



# Numbers aren't nasty: a workbook of spatial concepts

David J. Unwin  
Emeritus Professor in Geography  
Birkbeck London  
University of London

2010



# Table of Contents

|                                     |   |
|-------------------------------------|---|
| Acknowledgment                      | 5 |
| Copyright and intellectual property | 6 |

Chapter 1, **Purpose and Motivations** develops a scheme for spatial concepts and then discusses the educational use of the workbook.

|     |                                                         |    |
|-----|---------------------------------------------------------|----|
| 1.1 | Introduction:                                           | 7  |
|     | What is spatial literacy and what is a spatial concept? |    |
|     | Developing a schema                                     |    |
| 1.2 | Educational challenges and structure                    | 19 |
| 1.3 | References                                              | 24 |

Chapter 2, **Location, Spaces and Distance** deals with the spatial primitives of location, distance, space and projection/reference frame together with the standard geometric classification of objects.

|     |                                                                            |    |
|-----|----------------------------------------------------------------------------|----|
| 2.1 | Aims and Introduction                                                      | 27 |
| 2.2 | Exercise (1): Location - where do you live, and what do you live in or on? | 27 |
| 2.3 | Exercise (2): Scale and representation                                     | 30 |
| 2.4 | Exercise (3): Adjacency and relations between elements of a set            | 33 |
| 2.5 | Exercise (4): Conceptions of distance                                      | 37 |
| 2.6 | Exercise (5): Some complications with 'distance'                           | 39 |
| 2.7 | Exercise (6): Projection and location                                      | 43 |
| 2.8 | Exercise (7): Transforming locations                                       | 46 |
| 2.9 | References                                                                 | 48 |

Chapter 3, **Patterns of Point Objects**, provides exercises mostly based in the CRIMESTAT package that explore second order concepts such as dispersion, density and pattern using as it examples point located objects.

|     |                                            |    |
|-----|--------------------------------------------|----|
| 3.1 | Introduction                               | 50 |
| 3.2 | Exercise (8): Dotting the map              | 51 |
| 3.3 | Exercise (9): Drawing your own pin map     | 54 |
| 3.4 | Exercise (10): Proportionate symbol maps   | 59 |
| 3.5 | Exercise (11): Centrography                | 62 |
| 3.6 | Exercise (12): Nearest neighbor statistics | 66 |
| 3.7 | Exercise (13): Ripley's K statistic        | 70 |

|                                                                          |    |
|--------------------------------------------------------------------------|----|
| 3.8 Further Exercises: Other things you can do with <i>Crimestat III</i> | 75 |
| 3.9 References                                                           | 76 |

Chapter 4, **Lines on maps**, looks at line objects and the associated concepts of length, direction and connection.

|                                            |     |
|--------------------------------------------|-----|
| 4.1 Introduction                           | 77  |
| 4.2 Exercise (14): Lines on maps           | 78  |
| 4.3 Exercise (15): Measuring length        | 80  |
| 4.4 Exercise (16): Fractals                | 83  |
| 4.5 Exercise (17): Direction               | 87  |
| 4.6 Exercise (18): Analyzing tree networks | 92  |
| 4.7 Exercise (19): Analyzing networks      | 96  |
| 4.8 References                             | 100 |

Chapter 5, **Area: the devil's variable?** Explores the idea of area and uses the *OpenGeoDa*<sup>TM</sup> package to produce maps and autocorrelation measures of the concept of 'pattern'.

|                                                                                             |     |
|---------------------------------------------------------------------------------------------|-----|
| 5.1 Introduction                                                                            | 102 |
| 5.2 Exercise (20): Types of areas on maps                                                   | 103 |
| 5.3 Exercise (21): Colour maps for area objects                                             | 106 |
| 5.4 Exercise (22): Choropleth maps for area objects                                         | 109 |
| 5.5 Exercise (23): Measuring area                                                           | 113 |
| 5.6 Exercise (24): What do we mean by 'shape'?                                              | 117 |
| 5.7 Exercise (25): Mapping area data using <i>OpenGeoDa</i> <sup>TM</sup>                   | 119 |
| 5.8 Exercise (26): Using <i>OpenGeoDa</i> <sup>TM</sup> to compute a spatial weights matrix | 127 |
| 5.9 Exercise (27): Spatial autocorrelation and pattern                                      | 131 |
| 5.10 Exercise (28): Global spatial autocorrelation using <i>OpenGeoDa</i> <sup>TM</sup>     | 140 |
| 5.11 References                                                                             | 147 |

Chapter 6: **Fields** provides a series of exercises associated with self-defining continuous fields such as continuity, gradient and trend, making use of the *3DField* and *E (Z)-Kriging* public domain packages.

|                                                     |     |
|-----------------------------------------------------|-----|
| 6.1 Introduction                                    | 149 |
| 6.1 Exercise (29): Continuity and isolining a field | 150 |
| 6.3 Exercise (30): Isolining by machine             | 156 |
| 6.4 Exercise (31): Visualizing fields               | 161 |

|                                                                           |     |
|---------------------------------------------------------------------------|-----|
| 6.5 Exercise (32): Trends in fields                                       | 163 |
| 6.6 Exercise (33): Spatial structure and spatial interpolation by kriging | 167 |
| 6.7 Exercise (34): Spatial structure from the semi-variogram models       | 175 |
| 6.8 References                                                            | 179 |

Chapter 7: **Taking it Further** does not attempt to list all the resources that an instructor wishing to develop these exercise further might call in, since to do so would involve an enormously long list. Rather it provides pointers to some web sources that will in turn lead into these materials.

|                                                                                |     |
|--------------------------------------------------------------------------------|-----|
| 7.1 Introduction                                                               | 180 |
| 7.2 General Geography and Geographic Information Science projects of relevance | 180 |
| 7.3 Spatial concepts                                                           | 182 |
| 7.4 Software used in the Workbook                                              | 183 |
| 7.5 Linked Resources: Data Files                                               | 185 |

## Acknowledgement

This Workbook collects together materials I have used over many years and in many different educational contexts. I am grateful to Dr. Nick Tate, University of Leicester for his suggestion that they could be gathered together under a heading of 'spatial literacy' and to the SPLINT initiative for the Fellowship that enabled me to do this. In preparation, I am grateful for additional discussion with Sarah Witham Bednarz, Chris Brunsdon, Jason Dykes, Nick Tate, and Harvey Miller. Students at Universities in Leicester, London, Hamilton (NZ), Christchurch, and Redlands together with those enrolled on my course for <http://www.statistics.com> have often unwittingly helped in the development of almost all the ideas and exercises that are presented. Some of the exercises aren't possible without use of software created by Luc Anselin, Ned Levine, the team that created 3Dfield, and Jo Wood. Alex Szumski 'road tested' the exercises and checked the manuscript for typos and the like. Any remaining errors are of course entirely my own responsibility.

I am grateful to Karl Grossner for permission to use Figure 1.1, to John Davis for permission to reference and use data (STRIAE and NOTREDAM) from his text *Statistics and Data Analysis in Geology* and to Wiley & Sons (New York) for permission to redraw and reuse my own figures 4.1, 4.3, 4.6, 5.4, 5.13, and 6.2 from O'Sullivan and Unwin (2003 & 2010) *Geographic Information Analysis*.

**David Unwin**  
December 22<sup>nd</sup> 2010

## Copyright and Intellectual Property

In the post World Wide Web, e-learning, world, attribution of Copyright to materials such as those in this Workbook isn't easy. There are ideas and possibly short sections of text that could be claimed by John Wiley & Sons, Birkbeck London, University of Leicester, University of Redlands (USA), Karen K Kemp (with whom I worked in Redlands) and myself. As sponsors of the project that has assembled them and provided their *raison d'être* in a spatial literacy context, the SPLINT Centre for Excellence in Teaching and Learning might well also lodge a claim. Where there might be doubt, I have indicated this by use of an appropriate qualifying phrase such as 'source' or 'taken from.'

In this form and context these materials have been authored entirely by myself, so it is reasonable to claim them as © David J. Unwin (2010) and to assert the moral rights of authorship. They are made available under the understanding that they can be used in teaching and modified to suit local circumstances without and claim or charge provided that the source is recognized by the statement:

*These materials have been developed from the Workbook 'Numbers aren't Nasty' assembled by David J. Unwin under the auspices of the (UK) Spatial Literacy in Teaching Initiative, (2010).*

Any repurposing for commercial gain is subject to the usual 'all rights' © claim given above.

# Chapter 1: Purpose and Motivations

## 1.1 Introduction

What is spatial literacy and what is a spatial concept?

This little workbook provides a series of relatively simple, usually numerical or computer-based, exercises that together illustrate some of the basic spatial concepts whose mastery might be held to be a component of what has been termed *spatial literacy*. The exercises themselves have been drawn from experiences teaching geography at University level in a variety of institutions and have in most cases been tried and tested many times. A few are newly minted to fill some of the more obvious gaps in the coverage. In developing them I am aware that for those of a modern cultural-geographic persuasion the exercise, and the concerns to develop geography as 'spatial science' that they address, will seem both naive and irrelevant to the discipline as they define it. To compound this alleged sin the emphasis on numerical argument will also seem old-fashioned and, through some guilt-by-association argument, 'positivist'. I do not accept either of these critiques. Following Gatrell (1983, page 7) *I hope earnestly that the book will not be seen simply as 'spatial science', since I think the path charted here shows us that a concern for space transcends the artificial partitioning of geography into distinct research traditions and ultimately the artificial separation of disciplines in science.*

Work by psychologists (summarized in Golledge, 2002, page 3-4) seems to indicate that 'geographers' do 'think differently' from other academics, and suggests that this difference is characterized by an ability to reason about space and to represent its complexities graphically (mostly but not entirely by maps) in ways that are not matched elsewhere by any other discipline. I suggest that increasingly there is a need to include within the term 'geographer' not just those who have had the benefit of some instruction in the academic discipline, but also all those, scientists and the general public, who now routinely acquire and use spatial information through media such as GPS, GIS, satellite navigation systems, location based services, on-line mapping systems, virtual globes and even location-based games such as 'geocaching'. All of these activities require individuals to demonstrate some measure of *spatial literacy*, which can be defined as *the ability to think and act in any context that requires the recognition that location in space is important.*

Within this, *spatial thinking* has itself been defined by a group established by the (US) National Research Council (2006, page 12) as *a collection of cognitive skills*

*comprised of knowing concepts of space, using tools of representation, and reasoning processes.* These three abilities – knowing spatial concepts, representing them and reasoning from and about them - are not the same as what, in his Presidential Address to the Association of American Geographers, Golledge (2002) refers to as knowledge *of* space, which is *the accumulation of facts about the spatial arrangement and interactions comprising human-environment relations.* Rather it is knowledge *about* space, the *recognition and elaboration of the relations among geographic primitives and advanced concepts derived from these primitives (such as arrangement, organization, distribution, pattern, shape, hierarchy, distance, direction, orientation, regionalization, categorization, reference frame, geographic association and so on) and their formal linking into theories and generalizations* (Golledge, 2002, 1). Self-evidently, other disciplines make use of spatial concepts; a 'geographical' perspective on them is not the only one that could be taken. In what follows I have taken it largely on the grounds of convenience and familiarity, making no claims that the exercises I present and the concepts they explore have any wider utility.

All this is fine, but it begs a very serious question, which is the identification of the spatial concepts, representations and styles of reasoning that are in some sense uniquely 'geographic'. As a discipline in the UK we have not done well in this respect, but I doubt that this difficulty is at all unusual in any science. For example, the revised 2003 Quality Assurance Agency for Higher Education's *Benchmark Statement for Geography* (QAA, 2007) provides a long list of qualities that make up what the panel (of which the author was a member) considered to constitute 'geographical understanding'. The list includes terms such as 'spatial variation', 'scale', 'difference' and 'representation' but at no point are these terms defined. Given the highly 'political', contested nature of the task, and the need felt by the panel to incorporate at times very diverse views on the nature of academic geography, this is perhaps hardly surprising. As will be seen, this neglect in the UK contrasts markedly with concerns in the USA to develop both curricula and resources that in some sense promote spatial thinking (Beard *et al.*, 2008).

Any list is likely to be partial, incomplete, and highly contingent in time, but it further seems to me that, if as a discipline we claim to have some unique way of looking at the world, we owe it to our students and the general public to make some effort to say what this involves. Such an exercise can be justified for its own sake as an intellectual challenge, but, more practically, it can also be seen as an essential part of any development of curriculum that has educational aims related to spatial thinking. My view is that a clear identification and elaboration of the key spatial concepts is a necessary precursor to the development of curricula and materials. At best it might help; at worst it can do no harm.



Four past attempts are worthy of discussion. First, in the early days of geography's so-called 'quantitative revolution', John Nystuen published a short paper with the title *Identification of some fundamental spatial concepts* (Nystuen, 1963). This paper remains a landmark, not so much for its conclusions, but for the underlying and frequently misunderstood objective. This was not to provide a list of all the spatial concepts that we might want to include. His project was far more ambitious. Part of his first sentence reads: *to consider how many independent spatial concepts constitute a basis for the spatial point of view ...* and it goes on to talk about *the complete minimum set of concepts necessary to the spatial point of view of the geographer*. Interestingly, his claimed motivation for the exercise was to *clarify (his) objectives in studying geography* and the main example he uses relates to a class being taught in a mosque.

The three primitives to emerge from Nystuen's analysis were:

- Direction or orientation
- Distance
- Connection or relative position

with at least one extra 'notion not considered', which was

- Boundaries.

Nystuen then argued that given operational definitions these concepts represent the axioms of the spatial view with *other words, such as pattern, accessibility, neighborhood, circulation, [etc] (are) compounds of these basic terms*. In some respects this attempt is similar to work in dimensional analysis in which, for a given type of system, an attempt is made to define the minimum set of basic dimensions of which the mass, length and time (MLT) set often used in physical science is the best known. What is perhaps surprising about Nystuen's list is that, although site/situation and location/place are considered, his minimum set does not include the notion of *location* which is perhaps the most quintessentially geographical concept of all.

A second attempt to list the essential spatial concepts is that by Golledge (2002) in which he lists 19 'things' that constitute a *partial list of thinking and reasoning processes* that *should help* us determine what comprises *Geographic Thinking and Reasoning* (note his use of capitals). According to Golledge, a spatially literate person should be able to comprehend:

- Scale transformation;
- Transformation from one dimension to another;

- Superordinate and subordinate relations and frames of reference (cardinal, relational, local, global);
- Spatial alignment;
- Distance effects;
- Spatial association;
- Orientation and direction;
- Regionalization/spatial classification;
- Clustering and dispersion;
- Spatial change and spread;
- Spatial and non-spatial hierarchy;
- Density and distance decay;
- Spatial shapes and patterns;
- Locations and places;
- Overlay and dissolve;
- Integration of geographic features (points, networks, regions);
- Spatial closure/interpolation;
- Proximity and adjacency;
- Spatial forms.

This list can be structured in at least two ways. First, borrowing from the US National Research Council definition of spatial literacy, with some difficulty we can recognize those that are *primitive* spatial concepts, those that involve *representation*, and those that involve *reasoning* about space. Second, as Golledge recognized, they can be structured into an hierarchy in which identity and location (a 'something' 'somewhere') sit at the base and are then used to derive 'first order'/'simple spatial' concepts such as 'distance' and 'direction', from which can be derived 'second order'/'complex spatial' concepts such as 'distribution' and 'association'. The most complete attempt to do this that I know of is that by Injeong Jo and Sarah Bednarz (Jo and Bednarz, 2009) in their taxonomy of spatial thinking. This is summarized in Table 1.1, where in addition I have attempted to group like concepts together.

| <b>Level</b>         | <b>Concepts</b>                                                                                                                                          |
|----------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------|
| Non spatial concepts | Identity/quality, magnitude/quantity                                                                                                                     |
| Primitives           | Place-specific identity, location                                                                                                                        |
| Simple-spatial       | Distance, direction<br>Connection and linkage<br>Movement, transition<br>Boundary, region, shape<br>Reference frame, alignment<br>Adjacency<br>Enclosure |
| Complex-spatial      | Distribution, pattern, dispersion,                                                                                                                       |

|  |                                                                                                                                                           |
|--|-----------------------------------------------------------------------------------------------------------------------------------------------------------|
|  | clustering, density<br>Diffusion<br>Dominance, hierarchy, network<br>Association<br>Overlay<br>Gradient, profile, relief<br>Scale<br>Projection<br>Buffer |
|--|-----------------------------------------------------------------------------------------------------------------------------------------------------------|

Table 1.1: A hierarchy of non spatial and spatial concepts, based on Golledge (2002) and Jo and Bednarz (2009)

At the base of the hierarchy of spatial concepts is the primitive notion of a location, or, more properly, a *set* of locations and it is this notion that seems to me to be the most fundamental of all, since we can build from it most, if not all, of the remaining concepts in the list.

A third and more recent list of spatial concepts is that by Janelle (undated, see also Janelle and Goodchild (2009)) in his essay on *Spatial concepts and spatial reasoning in the social sciences: an agenda for undergraduate education* and in what the context makes clear is not intended to be either a complete or even a minimal set, he lists the following eight concepts, noting that

*These concepts are demonstrable at all levels of space and time (from sub-atomic to galactic, past through future, and microseconds to ions). They can be rendered understandable through simple illustrations to young children but they are also sufficiently engaging at advanced levels for thinking about scientific and social problems.*

- Location -- Understanding formal and informal methods of specifying "where";
- Distance -- The ability to reason from knowledge of relative position;
- Network -- Understanding the importance of connections;
- Neighborhood and Region -- Drawing inferences from spatial context;
- Scale -- Understanding spatial scale and its significance;
- Spatial Heterogeneity -- The implications of spatial variability;
- Spatial Dependence -- Understanding relationships across space;
- Objects and Fields -- Viewing phenomena as continuous in space-time or as discrete objects.

The website <http://www.spatial.ucsb.edu> refers to these as the eight foundational spatial concepts.

Finally, the team at UCSB that created the website <http://www.teachspatial.org> have taken an interesting empirical approach by examining the contents of some 20 sources (as of August, 2009) from which they extracted some 189 'unique terms' from over 300 references or 'assertions'. A frequency analysis shows that the emphasis given to each term varies strongly with discipline, whether geography, design, psychology, science education, linguistics, geosciences or social science, but an examination of the graphics used to report these differences also shows considerable agreement on some concepts. The 169 terms were classified by 'category' as shown in Table 1.2, but the rationale behind this categorization isn't explained. Nonetheless, the website represents by far the most complete and convenient analysis of the fundamental concepts of spatial thinking yet created.

| <b>Category</b>        | <b>Explanation</b>                                         | <b>Selected examples of terms</b>                                                   |
|------------------------|------------------------------------------------------------|-------------------------------------------------------------------------------------|
| General concepts       | ... concerning spatial and spatiotemporal context          | Space; place; field view; object view; continuity ...                               |
| Primitives of identity | The existence, nature and labeling of things in the world  | Object; attribute; objects and fields ...                                           |
| Spatial relationships  | Comparative locations of entities and their parts          | Direction; location; connection; distribution; adjacency ...                        |
| Measurement            | ... of objects and of relationships and related issues     | Shape; distance; gradient; area; volume ...                                         |
| Spatial structures     | ... as observed, and derived from measurement and analysis | Boundary; network; path; surface; region ...                                        |
| Dynamics               | Distinctly spatiotemporal concepts                         | Spatial interaction; diffusion; motion; force; frequency ...                        |
| Representation         | External tools and mental processes                        | Map; perspective; map projection; point; line; polygon; grid; coordinate system ... |
| Transformations        | ... on data                                                | Scale; spatial interpolation; overlay; buffer ...                                   |
| Spatial inference      | Products of analysis and conclusions drawn                 | Spatial dependence; spatial heterogeneity; distance decay; areal association ...    |

Table 1.2: Fundamental concepts of spatial thinking by category (after <http://www.teachspatial.org/fundamental-concepts-spatial-thinking>)

It seems to me that each of these past attempts includes elements that any schema should have but that all lack a consistent framework into which the concepts can be placed. Several attempts to do this, such as those by Golledge (2002) and Jo and Bednarz (2009) are shown graphically at <http://www.teachspatial.org/schemas> where they are referred to as schemas. Figure 1.1, due to Karl Grossner at UCSB, is taken from the website and shows how spatial concepts can be mapped onto the NRC elements of spatial thinking.

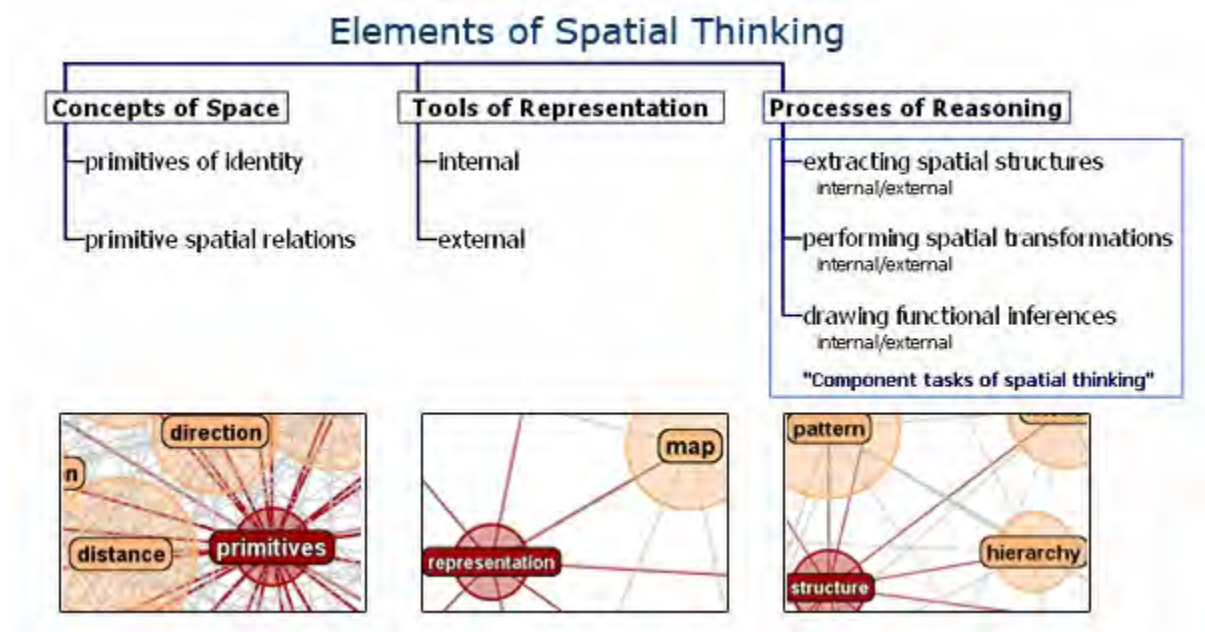


Figure 1.1: Karl Grossner's schema for spatial concepts (taken from <http://www.teachspatial.org/nrc-elements-spatial-thinking>)

## Developing a schema

It should be apparent that there are many ways by which such schemas can be developed, with different views being appropriate for different purposes. In what follows I develop a schema that helps clarify educational aims and intended learning outcomes that are appropriate for this workbook but that at the same time enables a 'mapping' into familiar ideas from the conventional geographic information science. The schema uses three organizing notions:

- A definition and view of 'distance' as a relation between primitives called 'locations' based on the work of Gatrell (1983);

- The familiar geometric classification of entities into points, lines, areas and fields; and
- Gollege's idea of a hierarchy of levels.

### **Towards a schema for spatial concepts (1): the relational view**

Following the excellent little book by Gatrell (1983, Chapter 2), this can be explained using set theoretic concepts. A set is any well-defined collection of 'objects', for example the set of F1 racing car teams that can be denoted:

$$F = \{\text{Brawn, Red Bull, Ferrari, BMW, ... etc}\}$$

Note that we can define the *elements* of this set easily and unambiguously, but more serious problems might occur if we were to attempt to define, say, the set of all motor car manufacturers, where ambiguity might arise because the set itself is 'fuzzy' (what is a 'car?'). Gatrell's book elaborates on this and provides numerous examples, but for a set to be uniquely 'geographic' it seems to me that it must consist of a collection of *locations*, for example:

$$T = \{\text{the set of cities in UK}\}$$

$$B = \{\text{the set of offices belonging to the CBD of a city}\}$$

$$M = \{\text{the set of mountain tops over 1000m in UK}\}$$

$$C = \{\text{the set of English counties}\}$$

Note that this implies that we can recognize and identify the entities we call 'cities', 'offices' and 'counties' but in themselves these are non-spatial concepts, simply 'somethings' that would occupy the first row of Table 1. Getting them onto the second row implies that we can add to this 'something' a 'somewhere' that creates a *set of locations*. Once defined, our locational sets allow us to build our geographies of interest by using the idea of a *relation on a set*.

With a set,  $C$ , consisting of, say,  $n$  elements can be defined its *Cartesian product*  $C \times C$  as another set of all the ordered pairs  $\{c_1c_1, c_1c_2, \dots, c_nc_n\}$ . Such a set can be visualized as a square matrix with  $n$  rows and  $n$  columns and has  $n^2$  elements. A *relation* on the set is any subset of the Cartesian product set and can be defined in numerous ways. Let us take a simple example. Suppose that we are dealing with the set

$$C = \{\text{the set of English counties}\}$$

This is a set of locations. Its Cartesian product gives an ordered list of all possible pairs of counties of which two elements are, for example, (Derbyshire, Yorkshire) and (Derbyshire, Essex). If the relation that we are interested in is that of adjacency, coded '1' if the two counties are adjacent (i.e. they share a

common boundary) and '0' if they are not, then we can replace the adjacent pair of counties (Derbyshire, Yorkshire) with a '1' and the pair (Derbyshire, Essex) that are not adjacent by a '0'. Notice that in this case we would almost certainly disallow a county from being adjacent to itself, making the relation *irreflexive*, and that in this example the relationship is *symmetric*, since (Yorkshire, Derbyshire) must also have the value '1'. Gatrell (1983, pages 15-24) provides further geographical examples.

Why is this important? The next step in Gatrell's argument is critical. In our example we used adjacency as an appropriate relationship between the set of areas that make up the counties of England. We could just as easily use some measure of their separation in space, making the chosen relation a 'distance'. Operationally, this 'distance' could be measured in any of a number of ways, for example from the nearest points on the county boundaries, from some centroids in each county, or by the time taken to drive from one county to the other. Viewed in his way, the adjacency relation we used in the example is also a measure of geographical 'distance'; as Gatrell observes there are many possible concepts of distance best summarized in the quotation he gives from the well-known columnist Katherine Whitehorn writing in the *Observer* newspaper many years ago (21<sup>st</sup> December 1980):

*'How far is it to Bethlehem? Not very far, we used to pipe as children. Depends on your point of view, if you ask me. How many shopping days to Christmas, how long is a piece of time and whether Bethlehem is £90, five hours flying time or just a prayer away is entirely a matter of opinion'*

Perhaps the most straightforward 'distance' we can recognize is the Euclidean straight line in a metric space between any two point locations,  $i$  and  $j$ , each defined by their  $(x, y)$  co-ordinates:

$$d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$$

The essential point here is that the complete set of distances, perhaps represented in matrix form as  $\mathbf{D}$ , formed by the Cartesian product of the original set of locations itself defines a 'space'. Euclidean distances are the shortest paths between the locations and have the underlying properties of non-negativity (no  $d_{ij}$  is less than zero), reflexivity (the distance to a location from the same location is 0), symmetry ( $d_{ij} = d_{ji}$ ), and obey the so-called triangle inequality. This states that the length of the longest side of the triangle is less than the sum of the lengths of the two shorter sides. The space they define is a Euclidean plane, which in turn is an example of a metric space. What follows from this is that *other types of location sets and other concepts of relation define other types of both metric and non-metric spaces*. Much of the rest of Gatrell's book shows how such spaces can be extracted from data on the relations between location sets and then examined and visualized. An obvious and relatively straightforward

example is where the scale of a study means that we have to consider the earth's curvature, such that or shortest path is now a Great Circle. Other metric spaces are also possible as are spaces that are decidedly non-metric. Perhaps the best known work exploring these spaces is that by Michael Worboys and his collaborators (Worboys, 1996; 2001; Worboys, Mason and Lingham, 1998; Worboys, Duckham and Kulik, 2004). Many of the concepts applied to area and networks (see below) relate primarily to spaces created by a 'distance' that is simply an adjacency or link, so the space is essentially that of a lattice or network.

These non-Euclidean spaces are important because they often are those in which people think and act in ways that require the recognition that location in some space is important. In other words they are a key component of spatial literacy, and analysis using them provides a way into what, following the naïve physics manifesto of Patrick Hayes (1978), Egenhofer and Mark (1995) outlined as its naïve geography equivalent. To see this link and give an idea of the flavour of their proposal, it is worth listing the admittedly incomplete list of fourteen elements of naïve geography they recognize:

- Naive geographic space is two-dimensional;
- The earth is flat;
- Maps are more real than experience;
- Geographic entities are ontologically different from enlarged table-top objects;
- Geographic space and time are tightly coupled;
- Geographic information is frequently incomplete;
- People use multiple conceptualizations of geographic space;
- Geographic space has multiple levels of detail;
- Boundaries are sometimes entities, sometimes not;
- Topology matters, metric refines;
- People have biases toward north-south and east-west directions;
- Distances are asymmetric;
- Distance inferences are local, not global;
- Distances don't add up easily.

This list makes it clear that analysis in spaces other than the simple Euclidean is not and will not be easy, and in some sense standard representations of space in GIS software and by conventional mapping do not help this. Notable exceptions are the work by Worboys and his colleagues cited above, work modifying standard spatial analytical methods for use across a network representation, and work using cartograms and related devices.

### **Towards a schema for spatial concepts (2): the geometric view**



The development thus far has led to the notion of sets of geographical primitives we call locations, together with the idea that the spaces to which the spatial concepts refer are generated as some sort of 'distance' relation between elements of a set of such locations.

This simple idea can be taken further by introducing the by-now familiar notion that geographic entities can be represented and classified by their fundamental dimension of length,  $L$ , into discrete objects called points (length dimension  $L^0$ ), lines ( $L^1$ ), areas ( $L^2$ ) and continuous, self-defining fields ( $L^3$ ). This classification has been used many times in the literatures of cartography and spatial analysis (see for example, Unwin, 1981). What becomes clear is that each geometry carries with it a type of space within which in turn there is a set of possible spatial concepts of interest.

### **Towards a schema for spatial concepts (3): hierarchies of concepts**

So, we now have the idea of different geometric/geographic object types, each generating a type of space within which different spatial concepts can be recognized. The final step, following Golledge's (2002) schema, is to recognize within each type of space a division of concepts that are appropriate for thinking about these spaces into *primitive/first order*, *complex/second order* and a loose group that might be called *analytical/ third order*. This distinction between levels is similar to that made by Golledge (see above), but it is not perhaps as easy to make in practice as it might seem in theory.

At the base sit some *primitive/first order notions*. For example if the entity is a point the appropriate primitives are its location in some reference frame/projection and its magnitude. If it is a line object then its magnitude is what we call 'length' but we also must add its direction/orientation. Similarly for an area object the primitives are its boundary and shape and for a field we have its 'height' at some location.

*Complex/second order* concepts take these primitives and combine them in some way to create new emergent concepts. Combinations of point objects allow us to think about spatial concepts such as distribution, dispersion, and pattern and bundles of line objects gives notions of linkage/connection into networks. Collections of area objects provide us with notions of adjacency, fragmentation, enclosure, pattern/ clustering (autocorrelation), hierarchy, and dominance. When we deal with more than one sample 'height' of a continuous field, second order concepts such as continuity, gradient, profile, relief, and trend appear.

Finally, at *the analytical/third order* we have a series of concepts that emerge as a consequence of some analysis such as point process models, the ideas of

stationarity and anisotropy/isotropy as applied to point objects, shortest/least cost paths and network generation models, areal association, spatial interaction models and tessellations as related to area objects, and equivalent vector fields, least 'cost' paths, and surface networks derived from fields. In a GIS environment many of the standard operations entail some form of association or transformation between objects of different spatial dimension. Examples include point in polygon determination (points and areas), overlay (two sets of area objects), density estimation (point to field), buffering (point line or area to area), and so on. Whether these operations should be included as 'concepts' is perhaps moot.

The complete schema is shown as Table 1.3.

| <b>Nature of Element</b> | <b>Relation</b> | <b>Space</b>                  | <b>Appropriate Concepts</b>                                                                                                                                                                                                                                                             |
|--------------------------|-----------------|-------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Point objects            | Distance        | Metric (especially Euclidean) | <b>Primitive/First order</b><br>location, magnitude<br>reference frame,<br>projection<br><b>Complex/Second order</b><br>distribution,<br>dispersion, pattern,<br>clustering, density.<br><b>Analytical/Third order</b><br>Point process models,<br>stationarity,<br>anisotropy/isotropy |
| Line objects             | Connection      | Network                       | <b>Primitive/First order</b><br>length, direction<br><b>Complex/Second order</b><br>linkage/connection,<br><b>Analytical/Third Order</b><br>network generation models                                                                                                                   |

|              |                        |                  |                                                                                                                                                                                                                                                                                                 |
|--------------|------------------------|------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Area objects | Contiguity/proximity   | Lattice          | <b>Primitive/First order</b><br>boundary, area, shape, region,<br><b>Complex/Second order</b><br>adjacency, fragmentation, enclosure, pattern/clustering<br>(autocorrelation), hierarchy, dominance, scale<br><b>Analytical/Third Order</b><br>models of areal association, spatial interaction |
| Field        | Distance and magnitude | Euclidean metric | <b>Primitive/First order</b><br>height<br><b>Complex/Second order</b><br>continuity, gradient, profile, relief<br><b>Analytical/Third order</b><br>vector fields, trend, surface networks, semi-variogram                                                                                       |

Table 1.3: Elements, relations, spaces and spatial concepts.

Evidently and perhaps obviously, several of the concepts, notably *pattern* occur more than once in the table and, perhaps significantly, the only three concepts in Golledge’s list that do not appear are *association*, which is essentially non-spatial, and *movement/transition* and *diffusion* which involve some notion of temporality.

## 1.2 Educational Challenges and structure

It is one thing to list and categorize these concepts, quite another to identify appropriate intended learning outcomes (ILO) for them, and to develop approaches and resources that might facilitate students in some sense ‘learning’ them. In part this may simply be that in the past we have not tried very hard to focus on them in our teaching but it is perhaps also related to their somewhat abstract nature. There is no single ‘correct’ or even optimal way to navigate through the entries in this table, since much will depend on the overall aims of

any curriculum that makes use of it. Table 1.4 summarises the offerings in this Workbook and can be used to suggest some possible route ways through them.

| <b>No</b> | <b>Exercise</b>                                              | <b>Geometry</b>                    | <b>Space</b>             | <b>Order</b>              |
|-----------|--------------------------------------------------------------|------------------------------------|--------------------------|---------------------------|
| 1         | Location - where do you live, and what do you live in or on? | Point, line<br>Area, even<br>field | All                      | Primitive                 |
| 2         | Scale and representation                                     | Point, line,<br>area and<br>field  | All                      | Primitive (?)             |
| 3         | Adjacency and relations<br>between elements of a set         | Area                               | Lattice/<br>Adjacency    | Complex                   |
| 4         | Conceptions of distance                                      | Line                               | Metric                   | Complex                   |
| 5         | Some complications with 'distance'                           | Adjacency                          | Metric<br>and non-Metric | Complex                   |
| 6         | Projection and location                                      | Points, lines and<br>areas         | Metric                   | Primitive                 |
| 7         | Transforming locations                                       | Points                             | Euclidean/<br>Metric     | Primitive                 |
| 8         | Dotting the map                                              | Points                             | Euclidean/<br>Metric     | Complex                   |
| 9         | Drawing your own pin map                                     | Points                             | Euclidean/<br>Metric     | Complex                   |
| 10        | Proportionate symbol maps                                    | Points                             | Euclidean/<br>Metric     | Complex                   |
| 11        | Centrography                                                 | Points                             | Euclidean/<br>Metric     | Complex                   |
| 12        | Nearest neighbor statistics                                  | Points                             | Euclidean/<br>Metric     | Complex &<br>Analytical   |
| 13        | Ripley's K statistic                                         | Points                             | Euclidean/<br>Metric     | Complex &<br>Analytical   |
| 14        | Lines on maps                                                | Lines                              | Metric                   | Primitive                 |
| 15        | Measuring length                                             | Lines                              | Metric                   | Primitive                 |
| 16        | Fractals                                                     | Lines                              | Metric                   | Primitive &<br>analytical |
| 17        | Direction                                                    | Lines                              | Metric                   | Primitive                 |
| 18        | Analyzing tree networks                                      | Lines                              | Network                  | Complex                   |
| 19        | Analyzing networks                                           | Lines                              | Network                  | Complex                   |
| 20        | Types of areas on maps                                       | Areas                              | Lattice                  | Primitive                 |
| 21        | Colour maps for area objects                                 | Areas                              | Lattice                  | Primitive                 |
| 22        | Choropleth maps for area<br>objects                          | Areas                              | Lattice                  | Primitive                 |
| 23        | Measuring area                                               | Areas                              | Euclidean/<br>Metric     | Primitive                 |
| 24        | What do we mean by<br>'shape'?                               | Areas                              | Euclidean/<br>Metric     | Primitive                 |

|    |                                                                          |       |                   |                                               |
|----|--------------------------------------------------------------------------|-------|-------------------|-----------------------------------------------|
| 25 | Mapping area data using <i>OpenGeoDa</i> <sup>TM</sup>                   | Areas | Lattice           | Primitive                                     |
| 26 | Using <i>OpenGeoDa</i> <sup>TM</sup> to compute a spatial weights matrix | Areas | Lattice/adjacency | Complex                                       |
| 27 | Spatial autocorrelation and pattern                                      | Areas | Lattice/adjacency | Complex                                       |
| 28 | Global spatial autocorrelation using <i>OpenGeoDa</i> <sup>TM</sup>      | Areas | Lattice/adjacency | Complex                                       |
| 29 | Continuity and isolining a field                                         | Field | Euclidean/Metric  | Complex/second order                          |
| 30 | Isolining by machine                                                     | Field | Euclidean/Metric  | Complex/second order                          |
| 31 | Visualizing fields                                                       | Field | Euclidean/Metric  | Complex/second order                          |
| 32 | Trends in fields                                                         | Field | Euclidean/Metric  | Complex/second order                          |
| 33 | Spatial structure and spatial interpolation by kriging                   | Field | Euclidean/Metric  | Complex/second order & analytical/third order |
| 34 | Spatial structure from the semi-variogram model                          | Field | Euclidean/Metric  | Analytical/third order                        |

Table 1.4: The workbook exercises classified by geometry, space and hierarchical level.

