Models of Scale and Scales of Modelling

Michael F. Goodchild

Levels of Detail

Describing the world is such an innate human skill that we scarcely ever think about the methods we use, or how successful we are, or how to be more successful. Humans have been describing the world to each other ever since hunter-gatherers began to share information on successes and failures with the rest of the family and band. Social insects such as bees have highly developed ways of communicating such basic geographical information as the direction and distance of important food sources. The travel diaries of von Humboldt, Lewis and Clark, or Powell contain detailed descriptions of biological and topographical resources, and the interpretations of observed phenomena that the writers formed while in the field. Cartographers developed accurate and succinct ways of communicating much of this information in the convenient form of paper maps, and surveyors developed methods for recording position extremely precisely. Today, Earth-orbiting satellites capture information about the Earth's surface in vast quantity and great detail, such that the amount of information about the planet's surface gathered every day and distributed to scientists and the general public through the Internet has the potential to substantially exceed the amount of information contained in all of humanity's accumulated corpus of printed text.

But despite the potential for accurate description, and the enormous capacity of today's (and tomorrow's) information technologies, the complexity of the Earth's surface is such that the most voluminous descriptions are still only coarse

---

Modelling Scale in Geographical Information Science. Edited by Nicholas J. Tate and Peter M. Atkinson.

generalizations of what is actually present. There are approximately 500 million km^2 of surface. If each square metre were described in the crudest terms, say by using only two characters to describe the class of land cover, and ignoring every other aspect of the square metre, the result would be a data set of approximately one petabyte, or 10^{15} bytes, roughly equal to recent estimates of the total amount of information accessible through the Internet (and enough to keep a typical 56 Kbps modem busy 24 hours a day for 10000 years). Thus early humans quickly developed the skills needed to generalize about the Earth's surface, to compress information into succinct forms, and to ignore all but the most important details. As Margaret Oliver writes in Chapter 11, 'we cannot describe (the phenomenon) at every location'. Geographical description became a sophisticated art, and then a science, as hunter-gatherers, early explorers, geographers and other scientists found ways of describing the important and interesting characteristics of places in the limited spaces of notebooks, film cassettes, paper maps, and most recently digital storage devices and geographical information systems.

The level of detail of a geographical description is clearly one of its most important properties. It allows the recipient of the description to judge how much is missing, and the degree to which what actually exists might differ from what is described. In modern environmental modelling, it allows the user of a data set to estimate the uncertainty associated with the results of the modelling. For example, weather maps show only the general patterns of circulation over an area; every user of weather maps knows that the weather at a given place can vary substantially from what is shown on the map.

Scale

Over the years a single term, 'scale', emerged to capture the sense of 'level of detail'. Unfortunately, 'scale' has many alternative meanings in English, from a balance to the covering of a fish or reptile (French, often more precise, uses '*échelle*' to translate this book's meaning of the word, but '*balance*' and '*écaille*' for these other meanings, respectively; German uses '*massstab*' to describe the scale of a map, from words meaning 'measuring stick'). *Webster's New World Dictionary* (Guralnik, 1974) gives 15 distinct meanings in English of 'scale' as a noun. Dale Quattrochi and his co-authors explore many of the more relevant meanings of scale in Chapter 1, and point out several of the potential semantic pitfalls: scale can refer both to the level of detail of a description, and to the scope or extent of the area covered (interestingly, not among the 15 dictionary meanings); and scale is large to a cartographer when there is plenty of detail, but large to most other scientists when there is little. A quick search for pages relevant to 'scale' on Altavista generated 13.2 million hits, far more than for 'GIS' (3.1 million) but well below 'environmental model' (23.7 million), 'map' (54.5 million), and of course 'sex' (55.6 million) (I am indebted to David Rhind for pioneering this form of analysis; see Rhind, 1999). Scale can refer not only to geographical detail, but also to the degree of detail in the recording of change through time (temporal scale or detail), or in the degree of detail in the spectral characteristics of data in remote sensing (spectral scale or detail). Throughout this book, however, and

unless otherwise specified or implied by the context, scale should be taken to refer to the degree of geographical detail in a description of the Earth's surface.

