is very often the case that changing the way we represent our spatial data types allows us to gain a new perspective on our problem and may in itself lead easily and directly to a solution. The next exercise gives an example.

Changing the Representation: Delimiting Town Centers

A town center (or downtown) is a good example of an area object with uncertain boundaries. We all know when we are in one, but precisely where we enter it is more difficult and the criteria we use vary from place to place. In the United Kingdom, the need to monitor the economic health of town centers led in the 1990s to a desire in government to develop a consistent set of town center boundaries relevant to all towns in the country. A paper by Mark Thurstain-Goodwin and David Unwin (2000) reports on the method that was adopted. It gives a good illustration of how changing the representation by a geometric transformation led to the development of a working system.

In the United Kingdom, the increasing availability of high-resolution spatial data using the so-called unit post code as its georeference not only makes this transformation useful, it also make it essential for analysis. At this very high level of spatial resolution in which data on a series of urban functions (retail, entertainment, commercial, and so on) are known as nearly exact (x, y) locations, use of these data as either point or area objects is not easy. What happens is that the intrinsic spatial ''granularity'' of these functions makes it difficult to apply any of the traditional point- or area-based methods. The alternative that was developed used kernel density estimation (see Section 3.6) to transform the data from point or area objects into continuous surfaces of spatial densities. Town centers could then be delineated by choosing appropriate contours on the density surfaces as their boundaries.

Viewed from our perspective as geographic information analysts, these transformations share a characteristic that can be worrying but that is often forgotten. With some exceptions, such as kernel density estimation and spatial interpolation using kriging, all are deterministic operations assuming that, since the input data are exact (x, y) coordinates and the processes are simple arithmetic manipulations performed by computer using many significant digits, the outputs must similarly be, to all intents and purposes, also exact. What this view forgets is that in any GIS these same coordinates are themselves a digital representation of the real world and that this representation cannot be exact. Locations are often found using error-prone semiautomatic digitizing or are badly recorded in some original field survey. The result of any geometric operations on such data will to a greater or lesser extent carry forward this uncertainty into any outputs. If there is further uncertainty introduced by the algorithms we apply to the data, as in kernel density estimation (Section 3.6), interpolation (Chapters 9 and 10), and overlay (Chapter 11), then the situation becomes even more complex. In Section 1.3 we encountered one consequence of such errors when discussing the true length of a line object, such as a coastline, thought to be of fractal dimension. Even if we know any line connecting two digitized locations to be straight, what is the impact of uncertainty in these coordinate locations on the true position of the line? Similarly, if we have a digitized outline of a wooded area, what is the impact of a similar error on our estimate of the true wooded area? And how are such errors propagated through into the results of a complex series of spatial geometric manipulations?

These questions have been addressed in the research literature (see, for example, Heuvelink et al., 1989; Heuvelink, 1993; Heuvelink, Burrough, 1993), but there is little evidence that their implications are being carried forward into routine work with spatial data and GISs. Over a decade ago, one of us (Unwin, 1995) reviewed the literature on error and uncertainty in GISs and concluded that systems need to be sensitive to these issues. The simple truth is that they (and the great majority of their users) are not.

1.6. THE ROAD AHEAD

In the remainder of this book, we take you on a tour of the field of geographic information analysis. We have organized the tour in what we hope you will find is a logical way. The next chapter looks at some of the big problems of spatial analysis—what makes spatial statistical analysis different from standard statistical analysis and the pitfalls and potential therein. Chapter 3 looks at methods by which spatial data can be visualized. Chapter 4 describes some fundamental issues in the analysis of spatial data, defining the important concepts of pattern and process, and Chapter 5 deals with the description and statistical analysis of point patterns. Chapter 6 looks at more recent approaches to this important topic. The critical property of spatial autocorrelation is introduced in Chapter 7, which deals with analysis of area objects. Chapter 8 brings together these ideas with some of the visualization materials from Chapter 3 to look at the relatively recently developed idea of local statistics. Chapters 9 and 10 deal with the analysis of continuous fields. In Chapter 11, we look at map overlay operations from a spatial analytic perspective. Finally, Chapter 12 describes some newer directions and developments in spatial analysis.