One might order a curriculum much as in this workbook, using the geometry employed from 'points' through 'lines' and 'areas' to 'fields', treating the nature of the space and the hierarchical level as secondary. This has the merit of simplicity and has been used many times by textbook authors covering these and similar materials. More ambitious approaches might be tempted to use the nature of the spaces progressing from simple adjacency lattices through networks to metric and Euclidean spaces. A third approach would be to work from primitive/first order concepts through complex/second order ones to those that involve analysis at the third order. A moment's thought will indicate that it would be a mistake to equate these hierarchical levels with specific ages/levels in the educational system. As Golledge, March and Battersby (2008, Table 10) recognize, many of these concepts can and should be taught over a wide range of 'levels' in education from kindergarten, through the US grade system and its equivalents elsewhere, to at least Masters degree. Differentiation in this is possible by the choice of learning outcomes appropriate to the level of the teaching and of the sort that constitute the greater part of the UCGIS Body of Knowledge (DeMers, 2009).

Spatial concepts can be listed and taught in the lecture theatre, but, consistent with a *constructivist* stance to learning, my view is that they are best introduced to students through simple, illustrative practical exercises in which they are operationalized with real data:

*Tell me, I forget  
Show me, I remember  
Involve me, I understand*

This view is to an extent counter to much practice in the discipline, which uses the practical class as a vehicle for teaching skills associated with various methods of investigation such as mapping, GIS, and statistical analysis. In their discussion of the use of 'practicals' and projects in higher education in geography, Gold *et al.* (1991, pages 36-58) note that confining use of active learning to teaching techniques is a bad teaching strategy that is of relatively recent origin. Traditional practical teaching at first followed the approach introduced into the universities in the nineteenth century, when students were encouraged to repeat many of the classic experiments of science with a view primarily to understanding the underlying concepts. Acquisition of the necessary technical skills wasn't the main objective but rather it was a bonus that accrued from following the experiments. If in a UK secondary school you followed a GCSE course in either physics or chemistry you will be familiar with this approach. At first, practical work in higher education in geography addressed similar objectives well into the 1960s with what was considered at the time to be core substantive materials such as cartography and land survey taught using the practical method. Since then a whole raft of new approaches to knowledge acquisition using statistical analysis, remote sensing, GIS and so on have been seen as necessary skills that students should acquire and the 'practical' class has had its emphasis changed from a method for teaching substantive geography to one whose primary intended outcomes are associated with learning technical and transferable skills. A similar change can also be seen in the use of the field class. All this is not to argue against such teaching, but against the artificial separation between teaching about techniques and teaching substantive geography; as is attempted in this Workbook, 'learning by doing' can and should be applied over the entire curriculum. Griffith (1987, 1992) would seem to agree.

What follows is a Workbook of educational materials that use simple numerical exercises with 'freeware' software systems to explore some of the fundamental spatial concepts identified in Section 1.1. Some are very short, more of the nature of what elsewhere are called 'thought exercises' (O'Sullivan and Unwin, 2010, page 5) in which students are asked to think through something and then reflect on what this illustrates. Others ask students to use the vast information source that is the World Wide Web (WWW) to find some illustrative materials. From this it is a short step to exercises that use specific software and data to

address and illustrate particular concepts. Where relevant sample data are suggested and should be readily available from standard WWW sources.

That most involve the use of 'the steel bars of numbers' to illustrate the concepts is deliberate and it explains the title 'Numbers aren't nasty' I have chosen for this Workbook. Regrettably, many in contemporary academic geography seem to think that numbers are nasty, with all sorts of associated guilt, not least of which are a lack of imagination and a false association with some caricature the people concerned have in their minds of 'positivism'.

Although most of the exercises have been 'road tested', sometimes many times, some are untested, created to give some sort of continuity to this workbook. Whether 'tried and tested' or new, instructors will almost certainly need to adapt them to address different intended learning outcomes or meet local circumstances and, as with any purchased goods, *caveat emptor* applies. In particular the suggested student briefings are only indicative of what might be used. No guidance or suggestions are made for how these materials might be used in specific curriculums or at what level. Any instructor, and hopefully there will be some, wishing to incorporate any of these materials into their teaching will be perfectly capable of assessing when and how best to use them. Although the Workbook covers a range of spatial concepts, there are some obvious gaps.

Each exercise has been structured in the same way, with a standard format that contains:

1. Aims and introduction
2. Geometry, space and level
3. Intended learning outcomes
4. Resources needed
5. Suggested student briefing
6. Comment/answers
7. Suggestions for modification
8. References

In this (1) outlines which of the spatial concepts are addressed by the exercise, (2) provides a key to where the exercise fits in the schema developed in this Chapter. Very importantly (3) narrows down the aims in (1) to some specific learning outcomes. (4) lists the necessary resources whilst (5) contains a suggested student briefing. In this part I have tried hard to avoid text that reads like a simple sequence of operations (press this, type that etc) that lead to a single desired outcome. Quite apart from the sheer boredom of assessing such work and the opportunities for lazy students to plagiarize the work of others, I fail to see what this type of briefing usually achieves. My preference is for what in the past has been called 'open-ending', leading to different results, either

through different pathways or through the use of different input information (Unwin, 1980). I realize that not all the suggested exercises manage to do this as effectively as I would have liked. Typically these latter examples are where software such as *CRIMESTAT III* and *OpenGeoDa*<sup>™</sup> are being used to calculate values for some test statistic and the suggested briefing is a blow by blow account of how to achieve the intended end product. My defence at the accusation of an apparent inconsistency with the ideas expressed in Unwin (1980) is that knowing what to do in such circumstances is pretty useless unless one also knows how to do it. 'Open-ended' alternatives can easily be developed once the work flow has been established. Finally, Sections (6) and (7) provide an opportunity for reflection on what use of the exercise has shown. Readers – if there are any – familiar with my texts published in 1981, 2003 and 2010 will recognize the origins of some of these materials. I can only apologise if they seem 'old hat' and express the hope that there is value in having them collected under the same, more convenient, roof.

Chapter 2 deals with the spatial primitives of location, distance, space and projection/reference frame together with the standard geometric classification of objects. Chapter 3 provides exercises mostly based in the *Crimestat III* package that explore second order concepts such as dispersion, density and pattern using as it examples point located objects. Chapter 4 looks at line objects and the associated concepts of length, direction and connection. Chapter 5 uses the *OpenGeoda*<sup>™</sup> package to examine concepts associated with area objects such as fragmentation and pattern (autocorrelation). Chapter 6 provides a series of exercises, some using *3Dfield*<sup>™</sup>, associated with self-defining continuous fields such as continuity and trend. Finally, Chapter 7 gives a few pointers to where additional supporting materials might be found.

### 1.3 References

Beard, K, and nine others (2008) *Symposium on a Curriculum for Spatial Thinking: Executive Summary*. Available (24/08/09) at <http://www.teachspatial.org/executive-summary-symposium-curriculum-spatial-thinking#attachments>

DeMers, M. (2009) Using intended learning objectives to assess curriculum materials: the UCGIS Body of Knowledge. *Journal of Geography in Higher Education*, 33: 370-377

Egenhofer, M. J., and D.M. Mark (1995) Naive Geography. In: Frank, A. U. and Kuhn, W. (eds) *Spatial Information Theory: A Theoretical Basis for GIS*, Berlin: Springer-Verlag, Lecture Notes in Computer Sciences No. 988, pages 1-15. See also <http://www.ncgia.buffalo.edu/i21/ng/ng.html>



Gatrell, A.C. (1983) *Distance and Space: a geographical perspective*. (Oxford, England: Clarendon)

Gersmehl, P (2006) Wanted: A concise list of neurologically defensible and assessable spatial thinking skills. *Research in Geographic Education*, 8:5-38

Golledge, R.G. (2002) The nature of geographic knowledge. *Annals of the Association of American Geographers*, 92(1): 1-14.

Golledge, R.G., March, M. and S. Battersby (2008) Matching geospatial concepts with geographic educational needs. *Geographical Research*, 46: 85-98

Gold, J.R., Jenkins, A., Lee, R., Monk, J., Riley, J., Shepherd, I and D. Unwin (1991) *Teaching Geography in Higher Education: a Manual of Good Practice* (Oxford, Blackwell, available in full at <http://www2.glos.ac.uk/GDN/gold/index.htm>)

Griffith, D.A. (1992) Teaching spatial statistics to geographers using Minitab. *Journal of Geography in Higher Education*, 16(1): 45-60

Griffith, D.A. (1987) Teaching autocorrelation by simulation. *Journal of Geography in Higher Education*, 11(2): 143-153

Hayes, P. (1978) The Naive Physics Manifesto. IN: D. Michie (ed.), *Expert Systems in the Microelectronic Age*. (Edinburgh: Edinburgh University Press), pp. 242-270.

Janelle, D.G. (undated) Spatial concepts and spatial reasoning in the social sciences: an agenda for undergraduate education. Available (24/08/09) at: <http://spatial.ucsb.edu/resources/teach-learn/docs/Janelle-Spatial-Concepts-in-the-Social-Sciences.pdf>

Janelle, D. G. and M. F. Goodchild (in press, 2009). Concepts, Principles, Tools, and Challenges in Spatially Integrated Social Science. IN: Nyerges, T., H. Couclelis, and R. McMaster (Eds.) *GIS & Society Research*. Sage Publications. (from summary at [www.teachspatial.org](http://www.teachspatial.org))

Jo, Injeong and S.W. Bednarz (2009) valuating geography textbook questions from a spatial perspective: using concepts of space, tools of representation, and cognitive processes to evaluate spatiality. *Journal of Geography*, 108:4-13

Lee, M.P. and J.B. Soper (1987) Using spreadsheets to teach statistics in geography, *Journal of Geography in Higher Education*, 11(1): 27-33

National Research Council (2006) *Learning to Think Spatially* (Washington DC: National Academy Press)

Nystuen, J.D (1963) Identification of some fundamental spatial concepts. *Papers of the Michigan Academy of Science, Arts, Letters*, 48: 373-384. Anthologised at least twice, of which the most influential was that in Berry, B.J.L. and D.F. Marble (1968) *Spatial Analysis: a reader in statistical geography* (Englewood Cliffs, NJ: Prentice Hall). The most accessible is probably now in Agnew, J., Livingstone, D.N., and A. Rogers (1996) *Human Geography: an essential anthology* (Oxford: Blackwell)

O'Sullivan, D. and D. Unwin (2010) *Geographic Information Analysis* (Second Edition, New York: Wiley)

QAA, Quality Assurance Agency for Higher Education (2007) *Geography 2007*. available at <http://www.qaa.ac.uk/academicinfrastructure/benchmark/statements/Geography.pdf>

[Unwin, D. J.](#) (1980) Make your practicals open-ended, *Journal of Geography in Higher Education*, 4 (2): 39-42.

Unwin, D.J. (1981) *Introductory Spatial Analysis* (London: Methuen)

Worboys, M.F. (1996) Metrics and Topologies for Geographic Space, In: Kraak, M.J. and Molenaar, M. (eds.) *Advances in Geographic Information Systems Research II: Proceedings of the International Symposium on Spatial Data Handling, Delft*, (London: Taylor & Francis) pages 365-376. (See also a pre-publication version at <http://www.spatial.maine.edu/~worboys/>)

Worboys, M. F. (2001) Nearness relations in environmental space, *International Journal of Geographical Information Systems*, 15(7): 633-651.

Worboys, M.F., Mason, K. and J. Lingham, J. (1998) Computational techniques for non-Euclidean planar spatial data applied to migrant flows, In: Carver, S. (ed), *Innovations in Geographical Information Systems: 5*, (London: Taylor and Francis) pages 35-45.

Worboys, M., Duckham, M. and L. Kulik, (2004) Commonsense notions of proximity and direction in environmental space. *Spatial Cognition and Computation*, 4 (4), 285-312.

# Chapter 2: Location, Spaces and Distance

## 2.1 Aims and introduction

Chapter 1 outlined a schema for spatial concepts leading to the notion that for point objects the appropriate relation is a distance, usually, but not always in a Euclidean metric space. First order concepts in such a space are location, magnitude and the reference frame and projection. In this chapter we develop student's appreciation of the 'first order' basic spatial concepts of *location*, *scale*, *adjacency*, *distance*, and *projection* as well as the variety of ways by which they can be 'measured' in different 'spaces'.

## 2.2 Exercise (1): Location - where do you live, and what do you live in or on?

### **Aims and introduction**

This is a simple, almost trivial, thought exercise that addresses the variety of ways by which we locate objects but it leads into such interesting and potentially rich discussion. I have used it many times as an *ice-breaker* at the start of courses in geographic information science.

### **Geometry, space and level**

With luck responses will use each of the geometries we recognize (point, line and area for certain, possibly even field) and so the idea that how we specify our location can generate different types of space. The notion lies at the base of our hierarchy as primitive/first order.

### **Intended learning outcomes**

After doing this exercise, students will:

- Be able to list many ways by which we locate ourselves and objects;
- Understand that different approaches are used for different purposes;
- Understand that the basic question *where?* Can be answered by specifying a point location, a line object (such as a street) or a named area object;
- Note how the resolution of these systems varies;
- Appreciate that each approach implies location in some different space;

- Understand that analyses will be constrained by the nature of these spaces.

### **Resources needed**

Pencil and paper/whiteboard/OHP to taste

### **Suggested student briefing**

1. Ask the class to take five minutes to write down as many ways as they can think of how they would answer someone who asks 'where do you live'?

### **Comment/answers**

The main purpose of the exercise is to lubricate a discussion that should address all the intended learning outcomes. There are many ways of locating objects and it is useful to collate a lost of all those suggested. Latitude and longitude, some grid coordinates in a National or State projection, house number with street name, a named area of a city and a description of how to get to the residence are examples among many possibilities. All may be considered aliases of each other, and all have their differing uses and potential resolution. Using county or city name, census tract, post/zip code, telephone area code, or regional names are also ways of identifying location, but with variable resolutions. Precise numerical location by latitude/longitude or even in a projection system such as UTM may be necessary for mapping and analysis but are not often used in everyday discourse.

Waldo Tobler (2002) provides an amusing anecdote that has a serious implication as follows:

*"Peter Gould and I conducted a little experiment in the late 1980s to demonstrate this. He sent envelopes on which he had typed my name along with the geographical coordinates of my house to thirty-four colleagues throughout the world. The envelopes contained only a blank piece of paper. The instructions to his friends were to add stamps and put the envelopes into the local postal system".*

Tobler reports that just four of these letters arrived but in every case they were delivered to his university office because of the 'Professor' title on the envelopes. He notes that some of the postal clerks added the name 'Santa Barbara' to them and that some of the letters were routed through unexpected places.

These differing methods can be converted from one to the other, or converted to latitude and longitude, with a precision that depends on their resolution.

## Suggestions for modification

Three simple additions to this exercise are:

- (1) Use a www location finding utility such as in UK the one at <http://www.streetmap.co.uk> to perform searches using each and every georeference the system uses (street name, telephone code, OS ( $x, y$ ), post code, place name and latitude/longitude);
- (2) Qualify the question by some imagined location, such as overseas, in your country, and in your locality. The first will demonstrate that the same place can be georeferenced in many ways, whilst the second shows that the one used is frequently context dependent;
- (3) Make use of a class set of GPS receivers, or, with an eye to the immediate future, location-aware mobile/cell phones or digital cameras capable of geo-tagging. Obvious extensions might include student field class exercises based around activities such as those reported at <http://confluence.org> to create a benchmark photographic record of a field area, some form of GPS drawing (see <http://www.gpsdrawing.com>), and an activity based on the sport known as geocaching (see <http://www.geocaching.com>).

## 2.3 Exercise (2): Scale and representation

### **Aims and introduction**

A real difficulty that geography has is the considerable range of spatial scales that are of interest, from the relatively small spaces of low-order drainage basins and towns through to the global. This exercise attempts to reinforce the idea of *representation* of different features on maps (and within GIS data bases) being contingent on the scale at which we chose to observe them. In introducing the exercise it is wise to make it clear that both a digital data base description and an analogue map of the same entities are alike in that both are representations of some underlying reality. Using on-screen or paper maps just makes life a little easier.

### **Geometry, space and level**

Again, all geometric object types should be addressed and the influence of scale on this representation noted. Using standard maps implies that we are in a metric, even Euclidean, space and, in so far as it relates to the reference frame used the level is, arguably, primitive.

### **Intended learning outcomes**

After completing this exercise students will

- Understand that in geography the scale at which we examine phenomena affects what is represented and this applied both in analogue and digital representations;
- Because of this, all representations generalize by selection, changing the nature of the entities, and how they are mapped;
- The intended use of the representation depend on the intended use of that information;
- The extreme difficulty of digital representation of entities that would sustain mapping and analysis over a range of scales.

### **Resources needed**

Web browser or series of appropriate paper maps

### **Suggested student briefing**

1. Go to <http://www.streetmap.co.uk>. This is one of several websites that provide map extracts over a range of scales. At the window enter either

the name or the postcode of a location within Great Britain with which you are familiar;

2. The result will be that you arrive at a topographic map of the area around your chosen point at the *LandRanger* scale of 1:50,000. The website calls this 'Zoom level 4'. There is an icon that enables you to view a wider area and the usual on-screen tools to allow you to pan around. Below and to the left of the map is a very useful little icon that allows access to pages that give the full key to each of the map series that are available;
3. You can now step up and down scale by clicking on the zoom tool to the right of the map display. Quickly ensure that you can access the 1:25,000 (Zoom level 3) and 1:5000 (Zoom level 2) larger scales as well as the 1:100,000 (Zoom level 5) and 1:200,000 (Zoom level 6) mapping;
4. The exercise is simple. For the table of features and map scales on Table 2.1, enter a code to say how the feature is represented. It will help if you use codes as P (point feature) L (line feature), A (area object), F (field) and M for features that are 'missing' at that scale. Some entities may be represented by more than one such geometric primitive. In each case is the representation true to scale, or has the selected entity been 'symbolized' in some way? At this stage you might also too describe the cartographic representation of the entity — what colour, line style or font is used and why?

| <b>Feature Type</b>         | <b>1:5000</b> | <b>1:25000</b> | <b>1:50000</b> | <b>1:100000</b> | <b>1:200000</b> |
|-----------------------------|---------------|----------------|----------------|-----------------|-----------------|
| Roads                       |               |                |                |                 |                 |
| Houses                      |               |                |                |                 |                 |
| Rivers                      |               |                |                |                 |                 |
| Land height above sea level |               |                |                |                 |                 |
| Public house ('pub')        |               |                |                |                 |                 |

**Table 2.1 Blank recording table**

### **Comment/answers**

The table can be developed to look something like Table 2.2.

| <b>Feature Type</b> | <b>1:5000</b>                    | <b>1:25000</b>                           | <b>1:50000</b>                          | <b>1:100000</b>                    | <b>1:200000</b>                              |
|---------------------|----------------------------------|------------------------------------------|-----------------------------------------|------------------------------------|----------------------------------------------|
| Roads               | L, named and coloured. ALL shown | L, with coloured status given. ALL shown | L, size exaggerated, coloured by status | L, size very exaggerated, coloured | L, size very exaggerated, coloured by status |

|                             |                                                        |                                                    |                                                  | by status                                                   |                                                           |
|-----------------------------|--------------------------------------------------------|----------------------------------------------------|--------------------------------------------------|-------------------------------------------------------------|-----------------------------------------------------------|
| Houses                      | <b>A</b> , Grey colour                                 | <b>A</b> , each building footprint correctly sized | <b>A</b> , with some point symbols for specifics | <b>P &amp; A</b><br>Point symbols and grey stipple in towns | <b>P&amp;A</b><br>Point symbols and grey stipple in towns |
| Rivers                      | <b>L</b> , blue, but only the largest shown            | <b>L</b> , Blue lines                              | <b>L</b> , Blue lines                            | <b>L</b> , blue lines, generalized                          | <b>L</b> , blue lines, generalized                        |
| Land height above sea level | <b>M</b>                                               | <b>F</b> , contours and spot heights               | <b>F</b> , contours and spot heights             | <b>M</b>                                                    | <b>M</b>                                                  |
| Public house ('pub')        | <b>M</b> , but some urban land use shown coloured as A | <b>M</b> , this is noteworthy!                     | <b>P</b><br>PH symbol                            | <b>M</b>                                                    | <b>M</b>                                                  |

Table 2.2 Suggested solution?

The key points here are related to the intended learning outcomes. The entities shown are a selection that depends on map scale and function. The geometric character of this representation (point, line, area, field) changes according to the mapping scale, with many being exaggerated and/or symbolized at smaller scale mappings.

### **Suggestions for modification**

As presented the exercise combines all the outcomes into the one analysis, but an obvious ploy is to create separate tables for the selection and character and cartographic representation.

The same exercise can use other similar websites, of which in UK that belonging to the Ordnance Survey is the most obvious. It is both harder to use and more volatile than the one suggested. The entire issue of representation in geography is explored in a series of essays edited by Fisher and Unwin (1995).



## 2.4 Exercise (3): Adjacency and relations between elements of a set

### **Aims and introduction**

How far apart are Derbyshire, Yorkshire and Essex? This exercise uses Gatrell's development of the spatial concept of *adjacency* as a relation between elements (objects) of a set (Gatrell, 1983, see Chapter 1).

### **Intended learning outcomes**

After completing this exercise students will:

- Understand that the spatial concept of adjacency is an example of a relation on a set of area objects and can be viewed as a measure of 'distance' between area objects;
- These distances themselves define a 'space';
- Understand that the relation can be described by a geographic structure matrix of the sort that frequently occurs in quantitative analysis;
- Be able to develop such a matrix from a map of planar enforced contiguous zones;
- Be able to state some of the properties of this matrix.

### **Geometry, space and level**

Here we deal with areas in a space generated by the adjacency relationship. Since we deal with numerous area objects, the level is complex/second order.

### **Resources needed**

Maps showing area objects such as States, Counties, and so on. Take care to ensure that no more than 10-15 zones are represented. For recording the results it is useful for students to use Excel or even the 'table' function in WORD as a way of keeping track. Dividers for measuring lengths.

### **Suggested student briefing**

1. Find a small pattern of zones, such as the Standard Economic Regions of England and Wales or a contiguous set of State of the USA or Counties in State and use it to create a table (matrix) of *adjacencies* in which each element is coded '1' if the zones are adjacent and '0' otherwise. This is done by listing the areas as both rows and columns of the matrix in which the same ordering is used for both the rows and the columns.

For example, Figure 2.1 below shows the eight 'Standard Statistical' regions used by UK government in the period 1945 to 1994.



Figure 2.1 The 1945-1994 Standard Statistical regions of England. 1: North 2: North West 3:Yorkshire and Humberside 4: West Midlands 5:East Midlands 6: East Anglia 7: South West and 8: South East

This yields a matrix of adjacencies as Table 2.3.

|                      | <b>1</b> | <b>2</b> | <b>3</b> | <b>4</b> | <b>5</b> | <b>6</b> | <b>7</b> | <b>8</b> |
|----------------------|----------|----------|----------|----------|----------|----------|----------|----------|
| <b>1</b> North       | *        | 1        | 1        | 0        | 0        | 0        | 0        | 0        |
| <b>2</b> North West  | 1        | *        | 1        | 1        | 1        | 0        | 0        | 0        |
| <b>3</b> Yorks       | 1        | 1        | *        | 0        | 1        | 0        | 0        | 0        |
| <b>4</b> West Mids   | 0        | 1        | 0        | *        | 1        | 0        | 1        | 0        |
| <b>5</b> East Mids   | 0        | 1        | 1        | 1        | *        | 1        | 0        | 1        |
| <b>6</b> East Anglia | 0        | 0        | 0        | 0        | 1        | *        | 0        | 1        |
| <b>7</b> South West  | 0        | 0        | 0        | 1        | 0        | 0        | *        | 1        |
| <b>8</b> South East  | 0        | 0        | 0        | 0        | 1        | 1        | 1        | *        |

Table 2.3 England and Wales Standard Region adjacencies

This sort of matrix is an example of what in the literature is called a 'geographic structure matrix' and it is usually symbolized by a bold capital **W**. Each *element* of the matrix is denoted  $w_{ij}$  in which the subscript 'i' is the row number and the subscript 'j' is the column. Thus the entry at row

5, column 6 is a '1', which tells us that zones 5 (East Midlands) and 6 (East Anglia) are adjacent. This type of matrix is much used in spatial statistical analysis, and we will encounter it again.

2. Why have we placed stars down the so-called principal diagonal of this matrix? Is the matrix symmetric (a mirror image) about this principle diagonal? If it is, why?
3. Can we improve on this definition? Use your map to develop a definition of adjacency that is a ratio-scaled number, such as the shared boundary length (measure this by stepping with dividers). Is the resulting matrix still symmetric?
4. Finally, develop this further for each zone by re-expressing this as a proportion of the total boundary length of that zone. Now, is this matrix symmetric and if not, why not?

### **Comment/answers**

See Section 1.2 for a discussion. The key issues to develop are

- a) That in this example we define the spatial concept of adjacency as a relationship between the elements (the zones) of a set (The Statistical Regions). In the first case the concept is expressed as a binary (yes/no) but as the second part illustrates we could expand this to an ordinal or ratio scaled number that expresses the strength of the adjacency;
- b) Adjacency is a sort of 'distance' between the zones in which the distance is defined in a 'space' generated by these adjacency values;

The stars indicate that we need to make an operational decision as to whether or not any zone can be considered adjacent to itself and of course the matrix is symmetric because of the way we define adjacency as sharing a common boundary. The same applied to the length of common boundary, but by expressing this as a proportion of each local area's total boundary length we lose the symmetry.

### **Suggestions for modification**

Once developed there is a great deal of exploration possible related to the properties of this type of adjacency matrix but these require more advanced mathematics:

- a) Using software such as MINITAB™, Excel, or a few lines of simple code, the **W** matrix can be 'powered' to explore the adjacency relationship at differing lags as described in detail by Unwin (1981), Tinkler (1977) and Garner and Street (1978);
- b) At more advanced level, students can be introduced by way of simple adjacency to the more general types of spatial weights matrices and their properties, including those of their eigensystems. The literature on this is summarized in O'Sullivan and Unwin (2010, pages 200-205). A key point to make here is that the **W**-matrix is really an expression of a hypothesis about what is important in the geography chosen. Exercise (26) uses some public domain software to compute and store a **W**-matrix.

## 2.5 Exercise (4): Conceptions of distance

### **Aims and introduction**

As we have seen adjacency is an example of a relation between elements of a set of locations and in the last exercise our locations were planar enforced named area objects. This exercise uses a journey in London (England) but could easily be modified for any other city with which you are familiar using similar resources.

### **Geometry, space and level**

The geometries of these journeys are essentially metric but the spaces in which they take place will not always be Euclidean. Distance is a complex/second order concept.

### **Intended learning outcomes**

After completing this exercise students will:

- Understand that a conventional distance, as measured by a Crow's flying, straight line is an example of a relation on a set;
- Realize that in human behaviour we seldom have access to these straight line distances and that a motor car would be constrained to a network distance which, if a taxi cab, would be the shortest path between the two locations;
- Be able to generalize this further into cost and time distances;
- Have a view as to which would be most appropriate in an applied problem;
- Again, realize that these various relations each produce a 'space';

### **Resources needed**

Web browser

### **Suggested student briefing**

Our objective is to get from Euston main-line railway station to Waterloo main-line station using different transport methods and measuring the 'distance' between these two locations in different ways. Throughout note that we are dealing with a set consisting of two locations {Euston station, Waterloo station} and exploring possible specifications of a relation on this set.

1. First of all, what is the *straight line distance* between these stations? A suitable map can be found at the *Transport for London* website: <http://www.tfl.gov.uk/assets/downloads/Central-London-Day-Bus-Map.pdf>
2. You will need to use a ruler and knowledge of the scale of the map to answer this. As an alternative use *Google Earth*<sup>™</sup> to find both railway stations and then use the 'ruler' in 'tools' to find the 'straight line' distance.
3. Now suppose you were to hire a taxi cab for the same journey. The driver has a choice of literally hundreds of ways through this network, each with its own *network distance*, but they will normally take the *shortest route*. This can also be found using a Google<sup>™</sup> system in this case Google<sup>™</sup> maps and the supplied 'get directions' function. You'd probably be concerned with how much this might cost you to get a *cost distance*. There is a schedule of taxi fares for London at <http://www.tfl.gov.uk/gettingaround/taxisandminicabs/taxis/1140.aspx> that you can use to estimate this;
4. Finally, of course, many Londoners would make the same journey by underground (the famous 'Tube') at a standard fare for a journey in what's called Zone 1. The 'distance' of concern would almost certainly be the *time the journey takes*. The same website has a trip choice aid that will estimate this for you at <http://www.tfl.gov.uk/gettingaround/default.aspx>, and you can see which tube line is involved by looking at <http://www.tfl.gov.uk/assets/downloads/standard-tube-map.pdf>
5. So in an applied problem – getting from station to station – which distance is appropriate: straight line, network shortest path, some other route through the road network, cost distance, or the 'cost' in time taken?

### **Comment/answers**

I make it about 3.0km in a straight line, but a taxi will need to go maybe 4.85km at a cost of around £10. Taking the 'Tube' costs a lot less and is a fairly easy trip on the so-called Northern Line with six intermediate stops. The Transport for London website suggests this would take about 10 minutes.

### **Suggestions for modification**

At some cost in extra student effort, why not build up a **W** matrix of a selected 'distance' between a selection of the main line termini in London (for example Liverpool Street, St. Pancras, Euston, Paddington, Victoria, Waterloo)? The easiest way is probably via Transport for London's journey planner and this will reinforce understanding of the concept.

## 2.6 Exercise (5): Some complications with 'distance'

### Aims and introduction

This exercise is taken fairly directly from papers by Michael Worboys in which he explores some of the basic properties of metric and non-metric spaces.

### Geometry, space and level

The geometry is that of adjacency in metric and non-metric spaces, with the relation measured in some notion of the complex/second order concept of distance.

### Intended learning outcomes

After completing this exercise students will:

- Understand that it is possible to change definitions of 'distance' and thus the properties of the spaces created;
- Realize that different 'spaces' are appropriate for different analyses and that 'network distances' are often more appropriate than straight line ones.
- Be able to develop a simple measure of the 'influence' of nodes in a network.

### Resources needed

Tables of road distances between a sample of places as often found in motoring Atlases.

### Suggested student briefing

1. Using a standard road atlas that includes a table/matrix of the road distances between towns for your country and select any five towns;
2. Use the supplied matrix to develop a **W** matrix showing the  $0.5n(n-1)=10$  unique distances between these places;
3. Using a suitable threshold distance, convert your matrix into a 0/1 binary matrix of *adjacencies*;

4. Examine each of the measures you have created in relation to the three properties that any *metric distance* should have. Simply stated these are (1) that the distance between points must be a positive number unless the points are the same, in which case the distance will be zero; (2) that the distance between two points is independent of which way round it is measured; and (3) the *triangle inequality*, which states that it must always be at least as far to travel between two points via a third point rather than to travel directly;
5. If for each place, we define adjacency by its two nearest neighbors, how does the derived binary adjacency matrix **A** differ and why?
6. Based on a simple inverse distance rule 'normalize' the distances for each location such that each row total sums to 1. For each place these are expressions of a *relative distance*;
7. Summing for each column and ignoring the infinities, which is the most 'influential' place in this system?

### Comment/answers

This is a sample set of result for a pretty arbitrary choice of places, although readers with knowledge of the *vitae* of a past Ordnance Survey Director General will understand why Berwick on Tweed is included. Table 2.4 provides the measured road distances.

| <i>From / To</i>   | <i>London</i> | <i>Aberdeen</i> | <i>Aberys</i> | <i>Ayr</i> | <i>Berwick</i> |
|--------------------|---------------|-----------------|---------------|------------|----------------|
| <i>London</i>      | 0             |                 |               |            |                |
| <i>Aberdeen</i>    | 503           | 0               |               |            |                |
| <i>Aberystwyth</i> | 211           | 445             | 0             |            |                |
| <i>Ayr</i>         | 394           | 177             | 314           | 0          |                |
| <i>Berwick</i>     | 338           | 182             | 311           | 134        | 0              |

Table 2.4: Road distances in Great Britain

Using a 220 mile threshold to define the spatial relationship 'near', we get the adjacency matrix shown in Table 2.5.



| <i>From/To</i>     | <i>London</i> | <i>Aberdeen</i> | <i>Aberys</i> | <i>Ayr</i> | <i>Berwick</i> |
|--------------------|---------------|-----------------|---------------|------------|----------------|
| <i>London</i>      | 0             |                 |               |            |                |
| <i>Aberdeen</i>    | 0             | 0               |               |            |                |
| <i>Aberystwyth</i> | 1             | 0               | 0             |            |                |
| <i>Ayr</i>         | 0             | 1               | 0             | 0          |                |
| <i>Berwick</i>     | 0             | 1               | 0             | 1          | 0              |

Table 2.5: An adjacency matrix from Table 2.4

This remains symmetrical because the metric distance AB is same as distance BA

If, for each place we define adjacency by its two nearest neighbors, we get the matrix shown in Table 2.6.

| <i>From/To</i>     | <i>London</i> | <i>Aberdeen</i> | <i>Aberys</i> | <i>Ayr</i> | <i>Berwick</i> |
|--------------------|---------------|-----------------|---------------|------------|----------------|
| <i>London</i>      | 0             | 0               | 1             | 0          | 1              |
| <i>Aberdeen</i>    | 0             | 0               | 0             | 1          | 1              |
| <i>Aberystwyth</i> | 1             | 0               | 0             | 1          | 0              |
| <i>Ayr</i>         | 0             | 1               | 0             | 0          | 1              |
| <i>Berwick</i>     | 1             | 1               | 0             | 0          | 0              |

Table 2.6: An alternative adjacency matrix

This matrix is asymmetrical; places do not share the same neighbors

Using these same distances to compute an interaction matrix **W**, based on a simple inverse distance rule and then normalizing this such that each row total sums to 1 gives the results in Table 2.7.

| <i>From/to</i>     | <i>London</i> | <i>Aberdeen</i> | <i>Aberystwyth</i> | <i>Ayr</i> | <i>Berwick</i> | <i>RowSum</i> |
|--------------------|---------------|-----------------|--------------------|------------|----------------|---------------|
| <i>London</i>      | $\infty$      | .1626           | .3877              | .2076      | .2420          | 1.000         |
| <i>Aberdeen</i>    | .1293         | $\infty$        | .1461              | .3674      | .3573          | 1.000         |
| <i>Aberystwyth</i> | .3548         | .1682           | $\infty$           | .2362      | .2407          | 1.000         |
| <i>Ayr</i>         | .1350         | .3004           | .1677              | $\infty$   | .3968          | 1.000         |
| <i>Berwick</i>     | .1546         | .2872           | .1681              | .3901      | $\infty$       | 1.000         |
| <i>Col_Sum</i>     | .7737         | .9184           | .8696              | 1.2013     | 1.2368         |               |

Table 2.7 : An interaction matrix based on Table 2.4

Summing for each column and ignoring the infinities, the most 'influential' place in this system is BERWICK. Those who know UK will know that this isn't a very useful result, largely because of the selection of places used. At least one very famous GI scientist might disagree!

### **Suggestions for modification**

As in the previous exercise software such as MINITAB™, Excel, or a few lines of simple code, the original **W** matrix can be 'powered' to explore the network distance at differing numbers of 'steps' (lags) lags as described in detail by Unwin (1981), Tinkler (1977) and Garner and Street (1978). In turn this leads to an appreciation of the very general notion of relations of almost any kind defining spaces of interest.

## 2.7 Exercise (6): Projection and location

### **Aims and introduction**

Students should by now appreciate that metric Euclidean distances define perhaps the simplest of the spaces that as geographers we analyze.

A very good way to introduce the distortion properties of map projections is to use the computer to examine the graticules (grids of parallels and meridians), continental outlines, and Tissot's Indicatrix for a variety of projections

Even creating a simple 'upside down' Mercator projection and moving the central meridian from the Atlantic to somewhere else radically alters one's perception of the planet.

### **Intended learning outcomes**

This exercise can be as long or as short as necessary to fit a variety of intended learning outcomes. If you teach map projections in any detail, the ability easily to display at least 25 standard projections and their associated graticules together with the world coastline and a map of *Tissot's Indicatrix* is absolutely invaluable.

- List the major types of map projections and developable surfaces and explain, with examples, how the Indicatrix enables their distortion properties to be visualized;
- Explain why the number of possible map projections is to all intents and purposes infinite.

### **Resources needed**

Software to support such an exercise can be found in many GIS. Alternatively Jo Wood's *Projector* is available as a Java applet at:

<http://www soi.city.ac.uk/~jwo/projector/>

Note that this needs Java to be installed on any machines using it.

### **Suggested student briefing**

It often comes as a surprise how the 'shape' of the continents and oceans changes when we change our map projection. Although promoted largely as

propaganda, the controversy surrounding the *Peters projection* UN is a good example of how our perception of the planet can be changed. This experiment uses some software to draw the outline of the coasts at global scale on a series of 25 different map projections.