Humans generalize constantly, not only about the form of the Earth's surface, but also about the characteristics of other humans, and about the processes that drive the physical and biological systems around us. Occam's Razor, or the principle of choosing the simplest among alternative explanations of some phenomenon, is a form of generalization. However, the precise forms of generalization that prove successful with respect to any given phenomenon depend on the nature of that phenomenon. Thus to make generalization work for geographical phenomena, or phenomena distributed over the surface of the Earth, it is necessary to know something about the general nature of those phenomena. The most important and striking of these is clearly spatial dependence, or the tendency of phenomena to vary slowly in space, such that variance over a short distance is generally less than variance over longer distances. Spatial dependence is often elevated to the status of a law and attributed to Waldo Tobler (1970): 'All things are similar, but nearby things are more similar than distant things.' Spatial dependence is often described as an option, and it is common to conduct significance tests that use a null hypothesis of its absence. But a world of no spatial dependence would be an impossible world, with every variable exhibiting a pattern of white noise; in this world the full range of variation observed over the Earth's surface would potentially occur in every part, no matter how small the part. Thus if the null hypothesis of no spatial dependence is ever accepted in practice, it is almost certainly because of a Type II statistical error.

Spatial dependence provides the key to effective description, because it allows us to ignore differences over short distances, since these are likely to be small. Many metrics of spatial dependence have been devised, and one of the most successful, the variogram, is the basis of several of the chapters in the third section of this book. The most fundamental contribution of spatial dependence, and Tobler's law, is that it explains the primary form of generalization applied to spatial data: the removal of detail over short distances. Climatologists need only measure weather conditions at widely spaced measuring sites, because variation between them will be slow and easily interpolated with adequate accuracy; useful pictures of the Earth can be obtained from space with picture elements as large as 1.1 km^2 (the AVHRR sensor discussed by Jennifer Dungan in Chapter 12) because averaging images over these comparatively wide areas still preserves many of the large features of interest in the image; and the census can still provide useful data even though people must be aggregated into comparatively large areas to protect the identities of individuals.

We use many different terms to describe the process of removing local detail. In the context of spectral or Fourier analysis, a *low-pass filter* removes variation with wavelengths less than a specified threshold. In image processing, a *convolution* of an image by applying weights described by a two-dimensional function will remove local detail if the weights are all positive. The distances over which variation is removed are defined by the function, and Peter Atkinson describes the functions characteristic of the sensors used in Earth observing satellites in his chapter on geostatistical regularization (Chapter 13).

We also use many different terms, in addition to scale, to describe the level of detail in geographical data. In Chapter 12, Jennifer Dungan follows common practice in spatial statistics by referring to data's *support*, or the domain 'within which linear average values of a [geographical] variable may be computed', and notes the similarity between this definition and that proposed for scale in an earlier paper by Quattrochi (1993). *Resolution* has similar a meaning. However, while scale may be defined by a variety of metrics, including the cartographer's *representative fraction*, or the ratio of distance on the map to distance on the ground, resolution is normally defined as a length measure, and support as a measure of length, area or volume. There seems to be little prospect of defining a standard lexicon over such a diffuse research enterprise, which includes participation by so many disciplines; instead, we seem likely to face a permanent degree of semantic confusion.

Geographical Ontology

While the potentially overwhelming volume of geographical data often dictates the need for compression and efficiency in description, there are many other factors that determine scale in scientific practice. The process of determining the spatial resolution of remotely sensed data is complex, involving issues of instrument design, communication speeds and scientific needs as well as data volume. Thus the advent of the IKONOS satellite, providing images at 1 m resolution, an order of magnitude more detailed than any previous instrument, is likely to have a major impact on many forms of environmental research (as well as on the ability of the news media to obtain clear commercial pictures of US aircraft temporarily stranded in China, as happened in April 2001). The ability to build instruments at much greater resolution allows us to ask what scientists really *need* with much greater interest than before, and in Chapter 13 Peter Atkinson provides a novel basis for making such decisions, and for weighing the benefits of finer resolution against the increased costs and increased data volume: in principle, data volume rises as the inverse square of resolution.

Greater power in computation and storage, and in the ability to transform one form of description to another, also creates greater interest in the *ontological* questions – basic questions about how the geographical world is to be described. Scale is one ontological question, because how the world is described, particularly in the case of scale, ultimately determines the kind of science that can be done with the description. Other representational choices, such as the choice between raster and vector data structures, or the description of change through time, or the variables to be measured, are also part of geographical ontology, an area that was recently added to the approved research agenda topics of the US University Consortium for Geographic Information Science (www.ucgis.org).

Scale is also critically important in understanding *what* is measured. Consider the property of slope, a readily understood and commonly used measure of the topographical form of the Earth's surface. Slope is often derived from a digital elevation model (DEM), a regular grid of point determinations of elevation, by examining the elevations of a 3 × 3 neighbourhood, and determining the slope at the central point through a process of least-squares fitting of a plane. One might assume that the effect of reducing the spacing of points would be to produce better and better approximations to slope. But slope is a property of differentiable surfaces with well-defined

tangents, and Mandelbrot (1977) long ago pointed out that natural topographical surfaces do not generally have well-defined tangents. There is thus no true slope to which estimates converge as point spacing falls. Instead, we must define slope as a function of point spacing (or scale). In effect, slope is a property of scale, and slope(100) is potentially as interesting, useful and richly informative as slope(10) or slope(1), where the argument in parentheses denotes the point spacing of the DEM (the *posting* distance).