Throughout, we have tried to keep the level of mathematics as low as possible, but there are places where we have to draw on what to some may be unfamiliar matrix algebra. To help you, we have included an Appendix that summarizes the basics you need to know. If your mathematics is a little rusty, we suggest that you have a look at this appendix now.

CHAPTER REVIEW

- *Spatial analysis* is just one of a whole range of analytical techniques available in geography. It should be distinguished from *spatial data manipulations*, on the one hand, and *spatial modeling*, on the other.
- For the purposes of this book, *geographic information analysis* is the study of techniques and methods to enable the representation, description, measurement, comparison, and generation of spatial patterns.
- *Exploratory*, *descriptive*, and *statistical* techniques may be applied to spatial data to investigate the patterns that may arise as a result of processes operating in space.
- Spatial data may be of various broad types: *points*, *lines*, *areas,* and *fields*. Each type typically requires different techniques and approaches.
- The relationship between *real geographic entities* and spatial data is complex and *scale-dependent*.
- Representing geographic reality as points, lines, areas, and fields is *reductive*, and this must be borne in mind in all subsequent analysis.
- These objects are frequently not as simple as this geometric view leads one to assume. They may exist in *three spatial dimensions*, move and change over *time*, have a representation that is strongly *scale-dependent*, relate to entities that are themselves *fuzzy* and/or have *indeterminate boundaries*, or even be *fractal*.
- Although we have emphasized the difference between spatial analysis and GIS operations, the two are interrelated, and *most current spatial analysis is carried out on data stored and prepared in GISs*.
- Simple geometric *transformations* enable us to change the way entities are represented, and this might be useful for analysis.
- Finally, in any analysis of geographic information, we need to develop a sensitivity to the likely *sources of error* in our results.

REFERENCES

Bailey, T. C. and Gatrell, A. C. (1995) *Interactive Spatial Data Analysis* (Harlow, England: Longman).

Batty, M. and Longley, P. A. (1994) *Fractal Cities: A Geometry of Form and Function* (London: Academic Press).

Burrough, P. A. (1981) Fractal dimensions of landscapes and other environmental data, *Nature*, 294: 240–242.

Chrisman, N. R. (1998) Rethinking levels of measurement of cartography. *Cartography and GIS*, 25: 231–242.

Chrisman, N. R. (1999) A transformational approach to GIS operations. *International Journal of Geographical Information Science*, 13: 617–637.

Cressie, N. (1991) *Statistics for Spatial Data* (Chichester, England: Wiley).

DCDSTF—Digital Cartographic Data Standards Task Force (1988) The proposed standard for digital cartographic data. *The American Cartographer*, 15: 9–140.

Diggle, P. (1983) *Statistical Analysis of Spatial Point Patterns* (London: Academic Press).

Duckham, D. and Drummond, J. (2000) Assessment of error in digital vector data using fractal geometry, *International Journal of Geographical Information Science*, 14: 67–84.

Fisher, P. F. and Unwin, D. J., eds. (2005) *Re-presenting GIS* (Chichester, England: Wiley).

Ford, A. (1999) *Modeling the Environment: An Introduction to System Dynamics Models of Environmental Systems* (Washington, DC: Island Press).

Fotheringham, S., Brunsdon, C., and Charlton, M. (1999) *Quantitative Geography: Perspectives on Spatial Data Analysis* (London: Sage).

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

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

Haynes, R. M. (1975) Dimensional analysis: some applications in human geography. *Geographical Analysis*, 7: 51–67.

Haynes, R. M. (1978) A note on dimensions and relationships in human geography, *Geographical Analysis*, 10: 288–292.