1. If the entire topic of map projection is unfamiliar, go to the *Wikipedia* entry [http://en.wikipedia.org/wiki/Map\\_projection](http://en.wikipedia.org/wiki/Map_projection) and read it through carefully.
2. Types of deformation
  - At your computer, go to the Java applet at <http://www.soi.city.ac.uk/~jwo/projector/>. If it does not load then you may well have also to install Java.
  - In *view* ensure that you have switched on all three possible displays (graticule, Indicatrix and coastline)
  - Select *cylindrical* then *Mercator* as an example of a conformal projection. Examine the Indicatrix and the graticule and confirm that the former is everywhere a circle not an ellipse. In doing this it is useful to go back to *view* and switch off the coastline display. What property of this projection does this signify? What differs across the map of circles and why?
  - Repeat for the *Lambert Equal Area* which is another cylindrical projection. How does the distortion shown by the plot of the Indicatrix differ and what property does this signify?
  - Repeat for the *Equidistant Conic* as an example of an equidistant projection.

It should be clear that all map projections distort some combination of area, shape, direction, bearing, distance and scale. It should also be clear that the Indicatrix is a simple graphical way of showing some of these properties by way of the Indicatrix area, orientation and, if an ellipse, the ratio of its two axes. Ensure that you understand how these relate to properties of each map projection.

3. Developable surfaces
  - Examine a variety of azimuthal projections;
  - Examine a variety of conic projections;

- Examine a variety of cylindrical projections;
- Finally, look at the Mollewieide and Eckert as examples of mathematical projections.

In all cases, try to relate the map of the global coastlines with the graticule and the Indicatrix

### **Comment/answers**

There are no 'correct' answers.

### **Suggestions for modification**

WWW has numerous other resources for teaching map projections in a relatively informal way, for example:

<http://www.csiss.org/map-projections/index.html>

This site has examples of numerous other projections and links to alternative software for creating them such as:

[http://www.uff.br/mapprojections/mp\\_en.html](http://www.uff.br/mapprojections/mp_en.html)

Other sites are:

<http://www.nationalgeographic.com/xpeditions/lessons/01/g912/projections.html>  
<http://www.youtube.com/watch?v=AI36MWAH54s>

The exercise can be extended almost without limit. Formally it is useful to quiz students on what projections are appropriate for a variety of mapping tasks. One of the more fascinating extensions is to introduce a discussion on how our familiar map projections influence our perceptions of our place on the planet. Once students have accepted that the planet looks different according to how we chose to display it, a further discussion on our mental maps can be introduced.

## 2.8 Exercise (7): Transforming locations

### Aims and introduction

A critical and sometimes neglected step in many data integration exercises is the co-registration of data from different sources onto the same co-ordinate system using a *grid on grid transformation*. This exercise can be done using GIS software, but it is very instructive to follow all the steps using more basic tools. This graphical exercise is intended to help fix ideas about co-ordinate transformation.

### Intended learning outcome

After completing this exercise students will:

- Understand that we can take one pattern of locations into a different co-ordinate system by means of an affine transformation that involves a change in scale, translation of the origin and rotation of the axes.

### Resources needed

You will need some lined graph paper, some tracing paper (ideally transparent lined graph paper), a pencil and, perhaps, a calculator or spreadsheet.

### Suggested student briefing

1. Create a grid on your graph paper with  $X$  and  $Y$  axes each going from 0 to 100.
2. On this grid mark eight randomly located points and read off their  $(x, y)$  co-ordinates.
3. Use the tracing paper to prepare an identical set of axes, but do not mark any points on it.
4. Place the transparent grid on your original one with its origin exactly on the origin of the original one and rotate it by a small known angle (say  $15^\circ$ ).
5. Next shift (translate) the origin by a known small amount, and then mark on the transparent paper the positions of your eight points.
6. Read off the co-ordinates of the eight points in this new system.

### Comment/answers

The exercise is intended to demonstrate how an affine transformation 'works'. A detailed blow by blow account is in O'Sullivan and Unwin (First Edition only, 2003, pages 290-301).

### **Suggestions for modification**

If mathematically inclined, compute and use the affine transformation matrix for this operation. The development of the single affine matrix is described in the reference provided.

## 2.9 References

- Fisher, P.F. and Unwin, D.J. (eds., 1995) *Re-presenting GIS* (Chichester: Wiley)
- Garner, B, and W.A. Street (1978) The solution matrix: alternative interpretations. *Geographical Analysis*, 10, 185-190
- Maling, D. (1992) *Co-ordinate Systems and Map Projections* (Oxford: Pergamon)
- Mark, D.M & A.U. Frank (eds.) 1991 *Cognitive and Linguistic Aspects of Geographic Space* (Dordrecht: Kluwer)
- O'Sullivan, D. and D. Unwin (2010) *Geographic Information Analysis*, (Second Edition, Hoboken, NY: Wiley), especially pages 41-49 and 200-205
- O'Sullivan D., and Unwin, D. (2003) *Geographic Information Analysis* (First Edition, Hoboken, NY: John Wiley & Sons), pages 290-301.
- Snyder, J.P. (1989). [pubs.er.usgs.gov/usgspubs/pp/pp1453 *Album of Map Projections, United States Geological Survey Professional Paper*]. United States Government Printing Office. 1453., see <http://pubs.er.usgs.gov/usgspubs/pp/pp1453>.
- Snyder, John P. (1987). *Map Projections - A Working Manual. U.S. Geological Survey Professional Paper*. United States Government Printing Office, Washington, D.C.. 1395. <http://pubs.er.usgs.gov/pubs/pp/pp1395>.
- Tinkler, K.J. (1977) *An Introduction to Graph Theoretical Methods in Geography*, CATMOG 14 (Norwich: GeoAbstracts), available in full at <http://www.qmrg.org.uk/catmog/>
- Tobler, W. R. (2002) Ma Vie: Growing Up in America and Europe, In: *Geographical Voices*, W. Pitts and P. Gould, eds, Syracuse University Press, pages 292-322
- Unwin, D. (1981) *Introductory Spatial Analysis* (London and New York: Methuen), pages 84-93
- Worboys, M.F., (1996) Metrics and Topologies for Geographic Space, In: Kraak, M.J. and Molenaar, M. (eds.) *Advances in Geographic Information Systems Research II: Proceedings of the International Symposium on Spatial Data Handling, Delft*, (London: Taylor & Francis) pages 365-376. The original paper is



at <http://www.spatial.maine.edu/~worboys/mywebpapers/sdh1996.pdf> and the same site has numerous other relevant papers

# Chapter 3: Patterns of Point Objects

## 3.1 Introduction

This chapter continues the examination and clarification of concepts relating to point objects, for which as argued in Chapter 1, appropriate, complex/ second order, concepts relate to words like 'distribution', 'dispersion', 'density', 'pattern' and 'scale' and, at higher level still, third order concepts relating to point process models, stationarity and isotropy/anisotropy. In this chapter we provide suggestions for exercises based mostly in the CRIMESTAT package that explore such second order concepts as dispersion, density and pattern in distributions of point located objects. The main exercise looks critically at familiar tests against the hypothesis of complete spatial randomness, otherwise perhaps known as 'no pattern'.

## 3.2 Exercise (8): Dotting the map

### **Aims and introduction**

Dot maps show differences in the location and density of point located 'events'. Although numerous methods of point pattern analysis have been developed, it is only seldom that in practical studies we have suitable data for these analyses. The aim of this exercise is to show why this is so, and to force students critically to examine any dot/pin maps that they see.

### **Geometry, space and level**

A set of located point objects when mapped in a metric space immediately presents complex/second order concepts referred to as distribution, dispersion, pattern, clustering, and density. This exercise uses visualization to address them.

### **Intended learning outcomes**

After doing this exercise, students will

- Be able to recognize a simple dot or pin map;
- Understand the characteristics of a true point pattern suitable for statistical analysis as distinct from a simple dot density map;
- Appreciate the importance of the 'art and science' of cartography in determining the look of a map.

### **Resources needed**

WWW browser with access to Google™.

### **Suggested student briefing**

1. In order to understand some of the issues in 'dotting' a dot map watch the tutorial from Sara Fabrikant at [http://www.csiss.org/streaming\\_video/csiss/fabrikant\\_dot\\_maps.htm](http://www.csiss.org/streaming_video/csiss/fabrikant_dot_maps.htm). Pay particular attention to the distinction she makes between a 'one to one' and a 'many to one' mapping and to the importance of exact locations vs. data that are aggregated over areas.

It should be clear that for a 'one to one' mapping the basic data have to be suitable, with perhaps five basic conditions being necessary. These are:

- The pattern should be mapped/projected on the plane such that distance between the points are preserved;
  - The study area should be determined objectively, with boundaries that are not arbitrary. In practice this is very hard to achieve;
  - The pattern should be an enumeration of all the defined point objects in the study area;
  - There should be a one-to-one correspondence between dots on the map and objects on the ground, one dot, one object;
  - Locations should be proper, not for example arbitrary points within areas chosen to be in some sense representative.
2. Now go to *Google*<sup>TM</sup> (or similar search engine) to find a proper dot map that meets all five conditions. If you search for 'dot map', ask yourself several questions:
- Is there a one to one between the dots and distinct 'events' such as the location of a crime, some facility or whatever? Often dotting is used as a cartographic symbol with a 'many to one' relationship to the phenomenon being mapped, for example '1 dot represents 2000 acres'. These are dot density maps of the 'choropleth' variety;
  - Are the locations 'proper'? Is each dot located at the correct place where the 'event' occurred or is to be found? Often dots are placed at the centroids of areas or in a stipple across an area, so the locations have no special meaning and can't be used in point pattern analysis;
  - If the two conditions above are met, is it a sample or a complete enumeration or census?
3. If searching using 'dot map' doesn't reveal anything, try instead a search using the post-GIS term for the same type of map which seems to be 'pin map'.

### **Comment/answers**

Almost all of the images returned using a 'dot map' search will actually be dot density maps that do not meet the five conditions. Searches using 'pin maps' seem to do better. It is probable that students will find genuine examples in some crime maps and/or maps in epidemiology. It is worth emphasizing exactly

what dot density maps show, which is area aggregated data and thus make the point that just because dots are used in the representation it does not mean that the data themselves relate to point objects. The simple conclusion is that we seldom have 'pure' point data at precise locations on the plane of the sort required by almost all the standard methods of point pattern analysis. This is a very important lesson!

Personally, I've never been sure that they are all that useful unless they show *rates of occurrence*. In crime pattern analysis, the dots might be useful because they tell the police where to deploy their resources, but in epidemiology and criminology surely it is the rate that matters, relative to some underlying factors?

In addition, a majority of the maps returned will be cartographically awful, hardly worth drawing in the first place.

### **Suggestions for modification**

Discussion of the conditions for point pattern analysis to be sensible can be extended further. For example, a case can be made that some methods of analysis do allow use of sampled data (most obviously using randomly distributed quadrats and/sampling nearest neighbour distances in ecology)

### 3.3 Exercise (9): Drawing your own pin map

#### **Aims and introduction**

There are two exercises here, the first of which simply uses *Google*<sup>™</sup> to produce maps, whilst the second uses Microsoft *Excel*<sup>™</sup> to produce maps of three supplied point data sets. The overall aim is to introduce the idea of patterns in point data that are revealed by the maps.

#### **Geometry, space and level**

As in Exercise (8), a set of located point objects when mapped in a metric space immediately presents complex/second order concepts referred to as distribution, dispersion, pattern, clustering, and density. This exercise uses visualization to address them.

#### **Intended learning outcomes**

After doing this exercise, students will be able to:

- Produce dot/pin maps of any facilities recorded in the Google<sup>™</sup> Maps data base and/or
- Map any point located data provide as (x, y) co-ordinate pairs in a Cartesian system;
- Criticize the cartography they employ;
- Suggest possible descriptions for the point patterns revealed.

#### **Resources needed**

Approach (a) requires a web browser and (b) needs Microsoft *Excel*<sup>™</sup> or, should you prefer it, whatever standard GIS you use. Approach (b) also requires access to the three supplied data sets called BOOK, BANK and SNOW.

#### **Suggested student briefing**

##### **a) The lazy way**

1. In your web browser got to <http://www.google.com> and select the Map option;
2. If you live in, or know of, any reasonably large city, enter text in the search box as 'coffee shops in xyz', where xyz is the name of your city. If coffee shops don't appeal then try some other suitable facility;

3. The result will be returned as a pin map of the type discussed in Exercise (8);
4. In 3.1 above, five criteria were suggested to use in testing whether or not such a map is of a genuine point pattern that might be analyzed using methods to be introduced in the next Section. Evaluate your result in the light of these five criteria;
5. Finally, in your own words how would you describe the patterns revealed? Are the point locations 'clustered', 'random' or 'regular'?

### **b) Doing it yourself using Microsoft Excel™**

Three text files of  $(x, y)$  co-ordinates of some point 'events' are provided:

**Book:** These are the 12 sample data taken from Table 5.2 page 131 of O'Sullivan and Unwin (2010)

**Bank:** This is a famous data set that has been analyzed many times, notably by the statistician Brian Ripley. These data were taken from the website associated with the text by Davis (2002). It gives 47  $(x, y)$  pairs giving the location (in a projected Euclidean co-ordinate system), of dark magnetite crystals in a polished cross-section of a rock called anorthosite. The interest is in whether or not this distribution is random within the section. Co-ordinates are on a 100x100 grid, with its origin at the bottom left, but no units are given (assume cm?) and the rock in question forms part of the doorway to one of the banks in the city of Cambridge, England. The origin of these data is simply to remind you that not all 'spatial analysis', even in a GIS, must be 'geo-spatial'.

**Snow:** This is probably the most famous point data set ever to be analyzed. It consists of the locations of 578 deaths from cholera recorded by Dr. John Snow in the Soho area of London during an outbreak of cholera in 1854. Snow mapped these data as a dot map and was able to show that they clustered around a single water pump (no piped water in those days!) in Broad (now Broadwick) Street. Acting on his advice, the authorities removed the handle from the pump and the epidemic ended soon after, although it may well have already been past its peak. The events are celebrated by a facsimile of the pump and in the naming of a nearby pub the 'John Snow'. All epidemiologists and all spatial analysts should at some time make a pilgrimage to the street and have a drink in the pub that now bears John Snow's name.

Snow's work is widely regarded as the birth of scientific epidemiology, and his demonstration that the vector for cholera was water-borne led to massive investment in UK during the second half of the nineteenth century to provide safe public water supplies. Of course, the story isn't as simple as it is sometimes

suggested. For a recent scientific account, see Brody, H. *et al.*, (2000). For a 'popular' account that by Steven Johnson (2006) is highly recommended.

These data were digitized at the request of Professor Waldo Tobler (UCSB) by Rusty Dodson of the US National Center for Geographic Information Analysis from a reprint of Snow's book *On Cholera* (Oxford University Press, London).

Although the origin is at (0, 0) these data have arbitrary co-ordinates that range on X from 8.280715 to 17.938930 and on Y from 6.090047 to 16.972760. In the real world, one full unit (e.g. 1.0000) represents about 54m on the ground, so the minimum enclosing rectangle has an area of about 0.3km<sup>2</sup>. Note that the coordinate system provides for a lot of unused 'white space' around these points, which you might judge should not be included in any analysis.

1. For BOOK, BANK and SNOW produce simple dot maps. This can be done in Microsoft *Excel*<sup>™</sup>, provided care is taken to scale the X, Y axes appropriately as follows. Go FILE>OPEN>FIND and navigate to where you have saved BANK.TXT. Then chose 'delimited' and set this to 'space' with the data type set as 'general'. This should incorporate the two columns into Microsoft *Excel*<sup>™</sup>;
2. In the chart wizard, it's a simple matter to chose the (X,Y) scatter plot. What I seem unable to do is to stretch the axes on these plot such that they have the same scale (you can set the range), and have usually done this before printing by clicking on the two horizontal axes of the display and pulling them out until I get the desired equal scale;
3. If you have access to *ArcGIS*<sup>™</sup> or similar, you should be able easily to draw 'proper' maps of these three distributions;
4. In Exercise (8), five criteria were suggested to use in testing whether or not such a map is of a genuine point pattern that might be analyzed using methods to be introduced in the next section. Evaluate your result in the light of these five criteria;
5. In your own words how would you describe the patterns revealed? Are the point locations 'clustered', 'random' or 'regular'? How do they differ?
6. How do you think the choices made for the 'frame' will affect the descriptions you have given?

## **Comment/answers**

### **Book**



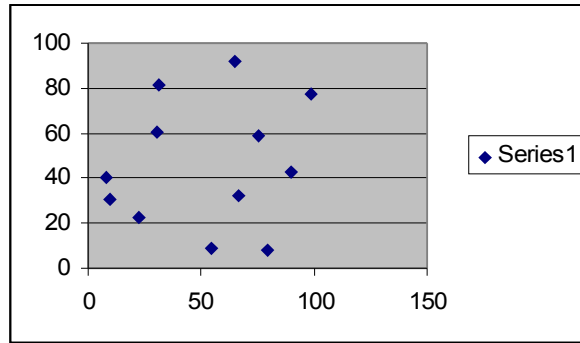


Figure 3.1 Dot map for Book

Figure 3.1 shows the distribution of events in BOOK. I have used Microsoft *Excel™*, copied into *WORD*. The box could be re-sized such that the scales on X and Y are the same, noting that the range of value on X is greater. Visually I would say that the pattern looks fairly random, but with only 12 events how can one tell?

### Ripley's BANK

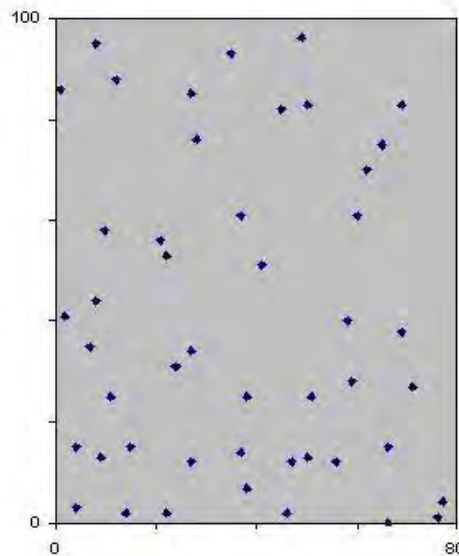
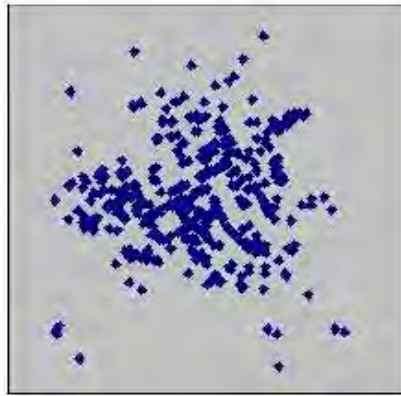


Figure 3.2 Dot map of BANK

Figure 3.2 shows the distribution of events in BANK. This is also from Microsoft *Excel™*, but scaled and transferred into *PAINT* where it has been edited a bit more. It ought to have a scale, as we are now well on the way towards a proper 'map'. Visually I'd describe it as 'more regular than random'.



**Snow's data**

Figure 3.3 Dot map of SNOW

Figure 3.3 shows the distribution of events in SNOW. The usual interpretation is that the cases 'cluster', that is they are more aggregated than random, with the clustering around a specific point, the Broad Street water pump. Note that in this hypothesis, we only have one cluster, so what is the value of the standard CSR model in this case?

One obvious point that the SNOW data show is the dependence of what we see on the 'edges of space' that we chose to use. We can make this look even more clustered by simply extending the frame. Zooming in to a subset of these same data might well make them look random or even dispersed. In the case of BANK, zooming out would gradually make them look more aggregated than random. In other words, the choice of frame is critical in the visualizations and what should be done is basically to proceed carefully unless there is a 'natural' frame.

### **Suggestions for modification**

An obvious extension is to ask students to run a kernel density estimate over these data, with three possible reasons for interest:

- a) As a means of locating 'hotspots' in the patterns;
- b) To initiate a discussion of band width, kernel function and even the appropriateness of the underlying 'geography' (For example, in the Snow case should we use street walking distances and not straight lines?);
- c) To show the value of a transformation from a pattern of discrete objects (the events) into a spatially continuous field of density estimates.

## 3.4 Exercise (10): Proportionate symbol maps

### **Aims and introduction**

Proportional symbol maps show differences in the location and magnitude of point located 'events' and are appropriate for visualizing what statisticians call a 'marked' point pattern. What we now have is a pattern of discrete point objects/events, but in each case we have an additional 'weight' attached to each event. Almost all the basic methods of point pattern analysis can be modified by use of such weights.

### **Geometry, space and level**

As Exercise (8) and (9), a set of located point objects when mapped in a metric space immediately presents complex/second order concepts referred to as distribution, dispersion, pattern, clustering, and density, but in this case we have two sources of variation in geographic space and but with a primitive/first order notion of magnitude added.

### **Intended learning outcomes**

After doing this exercise, students will

- Be able to recognize a true proportionate symbol map;
- Understand the difficulty of simultaneously associating both variation in magnitude and variation in geographic space;
- Be able to distinguish such maps and the data on which they are based from maps that use similar symbolism but to display area aggregated data;
- Appreciate the importance of the 'art and science' of cartography in determining the look of a map.

### **Resources needed**

WWW browser with access to *Google*<sup>TM</sup>.

### **Suggested student briefing**

1. Go to *Google*<sup>TM</sup> (or similar search engine) to find a proper proportionate symbol map that meets all the five conditions noted in Exercise (8). Ask yourself:
2. Is there a one to one between the dots and distinct 'events' such as the location of a crime, some facility or whatever?

3. What is the numerical variable that is attached to each event in the pattern?
4. Are the locations 'proper'? Is each dot located at the correct place where the 'event' occurred or is to be found?
5. If the conditions above are met, is it a sample or a complete enumeration or census?
6. Is the way the symbol used is related to the magnitude of the variable being displayed appropriate?
7. Can you make 'sense' of the distribution?

### **Comment/answers**

The results are likely to be much more satisfactory, cartographically speaking, than for dot/pin maps, but there is a real difference between maps in which the symbol refers to an exact spatial location (such as, for example, size-graduated circles to show the output from a series of point located factories) and those that refer to data that are an aggregate for a specified area and are usually located at some central point within the area (such as a population map of the Counties in a State). In fact, examples of the former will be hard to find. Almost always the maps found will actually have symbols (circles are favourite, but beware some of the bizarre symbols that were found) used as a form of area symbolism and their locations were at some arbitrary point (usually the centroids) within the areas to which the aggregate data refer. Exercise (22) on choropleth mapping makes some further points about this sort of data, especially the folly of mapping absolute totals when using area aggregated data. It is worth using the results to point out the problem of isolating effects related to the geography of the locations themselves at the same time as their magnitudes.

### **Suggestions for modification**

The main interest in this exercise is likely to be the weird and wonderful shapes used by some web cartographers to visualize the located quantities. It is well-known that use of even the simple circle with its area proportional to the value of the located datum can mislead. Human beings simply do not 'see' circle area in this way. The classic study and the suggested 'law' that corrects for it is by Flannery (1971).

Students may well also find maps that have as their symbols graduated pie charts showing the proportions of some constituent of the total. These can

display an enormous amount of data, but whether these visualizations are effective is moot, and might form the basis of an in-class discussion about the balance between map clarity and data volume/character.

## 3.5 Exercise (11): Centrography

### **Aims and introduction**

The best way to learn something about point pattern analysis is to do it. This exercise uses public domain software and three supplied data sets to go through typical analyses, but at the same time highlighting the practical implications of some of the difficulties. It will probably take students around 4-5 hours to complete all the tasks.

- To demonstrate computation of simple basic point pattern measures with different types of patterned data;
- To illustrate some of the problems and issues that might emerge in such use, notably the influence of the area used and the need to understand edge effects.

### **Geometry, space and level**

A set of located point objects when mapped in a metric space immediately presents complex/second order concepts referred to as distribution, dispersion, pattern, clustering, and density. This exercise uses simple arithmetic to address them.

### **Intended learning outcomes**

After doing this exercise, students will

- Realize that the boundaries of the space we choose greatly affect these types of measure;
- Understand that centrography does not explicitly capture the notion of pattern in a distribution of point events;
- Critically assess the situations in which these measures might be used to compare different distributions in the same area and/or the change in a distribution over a sequence of time slices.

### **Resources needed**

*CrimeStat III*, produced by Ned Levine Associates for use by police forces interested in the spatial distribution of crime, is available as a free download and can be used to analyze almost any point pattern, not just the distribution of crimes. Although it will compute many of the measures we have discussed, it doesn't compute a quadrat analysis or some of the more esoteric measures, such as  $G(d)$  and  $F(d)$ , nor does it have serious production graphical capabilities.

Likewise, if you have an *Apple*<sup>™</sup> machine you'll need to run it in Windows emulation mode. The graphic deficiencies can be overcome, either by use of *ArcGIS*<sup>™</sup> 'shape' (.shp) files as export and import, or by using ASCII text files (.txt) files imported into Microsoft *Excel*<sup>™</sup>.

### **Suggested student briefing**

1. Visit the website at <http://www.icpsr.umich.edu/CRIMESTAT/> and download the *Crimestat III* software. It is best to download all the associated files at the same time. Follow the instructions to install the program;

Almost all the problems you might have when using *Crimestat III* will be associated with errors made at the data description and entry stages, so it pays to take care.

In Data:

- Set file characteristics at ASCII
- Select file, say BOOK, and navigate to it
- Check the SPACE SEPARATOR and ensure that there are 0 header lines and 2 columns.

On the Data Set Up screen, take care to ensure that you:

- Set X as column 1
- Set Y as column 2
- The remaining fields should be either <none> or <blank>
- Set the type of co-ordinate system to 'Projected', units to 'm' (they are actually arbitrary).

Several of the *Crimestat* routines require either a 'reference' file, a 'measurement' file, or both. Experience suggests that although you can get some results without creating and saving these, it's often better to create do this right at the start. Note, too, that saving the parameters, available under the 'options' menu, saves both reference and measurement files. If you do this, you need to think hard about what area to input and about the density of estimates you want in, for example, the  $K(d)$  function and/or kernel density estimation routines, so you may need to revisit this step!

*You are now ready to analyze some data!*

2. Using BOOK, and to gain confidence, show that the mean center of the 12 events is at (52.575, 46.175). In addition, record the average density, and standard distance;
3. Use BANK and then SNOW to do the same things;
4. Do these numbers tell you very much? Do they help differentiate the patterns?
5. On your plot of the SNOW data, locate the mean center and confirm that it does indicate something useful.

### Comment/answers

Assuming projected data with co-ordinates in m Table 3.1 shows the results, but the exact numerical values aren't important:

| File | n   | Mean X | Mean Y | Density (m assumed) | Standard distance |
|------|-----|--------|--------|---------------------|-------------------|
| BOOK | 12  | 52.6   | 46.2   | 0.001571            | 43.98             |
| BANK | 47  | 36.6   | 40.1   | 0.006442            | 39.24             |
| SNOW | 578 | 13.0   | 4.7    | 5.500000            | 2.56              |

Table 3.1 Centrographic measures for BOOK, BANK and SNOW

Obviously one needs to convert the apparent density units into those appropriate for the particular data set. These centrographic measures tell us very little, at least in these applications. Answers will all lie close to the centre not simply of the data co-ordinates but of the frame chosen, around (50, 50). They are useful to compare patterns of 'events' when these events are of different kinds in the same geographic area, for example the locations of stores of differing types across a city area. They are also sometimes useful in tracking the evolution of a pattern over time. One minor use I can see is for the so-called 'standard deviational ellipse' (not circle) which can indicate a pattern of events that has some directional bias. Maybe, just maybe, the SNOW analysis shows a third useful thing they can do?

### Suggestions for modification

1. Read the PDF files from the manual Chapter 1 and then Chapter 2, sections I (*Data Setup*), II (*Spatial Description*) and III (*Spatial Modeling*);
2. There is no need to go further than Chapter 2, but you might also read appropriate bits of Chapter 4.1 – 4.17 (*Centrographic Statistics*,



Chapter 5.1, 5.7, and 5.40 and Chapter 8.1 – 8.14 on kernel density estimation);

3. You might also like to follow in its entirety the example given at the end of Chapter 3 (page 3.32 et seq.) using a supplied .dbf file and the 'general sample data' found in a ZIP file. If you do this take care to name the columns correctly.

## 3.6 Exercise (12): Nearest neighbor statistics

### **Aims and introduction**

In elementary texts such as Unwin (1981), the most often used measure of spatial pattern is the classic nearest neighbour statistic devised originally by Clarke and Evans (1954). The logic behind this statistic is described by O'Sullivan and Unwin (2010, pages 130-132 and 143-145). Basically the so-called *R-index* is the ratio of the mean of the observed distances from each event in the pattern to its nearest neighbour to the expected mean distance under the hypothesis that the pattern is random. A statistical significance test can be developed, since both the expected mean distance and its variance are readily obtained from simple mathematics. The approach has a number of possible traps for the unwary, not least of which are the choice of the 'frame' in which the events are considered to be present and the possible impact of unwelcome effects at the edges of the distribution when the number of events is low. This exercise uses *Crimestat III* and the same data as in Exercise (11) to illustrate these issues.

### **Geometry, space and level**

A set of located point objects when mapped in a metric space immediately presents complex/second order concepts referred to as distribution, dispersion, pattern, clustering, and density. This exercise introduces the analytical/third order concept of a spatial process.

### **Intended learning outcomes**

After doing this exercise, students will:

- Be able to conduct a nearest neighbour analysis using *Crimestat III*;
- Understand that, if the *R-index* is the ratio of the observed to expected mean distances to nearest neighbour, the 'expected' is relative to some hypothesis most obviously that of complete spatial randomness;
- Be able to assess the statistical significance of the departure of the computed *R-index* from 1.0 using Student's *t*;
- Discover how important the choice of frame is to the results obtained, preferably by showing how this is taken up into the calculations by way of the study region area used to find the expected mean distance under the null hypothesis that of complete spatial randomness;
- Note that a general test such as this might not be what is wanted in a specific case study such as John Snow's problem;
- Finally, understand that scale effects can be addressed using the same approach to examine distance to second, third, etc, nearest neighbours.

## Resources needed

*Crimestat III* together with the point pattern data sets used in Exercise (11).

## Suggested student briefing

For these exercises, unless specifically requested, DO NOT set the area in the measurement parameters section of the file set up.

1. Using BOOK use *>spatial description>distance analysis I>nearest neighbour analysis* to confirm that the mean distance to nearest neighbour for these 12 points is 21.62 as given on page ;
2. Using BANK confirm that the Clark and Evans R with no edge correction and using the *Crimestat III* default way of finding the area from the so-called minimum enclosing rectangle as  $(77-1)*(96-0) = 7296 \text{ cm}^2$  is:

$$R = \frac{d_{obs}}{d_{exp}} = \frac{7.81}{6.23} = 1.2539$$

With a t-value of 3.33, this is significantly different from random at  $p=0.001$ . Given that we have observed mean distance to nearest neighbor greater than expected, we infer that the distribution of points is 'more dispersed than random'.

3. However, there are two well-documented issues with this statistic, which make it not very 'GISable' and the consequences of both can be illustrated using these same data. First, there is a critical dependence on the area used in the calculation of the expected mean distance. In computing the index, *Crimestat III* defaults to use the area given by the range of the co-ordinates on the X and Y-axes which gives an area of  $7296 \text{ cm}^2$ . To illustrate this, repeat the analysis, but in this case in 'measurement parameters' set the area to be that of the entire frame, which is  $100 \text{ units} \times 100 \text{ units} = 10,000 \text{ units-squared}$ . Confirm that we now get:

$$R = \frac{d_{obs}}{d_{exp}} = \frac{7.81}{7.29} = 1.0710$$

With a t-value of 0.9317, this isn't statistically different from random. Note that there is no 'natural' boundary for these data: in fact I suspect

that the size was determined by Professor Ripley's unwillingness to measure any more crystals!

What does this tell you about the idea of 'randomness' in point patterns?

4. Second, there is also an effect at the edges of the distribution. For small numbers of events the effects can be quite dramatic, shifting the null value from 1 upwards into the 'more dispersed than random' range. This arises because points near the edges of the distribution are forced to find neighbors within the space, when in reality it is probable that their true nearest neighbors would be at some shorter distance outside the frame. This clearly biases the mean upwards. I know of four ways of handling this issue. The first is that used by *Crimestat III* in which such points are handled by taking the distance from a border event to the frame edge (assuming either a rectangular or circular frame) if this distance is less than any measured distance to the nearest event within the frame. Section 5.11 of the *Crimestat III* manual explains this in more detail. To see what happens, re-set the 'measurement parameters' area to zero and then run the program again, but this time ticking the 'rectangle' edge correction box. This will correct things as:

$$R = \frac{d_{obs}}{d_{exp}} = \frac{6.40}{6.23} = 1.0274$$

The second places a 'guard area' around the frame and proceeds as usual but allows points near the edge of the frame to find neighbors within the guard region. Of course, the nearest neighbour distances of these guard points are not themselves included in the analysis. The third approach uses a series of edge corrections, established by a combination of mathematics and experiment. Possibly the most elegant approach is the fourth, which wraps round both edges of the frame, to meet their opposite side and then proceeds in the usual way. Notice that we now have three possible values for the nearest neighbor statistic for these same data, dependent on what we assume about the area of the region and how any edge effects are handled by the software. If nothing else this should alert you to the need to take extreme care when using this approach, the more so if you don't actually know precisely how any GIS you use does its calculations.

5. Load SNOW and repeat the above analyses. It is instructive to experiment with different values for the region area. Does the nearest neighbor statistic help in any meaningful way in testing Snow's hypothesis?

6. A third problem with the distance to nearest neighbour is that, by taking only the nearest neighbour distances it only indicates the nature of any global patterning at this 'scale'. This deficiency can be circumvented by repeating the analysis (the mathematics is essentially the same) for successive 'orders' of neighbors and so examining the patterning at successively longer distance scales. *Crimestat III* lets you do this on the basic *distance analysis 1* screen by asking for as many neighbors as desired to be considered. Do this for both BANK and SNOW but when you have the results, use the GRAPH option to get simple plots of the change in  $R$  with order of neighbour.

### **Comment/answers**

See test above. The sensitivity of the test to the definition of the study area and with small  $n$  to edge effects comes as a surprise. It is useful to point out that Snow didn't really need to do any statistical analysis to get his point across. Steven Johnson's (2006) book should be referenced for the complete story.

### **Suggestions for modification**

In Snow's second map, the usual display of a point pattern that we have seen so far was supplemented by a line enclosing all the houses that from his local knowledge Snow knew to be closer to the infected Broad Street pump than they were to any other. In essence this was what future spatial analysts would call a Voronoi diagram or Thiessen network and it showed with astonishing clarity that not only did the cases cluster, they clustered around the suspect pump with only a few exceptions of deaths to people for whom the pump was not the nearest source of water. The pattern of streets in 1854 wasn't the same as it now is, but a rough approximation to the Snow's border can be obtained by computing and displaying the Voronoi/Thiessen network on top of a dot/pin map of the deaths.

If students have access to a GIS capable of computing the Voronoi tessellation a display of these point data with the Voronoi diagram for all 13 pumps in the area is a very convincing demonstration of the power of simple, almost geometric 'spatial analysis'. *3DField*, used in Exercise (30) will compute this as well.

## 3.7 Exercise (13): Ripley's K statistic

### **Aims and introduction**

In view of the problems with the single number approach, it is hardly surprising that spatial statisticians have tried to characterize pattern using distance functions such as those discussed in the text. Of these, Ripley's  $K(d)$  is the most satisfactory. This exercise is a simple introduction to the approach. The theory behind the approach is introduced in O'Sullivan and Unwin (2010, pages 135-137 and 146-148).

### **Geometry, space and level**

A set of located point objects when mapped in a metric space immediately present complex/second order concepts referred to as distribution, dispersion, pattern, clustering, and density. This exercise uses both the analytical/third order concept of a spatial process and visualization to address them.

### **Resources needed**

*Crimestat III* together with the point pattern data sets used in Exercise (11)

### **Intended learning outcomes**

After doing this exercise, students will:

- Be able to conduct a point pattern analysis using Ripley's  $K(d)$  function approach in *Crimestat III*;
- Understand the process by which a mean value of  $K(d)$  at some distance  $d$  is estimated;
- Be able to interpret a graph of computed values of  $K(d)$  against distance,  $d$ ;
- Understand how theoretical values for a random distribution can be derived and used to convert  $K(d)$  into the  $L(d)$  function in which the theoretical expectation for a random pattern is zero at all distances;
- Understand that this approach enables investigators to examine the scale at which a pattern of point events can be said to 'cluster';
- Be able to assess the statistical significance of the departure of the computed  $K(d)$  using the randomization approach.

### **Suggested student briefing**

Much modern work in point pattern analysis uses the  $K(d)$  function approach developed by Ripley (1976) which is based on all the distances between events in the pattern. Computation of Ripley's  $K(d)$  is easy to explain but very tedious to do except by computer. All we do is to place circles, of each radius  $d$ , centered on each of the events in the pattern and find the number of events that fall into that circle. Doing this with the same radius  $d$  centered on each and every event allows calculation of a mean number for this distance. All we then do is to repeat this for a series of distances. Each mean count is divided by the overall study area event density to give  $K(d)$ . Formally this is:

$$K(d) = \frac{\sum_{i=1}^n \#[S \in C(\mathbf{s}_i, d)]}{n\lambda}$$

$$= \frac{a}{n} \cdot \frac{1}{n} \sum_{i=1}^n \#[S \in C(\mathbf{s}_i, d)]$$

Remember that  $C(\mathbf{s}_i, d)$  is a circle of radius  $d$  centered at  $\mathbf{s}_i$  and the operation specified by the numerator is the 'number of' ( $\#$ ) 'events',  $S$ , 'included in' ( $\in$ ) that circle. Because *all* distances between events are used, over a range of distances, this function is much more informative about the patterning than any single number such as the  $R$ -index ever could be, but what values would we expect if the pattern is random? In fact this is easy to calculate, at least if there are no problems with edge effects and the definition of the area of interest. Since  $\pi d^2$  is the area of each circle, and  $\lambda$  is the mean density of events per unit area, the expected value of  $K(d)$  is simply

$$E(K(d)) = \frac{\lambda \pi d^2}{\lambda}$$

$$= \pi d^2$$

Because the expected function depends distance *squared*, both the expected and observed  $K(d)$ , this function can become very large as  $d$  increases and it is difficult to see small differences between expected and observed values when they are plotted on appropriately scaled axes. The usual way round this problem is to convert the expected value of  $K(d)$  to zero, by dividing it by  $\pi$ , taking the square root, and then subtracting  $d$ . as

$$L(d) = \sqrt{\frac{K(d)}{\pi}} - d$$

The result is another function of distance, this time called the  $L$  function. If the pattern is random performing the same operations on the observed values of  $K(d)$ , we should get values near zero. Where  $L(d)$  is above zero, there are more events at the corresponding spacing than would be expected under IRP/CSR; where it is below zero, there are fewer events than expected.