Similar arguments apply to population density, defined as the number of people living in a given area, divided by the area. Again it is impossible to define a true population density at a point, since any estimate depends explicitly on the scale at which density is measured. Methods of density estimation (Silverman, 1986) produce density estimates by convolving a point distribution of people with a filter or *kernel* function, as if every point were replaced by a continuous function, and the functions were then summed at every point. Suitable kernel functions have parameters that are measures of length, and the result of density estimation thus depends directly on the length measure associated with the kernel. Again, there is no true population density, only population density for a given parameter value and kernel, or more generally for a given scale.

It turns out that remarkably few geographical variables are scale-less. Almost all frequently mapped variables, such as soil class, vegetation cover class or land use class are scale-dependent in the sense that a length measure is inherent in the variable's definition, as are all variables captured by satellite remote sensing. Only such basic physical properties as atmospheric temperature and pressure, or soil pH, are amenable to scale-free definitions, though in practice the costs of data acquisition ensure that they will still be scale-specific when captured.

Changing Scale

This book is about *modelling* scale, and its title implies two distinct but interrelated needs: to deal with scale in modelling human and physical systems, and to model the effects of scale on description. The first occurs because scales are inherent in the ways processes operate. On the human side, our behaviour is often affected by characteristics that we perceive in the environment; for example, the choice of residential location can be affected by a perception of the socio-economic status of the surrounding neighbourhood. Such perceptions have inherent scales that determine the area over which they are defined, such that a potential resident might perceive socio-economic status averaged within a kilometre of the prospective residential location. In order to understand behaviour and to predict it correctly, we must therefore measure this causal variable of socio-economic status at the same scale as it is perceived. If it was measured at more detailed scales, appropriate transformations could create the correct causal variable, but if it was measured at scales that are too coarse, no such transformation is feasible, and successful modelling will be impossible.

Similarly, a physical process cannot be modelled successfully unless data are available at an appropriate scale. For example, if rainfall varies substantially over distances of less than 100 km, a map of rainfall that omits such variation as being below its defined level of spatial detail will be unsuccessful in modelling outcomes that depend on rainfall, such as agricultural yield or habitat type. Scale is important in modelling because input data must be of sufficient detail to capture the significant spatial variations important to the process being modelled.

The second sense of the title deals with changing scale. Often it is impossible to acquire data at the correct scale for modelling a given process, because the only available data are at scales that are too coarse. In such circumstances it would be useful to have a general measure of the amount of variation that is missing, even though the specific details of the variation are unknown. Models that allow us to predict the effects of scale on the properties of data, without actually acquiring data at those scales, are useful for a number of reasons:

- estimation of the inaccuracies that result from using data that are too coarse for modelling a given process;
- estimation of the costs and benefits associated with acquiring data at a range of scales;
- use of data acquired at one scale to simulate data with greater detail, such that the simulated detail matches expectations about its general characteristics.

We have several tools for modelling the effects of changing scale. The first section of this book deals with *fractals*, a conceptual and mathematical framework originated by Benoît Mandelbrot (1977). If a phenomenon exhibits fractal behaviour, then certain of its characteristics at a given scale can be predicted. For example, it is possible to predict how the length of a natural coastline varies with scale, based on analysis of its length at a selection of scales, if the coastline behaves as a fractal. Mandelbrot argued that many natural phenomena do indeed behave as fractals, and in Chapter 2 Nick Tate and Jo Wood review the literature on this topic and the mixed success that has resulted from efforts to show fractal behaviour through analysis of various classes of phenomena. The basic fractal model is clearly overly simplistic, and several methods, including the approach of multi-fractals described by Iqbal Adjali and Stephen Appleby in Chapter 4, are designed to provide a better fit.

Fractal models have been applied to both natural phenomena and social phenomena. Two of the chapters in the first section address social phenomena from a fractal perspective, and the entire second section of the book is devoted to issues of changing scale in social science. Here, scale is associated not only with the costs of acquiring data and with storage requirements, but also with the confidentiality of records. As a result, few investigators have any degree of control over the scale of available data, which is in most cases too coarse relative to the processes of interest. Many years ago, Robinson (1950) defined what has become known as the *ecological fallacy*, or the fallacy of inferring the characteristics of individuals from data about aggregates. Results obtained from data at scales that are too coarse not only lead to false inferences, but are also somewhat arbitrarily dependent on the particular details of the aggregation process, specifically on the zones used to produce aggregate statistics. This issue, first documented extensively by Openshaw (1983), has become known as the *modifiable areal unit problem* (MAUP).