Heuvelink, G. B. M. (1993) Error propagation in quantitative spatial modelling: applications in geographical information systems. *Netherlands Geographical Studies.* Utrecht: University of Utrecht.

Heuvelink, G. B. M. and Burrough, P. A. (1993) Error propagation in logical cartographic modelling using Boolean logic and continuous classification. *International Journal of Geographical Information Systems*, 7: 231–246.

Heuvelink, G. B. M., Burrough, P. A., and Stein, A. (1989) Propagation of errors in spatial modelling with GIS. *International Journal of Geographical Information* Systems, 3: 303–322.

Hearnshaw, H. and Unwin, D. J., eds. (1994) *Visualisation in Geographical Information Systems*, (London: Wiley).

Lam, N. and De Cola, L., eds. (1993) *Fractals in Geography* (Englewood Cliffs, NJ: Prentice Hall).

Langran, G. (1992) *Time in Geographic Information Systems* (London: Taylor & Francis).

Mandelbrot, B. M. (1977) *Fractals: Form, Chance and Dimension* (San Francisco: Freeman).

McHarg, I. (1969) *Design with Nature* (Philadelphia: Natural History Press).

Mitchell, A. (1999) *The ESRI Guide to GIS Analysis* (Redlands, CA: ESRI Press).

O'Sullivan, D. (2005) Geographical information science: time changes everything. *Progress in Human Geography*, 29(6): 749–756.

Raper, J. F. (2000) *Multidimensional Geographic Information Science* (London: CRC Press).

Richardson, L. F. (1961) The problem of contiguity. *General Systems Yearbook*, 6: 139–187.

Ripley, B. D. (1981) *Spatial Statistics* (Chichester, England: Wiley).

Ripley, B. D. (1988) *Statistical Inference for Spatial Processes* (Cambridge: Cambridge University Press).

Stevens, S. S. (1946) On the theory of scales of measurements. *Science*, 103: 677–680.

Thurstain-Goodwin, M. and Unwin, D. J. (2000) Defining and delineating the central areas of towns for statistical monitoring using continuous surface representations, *Transactions in GIS*, 4(4): 305–317.

Tomlin, C. D. (1990) *Geographic Information Systems and Cartographic Modelling* (Englewood Cliffs, NJ: Prentice Hall).

Turcotte, D. L. (1997) *Fractals and Chaos in Geology and Geophysics*, 2nd ed. (Cambridge: Cambridge University Press).

Unwin, D. J. (1982) *Introductory Spatial Analysis* (London: Methuen).

Unwin, D. J. (1994) Visualization, GIS and cartography, *Progress in Human Geography*, 18: 516–522.

Unwin, D. J. (1995) Geographical information systems and the problem of error and uncertainty. *Progress in Human Geography*, 19: 549–558.

Velleman, P. F. and Wilkinson, L. (1993) Nominal, ordinal, interval, and ratio typologies are misleading. *The American Statistician*, 47: 65–72.

Wesseling, C. G., Karssenberg, D., Van Deursen, W. P. A., and Burrough, P. A. (1996) Integrating dynamic environmental models in GIS: the development of a Dynamic Modelling language. *Transactions in GIS*, 1: 40–48.

Wilson, A. G. (2000) *Complex Spatial Systems* (London: Pearson Education).

Worboys, M. F. (1992) A generic model for planar spatial objects. *International Journal of Geographical Information Systems*, 6: 353–372.

Worboys, M. F. (1995) *Geographic Information Systems: A Computing Perspective* (London: Taylor & Francis).

Worboys, M. F., Hearnshaw, H. M., and Maguire, D. J. (1990) Object-oriented data modeling for spatial databases. *International Journal of Geographical Information Systems*, 4: 369–383.

Zeiler, M. (1999) *Modeling Our World: The ESRI Guide to Geodatabase Design* (Redlands, CA: ESRI Press).