*Crimestat III* will provide an estimate of  $K(d)$ , but to do so it has to have a defined reference file of a grid of locations. It also produces all the data needed to plot  $L(d)$ , together with a simulate envelope around this for use in evaluating the significance of departures from the random expectation for  $L(d)$ , which is zero.

1. Start *Crimestat III*
2. Using the *distance analysis 1* screen, set up and compute the  $K(d)$  function for both BANK and SNOW. The output is the  $L(d)$  (called  $t'$  on screen)

One problem with these functions is that at large  $d$  edge effects enter into consideration, where a substantial proportion of each circle is outside the study area. In these cases by definition there are no events outside the study region, so the number of events in the circles is lower than would be expected based on an assumption of uniform density. This is a problem in almost all work in spatial statistical analysis: we almost have to either to break some assumption made in the derivation of the theoretical values or to attempt some corrections that take them into account. Nowadays, plentiful computer power enables us to use a simulation approach to this problem. No matter which statistic we are interested in, the procedure is always very simple: use a computer to generate a large number of patterns according to some hypothesis we have about the process. In this case we'd simply use the computer's random number generator to give randomly located point 'events'. Next, for each pattern we measure the statistic to give an expected distribution of values against which the observed values of the same statistic can be compared. This approach lacks mathematical elegance, but it enables us to allow for things like edge effects, simply by using the same study region in the simulations as in our observed data. Such a simulation approach is known as a *Monte Carlo procedure*, and is widely used in modern statistics, but it is computationally very intensive especially when the number of events in the observed and hence also the simulated patterns is large. There is also controversy about how many simulated patterns should be used. Some purists recommend use of a very large number, say 999, whereas those willing to take a bigger risk in their assessment might only use 99, but it really rather depends on the extent to which the statistical; assessment is important.

3. For both BANK and SNOW use *Crimestat III* to compute and plot the  $L(d)$  function. In doing this, use the simulation routine with, say 99 runs to get and plot an estimate of a confidence envelope that can be used to assess the significance of the observed values. Note that use of the GRAPH button will plot the so-called simulation envelope and that it is then easy to see at what distances the observed  $L(d)$  is outside and thus



indicative of a distance scale at which the pattern is more/less regular than expected;

4. In doing this note that for SNOW with  $n = 578$  the simulation will take a perceptible length of time even on a fairly quick machine!
5. Describe how the two patterns differ and the extent to which these results confirm your previous analyses with the same patterns.

### Comment/answers

Figure 3.4 shows the results for BANK (left) and SNOW (right) using *Crimestat III* with no edge corrections and with simulation envelopes based on 99 runs.

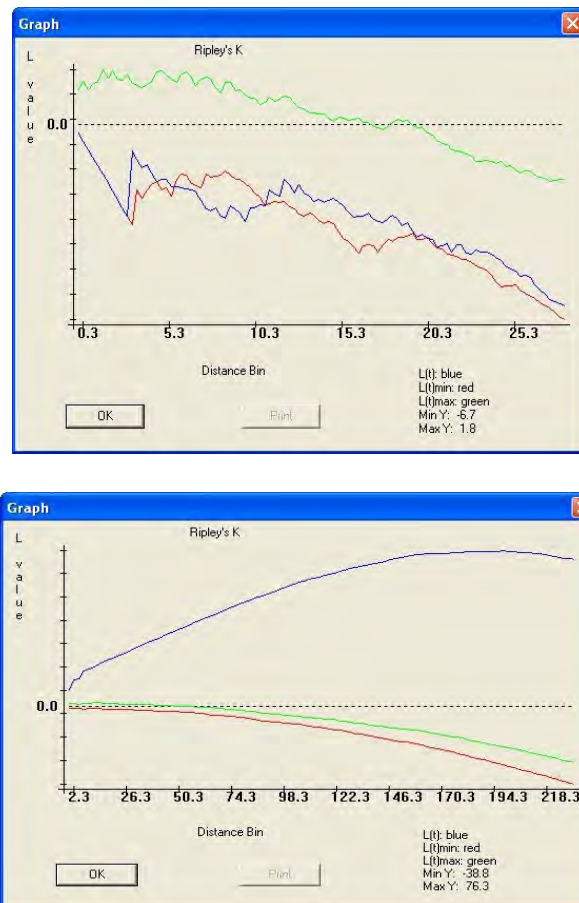


Figure 3.4 Results for Ripley's  $K(d)$  function for BANK and SNOW

On this display the observed L function is in blue and the green and red lines show the extreme values above and below in the simulations. It can be seen that the observed function for BANK is often within the simulation envelope but is always less than zero, indicating that at all distance scales there are fewer events at the specified distance than expected. This should confirm the idea that these data are more regular than random. For SNOW at every distance scale the L

function is well above zero, indicative of a pattern of events that is more aggregated (clustered) than random.

### **Suggestions for modification**

A variation on Ripley's  $K(d)$  function called the *O-ring statistic* (Wiegand and Moloney, 2004) or the neighborhood density function (Perry *et al.*, 2006), has been used. This is easily computed by noting that the original  $K(d)$  function is cumulative with the proportion of events from 0.0 to 1.0 at each circle radius plotted as a function of the radius,  $d$ . These more recent functions plot the actual proportion in a series of annuli centered on each event.

### 3.8 Further Exercises: Other things you can do with *Crimestat III*

This doesn't exhaust the options for work on point patterns based on *Crimestat III*. There are perhaps three additional options that you might like to explore using these same data at some future time:

1. Computing a kernel density estimation and then exporting the results for visualization into a GIS or other mapping program such as *3Dfield* (see Chapter 6);
2. So-called 'hot spot' analysis, which uses a modification of standard cluster analysis to find areas of above average local spatial point density, or 'hot spots'. In both criminology and epidemiology such concentrations have obvious practical significance;
3. Computing and exporting selected distance matrices for analysis in other software.

There is a useful 'how to do it' guide by Luc Anselin *An Introduction to Point Pattern Analysis using CrimeStat* that shows how to export from that package in *ArcView*<sup>™</sup> or *ArcGIS*<sup>™</sup> available at <http://geodacenter.asu.edu/system/files/points.pdf>.

### 3.9 References

Brody, H. *et al.*, (2000) Map-making and myth-making in Broad Street: the London cholera epidemic 1854. *The Lancet*, 356 (9223):64-68.

Davis, J.C. (2002) *Statistics and Data Analysis in Geology*, (3<sup>rd</sup> Edition, John Wiley & Sons)

Flannery, J.L. (1971) The relative effectiveness of some common graduated point symbols in the presentation of quantitative data' *Canadian Geographer*, 8: 96-109. See also Unwin, D.J. (1981) *Introductory Spatial Analysis* (London: Methuen), pages 33-36

Johnson, S. (2006) *The Ghost Map* (London: Penguin)

Levine, N. and Associates (2007) *Crimestat III: a spatial statistics program for the analysis of crime locations*, at <http://www.icpsr.umich.edu/CRIMESTAT/>

O'Sullivan, D. and D. Unwin (2010) *Geographic Information Analysis* (2<sup>nd</sup> Edition, Hoboken, NY: John Wiley and Sons)

Perry, G.L.W., Miller, B.P. and N.J. Enright (2006) A comparison of methods for the statistical analysis of spatial point patterns in plant ecology. *Plant Ecology* 187: 59-82

Ripley, B. D. (1976), The second-order analysis of stationary point processes, *Journal of Applied Probability*, 13, 255–266.

Wiegand, T and K.A. Moloney (2004) Rings, circles and null models for point patterns analysis in ecology. *Oikos*, 1104: 209-229

# Chapter 4: Lines on maps

## 4.1 Introduction

The schema of Chapter 1 developed the idea that line objects have as their basic relation the idea of connection in what was called a network space. This chapter deals with the first order concepts of length and direction as well as the second order concept of connection.

Line objects, that is, entities that have the single geometric dimension of length,  $L^1$ , are common in geography with perhaps the most significant examples being rivers, roads and so on. Usually, these objects are part of an interconnected set or *network* such as a drainage basin or communications network. This means that interest in their properties can be at several levels. First, we might be interested solely in the line objects themselves, for example, in their length and orientation in a defined space. Second, we might be interested in the mechanisms by which such line objects become connected into networks. Third, as in geomorphological analysis of drainage networks or in shortest path and related analysis, we might be interested in properties of some existing network taken as 'given'. Finally, interest might be in the flows of materials, energy and information along such lines and through any networks they create.

This makes the representation and analysis of such objects one of the most complex and challenging. Since the property of 'connection' is of course very general, the more so when one realizes that in essence it is the same as that of a relation on a set introduced in Section 1.1 and so is one of the most important general concepts in science. Surprisingly, the geographical literature hardly scratches the surface of this richness. The old text by Haggett and Chorley (1969) describes early work in the field. Keith Tinkler's 1977 monograph (Tinkler, 1977) contains some forward-looking ideas and the popular books by Watts (1999, 2003) give an idea of current work outside of geography. In GIS&T, there is a toolkit of methods for spatial analysis on a network (SANET) available at <http://sanet.csis.u-tokyo.ac.jp/index.html>. See also Okabe and Satoh (2009)

The exercises introduced in this chapter can only skim the surface of all this richness.

## 4.2 Exercise (14): Lines on maps

### Aims and introduction

In describing lines we have three new spatial concepts to add to the simple idea of location that sufficed to describe a point pattern. These are *length* or *distance*, *direction* and *connection*. Before dealing in detail with these, it is important to make a clear distinction between the lines used on maps to *represent* areal units or surfaces (boundary and iso-lines), and true linear objects in the sense that we introduced them in Section 1.1

### Geometry, space and level

This first exercise uses mapped examples of line objects that exist in a metric, usually Euclidean space but which if connected define a network space. Our concern is solely with the primitive/first order concepts of length and direction.

### Intended learning outcomes

After completing this simple exercise, students will be able to

- Differentiate between lines used as *cartographic symbols* and genuine *line objects*;
- List various ways by which such objects are represented on maps.

### Resources needed

Topographic map or map extract delivered over WWW as for example from <http://www.streetmap.co.uk>

### Suggested student briefing

1. Find a topographic map or map extract (from <http://www.streetmap.co.uk> or similar website) at a scale of around 1:50,000 . It will be a help if the map represents an area that you have visited or know well;
2. Make a list of all the true line objects shown on it. You should assemble a reasonably long list! In each case describe how the line object is represented cartographically;

3. To fix the idea further, list all the examples you can find on the same map of lines used to depict the edges of areas or the heights of surfaces.

### **Comment/answers**

Roads, railways, river courses, rights of way, electricity transmission lines, pipelines but not contours, or boundaries! This is an almost trivial exercise but it is one that experience suggests frequently exposes a failure to distinguish between representation of some area or surface entity by lines drawn on paper and a representation of a genuine linear object. For any given topographic map series the lazy way to complete the exercise is of course to use the map key. Surprisingly few students realize this.

### **Suggestions for modification**

A possible extension is to look at the nature of the line objects, on some scale of 'wiggleness' from very straight through to very irregular.

## 4.3 Exercise (15): Measuring length

### Aims and introduction

Many people, instructors and students alike, will find this exercise simply beneath their dignity, but it is a useful way to look at error in almost any measurement. Use of Gmap-pedometer gives a great deal of interest and flexibility.

In order to measure the length of a line object we have to break it down into a chain of points connected by straight line segments, given various names such as *arc*, *segment*, and *edge*. This means that visually obvious turning points along a line are recorded as points, and the length of each segment is calculated by simple Pythagoras as in Figure 4.1:

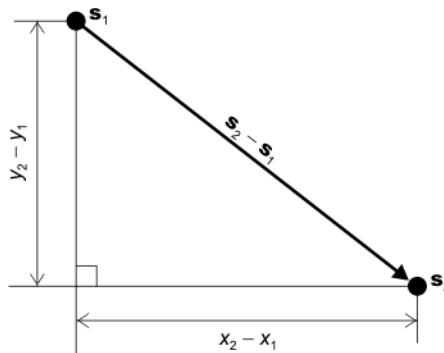


Figure 4.1 Straight lines and Pythagoras  
(Source: after O'Sullivan and Unwin (2003) Figure 6.4)

$$d_{12} = \|\mathbf{s}_2 - \mathbf{s}_1\| = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

The sum of a number of these straight line segments gives the total length of the line.

### Geometry, space and level

This exercise also uses mapped examples of line objects that exist in a metric, usually Euclidean space. Our concern is solely with the primitive/first order concept of length.

### Intended learning outcomes

After completing this exercise, students will understand:



- How all methods of measuring the length of a line object ultimately depend on a process of digitization into individual, straight line segments;
- The use of simple software in effect to perform what used to be called 'heads up' on-screen digitizing;
- That there are different ways of doing this and that each has its own characteristics, and that
- Different methods yield different results. Measurements should be accompanied by a statement of the conditions employed.

### **Resources needed**

Either a topographic map or access to the website at <http://www.gmap-pedometer.com>

### **Suggested student briefing**

This exercise is intended to illustrate how we can digitize line objects and measure their length. A simple practical example would be a recreational runner/jogger wishing to find how far he or she has run along a particular route.

1. Access the website at <http://www.gmap-pedometer.com>. You will see that this is an add on to Google Earth™ designed to allow you to measure distances along line objects (such as paths, rivers, roads and routes you have taken or intend to take. (Alternatively, but at some extra effort find a topographic map and do the exercise manually);
2. Navigate to somewhere known to you and select a line object whose length you wish to measure. This could be your walk to University, regular run, or some line feature such as a mapped stream. Note that you have access to several representations of the feature in different map projections and at different scales ('zoom level', but note that several are just on-screen enlargements of the same map);
3. Use the tool provided to digitize this line into short, straight line segments, measuring the length as you go along in either Imperial (miles) or metric (km) units. It is worth noting at this point that doing this with a semi-automatic digitizer or 'head up' using an on-screen map (q.v.) is analogous to the old fashioned method of using the edge of a piece of paper marking the turning points as you go along, which is the best way to do this exercise from a paper map. Exercise (16) has an even quicker way of doing this using dividers set to a constant scale distance;

4. Now repeat the exercise using the satellite image of the same route and at different zoom levels, recording your answer at each step;
5. Do the lengths you obtained differ and if so why?

### **Comment/answers**

There are several reasons why different attempts to measure the length will give different results of which the selection of points is but one. If using a map, whether on screen or not, it may be that the features used are represented in some way that is not planimetrically correct or has already generalized the feature. It is useful to point out that the 'wigglier' the line, the more likely is the measure to be in error.

### **Suggestions for modification**

An obvious extension is to measure the same line feature in the field using a standard hand held GPS unit or, if very pedantic, a full electronic distance measuring instrument. The exercise can also be used in class by asking the entire class to measure the same feature and using the pooled response to illustrate the classic 'curve of error'.

## 4.4 Exercise (16): Fractals

### Aims and introduction

Exercise (15) will have shown that there is some arbitrariness in how we represent a line object and measure its length. Adding up the lengths of line segments gives different answers, depending on which points are selected. There is a very important general result lurking in this problem. It has long been known that the irregular lines often studied in geography sometimes have lengths that appear to increase without apparent limit the more 'accurately' that we measure them: this is the well-known property of a 'fractal' object. The issue can be illustrated by the following example in O'Sullivan and Unwin (2003, page 142-150) and makes for an interesting student exercise.

### Geometry, space and level

This exercise uses mapped examples of line objects that exist in a metric, usually Euclidean space but which if connected define a network space. Our concern is solely with the primitive/first order concept of length, but of course one of the lessons learned is that the dimensionality of some line objects could well be fractional and greater than unity.

### Intended learning outcomes

After completing this exercise, students will:

- Understanding how sinuous lines can be thought of as 'fractal';
- Appreciate that the concept of a fractional dimension provides a means by which sinuosity can be measured;
- Be able to estimate the fractal dimension of a sinuous line using the so-called Richardson Plot;
- Say why this has important implications for some measurement in geography.

### Resources needed

Maps or other representations of very sinuous natural line objects such as a river or a linear boundary such as an indented coastline. Students familiar with the works of Douglas Adams might want to use map of Norway's 'Fjordland' with its 'crinkly edges'.

Dividers, pencil and paper, log tables and, ideally software able to complete a simple regression analysis (such as Microsoft *Excel's Analysis* add in).

## Suggested student briefing

1. Find a reasonably detailed topographic map at a scale something like 1:25,000 or 1:50,000 and select a length of river or coastline 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! Around a 20km (equivalent) length is about right and will not involve you in too much work;
2. Now set a pair of dividers at a large equivalent distance on the ground, say 1km and 'walk' them along the river counting the number of steps. We will call this set divider separation and its equivalent as a distance on the ground the 'yardstick'. Record the yardstick equivalent length and number of steps. It is permissible to record the final step as a fractional one;
3. First can you see that the number of steps multiplied by the ground equivalent of the yardstick separation is an estimate of the length of the line itself? Experience suggests that measuring length of a map using this approach is both easier and simpler than the usual methods with bits of string and/or the edge of a piece of paper;
4. Repeat using a halved yardstick, at say 500m equivalent; You should realize that this is also an estimate of the line length made at what in some sense is a higher resolution of analysis;
5. Repeat again and again until the yardstick becomes so short that the experiment is impractical;
6. You should now have a table of values for each yardstick of paired measure of the yardstick length and the apparent length of the line as mapped;
7. Convert both the numbers of steps and the yardstick length to their logarithms (Base 10) and plot the resulting numbers with the  $\log(\text{number of steps})$  on the vertical axis and  $\log(\text{yardstick length})$  on the horizontal;
8. If you have access to a spreadsheet program such as Excel™ you should be able to do this using that software. Hopefully, the points will fall roughly along a straight line but success isn't guaranteed!

9. Finally, use the spreadsheet (or a straight edge and a good eye) to fit a linear regression line to your data. The regression equation can be used to estimate the fractal dimension. It is of the form:

$$\log[L(s)] = (1-D)\log(s) + b$$

In which the 'Y' is  $\log[L(s)]$  (the total estimated length),  $b$  is the 'intercept constant', which isn't of concern,  $D$  is the fractal dimension and  $s$  is the step length (yardstick)

Note that since the length estimated length,  $L$ , gets less the longer the length you set for  $s$ , the slope of the line, shown by Mandelbrot to be  $(1-D)$ , is negative. You need to do a little arithmetic to get  $D$ .

### Comment/answers

Some results for a stretch of the New Zealand coastline from O'Sullivan and Unwin (2010, page 13) are shown in the Table 4.1.

| Resolution $L$<br>(km) | #segments, $N$ | Total length<br>(km) |
|------------------------|----------------|----------------------|
| 2.5                    | 132            | 330                  |
| 5.0                    | 52             | 260                  |
| 10.0                   | 18             | 180                  |

Table 4.1 Length measures for a stretch of New Zealand Coastline

On a double log plot the three points in this case lie roughly on a straight line, whose negative slope, estimated by linear regression, is 1.4372. This provides an estimate of the fractal dimension of this somewhat 'crinkly edged' coastline.

### Suggestions for modification

There are several issues related to this exercise that can usefully be introduced. First, note that the exercise and almost all the academic work on geography exploring the fractal concept uses a map representation of the feature whose fractal dimension is being estimated. This conflates the true fractal property with any cartographic generalization employed when the map was created. In reality, we should only do the experiment on the actual coastline. Any stored representation, such as the map we started with, has a limiting resolution such that ultimately it is composed of a large number of straight lines of geometric dimension 1. An interesting experiment in cartographic generalization is to use the same approach to

study the same line objects over using maps at different scales, for example in UK at 1:10,000, 1:25,000, and 1:50,000.

The second issue that can be explored is that of measurement and measurement error. What is the 'correct' length of this coastline? This may seem pedantic, but in fact it is a serious question when working with any digitally stored line data. As a sometime marathon runner, the example I always give is that of the famous blue 'running line' you see painted on the road at the London event. In fact, most runners will actually run further than the 42km they sign up to! How can you compare the lengths of curves whose lengths are indeterminate; or which really is the longest river in the world?

There is a superb website for all things fractal created by Frame, Mandelbrot and Neger at <http://classes.yale.edu/fractals/> that provides Java applets to illustrate lots of related ideas.

## 4.5 Exercise (17): Direction

### Aims and introduction

The previous two exercises should convince students that the fundamental spatial concept of length isn't as simple as perhaps they first thought. Line objects illustrate a second spatial concept, that of *direction/orientation* that turns out to be equally tricky. In every-day language and geography we use numerous terms to describe direction such as 'left/right', 'above/below' often rather loosely to refer to spatial direction that would be better described by some compass direction such as 'west/east', 'north/south'. Greater precision is provided if these directions are measured as angles in degrees ( $^{\circ}$ ) relative to some understood reference, most commonly taken to be north and measured clockwise from this reference direction. Angular measure gives two problems one of which is very apparent in this exercise, the other of which might well give less mathematically-inclined students problems in the calculations suggested. The first is that in angular measure the numbers repeat like the hands of a clock every  $360^{\circ}$ . Thus, for example the angular difference between the directions  $359^{\circ}$  and  $1^{\circ}$  isn't  $359-1 = 358^{\circ}$  but  $2^{\circ}$ . This gives periodic nature of the concept gives problems when we try to calculate summary measure such as the mean direction of a series of line objects. The second arises when we have to record the angle through which the minute hand of a clock passes each day. Obviously, this is not  $360^{\circ}$  since in 24 hours the minute hand will sweep through this angle 24 times giving an 'angle' of  $8640^{\circ}$  ! The alternative unit of angular measure that gets over this problem is the radian, defined such that there are  $2\pi$  radians in a circle and one radian is  $57.296^{\circ}$ . In a day, the minute hand of a clock sweeps  $24 \times 2\pi = 150.8$  radians. When using the computer in the following exercise it would be nice to shield students from this problem, but as will be seen the standard functions for necessary trigonometric functions have their arguments expressed in radian measure.

### Geometry, space and level

This first exercise uses measured pure directions. Our concern is solely with the primitive/first order concept of direction.

### Intended learning outcomes

After completing this exercise, students will understand:

- How direction can be measured in degree and radians;
- How to compute a directional average and its relative 'strength';

- That some problems arise in geography, and particularly physical geography, for which such analysis is necessary;

Optionally, the exercise can also be used as a simple introduction to using formulae in Microsoft *Excel*<sup>TM</sup> and as a revision class in basic trigonometry!

### Resources needed

Davis (2002, page 317) explores these issues very carefully and presents some data for the recorded directions of 51 glacial striae in a 35km<sup>2</sup> area of southern Finland. These are recorded as angular directions in degrees clockwise from north as in Table 4.2.

|    |     |     |     |     |     |
|----|-----|-----|-----|-----|-----|
| 23 | 100 | 126 | 134 | 155 | 212 |
| 27 | 105 | 126 | 135 | 157 |     |
| 53 | 113 | 127 | 137 | 163 |     |
| 58 | 113 | 127 | 144 | 165 |     |
| 64 | 114 | 128 | 145 | 171 |     |
| 83 | 117 | 128 | 145 | 172 |     |
| 85 | 121 | 129 | 146 | 179 |     |
| 88 | 123 | 132 | 153 | 181 |     |
| 93 | 125 | 132 | 155 | 186 |     |
| 99 | 126 | 132 | 155 | 190 |     |

Table 4.1 Directions of glacial striae in southern Finland  
(Source: Davis, 2002, page 317)

These data are in the text file *striae\_in\_finland\_1* that can be imported into a spreadsheet such as Microsoft *Excel*<sup>TM</sup> and a map showing their locations forms Figure 5.13 of Davis (2002). In addition you will need either access to Microsoft *Excel*<sup>TM</sup> or similar spreadsheet and/or trigonometric tables.

### Suggested student briefing

It is well-known that glacial striae can be used as a surrogate for the direction of ice flow in both valley glaciers and ice sheets. The data in *Striae\_in\_Finland\_1.txt* relate to 51 measure of the vector direction of striae in a small 35km<sup>2</sup> area of southern Finland. Each individual measure is recorded as an angle in degrees (°) clockwise from north and each reflects not only the general motion of the ice sheet, but also local deviations. A useful estimate of the overall ice sheet direction might be the mean of these directions. However, directions are tricky for two reasons that concern us here. First, they are relative to some assumed direction and in most



cases in geography we use one or other of the various 'norths' (grid, magnetic, true?) that would serve this purpose. We can then measure direction as an angle in degrees ( $^{\circ}$ ) clockwise from this from  $0^{\circ}$  to  $360^{\circ}$ . Of course these are the same direction, showing that angular measure is periodic. This gives a difficulty in calculating any standard summary measures such as the mean direction. Simply adding up the bearings of all the observed vectors and dividing by the number may give misleading results. The solution is shown in Figure 4.2 and is to resolve each direction of the vector OB on the figure into two 'components' one expressing its 'northerliness', which we'll call  $V(n)$  (the line OA on the figure) the other its 'easterliness', which we'll call  $V(e)$  (the line OC).

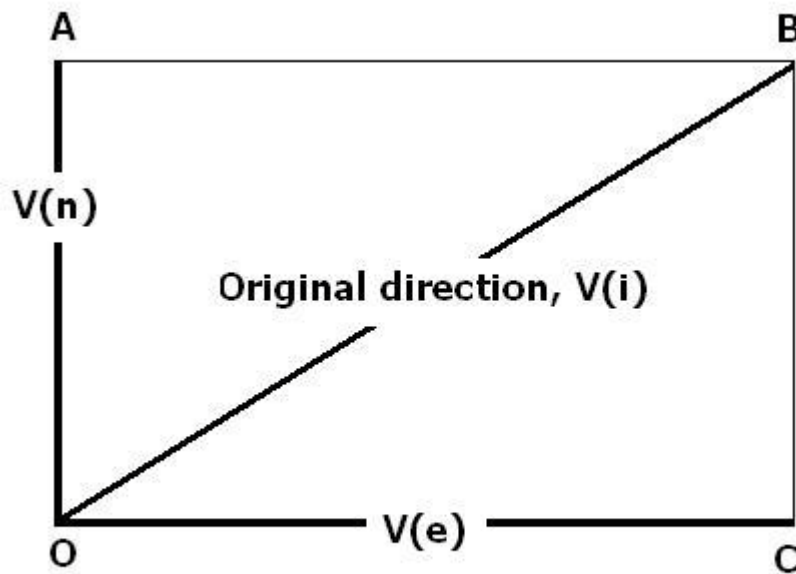


Figure 4.2 Resolving a direction into orthogonal components

These can then be separately averaged and the averages in each direction are then recombined to produce an average vector from the observations. The figure shows how this works. If we have a set of observations of path direction, then we set some arbitrary datum direction (typically north) we resolve each and every direction into its two components  $V(n)$  and  $V(e)$  by taking the sines and cosines of all the observations, and then add each set together to get two totals:

$$\begin{aligned}\sum V(n) &= \sum \cos \theta_i \\ \sum V(e) &= \sum \sin \theta_i\end{aligned}$$

The tangent of the mean direction is then given by:

$$\tan \theta_R = \frac{\sum V(e)}{\sum V(n)}$$

On Figure 4.2 the line OA points to the north the angular bearing that we use is given by AOB (and OBC).

We will illustrate this problem by determining the mean direction of a set of 51 measurements of glacial striae directions taken from a 35 km<sup>2</sup> area of southern Finland (see Davis, 2002, page 317 for a map).

1. The easiest way to do all this uses a spreadsheet program such as Microsoft *Excel*<sup>TM</sup>. The original data are in the text file *striae\_in\_finland\_1*. Incorporate these into Microsoft *Excel*<sup>TM</sup>;
2. Next create a second column in which you put the cosine of the each of the angular bearings. Obviously we are assuming here that you know how to do this, but there is a bit of a catch in that the so-called 'argument' of the cosine function in Microsoft *Excel*<sup>TM</sup> needs to be not in degrees but in an alternative measure of angle called radians. The appropriate formula first converts the angle into radian measure and then finds the cosine as in '=COS(RADIANS(A1))' etc. Sum this column to get  $V(n)$ ;
3. Repeat forming another column this time of the sines of the angles and sum to get  $V(e)$ . All that will be different is that the formula used will have SIN in it rather than COS.
4. If you are doing this by hand it's just a bit harder and you'll need to know how to look up sines and cosines for angles that are greater than 90°;
5. Finally divide  $V(e)$  by  $V(n)$  to get in radian measure the tangent of the mean direction. The angle in radians that has this tangent is returned by the ATAN function and this can be converted back into degrees using the formula '= DEGREES(ATAN(X))' where X is the location of your division;
6. Getting from this to the actual direction needs a further step, which in this case is to subtract the numerical value from 180 to give the mean direction. What answer do you get? What do you think it might say about the direction of flow of the ice sheet that created the striae?

This kind of analysis has been used to look at transportation routes, but by far the most common applications are in sedimentology, where the large

particles, or clasts, in a deposit show a preferred orientation that can provide significant environmental information

### Comment/answers

The first part is easy, provided the angles are converted into radian measure and we end with the expression:

$$\tan \theta_R = \frac{31.63672}{-25.7933}$$

Getting from this to a mean direction may well challenge students' knowledge of trigonometry and unfortunately Microsoft *Excel*<sup>TM</sup> doesn't help (Use of scientific calculator with the required functions can automate this step). The required arctangent is -50.8097, which is a direction of 129.18° clockwise from north. The full calculations are laid out in the Microsoft *Excel*<sup>TM</sup> sheet *striae\_in\_finland.xls*

### Suggestions for modification

What this mean does not do is account for the dispersion around the circle, in other words what sedimentologists studying the direction of clasts in river and glacial deposits call the *preferred orientation*. An obvious extension is to find this, using a recombination of the sin and cosine sums to get the average, or resultant vector magnitude:

$$V_R = \sqrt{\sum V(e)^2 + \sum V(n)^2}$$

whose mean turns out to be almost exactly 0.8. A further extension would be to ask how statistically significant different from random is such a preferred orientation. Davis's book (2002, page 322) has further directional data for some strange geomorphological features, the sub-parallel ellipsoidal depressions on the Atlantic Coastal Plain in North Carolina called the "Carolina Bays" that can be explored using similar methods. The classic early paper on the analysis of directional data is Pincus (1956).

## 4.6 Exercise (18): Analyzing tree networks

### Aims and introduction

Line objects seldom occur in isolation and are usually found as part of connected networks giving rise to the spatial concepts of network structure and connection. A great many geographical processes take place over a network of connected lines, such as transport and other communications networks, such that analytically it is usually more sensible to develop a network geography in which 'distance' is represented by the property of connection rather than by any metric measure. The analogy with the concept of adjacency explored in Sections 2.4 and 2.6 should be obvious.

A very important type of network, at least in geomorphology, is a branching *tree* of connected lines, where no closed loops are present and in the 1970s it was fashionable to examine the characteristics of river networks, which are probably the best example of this type of structure (see, for example Werritty 1972, Gardiner 1975, Gardiner and Park 1978). The classic analysis applied to tree networks is *stream ordering*, most of which is closely related to a scheme proposed by Horton (1945) and developed by Strahler (1952)

### Geometry, space and level

This exercise uses mapped examples of line objects that exist in a metric, usually Euclidean space but reduces the pattern of connection to a network that can be represented as a topological graph. Our concern is with the complex/second order concept of connection.

### Intended learning outcomes

After completing this exercise, students will be able to:

- Explain the difference between a *tree* and a *graph*;
- Describe how *connection* in a tree structure can be *ordered* using Strahler's 1952 modification of the original method due to Horton (1945);
- Understand that the order of a stream is a measure of its relative position in a connected network and is essentially a topological measure;
- Verify the so-called law of stream numbers for a natural river network and hopefully also understand that this is by itself not very useful 'law'.

## Resources needed

Source maps at 1:25,000 or 1:50,000 on which river networks are represented. Tracing paper and perhaps also spreadsheet program for display and analysis of the results.

## Suggested student briefing

1. From a topographic map at a scale of 1:25,000 or 1:50,000 (or nearest equivalents), trace off a drainage network represented by (usually) the 'blue lines'. Try to find a reasonably large basin with, say, about 50 or so 'sources' where streams begin;
2. Now 'order' this network using the so-called 'Strahler' method as shown in Figure 4.3. To do this, give all the fingertip tributaries the order '1'. Where two such 1<sup>st</sup> order streams meet, the stream that results becomes 2<sup>nd</sup> order, and should be labeled as such. Note that any 1<sup>st</sup> order tributaries that join a 2<sup>nd</sup> order stream will not change its order until it meets another 2<sup>nd</sup> order stream, when the stream becomes '3<sup>rd</sup> order';

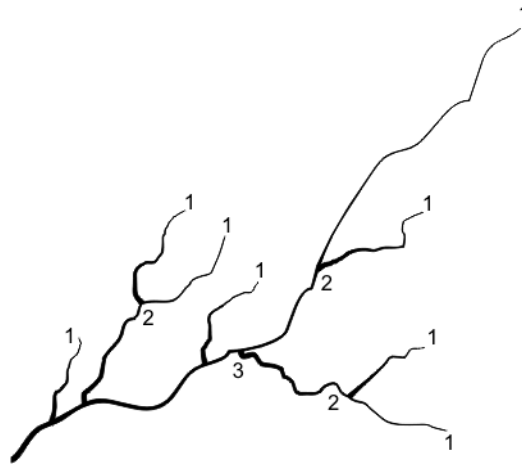


Figure 4.3 Strahler's stream ordering method  
(Source: after O'Sullivan and Unwin, 2003, Figure 6.8)

3. Continue until all the streams have been ordered in this way and there is just one stream with the highest order reached;
4. Now count the number of streams in each order. In the example this count is eight 1<sup>st</sup>, three 2<sup>nd</sup> and one 3<sup>rd</sup>;

5. Finally, plot a graph of the logarithm of the number of streams in each order on the vertical axis against stream order (1, 2, 3, ...  $n$ ) on the horizontal axis.
6. What does the resulting plot look like? Do you get a straight line?

### Comment/answers

Ordering schemes allow comparisons between stream networks. Horton himself found that when classified in this way drainage networks obey various empirical 'laws'. For example, the numbers of streams of different orders from the highest downward, closely approximates a geometric series. A geometric series is one like 1, 3, 9, 27, and 81 in which each term is obtained by multiplying the previous one by a fixed amount. In this scheme there is always just one stream of the highest order and the numbers of streams with other orders corresponds closely to such a series. If a plot of the  $\log_{10}$  of the stream number against stream order (1, 2, 3, 4, etc) is roughly a straight line it might be held to confirm this so-called 'law'. Taking logarithms of the numbers in a geometric series will produce a straight line whose gradient is a measure of what below is called the bifurcation ratio.

For example the river basin network shown in Figure 4.4 has 34 first, 10 second, two third and by definition just 1 fourth, order streams.

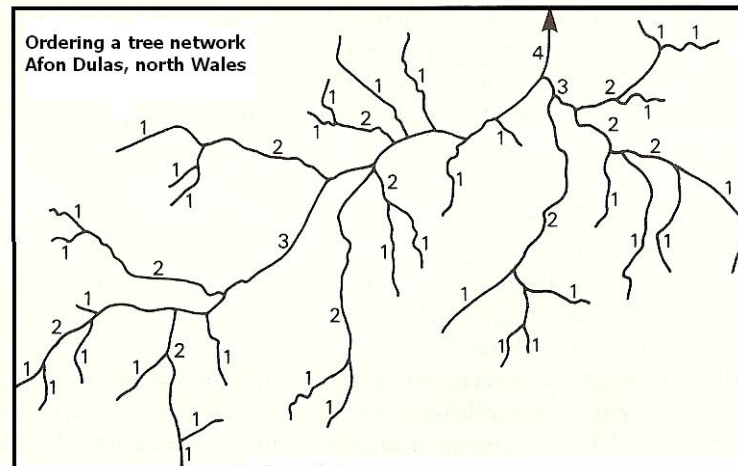


Figure 4.4 The Afon Dulas ordered  
(Source: after Unwin, 1981 Figure 4.5 b)

It can be seen from Figure 4.5 that a plot of the logarithms of these numbers falls on to what is very nearly a straight line:

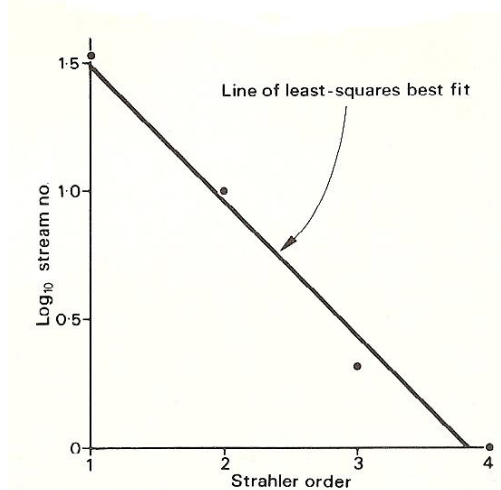


Figure 4.5 An illustration of the law of stream numbers  
(Source: after Unwin, 1981, Figure 4.6)

In the early days this was viewed as the *law of stream numbers* and was sometimes thought to have geomorphological significance but it turns out that many natural tree networks conform to this rule, so this information may not be very illuminating in itself. The *bifurcation ratio*, the ratio between consecutive terms in the geometric series, can be determined for different drainage networks and used as a direct comparison between cases. Natural drainage networks tend to vary in the range 3 to 5 according to rainfall, vegetation geology, and so on.