There have been many attempts to find solutions to the ecological fallacy problem. Among them, perhaps the best known is the work of King (1997). All rely in some way on assumptions, and although King's are plausible, there is no reason to expect them to be generally and precisely true (see, for example, Anselin, 2000). Instead, geographers and others have looked for novel ways of addressing the ecological fallacy and MAUP. In Chapter 7, Stewart Fotheringham and his co-authors describe geographically weighted regression and show how it is possible for the investigator to control scale explicitly, and to produce a richly interpretable set of results. In Chapter 8, Seraphim Alvanides, Stan Openshaw and James Macgill turn the question completely around, and look to aggregation zones as areas that can be designed to satisfy specific analytical needs.

Underlying almost all of the chapters in this book is the concept of a *field*, a geographical variable that is conceived as a function of the two spatial dimensions (see, for example, Worboys, 1995). Underlying remote sensing is the notion that the Earth's surface emits a field of radiation, which is convolved by the characteristics of the sensor to produce radiation estimates that are averaged over image pixels. Underlying work with population data is the notion that population density is a field, which is integrated over reporting zones to produce estimates of population counts. The alternative conceptualization, that the geographical world is an empty space littered with discrete objects that can be enumerated and may overlap, occurs in only a small number of chapters, but it particularly underlies the chapter by Elsa João on map generalization (Chapter 9). Much effort has gone into developing computer applications that emulate the process of cartographic generalization, in which objects are simplified as scale becomes coarser, but this effort is largely unsuccessful – automating the generalization process is a very hard problem to solve. By contrast, the techniques of generalization relevant to fields are straightforward – convolution and low-pass filtering are easily automated. Yet the discrete object view dominates in many areas of application of geographical information, particularly in the practical applications of geographical information systems.

Inventory and Prospect

Despite the growing number of books about scale, it is clear that the issue will be with us for a long time to come. This book is perhaps the first to attempt to address and integrate both human and physical dimensions, with half of the chapters focusing on the specific scale issues of social data, and with coverage of scale change from both field and discrete object perspectives. Chapter 1, by Dale Quattrochi and co-authors, includes a discussion of a tool designed specifically to facilitate analysis at multiple scales, although such tools remain rare in the mainstream of commercial, off-the-shelf (COTS) software.

Nevertheless, progress is being made. The proliferation of readily available geographical data, from an increasing range of remote sensing satellites, commercial data providers and governments, is rapidly changing the environment for researchers. No longer is it necessary to take the only data available, and worry about scale issues later. Today, investigators are likely to have at least some control over data scale, whether through involvement as designers of sensors and processing systems, or through the ability to search the WWW and compare the characteristics of available data sets through online metadata. Experience, the literature, better software tools, and more readily available data are helping to make us more sensitive to scale issues.

This book is appearing at an appropriate time, as interest in geographical information systems accelerates and diffuses through the environmental and social sciences. There is increasing recognition of the importance and value of a spatial perspective, from empirical problem-solving with GIS in areas such as anthropology, epidemiology and landscape ecology, to new spatially explicit extensions of economic and ecological theory. Across all of these areas scale is perhaps *the* most important topic of geographical information science, and an essential issue in any application of GIS.

References

Anselin, L., 2000, The alchemy of statistics, or creating data where no data exist, *Annals of the Association of American Geographers*, **90**(3), 586–592.

Guralnik, D.B. (ed.), 1974, *Webster's New World Dictionary of the American Language*, 2nd College Edition (Cleveland: W. Collins).

King, G., 1997, *A Solution to the Ecological Inference Problem: Reconstructing Individual Behavior from Aggregate Data* (Princeton: Princeton University Press).

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

Openshaw, S., 1983, *The Modifiable Areal Unit Problem*, CATMOG 38 (Norwich, UK: Geo Books).

Quattrochi, D.A., 1993, The need for a lexicon of scale terms in integrating remote sensing data with geographic information systems, *Journal of Geography*, **92**, 206–212.

Rhind, D.W., 1999, Keynote Presentation, International Cartographic Association, Ottawa, August.

Robinson, W.S., 1950, Ecological correlations and the behavior of individuals, *American Sociological Review*, **15**, 351–357.

Silverman, B.W., 1986, *Density Estimation for Statistics and Data Analysis* (New York: Chapman and Hall).

Tobler, W.R., 1970, A computer movie simulating urban growth in the Detroit Region, *Economic Geography*, **46**, 234–240.

Worboys, M.F., 1995, *GIS: A Computing Perspective* (London: Taylor and Francis).