### Suggestions for modification

An obvious and illuminative extension is to use a branch of a real tree. Does the same law hold? Another is to analyze a variety of drainage basins of the same overall 'order' to see how the bifurcation ratio varies with factors such as geology and climate. If the class is one of physical geographers it is useful to extend the discussion to consider the hydrological significance of the network of blue lines depicted on the source map and the rules/conventions adopted by the particular mapping agency in deciding whether a given channel is actually a 'river'. It should be clear that although the representation may be consistent, its hydrological significance is moot and this simple observation casts some doubt on the work of an entire generation of quantitative geomorphologists.

To generalize further, it is also useful to ask for other examples of tree networks. One that might occur to students is the typical hierarchical structure of files and folder on their computer's hard drive.

## 4.7 Exercise (19): Analyzing networks

### Aims and introduction

A second type of connection structure is a *network* in which there is no restriction barring closed loops. This is a much more general structure that is sometimes called a *graph* used to represent almost any relational structure. The theory of networks is called *graph theory* and is increasingly important in many disciplines. One type of graph that has recently been studied in detail is the *small world*, based on the surprise when we meet someone we know in an unlikely context and say 'small world, isn't it?'. A small world graph is one where most objects have only small numbers of neighbors, yet everywhere is relatively close to everywhere else (Watts, 1999, 2003) Watts and Strogatz, 1998). Watts and Strogatz (1998) speculate that small world graphs are a very common network structure. As examples, they discuss the neural network of a particular type of worm, the electrical distribution network of the Pacific Northwest, and the 'co-star' network of Hollywood actors!

More generally, because graphs express relations between the elements of a set, they can be used to represent almost anything where we have set of related objects in a *relational structure*. It follows that although the exercise starts from a road map, many of the simple analytical techniques available for graphs are equally applicable to *any* adjacency matrix (see Sections 2.4 and 2.6)..

### Geometry, space and level

This exercise uses mapped examples of line objects that exist in a metric, usually Euclidean space but reduces the pattern of connection to a network that can be represented as a topological graph. Our concern is with the complex/second order concept of connection.

### Intended learning outcomes

I would be the first to admit that this exercise involves some potentially difficult ideas from linear algebra and its relevance to the analysis of the concept of connection might not be immediately apparent, but it is a very useful way to introduce students to some key ideas. It is most useful if extended and modified in the suggested manner.

After completing this exercise students will be able to:



- Explain and show how a mapped network can be reduced first to a topological graph and then how that graph can be written as an connectivity matrix to describe the pattern of connection in the network;
- Understand that such matrices can be manipulated to develop numerous descriptive measures of the concept of connection.

### Resources needed

Pencil and paper and if extended, software such as MINITAB™ or MATLAB™ that will perform matrix multiplication.

### Suggested student briefing

1. In the same way that analysis by stream ordering reduces the stream segments to a simple topological pattern of connection, in which we are not interested in true distances or direction but simply in connection, most analysis of a network starts by reducing it to a matrix that records the pattern of connection between the various entities in the system. In this way of recording connection, the junctions or objects are called *vertices* and the *links* are the *edges*;
2. Consider the map of the main road network of the Isle of Skye as in Figure 4.6.

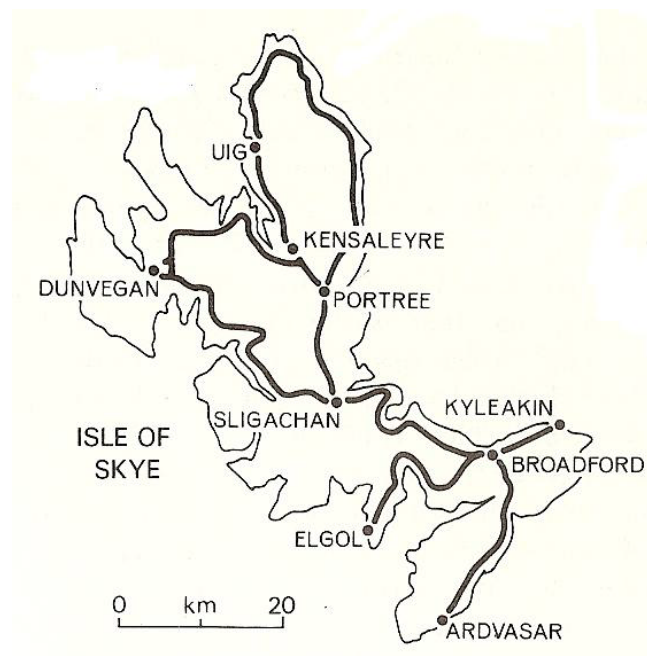


Figure 4.6 The main roads of Skye  
(Source: after Unwin, 1981, Figure 4.7 a)

There are many ways to represent the pattern of connection between the nine selected places on this map, but the simplest is to draw a picture similar to that in Figure 4.6.

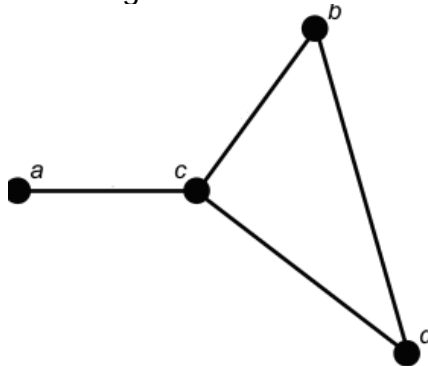


Figure 4.6 The graph of a network  
 (Source: after O'Sullivan and Unwin (2003) Figure 6.9 a)

The example has just four vertices, whereas your graph will have nine and you may well discover that there is more than one way of drawing this graph.

- The graph in our simple example can now be abstracted further to develop a topological *connectivity* matrix, let's call it **A**, in which for each vertex we set the element equal to '1' = 1 if for that row and column are connected and 0 if they are not. The matrix, **A**, corresponding to the little example above is:

$$\mathbf{A} = \begin{matrix} & \begin{matrix} a & b & c & d \end{matrix} \\ \begin{matrix} a \\ b \\ c \\ d \end{matrix} & \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix} \end{matrix}$$

In this the rows and columns are labeled with the vertex to which they correspond. Note that row and column order is arbitrary, but must both be the same. Note also that, since the main diagonal entries from top left to bottom right are all coded 0, we do not consider vertices to be connected to themselves.

- Create the connectivity matrix **A** for the Skye road network;

Now answer the following questions:

5. Which place is the most connected and why?
6. Is this matrix symmetric and if so why? Under what circumstances might a connectivity matrix not have this property?
7. On the so-called principal diagonal where the row and column vertex is the same, do we put a '1' or a '0'?

**Comment/answers**

A graph of the road network will look like Figure 4.7

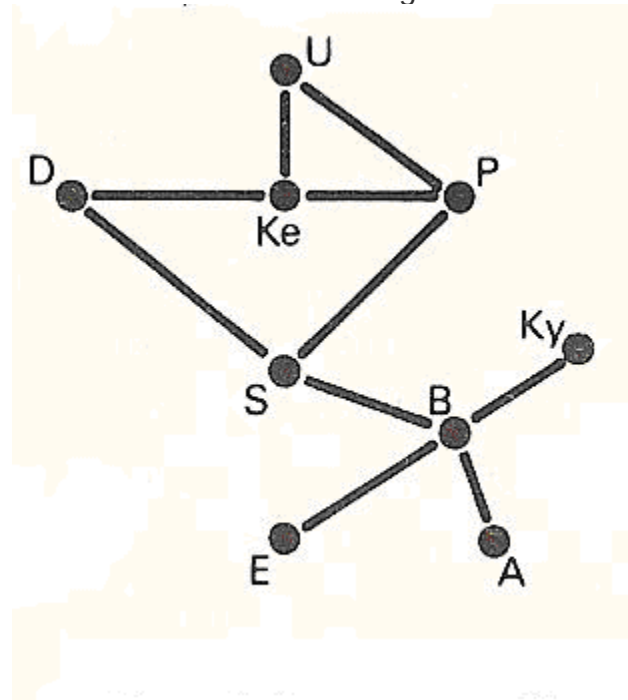


Figure 4.7 Skye's main roads  
(Source: after Unwin, 1981, Figure 4.7 b)

The matrix derived from it is:

| <i>From/To</i>   | <i>Ui</i> | <i>Ke</i> | <i>Po</i> | <i>Du</i> | <i>Sl</i> | <i>Br</i> | <i>Ky</i> | <i>El</i> | <i>Ar</i> |
|------------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| <i>Uig</i>       | 1         | 1         | 1         | 0         | 0         | 0         | 0         | 0         | 0         |
| <i>Kensalyre</i> | 1         | 1         | 1         | 1         | 0         | 0         | 0         | 0         | 0         |
| <i>Portree</i>   | 1         | 1         | 1         | 0         | 1         | 0         | 0         | 0         | 0         |
| <i>Dunvegan</i>  | 0         | 1         | 0         | 1         | 1         | 0         | 0         | 0         | 0         |
| <i>Sligachan</i> | 0         | 0         | 1         | 1         | 1         | 1         | 0         | 0         | 0         |
| <i>Broadford</i> | 0         | 0         | 0         | 0         | 1         | 1         | 1         | 1         | 1         |
| <i>Kyleakin</i>  | 0         | 0         | 0         | 0         | 0         | 1         | 1         | 0         | 0         |
| <i>Elgol</i>     | 0         | 0         | 0         | 0         | 0         | 1         | 0         | 1         | 0         |
| <i>Ardvasar</i>  | 0         | 0         | 0         | 0         | 0         | 1         | 0         | 0         | 1         |

Broadford has the most links (4), but this is only one possible measure of its position in the network. This adjacency matrix is symmetric about the main diagonal from top left to bottom right, because edges in the graph are *undirected*: if vertex  $v_i$  is connected to  $v_j$  then the opposite is also true. In a *directed graph* this is not necessarily true, and edges have a direction *from* one vertex and *to* another. This distinction can be important in how a street network is defined if there is a one-way system in operation. In our example we have put 1's down the principal diagonal to indicate that we consider a place to be connected to itself. Essentially this is a research decision, but it does have some 'interesting' mathematical consequences (see Unwin, 1981, pages 90-93).

### Suggestions for modification

If students have available software such as MINITAB and MATLAB that allow easy matrix manipulations it is instructive to multiply these matrices by themselves (that is, 'power' them). Unwin (1981, pages 86-93) has a discussion of what this can reveal. For example, the entries in the squared matrix record the number of different ways of moving between the corresponding vertices in two step and the entries in the principal diagonal will be the number of ways of going from a vertex to one of its neighbors and back again. It is simpler to think of these numbers as the number of its neighbours. If we record the *power* of  $\mathbf{A}$  at which the entry in each matrix position first becomes non-zero then we get a matrix that records the topological shortest paths in the graph. This is a sort of distance matrix (see Chapter 2). At considerably more advanced level, the eigensystem of this type of matrix is also useful (see Tinkler, 1977).

## 4.8 References

- Davis, J.C. (2002) *Statistics and Data Analysis in Geology* (3<sup>rd</sup> Edition, New York: John Wiley & Sons)
- Gardiner, V. (1975) *Drainage basin Morphometry*, Technical Bulletins, British Geomorphological Research Group, 14 (Norwich: GeoAbstracts)
- Gardiner, V. and C. Park (1978) Drainage basin morphometry: review and assessment, *Progress in Physical Geography*, 2: 1-35.
- Haggett, P. and R.J. Chorley (1969) *Network Analysis in Geography* (London: Arnold)
- Horton, R.E. (1945) Erosional development of streams and their drainage basins: hydrophysical approach to quantitative morphology, *Bulletin, Geological Society of America*, 56: 275-370
- Okabe, A. and Satoh, T. (2009) Spatial analysis on a network, In: Fotheringham, A.S. and Rogers, P.A. (eds.) *The SAGE Handbook on Spatial Analysis*, (London: SAGE), pages 443-464,
- Pincus, H.J. (1956) Some vector and arithmetic operations on two-dimensional orientation variates with application to geological data. *Journal of Geology*, 64: 553-557
- Strahler, A.N. (1952) Dynamic basis of geomorphology, *Bulletin, Geological Society of America*, 63: 923-938
- Tinkler, K.J. (1977) *An Introduction to Graph Theoretical Methods in Geography*. Concepts and techniques in Modern Geography, 14 (Norwich: GeoAbstracts), available at <http://www.qmrg.org.uk/catmog>
- Unwin, D. J. (1981) *Introductory Spatial Analysis* (London: Methuen)
- Watts, D. J. (1999) *Small Worlds: The Dynamics of Networks Between Order and Randomness*, (Princeton, NJ: Princeton University Press)
- Watts, D.J. (2003) *Six Degrees: the Science of a Connected Age* (London: Heinemann)
- Watts, D. J. and S. H. Strogatz (1998), Collective dynamics of 'small-world' networks, *Nature*, 393: 440-442
- Werrity, A (1972) The topology of stream networks. In: Chorley, R.J. (ed) *Spatial Analysis in Geomorphology* (London: Methuen), pages 167-196

# Chapter 5 Area: the devil's variable?

## 5.1 Introduction

As Chapter 1 indicated, area objects provide a very rich set of spatial concepts in the lattice space created by contiguity relationships. In the primitive/first order are concepts such as *boundary*, *shape* and *region*, which at second order generate ones such as *adjacency*, *fragmentation*, *pattern/clustering* (autocorrelation) and so on. In this chapter we develop student's appreciation of the basic spatial concepts of *area* itself, *shape*, *adjacency* and *pattern*. The latter is address through extended work mapping and analyzing pattern using a public domain software system developed and maintained by Dr. Luc Anselin that he calls *OpenGeoDa*<sup>TM</sup> (pronounce as Open 'gee-odder').

These concepts at first sight seems obvious and simple but, as these exercises will show, area objects come in many varieties, and are difficult to map and summarize. The simple fact that they can often be *modified* by changing the way they are defined leads to a classical issue, the *modifiable areal unit problem*, or MAUP. The complications that all this introduces justify the notion that area really is the *devil's variable*.

## 5.2 Exercise (20): Types of areas on maps

### Aims and introduction

A brief thought experiment or examination of a map will show that geographic areas come in numerous varieties. At high level there is a clear distinction to be made between those that are in some sense *natural* whose boundaries are unequivocally given by phenomena, such as the shoreline of a lake or island, the edge of a forest stand, or the mapped boundary of a particular soil type. The area objects so defined, respectively a 'lake', 'island', 'forest' and 'soil type' might be open to some argument (how dense do trees have to be before they are a 'forest?') but in principle at least their boundaries are fixed and immutable. Once recognised, this class of area objects is therefore self-defining.

We can contrast such natural areas with those *imposed* by human beings, such as countries, states, counties and census tracts. Such areas might or might not coincide with something natural, but their key property is that they are modifiable, not by changing anything in the underlying real world, but by simply changing their boundaries. Almost always such imposed regions have complex boundaries.

A third type of area frequently used in analysis arises where the space is divided into a regular shapes each of which is the same and which nest together and cover the region of interest without any overlap. The classic example is a grid of small square (usually and roughly) *pixels* ('picture elements') used in satellite remote sensing and some types of GIS. Clearly such areas are also imposed.

Although it will not be considered in this exercise, for completeness mention should also be made of a second type of area that is sometimes used. This is the Voronoi/Thiessen/Dirichlet region surrounding any point object and defining all those areas that are closer to that point than they are to any other.

At this level of generality we can thus distinguish between *natural*, *imposed*, and *gridded* types of area, but there are even more variations on the theme that occur in relationship to the nature of the object and its boundaries. First, is the area defined internally homogeneous, or are there 'inclusions' of other types of area that exist in reality but have been lost by the way the area is represented, either on a map or in a data base? The archetypical example of this is an area of a specified soil type that might well have inclusions of some other type that are too small to record, but a large lake might well have islands that are too small to map, and so on. Second, what is the nature of the boundary of the area? Is it a certain line that is *crisp* in the sense that, in principle at least, one could walk along it, or is the boundary in some sense *uncertain*? Again the archetypical

example of such boundary uncertainty is that of a soil, where one type might grade imperceptibly into another.

Finally, in this typology we can consider the relationship of each area with other similar objects. Some are simply *isolated* examples that stand alone (of course these can also be *planar enforced*). Others, such as a pattern of historical fire burn areas in which separate fires can have burnt over the same area, in principle can *overlap* spatially. Yet others are *planar enforced*, that is, they nest together like pieces in a jigsaw puzzle, completely covering the region with no gaps between them.

### **Geometry, space and level**

Area objects create a space that can perhaps best be called a *lattice* in which adjacency is the measure of distance. If we know their boundaries they can of course be displayed in a metric space such as the Euclidean. The hierarchical level is essentially primitive, concerned with the identification of individual examples of area objects.

### **Intended learning outcomes**

After completing this exercise students should be able to:

- Define and exemplify different types of area;
- Outline why there is a possible problem using information about areas that are modifiable (the MAUP).

### **Resources needed**

A selection of maps or on-screen equivalents. A suitable map series to pick would be the UK Ordnance Survey 1:50,000 *Landranger*<sup>TM</sup> series.

### **Suggested student briefing**

1. Familiarize yourself with the area concepts related to the type of area ( Natural/Imposed/Grid?), their internal nature (homogeneous/inhomogeneous/), boundary (Crisp/uncertain?) and relationship with other areas (isolated/overlap/planar enforced);
2. Select a sample Ordnance Survey 1:50,000 scale map and examine it carefully (you can of course do this on screen, from, for example <http://www.streetmap.co.uk>) to find at least ten examples of areas that have been mapped. In each case classify the nature of the area according to its type, internal character, boundary and relationship to other similar areas. In each case record a name for the area and a grid reference;



3. Record your results in a table similar to Table 5.1;

| Area object | Type? | Inside? | Boundary? | Relationship? |
|-------------|-------|---------|-----------|---------------|
|             |       |         |           |               |
|             |       |         |           |               |

Table 5.1 recording types of area object

4. With three recognize 'types', two sorts of internal characteristic, two type of boundary and three possible relationships, this allows for the possibility of  $3 \times 2 \times 2 \times 3 = 36$  differing types of area. You could try to see if you can find an example of each and every one of these!

### Comment/answers

The simple way to do this exercise is to work from the map key to see what types of area are mapped and then search for an example on the map. Using OS 1:50,000 Sheet 115 *Snowdon*, Table 5.2 presents a selection of possible candidates. It is NOT exhaustive.

| Area object              | Type?   | Inside?       | Boundary? | Relationship?   |
|--------------------------|---------|---------------|-----------|-----------------|
| 675705 wood              | natural | homogeneous?  | crisp     | isolated        |
| 605735 rocks             | natural | inhomogeneous | uncertain | isolated        |
| 675835 grid square       | imposed | homogeneous   | certain   | Planar enforced |
| 665675 National Trust NT | imposed | homogeneous   | crisp     | isolated        |
| 560550 Llyn Cwellyn      | natural | homogeneous   | crisp     | isolated        |
| 580628 Arfon District    | imposed | homogeneous   | crisp     | Planar enforced |
| 696665 Craig Y Dulyn     | natural | inhomogeneous | uncertain | isolated        |
| etc                      | etc     | etc           | etc       | etc             |

Table 5.2 Some Snowdonian area objects

### Suggestions for modification

A very simple question to ask is what the implications are for determining the area of the entity concerned. A useful extension would be to repeat the exercise with a simple  $k$ -colour map such as one of geology or land use.

## 5.3 Exercise (21): Colour maps for area objects

### **Aims and introduction**

Chorochromatic or *area-colour* maps are fairly commonly used to show phenomena such as geology and land use. Frequently these are the result of what has been called an interpreted mapping whereby a surveyor has classified what is seen on the ground into some existing, consistent, externally-derived scheme, such as a list of land use categories or soil types. Such maps are often very uncertain, through a combination of fuzziness in the phenomenon being mapped, inclusions of some characteristics that are too small to represent in either the data base or on the map, and boundary uncertainty. A good example is a soils map. We might be uncertain about what is being mapped, which itself might be a fuzzy object. There might be inclusions of some characteristic that isn't the one mapped but that are too small to show at the map scale, or we might have bands of transition between categories, all of which make for uncertainty about exactly what soil there is at any particular location.

### **Geometry, space and level**

Area objects create a space that can perhaps best be called a 'lattice' in which adjacency is the 'distance'. If we know the boundaries they can of course be displayed in a metric space such as the Euclidean. The hierarchical level is essentially primitive but the implications of what is discovered for more complex concepts and analysis are important.

### **Intended learning outcomes**

After completing this exercise students should be able to:

- Recognize an area-colour map;
- Critically evaluate its cartography;
- List the sources of uncertainty in such maps;
- Understand some of the complications in accepting such maps at their face value.

### **Resources needed**

Internet browser

### **Suggested student briefing**

1. Use Google Images™ or a similar search engine to find examples of simple colour maps. Finding the right search terms might well be your problem here!
2. In each case, ask yourself how well the map actually represents the phenomenon being mapped. Questions to do with representation that you might ask relate to issues of fuzziness, the importance of scale, and how well the boundaries of the areas mapped are known;
3. Search for a copy of the map that Simon Winchester (2002) claims 'changed the world', which is William Smith's superb first geology map dating from 1815. Lots of sites have this map, try:  
<http://earth.unh.edu/Schneer/map2.jpg>;
4. How many 'colours' did Smith use? Having looked at Smith's map, next find a map of the political 'colour' of the States of the USA in the 2008 Presidential election. An example of many is at  
<http://www.npr.org/news/specials/election2008/2008-election-map.html#/president?view=race08>;
5. In what important ways does this map differ from Smith's? .

### **Comment/answers**

There are numerous examples of soils, geology or land use maps on WWW. Figure 5.1 is a very attractive example of Canadian land cover which has over thirty categories:

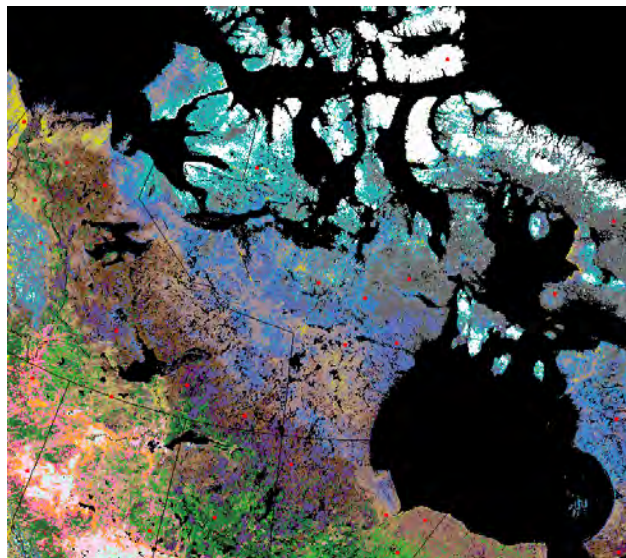


Figure 5.1: A chorochromatic map of Canadian Land use

A much simpler map of cell/mobile telephone coverage in USA is at:

[http://www.cellonenation.com/images/coverage\\_map\\_home.png](http://www.cellonenation.com/images/coverage_map_home.png)

and can be used to illustrate uncertainty in the classification and probable boundary issues.

William Smith's map was one where in principle at least the boundaries between colours are natural whereas those between the states of the USA are obviously imposed. The difference is real and important.

### **Suggestions for modification**

None

## 5.4 Exercise (22): Choropleth maps for area objects

### **Aims and introduction**

Choropleth or *area-value* maps are honest in the sense that they are true to the data, but frequently dishonest when used to draw conclusions about the real geography of the phenomenon being mapped. All this exercise does is to ask students to look carefully at some published examples.

### **Geometry, space and level**

Area objects create a space that can perhaps best be called a 'lattice' in which adjacency is the 'distance'. If we know the boundaries they can of course be displayed in a metric space such as the Euclidean. The hierarchical level is essentially primitive but the implications of what is discovered for more complex concepts and analysis are important.

### **Intended learning outcomes**

After completing this exercise students will hopefully:

- Never look at a choropleth map again without taking a very large pinch of salt;
- Be able to list the complications that need be considered when using such maps to describe regional patterns;
- Appreciate the importance of choice in the mapping techniques used;
- Understand why the addition of colour can reduce rather than enhance the legibility of these maps.

### **Resources needed**

Internet browser

### **Suggested student briefing**

1. As in Exercise (21) use a search engine to find an example of a good choropleth map and comment on what it really shows if we consider any or all of the following sets of issues:
  - The areas used. Are these natural or imposed? If the former, how were they defined and by whom? If the latter, are the areas used appropriate? Do large areas dominate the way the map looks? Are

the 'steps' at the edges of each block of color likely to reflect variation in the underlying phenomenon?

- The data. Are they counts of some sort? If so, there is an in-built tendency for larger areas to have bigger values and the map may be worthless. Choropleth maps only make sense if the numbers being mapped are ratios, either areal densities (such as number of people per unit of area) or population rates (such as the number of births per thousand population in the area). If the data are ratios, are these based on low numbers? If so the mapped values may very unstable to quite small changes. For example, if we add one person to an area in which there is already just one person, we double the population density, whereas adding one to several thousand people makes virtually no difference at all;
- The classification used. Prior to the use of computers almost all choropleth maps were *classed*, that is, each datum was assigned to one of a small number of classes. Experience suggests that from five to seven classes is appropriate, but the appearance of a map can be changed dramatically by varying the number of classes. Evans (1977) describes a large number of possible schemes. His conclusion is that you must examine the *statistical frequency distribution* of the data before deciding on a classification scheme. Tobler (1973) pointed out that modern displays don't actually need to class choropleths since they are capable of showing as many shades of 'color' as needed. His point was debated but, nowadays it seems that classless choropleths are rapidly becoming standard;
- Finally, the most obvious thing about a choropleth is the symbolism used. Traditionally, choropleths were created by shading each area using a 'screen' pattern of lines such that the more lines, the darker the area looked and the higher the value. It is now more usual to use a graded series of a single color to show the increase in intensity. Either way, how you choose to shade the map greatly affects its final appearance.

### **Comment/answers**

Are the areas mapped of sensible size or would a table of numbers have done just as well? Does it make the very common error of using raw numbers? Is the frequency distribution of values sensibly used in the classification or is it 'classless'? Is the symbolism sensible and appropriate?

A very clear example of the most common error is this choropleth from an agency that ought to know better is shown as Figure 5.2.

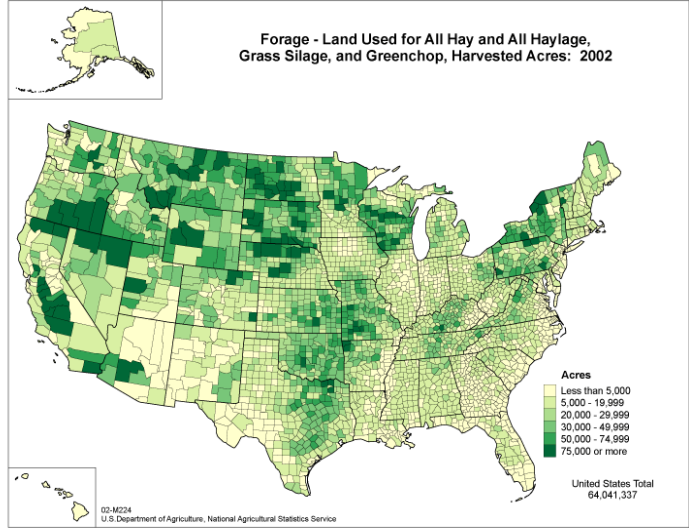
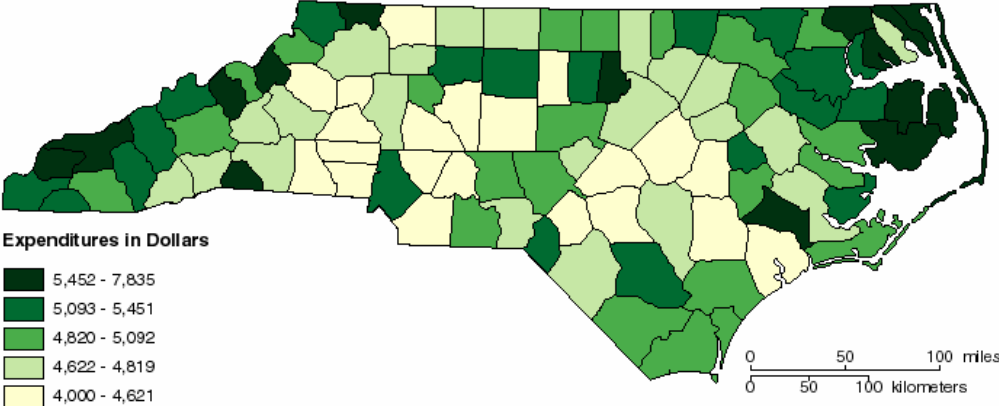


Figure 5.2 Good cartography, pity about the numbers!

The error is that of showing absolute values of the acreage in each of the US counties. It follows inexorably that the large counties (check out the mountain states) will tend to have more area devoted to almost anything than will a small county (such as in the east). Despite being well-drawn the map is worthless, and good cartography does not guarantee a good map, the numbers matter! Figure 5.3 shows another example using a similar colour ramp that does make some sense:

Per Pupil Expenditure for Public Education in North Carolina, 1994-1995



Source: NC Department of Public Instruction, *Statistical Profile*, 1996.

Figure 5.3 This is better!

This is standardized by the 'at risk' population, the number of school children in each county. Colour on these maps must be used with extreme care, see Brewer, C.A. (1994).

### **Suggestions for modification**

An exercise that has been useful in the past is to give students a series of contrasting variables to be mapped in summary form as frequency counts and then ask them to devise an appropriate class interval scheme and symbolism. There is no real need to create maps to make the point about the choice of mapping method!



## 5.5 Exercise (23): Measuring area

### **Aims and introduction**

This simple exercise asks students to estimate the area of some area object, such as the continent of Australia using the standard method in GIS based on Simpson's Rule method of integration

### **Geometry, space and level**

Measurement of area by the method listed assumes a Euclidean metric space with a string of boundary coordinates, but the concept is essentially primitive/first order.

### **Intended learning outcomes**

After completing this exercise and reflecting on what it teaches, students will:

- Know how to compute the area of a closed polygon representation of a real world area;
- Understand that almost all measurements made from maps or their digital representations need to be qualified by the method used to measure them;

### **Resources needed**

Source map on an equal area projection, pencil and paper, spreadsheet program. The collection of the (x,y) co-ordinates can of course be done using a digitizer/graphics tablet or via an on screen map and mouse.

### **Suggested student briefing**

A simple way to find the area of any area object (natural or imposed) uses a tried and tested approach known in mathematics as Simpson's Rule for numerical integration. The approach is shown in Figure 5.4 below:

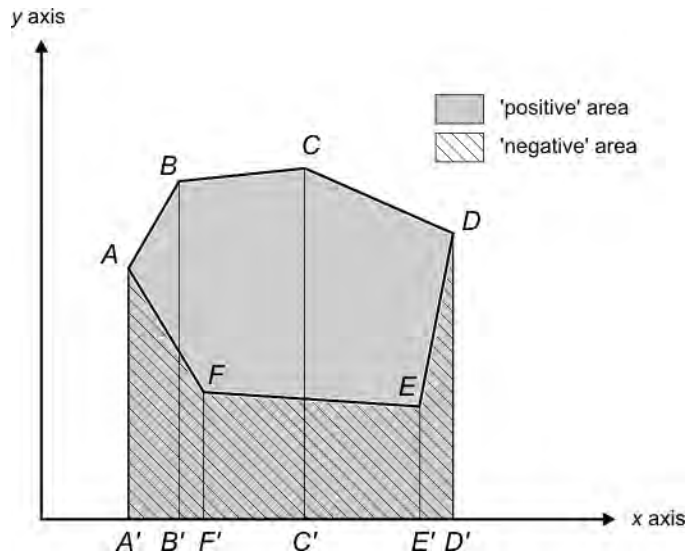


Figure 5.4 Measuring area from boundary co-ordinates (Source: after O'Sullivan and Unwin, 2010 Figure 7.2)

By dropping lines from each vertex of a string of  $(x,y)$  co-ordinates defining the boundary of the polygon that represents an area to the X-axis, we define a series of trapezoids, such as for example the one defined by the string  $ABB'A'$ . Simple mathematics gives the area of this little trapezoid as the difference in  $x$  co-ordinates multiplied by the average of the  $y$  co-ordinates:

$$\text{Area of } ABB'A' = (x_B - x_A) \cdot (y_B + y_A) / 2$$

In this case  $x_B$  is greater than  $x_A$ , so this area will be a positive number;

If we move round the polygon repeating the same calculation for each trapezoid in a clockwise direct at first all these areas will be positive as we move away from the Y-axis and  $x$  always increases. However, when we turn the corner and move back towards the Y-axis, so the reverse is the case, and the computed areas become negative numbers that are subtracted from the grand total. What we are left with at the end is actually the required area of the polygon. Provided we work clockwise and come back to the starting vertex, the general formula is simply:

$$\text{Polygon area, } A = \sum_{i=1}^n (x_{i+1} - x_i)(y_{i+1} + y_i) / 2$$

where  $(x_{n+1}, y_{n+1})$  is understood to bring us back to the first vertex  $(x_1, y_1)$ .

1. Trace the shoreline of some area object taken from a map (such as the continent of Australia) taking care to ensure that the source is drawn on

- an equal area map projection and record its boundary as a series of (x, y) co-ordinates;
2. Use the method outlined above to compute its area. This is easily done using any spreadsheet program. Enter your co-ordinates into the first two columns and copy these from row 2 onwards into the next two columns, displacing them upwards by one row as you do so. Make a copy of the first co-ordinate pair into the last row of the copied columns. The next column can then be used to enter and calculate the trapezoid formula. The sum of this column then gives your estimate of the continent's area. You will have to scale the numbers from co-ordinate units into real distances and areas on the ground.

### **Comment/answers**

For Australia, using a 1:30 000 000 map as a source and with just 45 co-ordinates for the shoreline gives an area of 7,594,352 km<sup>2</sup>. A semi-official value is 7,686,850km<sup>2</sup>. The closeness of this estimate to an official value is clearly an accident.

### **Suggestions for modification**

The real value of this exercise is as stimulus for several possible discussions, similar in some ways to those that might have been followed about 'distance'. Topics that might be introduced are:

- a) The uncertainty/error in area estimates. Which is 'right' and how could we tell? Almost all exact 'geometric' calculations on spatial data give answers that are really estimates of some true, but unknown value;
- b) What does this say about other calculations does an area estimate underpin (e.g. spatial density, most of the point pattern measures, shape indices);
- c) This assumes, reasonably in this case, that the inside of the area is homogeneous and the boundary is 'crisp'. What are the implications if, as in the case of the area of some defined soil type, neither of these conditions are met;
- d) How would you measure the total area of some phenomenon, such as a land use type that at the study scale is made up of hundreds of very small parcels? This could well introduce so-called 'dot planimetry' or even, post-GIS, counting the alike cells in a colour raster representation of the phenomenon;

e) Lastly, one to think about is what if the edge of the area (such as a coastline) is itself fractal?

There is a literature on area measurement, see for example Gierhart (1954), Frolov and Maling (1969).

## 5.6 Exercise (24): What do we mean by shape?

### **Aims and introduction**

All areas have a *shape* and sometimes this can be of interest (examples include the hexagons of central place theory, the streamlined forms of drumlins, and so on), but the concept isn't an easy one to define and attempts to 'measure' it have a long history. A simple definition is the set of relationships of relative position between points on the edge of the object which is unchanged by any changes in scale or size. Like Exercise (23), this exercise is intended primarily to get students to think about the concept.

### **Geometry, space and level**

Describing shape by the method listed assumes a Euclidean metric space with a string of boundary coordinates, but the concept is essentially primitive/first order.

### **Intended learning outcomes**

After completing this exercise students will:

- Understand that areas have shapes and that visually these can be seen to differ;
- Have refreshed their memories on how to measure both area and distance;
- Be able to derive the *compactness ratio*, comparing the observed shape with an area having the same perimeter as the chosen area object.

### **Resources needed**

Pencil and paper, plan outline of some chosen area object. Drainage basins as defined by their watersheds from topographic maps make good examples.

### **Suggested student briefing**

1. On the map supplied find a small drainage basin and define and trace its perimeter, upstream of some chosen point on the main stream;
2. Now consider the two-dimensional shape of this basin. How would you describe it? Is it elongated? Simple or complex? Compact? Streamlined? Clearly we have a vocabulary to describe the concept of the shape of an

area, but not everyone would agree about the words to use to describe any particular example. What environmental factors might have a role to play in determining the shape of this basin and why?

3. Next, try to characterize the shape of your drainage basin using a numerical *shape index*. There is a huge choice, and most depend on some comparison with a mathematically defined simpler shape such as a circle. Since a circle is the most compact shape of a given area that can exist, a reasonable measure is a *compactness ratio*, sometimes called the isoperimetric quotient. This is simply the ratio of the observed area to that of a circle having the same perimeter:

$$S = \frac{4\pi A}{P^2}$$

In this  $A$  = the measured area and  $P$  is its perimeter. This isn't hard to derive from elementary geometry and the so-called iso-perimetric inequality says that  $S$  will be less than or equal to 1;

4. Use whatever method you think best suited to measure the length of the perimeter of the watershed of your drainage basin and its area and use these to compute this index;
5. Now chose another drainage basin that seems visually to have a more/less compact shape and repeat the analysis. Does the compactness ratio help?

### **Comment/answers**

This exercise can be made more valuable as an in class exercise in which the results are shared and comparisons made over a wide range of drainage basins. Interpreting the results for a variety of landscape, may well tax your knowledge of geomorphology and geology. Handle with care.

### **Suggestions for modification**

As above.

## 5.7 Exercise (25): Mapping area data using *OpenGeoDa*<sup>™</sup>

### **Aims and introduction**

This is essentially a guided exercise in choropleth map creation using Dr. Luc Anselin's *OpenGeoDa*<sup>™</sup> package. It can also be undertaken in a GIS such as *ArcGIS*<sup>™</sup>. It is a great help to have completed Exercise (22).

### **Geometry, space and level**

As Exercise (22).

### **Intended learning outcomes**

After completing this exercise students will:

- Be able to use *OpenGeoDa*<sup>™</sup> with supplied data as a shapefile to produce choropleth maps;
- Appreciate the clear need to map either some spatial density or population ratio and not the absolute values;
- Understand how choice of different class intervals and color ramps can greatly alter the 'look' of this type of map.

### **Resources needed**

Computer running the *OpenGeoDa*<sup>™</sup> package with the supplied data set on sudden infant death syndrome (SIDS) in North Carolina. If 'presentation' maps are required with scale bars, north points and the like, then access to a standard drawing package such as Microsoft *Paint*<sup>™</sup> will also be needed.

### **Suggested student briefing**

1. If this has not already been done for you, or you are using your own machine, visit the website at <http://geodacenter.asu.edu> and download and install the *OpenGeoDa*<sup>™</sup> software. This is the latest incarnation of software for spatial statistical analysis created by Dr. Luc Anselin. You will need to go through a registration process to do this. The response from the Center is almost always very rapid, but allow time for this to take place. For reference, also download a copy of the pdf of a 244 page *Exploring Spatial Data with GeoDa, a workbook*, produced by Anselin for the *Center for Spatially Integrated Social Science (CSISS)*. Find it at the

same website by clicking on TRAINING and SUPPORT>GEODA  
HELP>TUTORIAL;

2. *OpenGeoda™* uses ESRI™ *shapefiles* to define the pattern of area objects in use. If you have *ArcGIS™* available, this is a useful feature, and it permits integration with other GIS packages that have shapefile filters. Now get the data, which are under DATA>SAMPLE DATA and are in the files *SIDS.zip* and *SIDS2.zip* that you can find towards the bottom of the list when you scroll down. Unzip them into a folder you have named and remember where you have put them;
3. These contain data for Sudden Infant Death Syndrome (SIDS) in the 100 counties of North Carolina (NC). The county is a basic administrative unit in the USA that is often used in mapping. These same data have been used in a series of papers in spatial statistical analysis by Cressie and his collaborators (Cressie and Read, 1985, 1989; Cressie and Chan, 1989). Their source is a table in Cressie (1993);
4. Before you do any mapping using *OpenGeoDa™* with these data, find an atlas and have a look at the basic geography of the State noting especially which counties contain the major urban areas;
5. The complete data in SIDS2 has records for 100 counties as:

AREA  
PERIMETER  
County NAME  
County ID  
FIPS number  
Cressie ID as in his book, Table 6.1, pages 386-389  
BIR74: number of births 1974-1978  
SID74: Number of Sudden Infant Deaths, 1974-1978  
NWBIR74: number of non-white births i1974-1978  
BIR79: number of births 1979-1984  
SID79: number of Sudden Infant Deaths, 1979-1984  
NWBIR79: number of non-white births 1979-1984  
SIDR74, SIDR79, NWR74 and NWR79 are derived rates in the two periods

Note that these SIDS totals are for a five year summation not, as perhaps is the impression given in the CSISS Workbook, for the years 1974 and 1979.



6. Now use *OpenGeoDa*<sup>TM</sup> to draw a simple choropleth map on screen. Proceed as follows. First, open *OpenGeoDa*<sup>TM</sup>, which will lead to a simple on-screen menu bar. Next, select FILE>OPEN SHAPE FILE and navigate to where you have stored the unzipped SIDS2 data. Now double click on the SIDS2 shapefile (note the icon used). The result will be a map of North Carolina's counties in which each and every county is shaded a monotone green. This is because at this stage you haven't specified a variable to be mapped;
7. Create a choropleth map using >MAP>QUANTILE with the default of four classes and selecting the BIR74 total births data. Your reward will be a pretty four class choropleth map with a rather pleasing color ramp from yellow to dark brown, but does it tell you anything about births in NC? Hopefully you realize that the high value counties are all those that include major cities such as Charlotte, Winston-Salem, Greensboro, Raleigh, Fayette and Wilmington. Basically, we are looking at a map of the population and the reason is, of course, our decision to map the actual totals and not some population ratio or areal density. Should you wish you can SAVE this map at the EDIT menu by using COPY TO CLIPBOARD and then pasting into, say, Microsoft *WORD*<sup>TM</sup>. A right click on the map gives an option to save in some standard graphics formats;
8. Go right back to the start using MAP>RESET and repeat these steps but try to map the SID74 data. Using MAP enables a range of possible class interval selection schemes to be applied, but I suspect that you won't easily create a sensible map. In part this is because we are still using raw totals, and so have the population factored into these numbers, but in addition, and as can be confirmed using EXPLORE>HISTOGRAM or EXPLORE>BOXPLOT and selecting the SID74 data, the frequency distribution is dominated by zeros and is basically Poisson. One way round this problem is to plot the data on an area cartogram base and a version of this is available as an option in *OpenGeoDa*<sup>TM</sup>;
9. Clearly, you need to express the SID74 totals as a rate of incidence,  $R_i$ , and the most obvious variable we have to act as denominator in such a calculation is the total births over the same time period found in the column BIR74. *OpenGeoDa*<sup>TM</sup> allows some very basic operations on its data tables. To calculate the proportion of SIDS deaths in the first period as a proportion of the total number of births, first

FILE>CLOSE ALL to start a new analysis.

Use FILE> OPEN SHAPE FILE and navigate to and select the SIDS shapefile and data (not SIDS2). You will get the familiar green map.

Now add a column to the data table using TABLE>FIELD CALCULATION and select RATE OPERATIONS. The table itself is shown and it is useful to see how the data are held as a simple matrix of numbers. The appropriate settings are:

Result : call this RATE  
Method: set as raw rate  
Weight file: leave blank  
Event variable: select SID74  
Base Variable: select BIR74

To do the computations click on APPLY and then CLOSE to get a new column, called RATE, into the data table. Confirm that you have some rate estimates using the viewing option in TABLE and by use of EXPLORE>HISTOGRAM and selecting RATE. You can get to this map using the supplied rate data (expressed as a rate per thousand) in SIDS2, it is useful to have gone through the logic step by step and also to be able to compare the raw numbers map with the rates one;

10. The rest is easy. Use MAP>QUANTILE and select RATE to get a map of the birth total standardized rates of SIDS deaths in NC. To save the map in a standard graphic format that can be exported and used in a report in WORD or have its graphics improved in a standard drawing package, right click on the map and then use the 'save image' command to save the map as a .bmp file. To save the legend, resize its right margin in the usual way, and then right click and use SAVE TO CLIPBOARD;
11. If this map doesn't appeal, then experiment with different classification schemes, data transformations and the 1979 data. You might also like to look at the SIDS rates relative to the numbers of non-white births in the same period, calculated as NWR74.

### **Comment/answers**

Figure 5.6 is the unimproved map of the total Births BIR74 in the five years to 1974. Comparing it with the locations of the counties that include the cities of NC it can be seen to be really simply a map of the population.

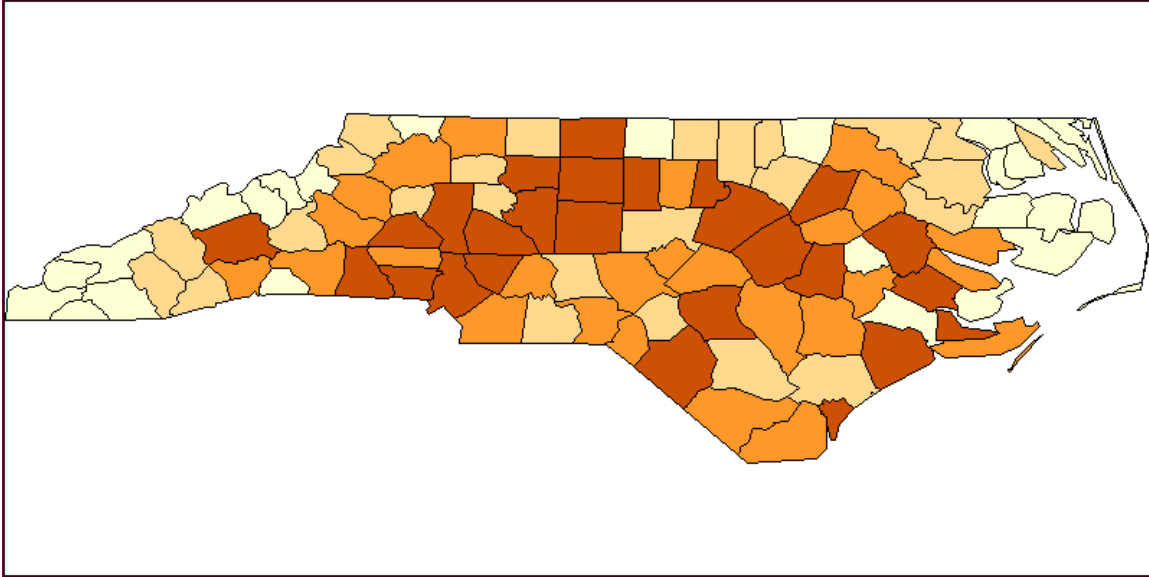


Figure 5.6 Raw total births to 1974, BIR74

Figure 5.7 shows the absolute totals in the SID74 data which give almost the same map:

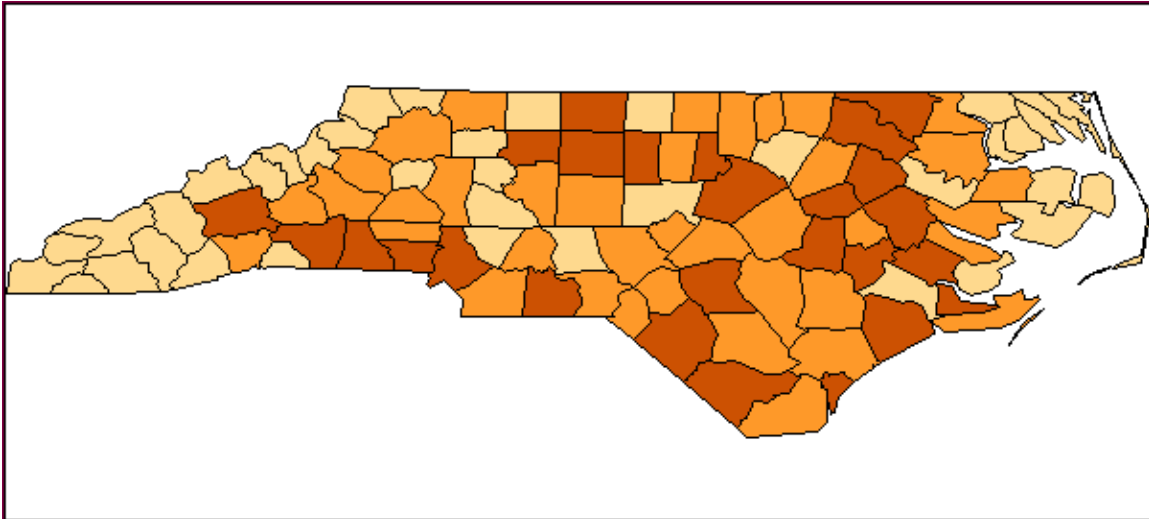


Figure 5.7 SIDS totals for 1974, SID74

These data illustrate quite well the difficulties in any choropleth mapping (see Exercise (22)). Specifically, we have:

- a) Small numbers, including several of the 100 counties with a zero score. You can confirm this by TABLE>PROMOTION and scrolling across to the

SID74 column. Small numbers mean that computed rates are very sensitive to quite small changes in the data;

- b) The frequency distribution as shown in Figure 5.8 is Poisson, not normal, which makes the selection of class intervals problematic:

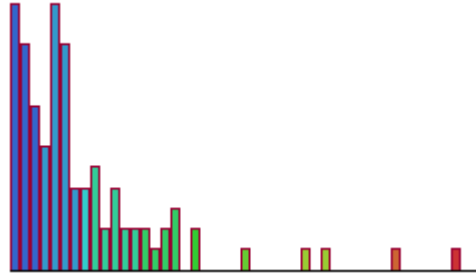


Figure 5.8 Frequency distribution of SIDS numbers

- c) Counts, not population (or maybe area) corrected rates. To circumvent this issue, one option is to compute a SIDS *rate* for each of the time periods as:

$$R_i = 1000 \left( \frac{s_i}{n_i} \right)$$

In which  $s_i$  is the count of SIDS deaths for the  $i$ -th county and  $n_i$  is the total births in the same period. The 1000 multiplier simply gives us the rate per 1000 births and is conventional when using this type of data. In fact all the Cressie papers modify this rate estimate by adding 1 to each of the SIDS scores to give a rate estimate as:

$$R_i = 1000 \left( \frac{(s_i + 1)}{n_i} \right)$$

This has the effect of allowing any maps to discriminate (slightly) among the counties that have zero total SIDS deaths. The best maps should show a clustering of counties with unusually high SIDS incidence in the north-east and south of the State, which is a linear function of the non-white births for the first time period (1973-1978).

Students should see why almost invariably we must map some rate or other and that by changing the 'classing' scheme we can change the 'look' of our maps. For example, Figure 5.9 shows the effect of using the SID74 numbers in a *percentile* map;

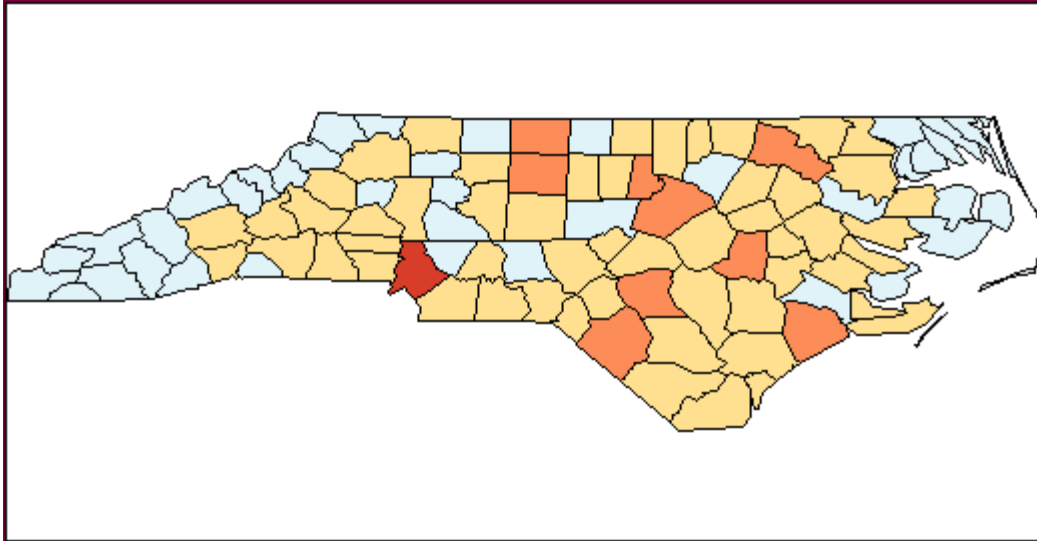


Figure 5.9 A map based on percentile class intervals, SID74 totals

At the end of the briefing we finally get to calculate and draw a proper map of the rates of SID as a proportion of the total number of births, the histogram for which is shown as Figure 5.10.

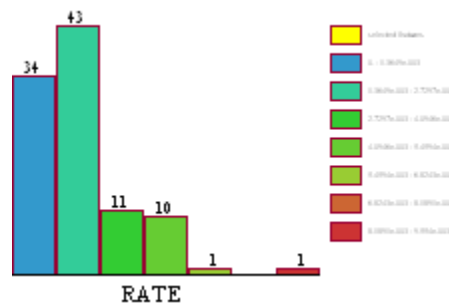


Figure 5.10 Frequency distribution of the rates of incidence of SIDS relative to total births

The final simple quantile map is shown as Figure 5.11. Whether it helps understand the phenomenon is perhaps moot.

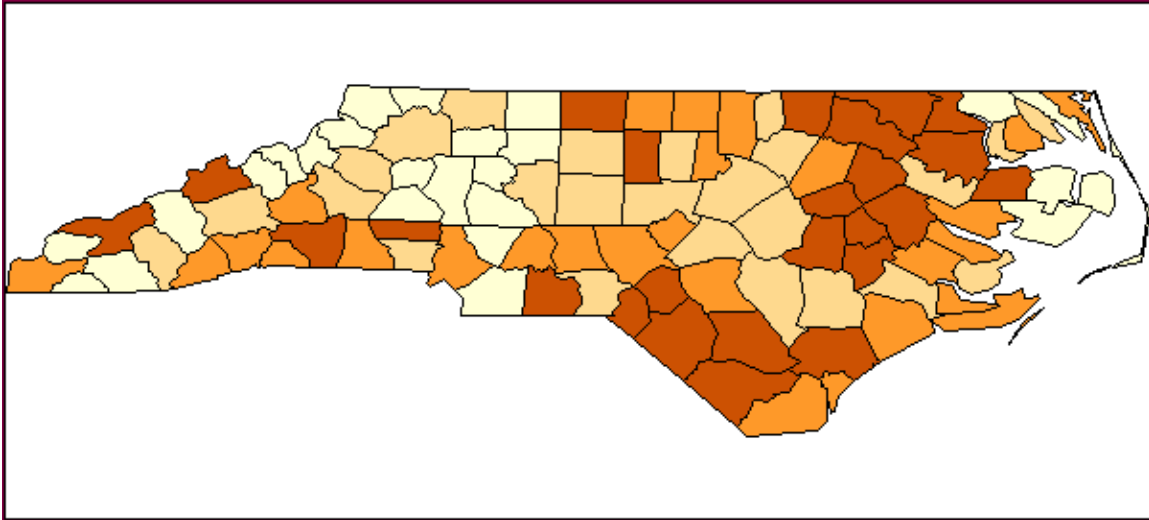


Figure 5.11 Map of the calculated RATE of SIDS 1974

### **Suggestions for modification**

By now students might be interested in using *OpenGeoDa*<sup>™</sup> more generally. They could do worse than to work through Exercises 5 to 14 and then 15 to 18 in the CSISS Workbook, which cover a variety of analysis, visualization and housekeeping functions. (*OpenGeoDa*<sup>™</sup> is slightly different from the version to which the Workbook makes reference, but the differences are nothing to worry about.

## 5.8 Exercise (26): Using *OpenGeoDa*<sup>™</sup> to compute a spatial weights matrix

### **Aims and introduction**

As we saw in Section 2.4, the concept of adjacency can be summarized in a spatial weights matrix, **W**. The definition and assembly of such a matrix is a critical step in the computation of all the standard global autocorrelation statistics. *OpenGeoDa*<sup>™</sup> offers two approaches to creating a **W** for a given system of areas, based on either contiguity or distance.

*Contiguity based measures* record whether or not zones share a common boundary, recognizing two cases. One, called the *Rook's case* is where the areas share any boundary. The second, or *Queen's case*, extends the definition to allow areas that share a single corner also to be counted as adjacent.

This type of **W** matrix can also be used to record connection in networks and, more generally still, links in *any* network of relationships. Given an initial matrix of first order adjacencies, simple powering can generate the 2<sup>nd</sup>, 3<sup>rd</sup>, etc., order adjacencies if these are required. In the spatial case discussed here, necessary conditions are that there are no 'islands' in the system that have no common boundary with any other area, that the areas do not overlap, and that they exhaust the plane, with no 'holes'. These later conditions define what in GIS terminology is called *planar enforced area 'geography'*

*Distance based measures* 'collapse' the area data onto a single reference point and then 'slice' the distances between all these points at some defined threshold to define adjacency using a simple binary scheme (1 = adjacent, 0 = otherwise) to populate a **W** matrix as before. Alternatively, the inverse of these same distances defines a ratio-scaled measure of adjacency which, when row-standardized such that all rows sum to unity, can also be used. It should be clear that this distance approach can accommodate 'islands' and again, by choosing increasing distances can create a correlogram that describes the autocorrelation structure over a range of distances. The disadvantage of the approach is the arbitrary nature of the reference points selected. For example, our North Carolina data uses the locations of the individual county seats as its point reference, which may, or may not, be sensible.

### **Geometry, space and level**

Area objects create a space that can perhaps best be called a 'lattice' in which adjacency is the 'distance'. This exercise emphasizes the concept of complex/second order concept of adjacency.

### **Intended learning outcomes**

After completing this exercise students will understand that:

- A spatial weights matrix **W** is a description of a 'geography' of interest based on either adjacency of areas in a planar enforced pattern, or on distance from some arbitrary reference point within each area;
- Such matrices represent in effect hypotheses about what is important in the 'geography of interest and can/should be formulated with this in mind;
- All indices of spatial pattern (autocorrelation) rely on having such a formulation;
- Technically, **W** matrices are 'sparse', containing a lot of zeros and are often represented using a specific structure that omits the zeros.

### **Suggested student briefing**

1. In *OpenGeoDa*<sup>TM</sup> open a new project using the SIDS2 shapefile, and access the weights creation routines using TOOLS>WEIGHTS>CREATE
2. In the dialogue boxes that follows, set the input file = SIDS2.shp (best to use the browse icon to find the full path for this). I find it useful also to set the Weights ID variable as the FIPSNO, which is the US Census Bureau's own identifier for the county.
3. Having done this set QUEENS CASE CONTIGUITY and click on CREATE. You will be prompted for a name for the file that records the **W** matrix for these data. Call it NC\_queen. This file has a .gal extension, but is otherwise a text file that can be examined using any text editor such as Microsoft *NotePad*<sup>TM</sup>
4. At this point examine the file and work out how it codes the 'geography' of the counties of NC.
5. In *OpenGeoDa*<sup>TM</sup> you can also examine properties of the chosen **W** using TOOLS>WEIGHTS>PROPERTIES to display a histogram of the 'contact numbers', that is the numbers of adjacent areas. Ideally, the histogram should be uni-modal and, of course contain no areas with a zero contact number indicative of its being either an island or an error. If the area was tessellated randomly, it can be shown that the mean number of adjacent areas will be six.



## Comment/answers

In (4) above Nc-queen.gal is coded:

```
100
1 3
19 18 2
2 3
18 3 1
```

etc.

This says that we have 100 records (line 1), each made up of a zone sequence number 1, 2, ..., 100 followed by its contact number (3), followed by the sequence numbers of the three 'adjacent' zones. This is a standard way of coding a large sparse matrix in which the majority of the entries are zero (for no adjacency).

In (5) a histogram of the *contact numbers* looks like Figure 5.12.

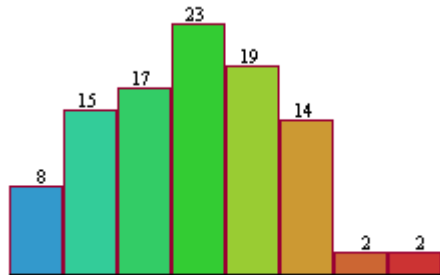


Figure 5.12 Contact numbers for the 100 counties of North Carolina

In a random tessellation of the plane the expected number of zone neighbors is 6 (hexagons), but in this case we get a modal class of 5, which I think is probably something to do with the way the US was surveyed and settled. By and large much of the US was laid out by the surveyors ahead of the people and so has greater regularity in its county boundaries. In UK where the boundaries came after the settlement our small area statistics give number around 6!

## Suggestions for modification

This exercise is needed to introduce the idea of spatial autocorrelation, but instructors may actually find it better to introduce a standard measures, such as the Joins Count and Moran's I dealt with in Exercises (27) and (28) first, and then refine student understanding using it.

## 5.9 Exercise (27): Spatial autocorrelation and pattern

### **Aims and introduction**

There are several ways by which the idea of global spatial autocorrelation as a measure of 'pattern' in attributes attached to area objects can be introduced. Many instructors will prefer to go directly to the Moran's *I* statistic and hence approach the subject by way of correlation using interval or ratio scaled variables. An alternative is to approach the same idea by way of the runs test for autocorrelation in a one-dimensional series. This approach and sections of the text that follow is taken by O'Sullivan and Unwin (First Edition 2002, pages 180-196) but is in turn based heavily on ideas presented by Silk (1979).

### **Geometry, space and level**

This exercise uses the concept of adjacency in a lattice space very directly to generate measures that relate to the complex/second order concept of clustering of like values. Following the exercise in full allows for exploration of the analytical concept of a spatial process as it applies to area objects.

### **Intended learning outcomes**

After doing these two linked exercises students will understand:

- The idea of a random process as selection from an underlying probability distribution;
- That any observed pattern is a realization of that process;
- Some summary measure, such as the number of runs in a sequence or the nature of adjacencies in a B/W coded 'map' can be used to develop a test statistic;
- Under appropriate assumptions the test statistic itself has a sampling distribution (in this case normal), such that the probability of any observed realization can be assessed by using a z-score referenced to standard tables.

Along the way they will also learn how to conduct two standard tests:

- A runs test for a one-dimensional sequence of a binary (0/1) variable;
- The equivalent in two dimensions, which is the classic joins count test as developed by Cliff and Ord (1973).

### **Resources needed**

Pencil and paper, calculators and standard statistical tables of percentage point of the normal distribution. Deck or decks of cards. A standard coin with heads and tails.

## Suggested student briefing

### Part 1 The Runs Test

1. If you took half a deck of cards, say the red diamonds (♦) and the black clubs (♣), shuffled it thoroughly and then drew the cards in order, what would the sequence of diamonds and clubs look like? Something like:

♣♦♣♣♦♣♦♣♦♦♦♣♦♣♣♣♦♣♦♦♦♦♦♣♦♣♣

Is this sequence random? That is, could it have been generated by a random process? If the sequence was

♣♣♣♣♣♣♣♣♣♣♣♣♣♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦

or

♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦♣♣♣♣♣♣♣♣♣♣♣♣♣♣

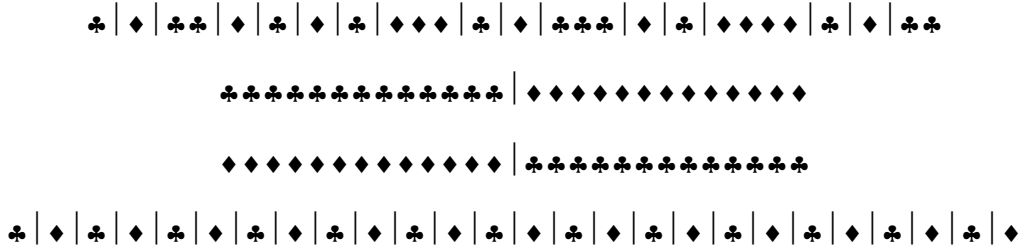
would you consider these likely outcomes of a random process? These both look highly unlikely, and in much the same way. Again, does the sequence

♣♦♦♦♦♦♣♦♦♦♦♦♣♦♦♦♦♦♣♦♦♦♦♦♣♦♦♦♦♦♣♦♦♦♦♦♣♦♦♦♦♦

seem a likely outcome of a random process?

The second, third and fourth all seem very unlikely, but for different reasons, yet all four sequences contain 13 of each suit, so it is impossible to tell them apart using simple distributional statistics. What you need to focus on is how each card drawn relates to the one before and the one after—to its *neighbours* in the sequence, in effect. Note that each of these sequences might have been generated by the same process and that we call each a realization of that process. The element of chance means that we have a lot of possible realizations and it is the *process* that involves randomness not the outcomes;

2. One way to quantify how unlikely these sequences are is by counting *runs* of either outcome (i.e. unbroken sequences of only clubs or only diamonds). For the examples above we have



making 17, 2, 2 and 26 runs respectively. Intuitively, you should be able to see that *either* very low *or* very high numbers of runs suggest that the sequence is not random. So, now conduct an experiment for yourself, with or without others in the class;

3. Take the diamonds and clubs from a deck of cards, shuffle them thoroughly and then draw them one at a time and record the sequence (occurrences) of diamonds and clubs obtained. This is clearly an example of a single realization of a random process that has two possible outcomes;
4. Either repeat this several times, or do this as an in-class experiment in which you share your results, and produce a histogram of the results. This is an experimental/empirical estimate of the sampling distribution of the number of runs. Compute the mean and standard deviation of these realizations;
5. The theoretically expected values are easily computed from the numbers of each used ( $n_1$  and  $n_2$ ) as

$$E(\text{number of runs}) = \frac{2n_1n_2}{(n_1 + n_2)} + 1 = \frac{2 \times 13 \times 13}{13 + 13} + 1 = 14$$

$$E(s_{\#runs}) = \sqrt{\frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)}} = 2.4980$$

Use these results for the means and standard deviation to find z-scores, and use standard tables of the percentage points of the normal distribution (or a website such as that at <http://www.fourmilab.ch/rpkp/experiments/analysis/zCalc.html>).

For example the first realization above gives us:

$$z = \frac{17 - 14}{2.4980} = 1.202$$

6. How unusual is your result? This approach is a standard statistical test for 'randomness' in a sequence of binary (0/1, R/B etc) numbers called the Runs Test.

## Part 2 Joins counts (on a grid)

1. Draw a 6 by 6 chess board and then 'color' each cell 'black' or 'white' using the same random mechanism that was used in Part 1;
2. However, in this application we will toss a coin 36 times, once for each grid cell and record the result 'black' if it's heads and 'white' if it's tails'. The result will be a crude map showing a grid of 36 cells, each coloured either black or red;
3. Now, using the Rook's case definition of adjacency (i.e. just adjacencies along rows and down columns, not the diagonal ones) count the numbers of joins of each type, calling the results  $J_{BB}$ ,  $J_{WW}$  and  $J_{BW}$ . Trying to do this will convince you that it's easier and less error prone to count on a cell by cell basis, counting *all* the joins to that cell and then halving the result. Check your results by noting that for this 6 by 6 grid, these numbers should sum to  $k= 60$  (why?);
4. By now you should just about see the connection of the runs count test to spatial autocorrelation. Our attribute in each cell is a binary variable (0/1, B/W) and instead of counting runs as our measure of pattern, we are using the joins in two spatial dimensions as the measure. The analogous test for a spatial pattern is called the *Joins count test for spatial autocorrelation*, developed by Andrew Cliff and Keith Ord in their classic book *Spatial Autocorrelation* (Cliff and Ord, 1973);
5. As you might expect, the equations that give the theoretical results for the expected outcome of the joins count statistics for an independent random process. The expected mean values of each count are given by:

$$E(J_{BB}) = kp_B^2$$

$$E(J_{WW}) = kp_W^2$$

$$E(J_{BW}) = 2kp_Bp_W$$

where  $k$  is the total number of joins on the map,  $p_B$  is the probability of an area being coded B, and  $p_W$  is the probability of an area being coded W. Note that at each colouring we could have either outcome, which means

that the numbers of black and white cells don't have to be equal. Because of the mechanism used we know that the underlying probabilities  $p_B = p_W = 0.5$ . This is called *free sampling*. Typically in a real map this isn't the case and we are forced to use the observed numbers of black and white cells to estimate the underlying probabilities. This is called *non-free sampling* and it gives rather more complex equations for the standard deviations (see below). The expected standard deviations are given by the more complicated formulae:

$$E(s_{BB}) = \sqrt{kp_B^2 + 2mp_B^3 - (k + 2m)p_B^4}$$

$$E(s_{WW}) = \sqrt{kp_W^2 + 2mp_W^3 - (k + 2m)p_W^4}$$

$$E(s_{BW}) = \sqrt{2(k + m)p_B p_W - 4(k + 2m)p_B^2 p_W^2}$$

where  $k$ ,  $p_B$  and  $p_W$  are as before, and  $m$  is given by

$$m = \frac{1}{2} \sum_{i=1}^n k_i(k_i - 1)$$

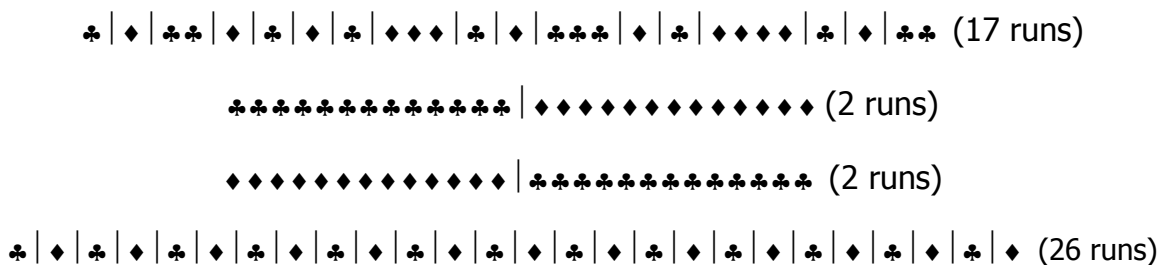
where  $k_i$  is the number of joins (or neighbors) to the  $i$ -th area. Computing  $m$  is just tedious (you should get  $m = 148$ , but don't take this on trust).

### Comment/answers

These exercises are deceptively simple, but they introduce some very fundamental ideas that can and should form the basis for discussion.

### Part 1: Runs test

Typically, students will count between 11 and 17 runs, and the extremes are obviously either 2 or 26. This exercise illustrates in a simple and intuitive way some important concepts and it should not be completed without summarizing them. These concepts are a realization of a random process, the development of a simple test statistic, the sampling distribution of this if the null hypothesis is correct, and the use of  $z$ -scores to assign a probability based on the assumption of normality.



In the cases above with 13 ♣'s ( $n_1=13$ ) and 13 ♦'s ( $n_2=13$ ), we have:

$$E(\#\text{runs}) = \frac{2 \times 13 \times 13}{13 + 13} + 1 = 14$$

$$E(s_{\#\text{runs}}) = \sqrt{\frac{2 \times 13 \times 13 \times (2 \times 13 \times 13 - 13 - 13)}{(13 + 13)^2 (13 + 13 - 1)}}$$

$$= \sqrt{\frac{105456}{16900}}$$

$$= \sqrt{6.24}$$

$$= 2.4980$$

So for the first example we have a z-score of:

$$z = \frac{17 - 14}{2.4980} = 1.202$$

This is unusual, but not massively so. For the other, extreme examples the z-score is -4.8038.

In class make sure that students understand the notions of each sequence being a realization of a process involving random selection from an underlying (in this case binomial) probability distribution. You can of course also change the original numbers of cards of each colour to use distributions other than  $p = q = 0.5$ .

In class I have usually compiled a histogram of the numbers of runs obtained on the board/OHP to show how as  $n$  increases so we start get something that can look vaguely normal centered on the expected theoretical value. Almost always you get some extreme values, which is pedagogically useful. I have less joy with the observed standard deviation in relation to its expected value, but students will easily see that the same process yields a number of outcomes...

**Part 2: Chess Board**



Again typically, students will usually get low z-scores, but as an in-class experiment with a reasonable number of students it is by no means unusual to get some extreme values. Since students know that the process used was 'random' this fact can be used to introduce critical ideas in classical statistical inference. If you ran Part A, a very useful thing to do is to show that the same key concepts apply here.

A sample result is given in Figure 5.13.

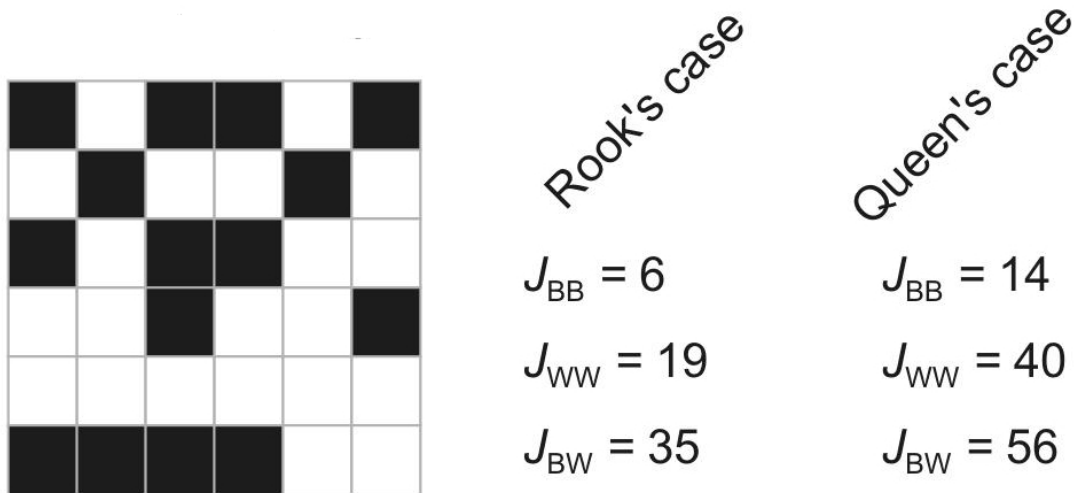


Figure 5.13 A realization of a random process and its Joins Counts  
(Source: after O'Sullivan and Unwin (2003) Figure 7.5 b)

Since we used *free sampling* by flipping a coin the appropriate probabilities for black and white are  $p_B = p_W = 0.5$  and with the 'Rook's Case' neighbors, the number of joins in the grid is  $k = 60$ . Finding  $m$  is simply tedious, but it comes to 148 as:

$$\begin{aligned}
 m &= 0.5 \left[ \underset{\text{corners}}{(4 \times 2 \times 1)} + \underset{\text{edges}}{(16 \times 3 \times 2)} + \underset{\text{center}}{(16 \times 4 \times 3)} \right] \\
 &= 0.5(8 + 96 + 192) \\
 &= 148
 \end{aligned}$$

Expected values for the means under the null hypothesis of random colouring are:

$$\begin{aligned}
E(J_{BB}) &= kp_B^2 = 60(0.5^2) = 15 \\
E(J_{WW}) &= kp_W^2 = 60(0.5^2) = 15 \\
E(J_{BW}) &= 2kp_Bp_W = 2(60)(0.5)(0.5) = 30
\end{aligned}$$

The standard deviations are:

$$\begin{aligned}
E(s_{BB}) = E(s_{WW}) &= \sqrt{kp_B^2 + 2mp_B^3 - (k + 2m)p_B^4} \\
&= \sqrt{60(0.5)^2 + 2(148)(0.5)^3 - [60 + 2(148)](0.5)^4} \\
&= \sqrt{29.75} \\
&= 5.454
\end{aligned}$$

and

$$\begin{aligned}
E(s_{BW}) &= \sqrt{2(k + m)p_Wp_B - 4(k + 2m)p_B^2p_W^2} \\
&= \sqrt{2(60 + 148)(0.5)(0.5) - 4[60 + 2(148)](0.5)^2(0.5)^2} \\
&= \sqrt{15} \\
&= 3.873
\end{aligned}$$

This leads to a z-score for BB joins of -1.650, for WW of 0.733 and for BW of 1.291 none of which exceeds the 95% threshold.

### Suggestions for modification

In the context of spatial autocorrelation, it is useful to repeat the experiment using a deck of cards (as in the runs test) to illustrate the idea of non-free sampling. Joins count statistics can be very awkward and a number of issues will arise:

- a) A large *negative* z-score on  $J_{BW}$  indicates *positive* autocorrelation since it indicates that there are fewer BW joins than expected. The converse is true;
- b) It is possible for the three tests to appear to contradict one another, in which case there is an interesting problem in logic. By simulation Cliff and Ord (19763) suggest that the BW test is the most reliable;
- c) Generally the choice of neighborhoods should not affect the scores and the decision about the overall structure, but this is not always the case;
- d) In this example, the assumption is that  $p_B$  and  $p_W$  are known in advance. Normally they are estimated from the data, and, properly speaking, much more complex expressions should be used for the expected join counts

and their standard deviations. These are given in Cliff and Ord (1973). The difference is equivalent to that between the card drawing and coin flipping examples described in the discussion of simple Runs Test;

- e) Finally, to make the calculations relatively simple, the exercise used a regular grid. For some US presidential election data treated at the State level over an irregular tessellation of areas see O'Sullivan and Unwin (First Edition, 2002, pages 192-196). Attempting to assemble the **W**-matrix of adjacencies used in this example by hand will quickly convince students of the usefulness of some automated approach as in Exercise (26), Section 5.8.

## 5.10 Exercise (28): Global spatial autocorrelation using *OpenGeoDa*<sup>TM</sup>

### **Aims and introduction**

Working through any realistic calculation of spatial autocorrelation using the Joins Count approach as in Exercise (27) quickly reveals two difficulties. The first is that use of binary data is very restrictive and throws away potentially important information. The second is the sheer labour involved and the chances of error. In practice, most workers in the field now use an approach using Moran's *I* statistic and just about everyone will use appropriate software to do the calculations. This exercise returns to the SIDS data for North Carolina and functionality provided in *OpenGeoDa*<sup>TM</sup> to compute a different measure of spatial autocorrelation, Moran's *I*. It is assumed that Exercises (25) and (26) have been completed along the way.

### **Geometry, space and level**

As Exercise (27).

### **Intended learning outcomes**

After completing this exercise students will be able to:

- Outline the construction of Moran's *I* index of spatial autocorrelation and explain why this is a measure of 'pattern' in area object attribute data;
- Use *OpenGeoDa*<sup>TM</sup> to do the necessary calculations.

### **Resources needed**

Computer running *OpenGeoDa*<sup>TM</sup> with the SIDS data available.

### **Suggested student briefing**

At the end of Exercise (25), you had a map of the rates of incidence of SIDS in the 100 counties of North Carolina, relative to the total number of births. Does this map show a 'pattern' that is different from that which we might have got had the 100 values for the rate of incidence of SIDS been allocated randomly to the 100 counties? To answer this question requires a test for global spatial autocorrelation of the type discussed in Exercise (27), Section 5.9. In that exercise we used a test, called the Joins Count, adapted for two spatial dimensions from the standard Runs Test of randomness in a one-dimensional

series and using an attributed that was coded as a binary variables (black/white, on/off etc). Even with just a few zones the labour in making the counts and performing the test was high, with a strong possibility of making some simple arithmetic error. This exercise introduces a different approach to detecting and measuring spatial autocorrelation based on an analogy with the standard correlation coefficient and applicable to interval and ratio scaled variables, such as the set of rate value that we have for the counties of North Carolina.

Moran's  $I$  spatial autocorrelation statistic is usually applied to areal units where numerical ratio or interval data are available, and is defined by a pretty terrifying equation as:

$$I = \left[ \frac{n}{\sum_{i=1}^n (y_i - \bar{y})^2} \right] \times \left[ \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} (y_i - \bar{y})(y_j - \bar{y})}{\sum_{i=1}^n \sum_{j=1}^n w_{ij}} \right]$$

The important part of the calculation is the *second fraction*, in which the numerator (top bit) is:

$$\sum_{i=1}^n \sum_{j=1}^n w_{ij} (y_i - \bar{y})(y_j - \bar{y})$$

which is actually a *covariance*. The subscripts  $i$  and  $j$  refer to different areal units or zones in the study, and  $y$  is the data value in each. By calculating the product of two zones' differences from the overall mean ( $\bar{y}$ ), we are determining the extent to which they vary together (co-vary). If both  $y_i$  and  $y_j$  lie on the same side of the mean (above or below) then this product is positive; if one is above the mean and the other below, then the product is negative and the absolute size of the resulting total will depend on how close to the overall mean the zone values are.

These covariances are then multiplied by  $w_{ij}$ . This is an element from a spatial weights matrix  $\mathbf{W}$  of the sort that we looked at in Exercise (26), Section 5.8. You will recall that each element in the matrix says whether or not we regard the zones represented by its position in the matrix are 'adjacent' or not. Its role here is to switch on/off the calculation. If the element  $w_{ij}$  is set equal to 1, the rest of the calculation is performed for the values in the zones  $i$  and  $j$ . and this will only happen if these two zones are considered adjacent. If it is set at 0, then the calculation results also in a zero. After summing for all possible pairs of zones the result will be a measure of the covariance shown by all the zones relative to their neighbours.

The rest of the formula standardizes the value of  $I$  for the number of zones being considered, the number of adjacencies involved, and the range of data values. Dividing by the sum of the entries in the  $\mathbf{W}$  matrix  $\sum \sum w_{ij}$  accounts for the number of joins in the map. Multiplying by

$$\frac{n}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

is actually *division* by the overall dataset variance, and corrects the value of  $I$  for the overall variance and so ensures that  $I$  is not large simply because the values and variability in  $y$  are high.

1. You are now ready to use *OpenGeoDa*<sup>™</sup> to do the hard work for you. First close any session you have running and access the SIDS2 data

```
FILE>CLOSE ALL
FILE>OPEN SHAPEFILE as SIDS2
```

These rate of SIDS as a proportion *relative to the total non-white births* data at the County level and aggregated over the five years from 1974-78 (note that I am being *very* specific here) are pre-computed for you as NWR74 in the SIDS2 data set, so we have our values for  $y$ . It might be useful at this point to re-draw the map and remind yourself about the basic geography of these rates:

```
MAP>QUANTILE and select NWR74 (remember these are the rate of sudden infant deaths relative to the non-white births in these counties).
```

Your map should show what appears to be a clear spatial concentration of counties with high rates;

2. Next, you need a spatial weights matrix,  $\mathbf{W}$ , to be assembled for the North Carolina Counties. Access the weights creation routines using TOOLS>WEIGHTS>CREATE and in the dialogue boxes that follows, set the input file = SIDS2.shp (best to use the browse icon to find the full path for this). I find it useful also to set the Weights ID variable as the FIPSNO, which is the US Census Bureau's own identifier for the county. Having done this set QUEENS CASE CONTIGUITY and click on CREATE. You will be prompted for a name for the file that records the  $\mathbf{W}$  matrix for these data. Call it NC\_queen. This file has a .gal extension, but is otherwise a text file that can be examined using any text editor such as Microsoft *NotePad*<sup>™</sup>. You now have the necessary elements to complete the calculation of Moran's  $I$ ;

3. To compute Moran's  $I$ , proceed as SPACE>UNIVARIATE MORAN and select NWR74 from the dialogue box and in the 'select weight' dialogue navigate to and select NC\_Queen as the weights matrix  $\mathbf{W}$ ;
4. The result is a display with two elements, a slightly mysterious  $(x, y)$  plot and, at the top in blue lettering, a computed Moran's  $I$ . You should get a value of  $I=0.7226$ . Ignore the plot and concentrate on this value. If this was a normal correlation coefficient, its value must be in the range from -1 to +1 and we'd think that 0.7226 is pretty high. As with a standard correlation coefficient, values above zero indicate positive spatial autocorrelation and values below it show chess-board like negative spatial autocorrelation. We might also use a standard  $t$ -test to discover whether or not this value exceeds some specified threshold at a specified  $p$ -level;
5. Unfortunately it isn't this simple with spatial autocorrelation. We suspect that we have evidence of a high autocorrelation, but is it significantly different from what we expect if the rates were randomly distributed? There are two complications. First, the expected 'null' value isn't zero, but for a reasonably 'well behaved' lattice of areas depends on the number of areas in the lattice as:

$$E(I) = - \frac{1}{(n - 1)}$$

Note that for even quite moderate  $n$  this will be close to zero. Second, the variance of  $I$  is a complicated function of the  $\mathbf{W}$  used and the assumptions one makes about the way the area values were generated. For a good account of the arcanums of all this, see Fotheringham, Brunsdon, and Charlton (2000, pages 201-209). Fortunately, there is an easier and in some respects more realistic option which is to ignore the theory and rely on a random permutations test. With  $n$  areas we have  $n!$  possible permutations of the values between the areas and, of course, we have observed and mapped just one of these. It is usually impractical to compute  $I$ -values for *all* possible permutations, so instead we use a *Monte-Carlo* sampling to select a reasonable number from them as the basis for our null values. In *OpenGeoDa*<sup>TM</sup> do this by selecting OPTIONS>RANDOMIZATION. This gives a menu of possible sample numbers. Clearly, the more the better, but at some cost in computation time. In this case, experiment will show that the result is so clear cut that 99 runs is adequate. Your value for the mean is likely to be around the theoretical expectation of -0.010101 and have a standard deviation around 0.06. At 0.7226 the observed value is thus around 11 standard deviations 11 standard deviations away from the mean, so you can be pretty confident that the pattern of rates is significantly different from

being spatially random. In other words there is a spatial pattern to these data: they are not random.;

6. Repeat the analysis using the '1979' aggregated rate data, listing and comment on the results. Is the pattern the same?

### Comment/answers

NWR74 looks like Figure 5.14:

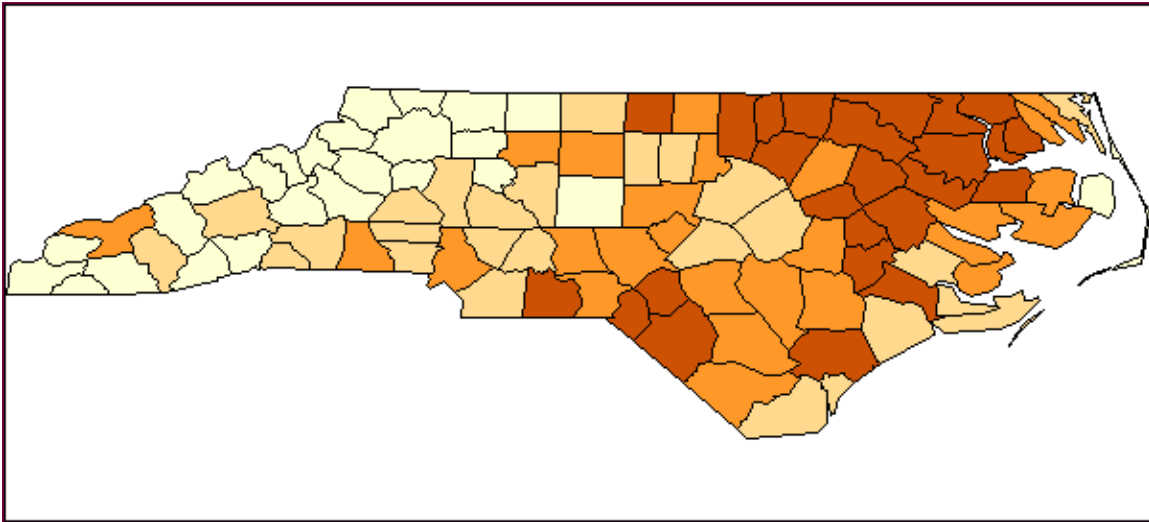


Figure 5.14

With the result for Moran's  $I$  as in Figure 5.15

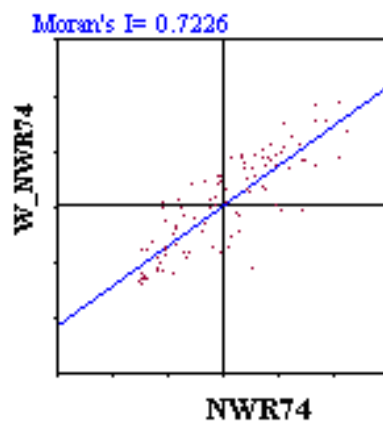


Figure 5.15 Moran's  $I$

The experimental null value will be around -0.010101 with a standard deviation so small that we can be pretty sure that the pattern is different from random. There is a pattern that shows positive spatial autocorrelation in which area with



like values tend to be closer together. This is an illustration of what has sometimes been called Tobler's law, sometimes called the 'First Law of Geography'. Like most 'first laws' in science it is in some sense definitional: if spatial autocorrelation did not exist there would be little point in studying any geography. In its original form the law was stated: "everything is related to everything else, but near things are more related than distant things" (Tobler 1970). Pedagogically, it is vital that students are shown how this concept is reflected in the way Moran's  $I$  is defined and calculated. That the pattern of SID over these five years shows a 'geography' is only the first step; explaining it is a different matter and would need some work with possible covariates, taking care to avoid autocorrelation issues.

The 1979 rate data (NWR79) will look something like Figure 5.16.

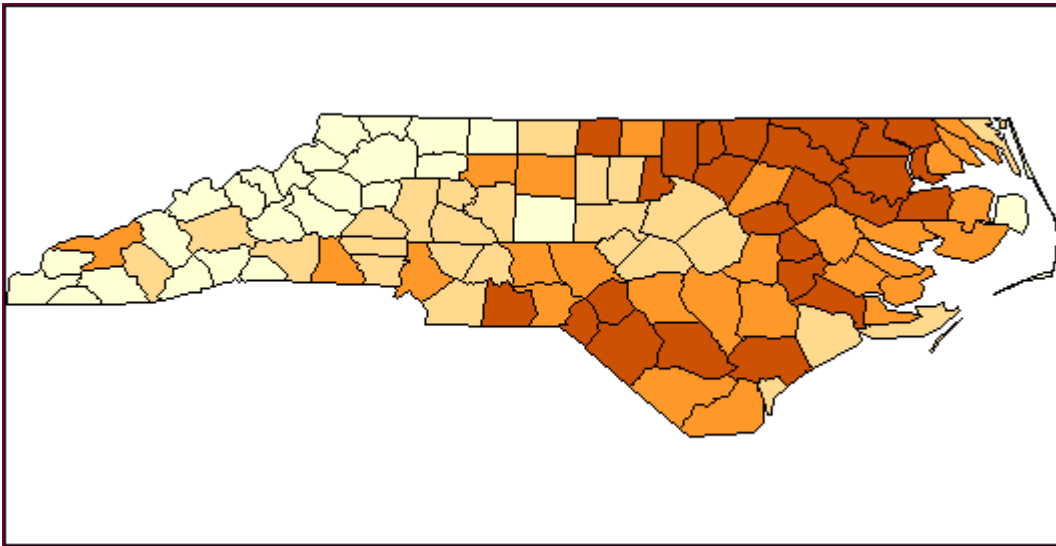


Figure 5.16 Map of the 1979 data rates

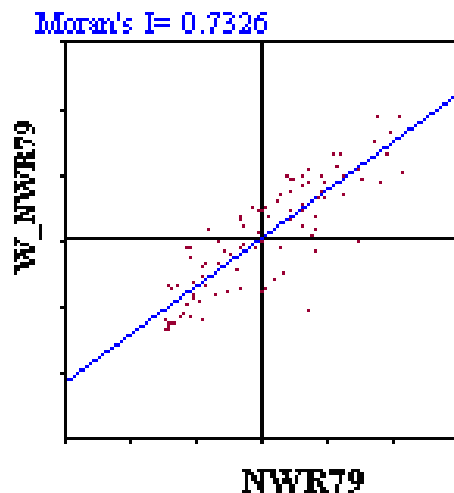


Figure 5.17 Moran's I for the 1979 data

Finally, Figure 5.17 shows the results for Moran's I.

### **Suggestions for modification**

The most obvious extension is to use this exercise to introduce the idea of a local statistic and illustrate this with the Moran Scatter Plot, a *Local Indicator of Spatial Association* produced by *OpenGeoDa*<sup>TM</sup> (see Anselin, 1995, 1996). Interpretation of these scatter plots isn't always obvious, needing a good appreciation of the actual geography of the studied area in order fully to understand them. Since this is exactly why the plot was produced in the first place there is an evident circularity in the process! Your

## 5.11 References

- Anselin, L. (1995) Local indicators of spatial association – LISA. *Geographical Analysis* 27: 93-115
- Anselin, L. (1996) The Moran scatterplot as an ESDA too to assess local instability in spatial association. In: Fischer, M. Scholten, H.J and D. Unwin *Spatial Analytical Perspectives on GIS*, (London: Taylor & Francis) pages 111-125
- Boyce, R. and W. Clark (1964) The concept of shape in geography. *Geographical Review*, 54: 561-572
- Brewer, C.A. (1994) Colour use and guidelines for mapping and visualization. In: MacEachren, A.M. and D.R.F. Taylor (eds.) *Visualization in Modern Cartography*, (Pergamon, Oxford), pages 123-147. See also 'ColorBrewer' at <http://www.personal.psu.edu/faculty/c/a/cab38/ColorBrewerBeta.html>
- Clark, W., and G.L. Gaile (1975) The analysis and recognition of shapes. *Geografiska Annaler*, 55B: 153-163
- Cliff, A.D. and J. K. Ord (1973) *Spatial Autocorrelation* (London: Pion)
- Cressie, N. and T.R.C. Read (1985) Do sudden infant deaths come in clusters? *Statistics and Decisions*, Supplement 2, 3:333-349
- Cressie, N. and T.R.C. Read (1989) Spatial analysis of regional counts. *Biometrics Journal*, 31: 699-719
- Cressie N., and Chan, N.H. (1989) Spatial modelling of regional variables, *Journal of the American Statistical Association*, 84:393-401
- Cressie, N. A. (1993) *Statistics for Spatial Data*, (Revised Edition, Wiley: Hoboken, NY), see pages 385-393.
- Dykes, J.A. and D.J. Unwin (2001) *Maps of the census: a rough guide*, at <http://www.agocg.ac.uk/reports/visual/casestud/dykes/dykes.pdf>
- Evans, I.S. (1977) The selection of class intervals. *Transactions, Institute of British Geographers*, 2: 98-124
- Fotheringham, A.S., Brunsdon, C and M. Charlton (2000) *Quantitative Geography; perspectives on spatial data analysis* (London: Sage) pages 201-209

Frolov, Y.S. (1975) Measuring the shape of geographical phenomena: a history of the issue. *Soviet Geography*, 16: 676-687.

Frolov, Y.S. and D.H. Maling (1969) The accuracy of area measurement by point counting techniques. *The Cartographic Journal*, 6: 21-35

Gierhart, J.W. (1954) Evaluation of methods of area measurement, *Survey and Mapping*, 14: 460-469.

Silk, J. (1979) Use of classroom experiments and the computer to illustrate statistical concepts, *Journal of Geography in Higher Education*, 3(1):13-25

Tobler, W. (1970), A computer movie simulating urban growth in the Detroit region, in *Proceedings of the I.G.U. Commission on Quantitative Methods*, Vol. 46, Supplement, June 1970 of *Economic Geography*.

Tobler, W.R. (1973) Choropleth maps without class intervals. *Geographical Analysis*, 5: 26-28

Wentz, E.A. (2000) A shape definition for geographic applications based on edge, elongation and perforation. *Geographical Analysis*, 32: 95-112.

Winchester, S. (2002) *The Map that Changed the World* (London, Penguin Books)

Zhang, D. & G. Lu (2004) Review of shape representation and description techniques, *Pattern Recognition*, 37(1): 1-19

# Chapter 6 Fields

## 6.1 Introduction

This chapter provides a series of exercises associated with self-defining continuous *fields* where the entity is an attribute that is *continuous* across space and thus is in some sense *self-defining*. The relations are those involving distance and magnitude and there are a number of associated spatial concepts such as the first order primitive we call *height*, and second order notions of *continuity*, *gradient* and *trend*

## 6.2 Exercise (29): Continuity and isolining a field

### **Aims and introduction**

Many students assume that it is easy to create a continuous isoline map of a field from a scatter of point attribute values. This exercise, adapted from O'Sullivan and Unwin (2010, pages 251-253) will show them otherwise. The key concepts that it illustrates are those of continuity, single value and the presence of spatial autocorrelation in the field variable. It is also a useful introduction to automated contouring.

### **Geometry, space and level**

At this level fields are self-defining and create a metric, usually Euclidean space. The primitive/first order concept is that of 'height'. This exercise explores the consequence of complex/second order concept of continuity in a very direct and obvious way.

### **Intended learning outcomes**

After completing this exercise students will realize that:

- Threading lines of equal value through a sample of control points is not a trivial problem and that different people will produce different maps;
- Almost always, the objective truth does not exist, but some interpretations (models?) will be better than others;
- Spatial interpolation relies on the assumptions of continuity and single-value as well as on the fact that the variable is spatially autocorrelated;
- In drawing isolines additional information and prior knowledge often play a large part in determining the result.

Along the way, they might also have learnt how to

- Draw isolines from a scatter of data control points.

### **Resources needed**

Base maps as below, pencil, paper, erasers.

### **Suggested student briefing**

1. The point attributes shown in Figure 6.1 are 'spot heights' of the average January temperature (°F) in a part of Alberta, Canada

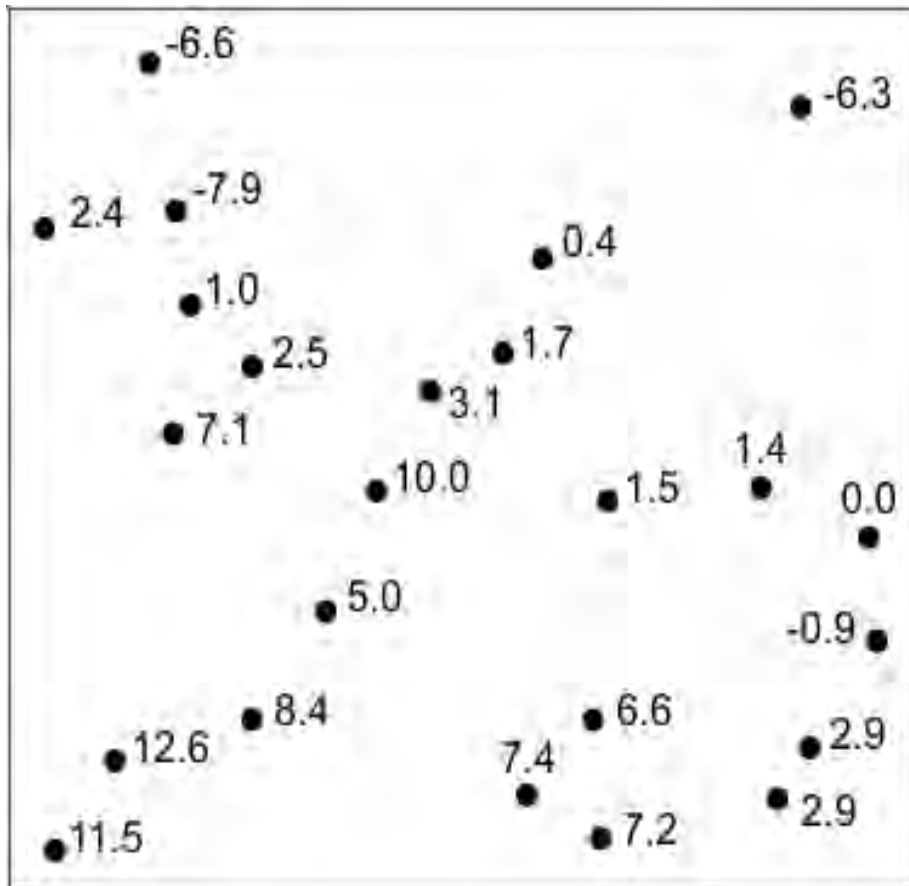


Figure 6.1 Average January Temperatures ( $^{\circ}$ F) over a part of Alberta  
 (Source: O’Sullivan and Unwin (2010), Figure 9.4)

2. Self-evidently, temperature is a continuous phenomenon. Everywhere has a value, so what we have here is a sample drawn from a continuous field of values. We can reconstruct the entire field by means of the process known as spatial interpolation in which we thread lines of equal value (generally these are called isolines, in this case they are iso-therms) through the numbers;
3. Using a pencil (you may also need an eraser!) and produce a continuous surface representation of these data, by drawing isolines of equal temperature This task is NOT as ‘easy’ or ‘trivial’ as it might at first appear. In locating the lines of equal value you need to bear in mind three things. First, don’t ‘join the dots’. The data are unlikely to be exact and each isotherm is likely to have substantial spatial width. Second, the best way to work is to select a value in the middle of the data range and locate and draw this first. For this map a mean of  $0^{\circ}$ F isn’t a bad starting point

(This part of the world is very cold in January). Having located this isotherm you can now work upwards and downwards away from it as some suitable interval. Third, perhaps a little more arguably, perhaps you should also try to make the resulting surface of average temperatures as smooth as you can, consistent with it honouring all the data. What this means is that the interpolated surface should pass through all the measured data exactly. In practice, this means that there should be no inconsistencies where measured temperatures lie on the 'wrong' side of relevant isotherms;

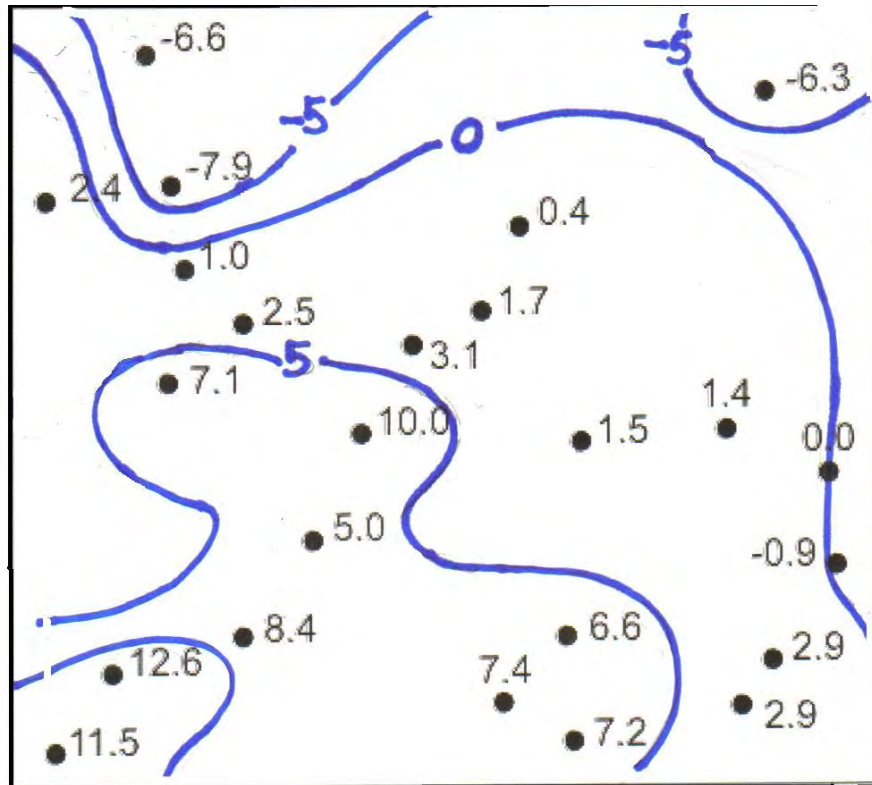
4. Now reflect on how what you have done relies entirely on three very fundamental assumptions. First, you assume *continuity*, that everywhere has a value and there are no very sudden, abrupt leaps or falls in these values. Second, you have also invoked Tobler's law that 'everything is related to everything else, but near things are more related than distant things' (Tobler 1970). This is in effect an assumption that the variable being mapped is *autocorrelated* over a distance greater than the spacing of the data points. Third you have assumed that the resulting field is *single-valued*, that is there is just one value for the variable at each and every point in the area. Think about all three assumptions. Under what circumstances might they not hold?
5. Having drawn your map consider the following problems:
  - a) Self-evidently, different people will draw different maps depending on how they interpret the field and some will be 'better' than others, so how might you assess the quality of any one interpretation?
  - b) Deliberately, the map of the sample points gives nothing else away. What additional information might have helped you locate the isotherms?
  - c) Finally, to what extent might any prior knowledge of the problem have helped in the process?

### **Comment/answers**

I confidently predict that a fair proportion of any class will make a complete mess of this exercise. Some simply don't 'get it', producing invalid surfaces that couldn't exist in reality (e.g. lacking continuity with contours crossing a lot, even bifurcating).

I like to think that hand drawn my effort isn't bad:





There is an interesting problem in these data. As a test of comprehension, students might be encouraged to think about the circumstance in which you could have an apparent cross in a contour pattern -- there is one possibility that might be seen here. Look at the isotherm for 5 degrees and the point that has the same value center left of the area. This could be a 'col' on the surface where the contours would touch and thus look as if they 'crossed'. To avoid this I have not extended them to touch at the data point but perhaps I should have? Many people put in a 'ring contour' hereabouts.

Assumptions of continuity don't hold for if we are dealing with depths of a faulted geological stratum and if there is an overfold nor will the single value assumption hold. Some things we might measure can vary at scales below that of the sampling interval and so should not be isolined.

Discussion might start with the proposition that for these data, the 'truth' simply does not exist. We can't go back to Alta and set up past weather stations even if we wanted to -- and this is very common in most science. An isoline pattern is usually a 'model' for some data, and in isolining by hand we impose our individual models on these data. The question is then what criteria can we have for quality in such an isoline pattern? A list of possibilities is:

- a) The field must be continuous;
- b) The field must honour all the data;

- c) Arguably, it should also be as smooth as possible consistent with the above;
- d) It should make maximum use of the available information given by the data about the spatial structure and the resulting isolines should be consistent with this;
- e) An 'objective' test is to separate the data into two sets and contour just one of them, looking at how well the result honors the data points that were left un-used;
- f) Although it is subjective, 'plausibility' seems to me to be a sensible criterion!

It follows that almost any additional data will help in the process, in the present case the main co-variate could well be station height, but there are others. I have repeated this and similar exercises with student classes many times and it is clear to me that preconceptions of the way we think the particular variable should behave spatially, or prior knowledge, is often what we use in drawing the isolines. A classic experiment by Dahlberg many years ago (Dahlberg, 1975) pitted experienced geologists against a computer interpolation routine using the same test data. The result was that all the maps produced were similar close to the data points, but differed, often wildly, in the uncontrolled areas where preconception played a bigger role. Very rarely in the real world are we required to isolate fields about which we have no prior knowledge.

### **Suggestions for modification**

You could of course ask students to find their own sample data, or develop a different case. Davis (2002), pages 370-397 is as comprehensive an account of issues in contouring as you are likely to see and the linked website has at least one suitable if large ( $n=360$ ) data set (LEDUC.TXT, see <http://www.kgs.ku.edu/Mathgeo/Books/Stat/index.html>)

It is also useful to introduce a little formalism into things by referring to the concept of a scalar field in which the field height is some function of location and can be described by the very general equation:

$$z_i = f(\mathbf{s}_i) = f(x_i, y_i)$$

In which  $\mathbf{s}_i$  is any location with locational co-ordinates  $(x_i, y_i)$  and which we express the ideas of continuity and single value. It is a small step from this to suggest that many of the ways we have of modeling and representing field data are different ways of expressing the same function.

As an aside, running this little experiment will almost certainly enable you to suggest that the consistency of method provided by a computer algorithm to interpolate is a powerful reason for automation.

## 6.3 Exercise (30): Isolining by machine

### Aims and introduction

This exercise adds no new concepts, but after doing Exercise (29) it would be a pity to miss it out and I would never use it without first trying to isolate by hand.

Over the years I have looked in vain for an easy-to-use freeware program that offers random to grid interpolation by a variety of methods together with some decent visualization capabilities. If what you want is visualization of some ready gridded data, then Jo Wood's *Landserf* (see <http://www.soi.city.ac.uk/~jwo/landserf/> ) is state of the art, but it does not offer any serious interpolation algorithms. If you have *ArcGIS*<sup>™</sup>, or are familiar with '*geoR*', then well and good, you will be able to do this exercise using those systems. *3DField* is a useful resource that is provided as limited freeware system and a more powerful package that costs very little.

### Geometry, space and level

As Exercise (29).

### Intended learning outcomes

After completing this exercise students will:

- Be able to run a simple computer-based interpolation routine;
- Understand that such routines have the advantage of consistency, but can produce poor results if used uncritically;
- Understand that a variety of visualization methods can be adopted.

### Resources needed

Machine with the freeware version of *3DField* running and with a data file containing the temperature data from Alberta used in Exercise (29) as:

|   |      |      |      |
|---|------|------|------|
| 1 | 1.8  | 0.8  | 11.5 |
| 2 | 5.7  | 7.1  | 12.6 |
| 3 | 1.2  | 45.3 | 2.4  |
| 4 | 8.4  | 57.1 | -6.6 |
| 5 | 10.2 | 46.7 | -7.9 |
| 6 | 11.4 | 40   | 1    |
| 7 | 15.9 | 35.4 | 2.5  |
| 8 | 10   | 30.9 | 7.1  |
| 9 | 15.7 | 10   | 8.4  |

|    |      |      |      |
|----|------|------|------|
| 10 | 21.1 | 17.5 | 5    |
| 11 | 24.5 | 26.4 | 10   |
| 12 | 28.5 | 33.6 | 3.1  |
| 13 | 33.5 | 36.5 | 1.7  |
| 14 | 36.4 | 42.9 | 0.4  |
| 15 | 35   | 4.7  | 7.4  |
| 16 | 40.6 | 1.6  | 7.2  |
| 17 | 39.9 | 10   | 6.6  |
| 18 | 41.2 | 25.7 | 1.5  |
| 19 | 53.2 | 4.4  | 2.9  |
| 20 | 55.3 | 8.3  | 2.9  |
| 21 | 60   | 15.6 | -0.9 |
| 22 | 59.1 | 23.2 | 0    |
| 23 | 51.8 | 26.8 | 1.4  |
| 24 | 54.7 | 54   | -6.3 |

Column 1 is an identifier, 2 contains the **X** co-ordinates, 3 the **Y**'s and 4 the **Z**'s. These data are in the text file *Alta\_temperatures.txt*

### Suggested student briefing

1. If it has not already been done for you, go to <http://field.hypermart.net> and download and install a copy of the free version of *3DField 3.1.0*. You will see that the 'unregistered' version limits you to  $\leq 50$  sample points and has reduced set of functions with no 3D views;
2. The  $(x, y, z)$  data for our January temperatures over Alberta are in the file *Alta\_temperatures.txt*;
3. Our aim is to create isotherm maps of these data using a selection of the standard algorithms. *3DField*'s interface takes a bit of getting used to but once up and running turns out to be fairly easy to use;
4. First, examine the *Alta\_temperatures.txt* data using Microsoft *Notepad*<sup>™</sup> to see how they are organized. The column order is Number, X co-ordinate, Y co=ordinate and then the temperature, Z;
5. Start *3DField*. First, you will need to ensure that the system understands how these data are formatted. Use EDIT>FORMAT SCATTERED DATA to set column 1 as the Label, column 2 as the **X**, columns 3 as **Y** and column 4 as **Z**;
6. Now incorporate the temperature data into the system using MAP>OPEN DATA FILE and navigate to where you have stored these data. Your reward should be a map of these sample data points. Should you wish it, you can improve the look of these using EDIT>POINT>MARKER and using a bolder font;

7. Like most interpolation systems *3DField* takes a two-stage approach. First, it uses one of a number of algorithms to take the data from a semi-random scatter to a very fine grid of estimated values. Second, it offers a variety of visualization options to map this grid. The density of the grid can be varied using options in GRID, but the business part of the system is either in the text window or hidden in OBJECTS>GRIDDING METHOD. In this first analysis go via OBJECTS>GRIDDING METHODS and step up to INVERSE DISTANCE METHOD, then chose OPTIONS and check that the distance exponent (POWER) is equal to the default value of 2.0. OK this and then SELECT the method. *3DField* calculates a fine grid of estimates using this as its interpolator;
8. Visualize this grid using MAP LIST>SIMPLE CONTOURS and customize these to be the same as those you used in your hand drawn effort in Exercise (29). The result is likely to be a shock and perhaps even a blow to your self-esteem. *3DField* has numerous options that let you change the cartography and you'll probably want to experiment with them, add text and so on, until you get something that of which you might be proud;
9. Having familiarized yourself with at least some of what the system can do, experiment with different interpolators (the choice is very, very wide!), recording your results as you go along. As each map is created you can save it in a selected standard format using FILE>SAVE IMAGE.

### **Comment/answers**

A *3DField* generated map using the inverse distance weighting (IDW) algorithm with a distance exponent of 2.0 gives the interpretation shown in Figure 6.2.

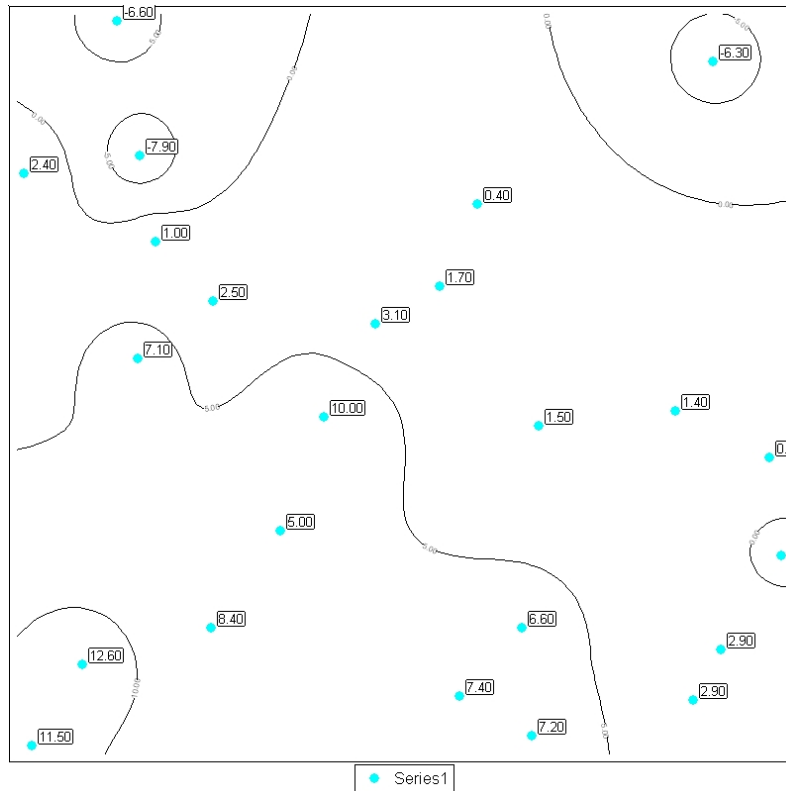


Figure 6.2 The Alta data isolated using IDW with an exponent of 2.0

I note that this isn't all that dissimilar from my hand-drawn effort, but students using the system will undoubtedly find that the computer-produced results are almost always better than their own efforts. Many years ago I did a simple study in which I matched student hand drawn maps of these same data to inverse distance weighted machine examples, discovering that choice of a distance exponent of 2.0 is a pretty good average for a human interpolator. Those with a very, very long memory will recall that the venerable and venerated SYMAP program of the mid-1960s used an exponent with the same value in its version of inverse distance weighting.

Once you have the data in a system that offers numerous methods of interpolation, each with numerous possible settings, the temptation is to generate lots and lots of maps. What of course you discover is that you can get plausible results from many (but emphatically NOT all) of these and yet we have no clear metrics for deciding which is 'best'.

### **Suggestions for modification**

If your interest is in automated contouring or visualization, the obvious extension is to develop an exercise in which students use *3DField* to experiment with a variety of interpolators and visualization methods.



## 6.4 Exercise (31): Visualizing fields

### **Aims and introduction**

Contour-type maps with various enhancements to help visualization are often found on the WWW. This exercise uses a search engine to find examples of maps that illustrate some of the ways used to represent field data cartographically. Examples include 'spot heights', contours, hypsometric tints, relief shading, and hachuring.

### **Geometry, space and level**

As Exercises (29) and (30), but perhaps concentrating on complex concepts such as relief and gradient. A careful search will find examples of the concept of a vector field, but note also that many method of relief enhancement used by cartographers use some version of 'hill shading' that makes use of the same concept.

### **Intended learning outcomes**

After completing this exercise students should be able to:

- Identify a range of methods used by cartographers to display fields;
- Critically evaluate these methods.

### **Resources needed**

Internet browser and/or access to a good map library.

### **Suggested student briefing**

1. Use Google Images™ or a similar search engine to find examples of contour type maps used (a) to represent earth surface relief and (b) to represent the field generated by some other spatially continuous variable such as temperature and mean annual rainfall.
2. In each case, list the methods used to indicate both the height of the field and its shape. Don't hurry this task: the list is almost certain to be longer than you might at first have thought;
3. In each case comment on the effectiveness of the method and its combination with other approaches.

## **Comment/answers**

A good, if dated, source is the various old *Tourist Editions* of the Ordnance Survey of Great Britain, based on their old one inch to the mile mapping. These sheets were used as test beds for sometimes very innovative cartography involving numerous methods for relief representation. My own favourite is the original early 1960s map produced for the Cairngorm mountains of Scotland that coupled spot heights, contours, hypsometric colouring and hill shading to produce one of the most beautiful maps of all time.

Standard methods used include spot heights and bench marks, contours, layer colouring/hypsometric tinting, oblique hill shading and even hachuring. There is a dated review in Unwin (1981, pages 153-161) and virtually every cartography text ever written. Computers have greatly extended the range of available methods, such that what is clearly intended as a 'map' seems almost to merge with displays that could justifiably be cited as examples of virtual reality. For a discussion of some the issues that this ability raises, see Wood (2002).

## **Suggestions for modification**

This exercise is useful as precursor to a discussion of how we get from raw height data, themselves a *sample* of the field, through some sort of *model* of the field, to the *representation* of that model in what essentially consists of three steps: sampling, describing by a continuous model and representing as a map (see O'Sullivan and Unwin 2010, 243-250). In turn this perhaps should lead to a critical assessment of the contour. It can also lead to discussion of the notion of the vector field given by the gradient of the field.

## 6.5 Exercise (32): Trends in fields

### **Aims and introduction**

A spatial concept that is often of interest is any regional trend in the height of a surface that can be explored using one of the collection of methods and approaches called *trend surface analysis*. GIS such as *ArcGIS*<sup>™</sup> and *IDRISI*<sup>™</sup> have this capability built in but it is a very simple matter to fit trend surfaces using any software that offers basic statistical analysis such as *SPSS*<sup>™</sup> and *MINITAB*<sup>™</sup>. This exercise uses Microsoft *Excel*<sup>™</sup> to do the required calculations for two simple problems.

### **Geometry, space and level**

Again we are in a Euclidean metric space, but dealing with an analytical/third order concept, that of trend.

### **Intended learning outcomes**

After completion of this exercise students will:

- Understand what is meant by the 'trend' of a surface;
- Be able to use Excel<sup>™</sup> to model linear trends across a surface;
- Be able to assess the statistical significance of any fitted model;
- To produce simple maps of a fitted linear surface;
- Know how to extend this analysis to fit quadratic trends.

### **Resources needed**

The data file (derived from the spot heights shown in the figure below and the data for January temperatures in a part of Alberta, Exercise (29) above held in *Alta\_temperatures.txt*. Access to Microsoft *Excel*<sup>™</sup> with the *Analysis* routines loaded.

### **Suggested student briefing**

1. First, and very important, ensure that on your machine you have the Excel<sup>™</sup> Analysis routines loaded. Do this by going to >TOOLS, >ADD INS, then tick ANALYSIS TOOLPAK and ensure that it loads;
2. To gain confidence, fire up the analysis-enabled Excel<sup>™</sup> and load in the sample data given in Figure 6.2 and in Table 6.1

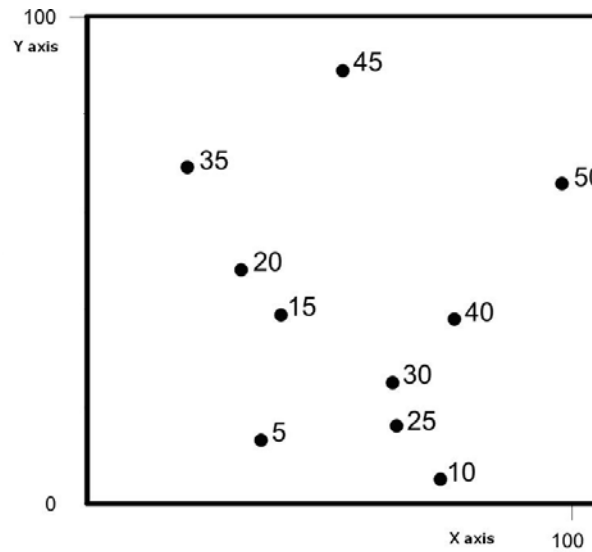


Figure 6.2 Some simple data for trend surface analysis  
 (Source: after O'Sullivan and Unwin (2003), Figure 9.5)

| Point | $x$ | $y$ | Height<br>$z$ |
|-------|-----|-----|---------------|
| 1     | 36  | 13  | 5             |
| 2     | 73  | 5   | 10            |
| 3     | 40  | 39  | 15            |
| 4     | 32  | 48  | 20            |
| 5     | 64  | 16  | 25            |
| 6     | 63  | 25  | 30            |
| 7     | 21  | 69  | 35            |
| 8     | 76  | 38  | 40            |
| 9     | 53  | 89  | 45            |
| 10    | 98  | 66  | 50            |

Table 6.1 Locational ( $x, y$ ) and height ( $z$ ) co-ordinates for the data in the above figure  
 (Source: after O'Sullivan and Unwin (2003), Table 9.1)

3. In TOOLS select DATA ANALYSIS and then use the slider to choose Regression. The simple way to compute a linear trend surface is use the mouse to select the height, or  $z$  values into the window labeled 'Input Y-range' and then repeat this for the two columns containing the ( $x, y$ ) locational co-ordinate that are our 'independent' variables. Note also that

you can switch on or off various output options. Clicking GO should produce a linear model with the constants:

$$\hat{z}_i = -9.6949 + 0.3373x_i + 0.4519y_i$$

In which  $\hat{z}$  is the estimated height at location  $\mathbf{s}_i$  with coordinates  $(x_i, y_i)$ ;

4. It is useful to draw out this surface. Draw out a square map frame with  $x$  and  $y$ -axes ranging from 0 to 100 on each axis and then produce contours of this linear trend surface at  $z = 0, 15, 30$  and  $45$ . It is easier to work 'backwards' using the above equation. Set  $z$  at the required contour value and then in turn set each of  $x$  and  $y$  at zero, in each case finding either the appropriate  $x$  or  $y$  value where the contour crosses that axis. You should end with a simple, regular inclined plane rising from the origin in the 'south west' towards the 'north east' of the frame. The  $R^2$  is 0.851 indicating a very close fit between the trend surface and the observed data;
5. Whether or not this fit is statistically significant can be tested using an F-ratio statistic

$$F = \frac{R^2 / \text{df}_{\text{surface}}}{(1-R^2) / \text{df}_{\text{residuals}}}$$

where  $\text{df}_{\text{surface}}$  is the degrees of freedom associated with the fitted surface, equal to the number of constants used, less one for the base term  $b_0$  and  $\text{df}_{\text{residuals}}$  is the degrees of freedom associated with the residuals, found from the total degrees of freedom  $(n - 1)$ , less those already assigned, that is,  $\text{df}_{\text{surface}}$ . In the example  $\text{df}_{\text{surface}} = 3 - 1 = 2$ , and  $\text{df}_{\text{residuals}} = 10 - 1 - 2 = 7$ , so that

$$\begin{aligned} F &= \frac{0.851/2}{0.149/7} \\ &= \frac{0.4255}{0.0213} \\ &= 19.986 \end{aligned}$$

This  $F$ -ratio indicates that the surface is statistically significant at the 99% confidence level and we can assert that the trend is a real effect and not due to chance sampling from a population surface with no trend of the specified linear form;

6. Can you suggest how we might modify this workflow to enable computation of a quadratic trend surface?

7. Now return to the data for January temperatures in Alberta that were used as part of Exercise (30) and use the same approach to compute trend surfaces for the Albertan temperature data and comment on the results you obtain.

### **Comment/answers**

Even with the guidelines provided, Part (7) is quite challenging. The Albertan temperature data give a linear trend surface as  $z = 13.164 - 0.120x - 0.259y$ , with RSS of 73% (equivalent to a correlation coefficient of  $r=0.85$ ) and F ratio of 28.6. There is a strong regional trend with values declining to the 'north east' (larger  $x$  and  $y$  values). For more detail of this example, see O'Sullivan and Unwin (2010, pages 281-287). That there is such a drift in these data should also add a note of caution about the assumptions made when interpolating them using ordinary kriging.

Fitting a quadratic is easy, simply create extra columns in Microsoft *Excel*<sup>TM</sup> for the additional terms  $x*y$ ,  $x^2$  and  $y^2$  and repeat the analysis with five predictors ( $x$ ,  $y$ ,  $x*x$ ,  $y*y$  and  $x*y$ ). Doing this to add the quadratic terms gives an RSS of 83%, which is probably 'worth' the three extra parameters and really only leaves a couple of data points badly fitted. There are maybe two notes of caution here. First, simply adding terms will always improve the fit, so care needs to be taken to ensure that the additions are worthwhile, both by being scientifically sensible and statistically significant. Giving a surface too much room for maneuver can generate severe edge effects including a marked tendency to create contours parallel to the frame used. Similarly, fitting very high order surfaces involves creating matrices for which the inverse can be hard to find, resulting in the possibility of severe numerical error.

There is a very basic guide to trend surface analysis, Number 5 in the old CATMOG series, now available at <http://www.qmrg.org.uk/catmog>.

### **Suggestions for modification**

Given that we have only  $n=24$  data points, it might be useful at this point to examine some other regression diagnostics, such as the effects of each datum on its own fitting (or *leverage*). Over twenty years ago Neil Wrigley and I showed that plots of the leverages against the residuals can be very useful aids in the interpretation of polynomial trend surface results Unwin and Wrigley, 1987a, 1987b)

## 6.6 Exercise (33): Spatial structure and spatial interpolation by kriging

### Aims and introduction

If they can be mapped, and are not just a series of random numbers, fields must exhibit spatial autocorrelation that can be characterized by developing a semi-variogram (SV) model. Using such a model to produce an interpolation by kriging is in some sense *optimal*, because it uses the autocorrelation in the observed data to inform the choice both of the range and the weightings to be applied. Many GIS simply compute the SV, summarize it in some standard default way, and then offer a limited choice of models to be used in the interpolation. My view is that, although it might be acceptable as a means of producing interpolations that look reasonable, this approach sweeps under the carpet not only a lot of detail that should not be treated in this way, but also hides the utility of a fitted semi-variogram model as a summary of the structure of the field.

### Geometry, space and level

As Exercise (29), but introducing the analytical/third order concept of the semi-variogram.

### Intended learning outcomes

After completing this exercise students will:

- Be able to use *3DField* to produce an interpolation of a field using a suggested model of the SV;
- Understand that different models will produce different maps;
- Have at least some insight into the complexity of the SV modelling process.

### Resources needed

Computer running *3DField*. The data to be used are in the file *topo\_data.txt*, which has  $(x, y, z)$  earth surface relief data taken directly from Davis (2003, pages 373) made available on the website at <http://www.kgs.ku.edu/Mathgeo/Books/Stat/index.html> as the text file NOTREDAM.TXT. To get these into the freeware version of *3DField* I have reduced their number from the 52 to 50 by deleting two that didn't seem to help the interpolation very much. The height,  $z$ , co-ordinates are in the file and are in feet above the datum used. The  $(x, y)$  are such that one map unit is 50 feet, so the entire

area is quite small, around 300 x 300 feet. Students will also need a display showing the plot of the experimental SV data (see below).

### Suggested student briefing

1. Bring the  $(x, y, z)$  data in the file `topo_data.txt` into *3DField* (you will need to use `EDIT>FORMAT SCATTERED DATA` and set the **X** in column 1, **Y** in column 2 and **Z** in column 3. The objective is to create an isoline map (in this case they really are relief 'contours') using kriging;
2. Using *3DField*, produce a simple plot of these point data, without attempting any interpolation. This arrives more-or-less automatically as you bring in the data. Make sure it makes sense, and that you have the data format correct;
3. Revise how we get from this scatter to the following experimental semi-variogram shown in Figure 6.3:

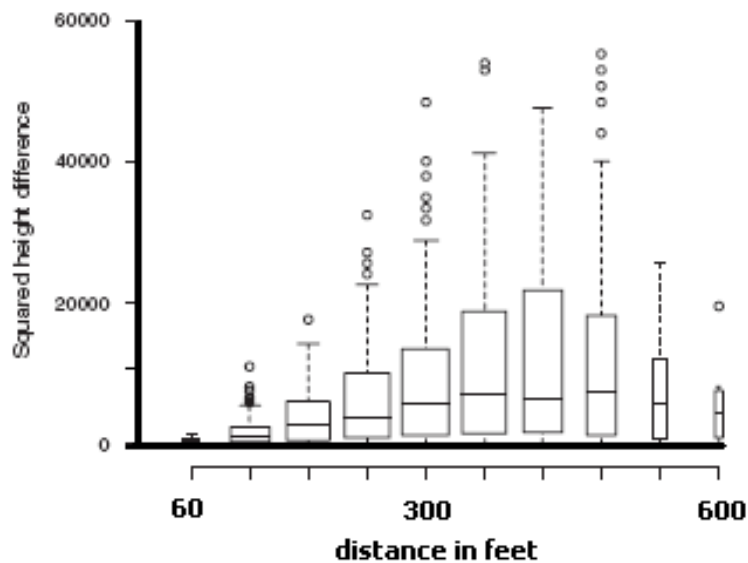


Figure 6.3 Experimental semi-variogram for the topography data

On the y-axis the quantity involved is a measure of spatial dependence, the difference in heights squared, plotted against distance apart on the x-axis. This was found by first calculating for every one of the  $0.5n(n-1)$  possible pairs of  $(x, y, z)$  data in the file their distance apart and the square of their difference in height. Since we have  $n=50$  points, the initial calculation gives  $0.5 \times 50 \times 49 = 1225$  pairs of values of:



$$2\hat{\gamma}(d) = \frac{1}{n(d)} \sum_{d_{ij}=d} (z_i - z_j)^2$$

The symbol  $\gamma$  and the '2' are conventionally used for this quantity which is called the *semi-variance*. For example the first two points in the data file have  $(x, y, z)$  co-ordinates as  $(0.3, 6.1, 870)$  and  $(1.4, 6.2, 793)$ . Confirm that these are 1.1045 distance units apart by using Pythagoras on the values for  $x$  and  $y$  and have a semi-variance of 5929 units of height squared. A plot of all 1225 of these values is called the variogram cloud, which we have then summarized by 'binning' these data into the ten intervals for which the box plots are as shown. The average value of the semi-variance at that distance lag is given by the small horizontal bar in each box;

4. So, what these means give is *estimates* of the semi-variance at a series of distances (lags) and step two is to summarize this with a continuous mathematical function that describes the way in which the height of the field changes with distance. In other words it expresses, not just Tobler's law, that 'everything is related to everything else, but near things are more related than distant things' (Tobler 1970), but also gives it a precise mathematical form for the field that our spot heights represent;
5. Not every mathematical function is suitable to describe these semi-variances but the choice is usually between the four that *3DField* offers (linear, spherical, exponential and Gaussian). For your first experiment, use the spherical model shown in Figure 6.4 that experience suggests will describe many experimental semi-variograms:

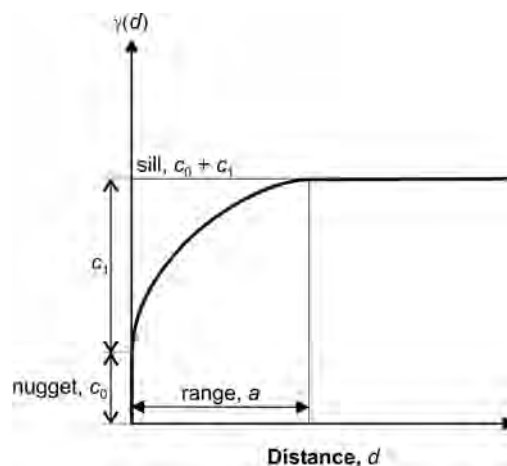


Figure 6.4 The spherical model for the semi-variogram  
 (Source: O'Sullivan and Unwin (2003) Figure 9.7)

This has three parameters that describe its shape and that need to be supplied to *3DField*. The first is the *nugget*, denoted  $c_0$ , and is the variance at zero distance. At zero distance all  $z$ -values should be the same and hence their difference should also be zero but this is rare for functions based on experimental data. A non-zero nugget may be thought of as indicating that repeated measurements at the same point yield different values and so the nugget value is indicative of uncertainty or error in the attribute values. Second, there is the *range*, denoted  $a$ , which is the distance at which the semi-variogram levels off and beyond which the semi-variance is constant and, third the *sill* is the constant semi-variance value beyond the range. The sill value of a data set may be approximated by measuring its total variance. The precise mathematical form of this model is:

$$\gamma(d) = c_0 + c_1 \left( \frac{3d}{2a} - 0.5 \left( \frac{d}{a} \right)^3 \right)$$

for variation up to the range,  $a$ , and then

$$\gamma(d) = c_0 + c_1$$

beyond it;

6. Next, you need to use the experimental data in step (3) to estimate the nugget, and sill variances and the range (feet) for these data. To begin with try a nugget of zero (which makes sense, why?), a range of 400 feet and sill variance of 6,000 units;
7. Finally, use *3DField* to produce the interpolation using OBJECTS>GRIDDING METHOD, select KRIGING and under OPTIONS pick CUSTOM. Now double click on the map list your chosen map type (SIMPLE CONTOURS?), which brings up a screen that allows you to set the kriging type (Spherical) and our chosen parameters (nugget  $c_0 = 0.0$ , range,  $a = 400$  (feet) and sill  $c_1 = 6000$ ). On the next screen chose a sensible value for the lowest contour to be displayed and the contour interval. Your reward should be a fairly 'good' map for these data. The mathematics of how the estimates are produced from the semi-variogram are outlined in O'Sullivan and Unwin (2010, pages 302-310);

8. Having got this far, it would be a pity not to experiment with other values for the spherical model parameters (do they make much difference?) and function used (likewise).

### **Comment/answers**

You might like to look at what others have made of these same data. They have been analyzed many times:

- a) Their source is the excellent text by Davis (2002, page 373). Davis subjects them same data to a variety of isolining procedures, reported on pages 374-387 of the same text;
- b) In Ripley (1981, pages 58-75) a statistician fits a variety of cleverly-derived surfaces to these same data, but seems unaware of the fact that they represent surface relief.

Figure 6.5 is a possible answer that I quite like from *3DField* using spherical model kriging with no nugget (no reason for one with topographic heights?), 390 feet range. Usually, people find kriging with a spherical model gives decent answers, but then so can the much simpler inverse distance weighting (IDW) if one is careful to chose the exponent. I confess to a suspicion here that things aren't quite as clear cut as one might imagine from reading the literature, and this is because our final interpolator is a weighted sum of the data values in some defined neighborhood around the location whose value is being estimated. In both kriging and IDW the weights decline with distance to some limit. It may well be that almost any function that decreases the data weighting with distance from the location and sums to unity will give a 'reasonable' result.

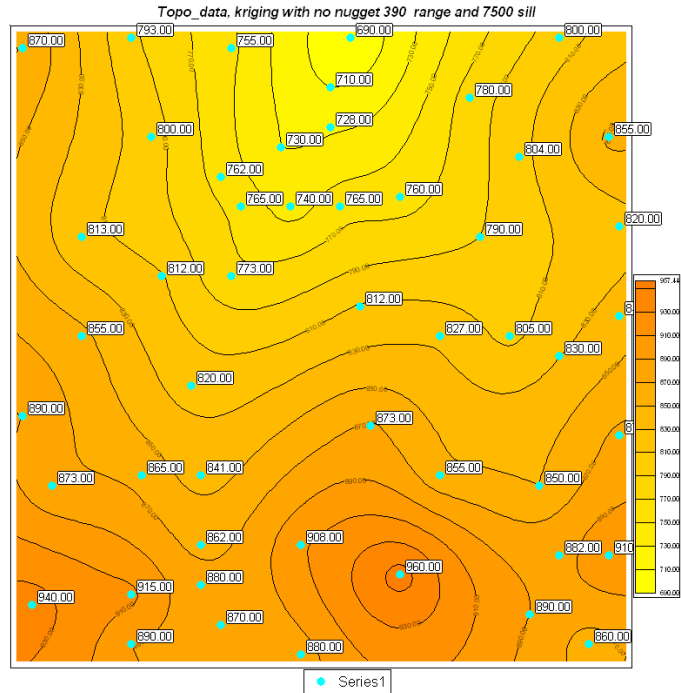


Figure 6.5 A kriged interpolation using a spherical model

For comparison, Figure 6.6 shows a standard inverse distance interpolation with a distance exponent of 2.0.

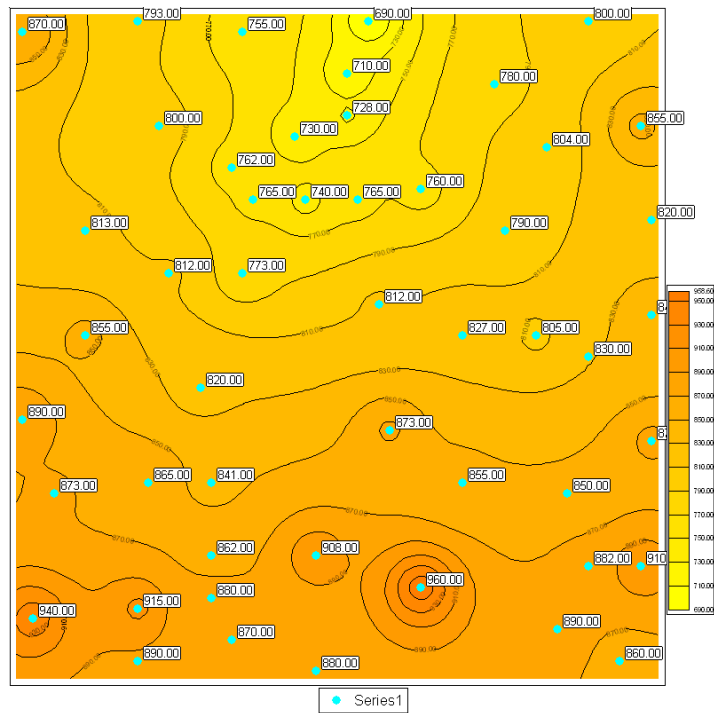


Figure 6.6 IDW interpolation of the same data with  $e=2.0$

For fairly obvious reasons 'Bull's eye' ring contours are a frequent outcome in IDW. There are a few in this example. Too many almost always indicate a poor quality answer where the method is struggling to honor a single data point.

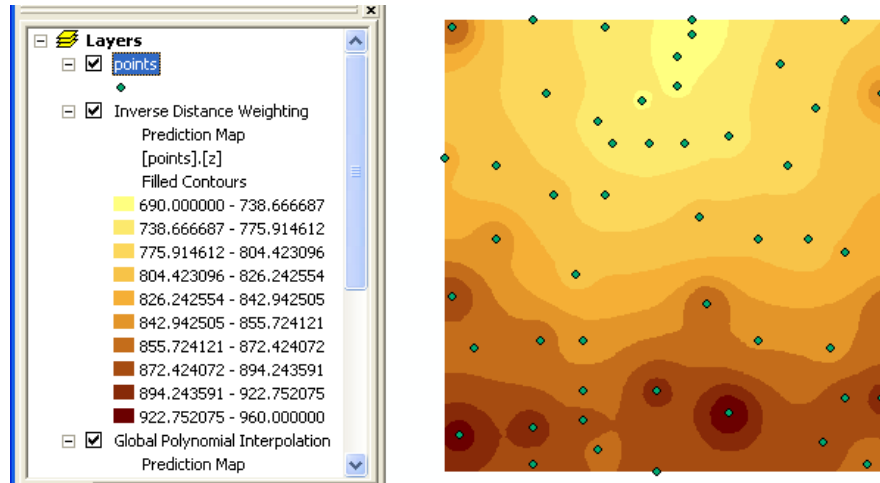


Figure 6.7 Interpolation in *ArcGIS™* using IDW

The same issue is shown in Figure 6.7, which is another using IDW but this time using *ArcGIS™*. Note that we have a lot of 'ring contours' that reflect the need to honour all the data. We also get some enclosed depressions that, limestone 'sink holes' excepted, don't exist in real terrain.

Given that we can produce hundreds of different maps, the temptation is to assume that the one that took most computational effort is likely to be the best, but this seems questionable. Many people seem content to use standard kriging even when there is a clear 'drift' in the mean of the data, as there is in this case, with a linear trend surface defined as

$$z = 913.8 - 1.695x - 25.252y$$

giving quite a good fit with an RSS of 66%. We should remove the drift, kriging the residuals and recombine the kriged estimates with the predicted trend values. Figure 6.8 is a perspective view from the 'north' of a surface produced by a careful kriging analysis using a spherical model after first removing the linear drift in the mean. Its fault, if fault it be, is the use of a colour 'ramp' from blue to red that isn't what we'd normally expect for the land surface relief (usually green, through browns to white?).

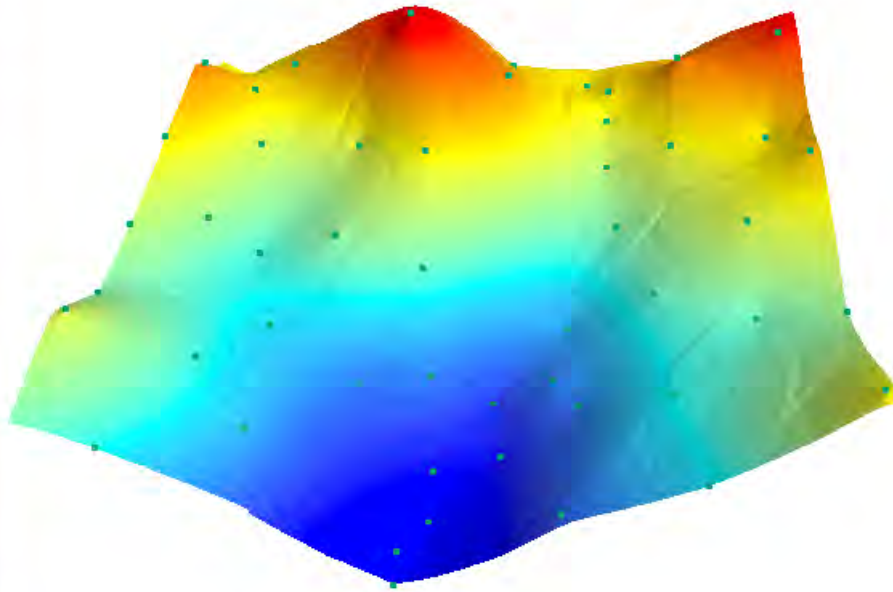


Figure 6.8 Perspective view of the surface produced by universal kriging

### **Suggestions for modification**

In an ideal world what is really needed is some user friendly software that takes the student through the entire semi-variogram modelling process, starting with computation of the cloud, showing the effect of differing binning strategies on its summary and then estimating the parameters of any chosen model. Lot's of systems for semi-variogram estimation exist, but to date I have not found one that actually meets these desiderata. Most (see <http://ai-geostats.org> for a current list), although good for research, are not designed for teaching and are simply too difficult for students to use. *Variowin*, originally a DOS program developed in the 1990s by Pannatier (1996), is awkward to use but there is an excellent step by step introduction to its use by Anselin to be found at <http://geodacenter.asu.edu/system/files/variowin.pdf> that could form the basis of such an exercise. Exercise (34) shows what can be done by careful design for teaching.

## 6.7 Exercise (34): Spatial structure from the semi-variogram models

### Aims and introduction

Completion of Exercise (33), might lead to the conclusion that there is nothing to choose between kriging and IDW interpolation. This exercise uses *E(Z)-Kriging*, a simple piece of software created by Dennis Walvoort of Wageningen (Holland) to demonstrate some of the theoretical advantages of using kriging. I find this a superb resource for teaching and learning about interpolation by kriging and the message students should take from what follows is much the same as that elsewhere: the phrase 'engage brain before clicking mouse in your GIS' summarizes it neatly! The exercise takes the form of a series of simple tests using data supplied by the system.

### Geometry, space and level

As Exercise (29), but concentrating on the analytical/third order concept of the semi-variogram.

### Intended learning outcomes

On completion of this exercise, students will have an intuitive understanding of some of the properties of interpolation by kriging:

- The effect of distance on the weights assigned to neighbouring control points;
- The ability to 'mask' data points;
- The ability to decluster clustered distributions of control;
- The impact of the nugget on the variance of the estimates and thus on the certainty of the interpolation values;
- The effect of different models for the semi-variogram.

### Resources needed

Computer running *E(Z)-kriging*.

### Suggested student briefing

1. If this has not already been done, download a ZIP file for *E(Z)-kriging* from:

<http://www.ai-geostats.org/index.php?id=114>

Unzip it and then run the program. You should also download and examine the associated manual, which has suggestions for some experiments that you can conduct for yourself using the system;

2. The basic display consists of three panels. On the left is a simplified data configuration, showing seven sample points arranged in a circle, which are thus at equal distances from the central location, the red dot, whose  $z$ -value is to be estimated. With the basic model all are just inside the specified range of the SV model. At the bottom is a data table giving the sample values. Note that you can use the mouse to move these points around, including the location to be interpolated. You can also change the number of points initially in the circle. The central pane allows selection of models for the semi-variogram, offering three different functional forms. The final panel shows a bar chart of the kriging weights, the predicted value and its error variance. There is also a very useful button from which to access the matrices used;
3. Now undertake the following experiments and answer the various questions that arise:

### **Experiment 1**

- What is the predicted value at the red dot's location?
- With the initial spherical model and data configuration, what weight is attached to each of the data point (The matrices section is the best place to look)?
- Why are these values all the same and on what do they depend?

### **Experiment 2**

Now move data point 6 inwards towards the circle center, stopping around 50 distance units away.

- What is the new predicted value and why has it changed?
- What is the weight now associated with data point 6, and what has been the effect on the variance of the interpolated value?
- Why has the pattern of weights on the other data points changed, especially that for data point 5?

### **Experiment 3**

Reset the system. Now move data point 5 around the circle circumference until it is as close to data point 6 as you can get it.



- What is the effect on the predicted value at the red dot location, its variance?
- From an inspection of the weights can you say why this has happened?

#### Experiment 4

Reset the system. Retaining the spherical model, use the slider to increase the nugget from zero upwards to 30 units.

- What is the effect on both the prediction and its variance?
- How does this variance change as a function of the assumed nugget?

#### Experiment 5

Finally, reset the system again, but this time fit both Exponential and Gaussian models to the same data configuration.

- What effects do these alternative SV models have on both the predicted values and their variances?
- Which model is 'best'?

#### Comment/answers

Students should acquire what at least is an informal understanding of the properties of various kriging solutions. Because of uncertainties in the exact positioning of the points, numerical values won't quite match these, but should be close.

Ex1:  $\hat{z}_0 = 33.4$  with a variance of 115.0 (i.e. +/- 10.7 units)

$w_i = 0.143$  for all  $i$ . There are two conditions here. First the distances from the red dot are all the same and, second, the points are evenly distributed around the circle. You can check that this latter condition is important by moving a point and this property distinguishes kriging from IDW. Note also that the weights do NOT depend on the data point  $z$ -values.

Ex 2:  $\hat{z}_0 \approx 40.1$  with a slightly reduced variance of  $\sim 106.4$  (i.e. +/- 10.3 units)

$w_i = 0.314$ . Data point (5) is being 'masked' by (6) so its weight drops to around 0.090. This masking property is another that differentiates Kriging from IDW

Ex 3:  $\hat{z}_0 \approx 31.5$  with variance 116.8. This illustrates the de-clustering property of kriging, yet another improvement on IDW, which would give both points the same weight. The weights on (5) and (6) are reduced to around 0.105 (5) and 0.092(6).

Ex 4:  $\hat{z}_0 \approx 33.4$  which is the same, but the variance increases to 149.3 (+/- 12.2). This reflects the uncertainty that a non-zero nugget introduces. It is easy to show how the prediction variance increases as the nugget is increased to a point at which the model is 'all nugget'.

Ex 5: With the same settings (where appropriate) the exponential model gives the same prediction but decrease the variance to 62.6 (+/-7.9) and so might be held to be better. The Gaussian with a variance at 56.9 (+/- 7.5) does even better. I guess one might declare the Gaussian the winner. In most applications it is the behavior at close range that matters most and it is often the case that these models improve on the standard ready-to-run spherical.

### **Suggestions for modification**

Walvoort suggests a series of other things you can investigate, for example:

- The distance effect of points, especially those outside the range, making a comparison with inverse distance weighting 'all or nothing' approach;
- The effects of using differing data values;
- Experimentation with varying block sizes in block kriging is also possible if you cover it in a more formal setting.

## 6.8 References

Dahlberg, E.C. (1975) Relative effectiveness of geologists and computers in mapping potential hydrocarbon exploration targets. *Mathematical Geology*, 7(5/6): 373-359.

Davis, J.C. (2002) *Statistics and Data Analysis in Geology* (Third Edition, New York: Wiley)

O'Sullivan, D. and D.J. Unwin (2010) *Geographic Information Analysis* (Second Edition, New York: Wiley)

**Pannatier, Y. (1995) *Variowin: Software for Spatial Statistics. Analysis in 2D* (New York, Springer Verlag)**

Ripley, B.D. (1981) *Spatial Statistics* (New York: Wiley)

Tobler, W. (1970), A computer movie simulating urban growth in the Detroit region, in *Proceedings of the I.G.U. Commission on Quantitative Methods*, Vol. 46, Supplement, June 1970 of *Economic Geography*.

Unwin, D.J. (1981) *Introductory Spatial Analysis* (London: Methuen)

Unwin D., and N. Wrigley (1987a) Control point distribution in trend surface modelling revisited: an application of the concept of leverage, *Transactions, Institute of British Geographers*, 12: 147-160

Unwin D. and N. Wrigley (1987b) Towards a general theory of control point distribution effects in trend surface models, *Computers & Geosciences*, 13: 351-355.

Wood, J. (2002) Visualizing the structure and scale dependency of landscapes. In: Fisher, P and D. Unwin, eds.) *Virtual Reality in Geography* (London: Taylor & Francis) pages 163-174.

# Chapter 7: Taking it further

## 7. 1 Introduction

The instructor who has got this far will no doubt have spotted gaps in the coverage of spatial concepts that are offered. A comparison of the content with the various concepts listed in Table 1.3 will show that no exercises have been included that examine concepts (typically at the analytical/third order level) such as anisotropy/isotropy, network generation models, enclosure, hierarchy, dominance, spatial interaction, vector fields, and surface networks. I have no doubt that exercises similar in spirit and structure to those provided could readily be developed for these concepts and, for that matter, any of the others in what might be a long list.

The internet and WWW have changed almost everything, making it easy to find and to share teaching resources of this sort. Indeed one of the more remarkable outcomes of the internet times in which we live has been the willingness of instructors to share teaching resources. Any list of available materials for teaching basic geography would be very long and anyhow rapidly become outdated. All that can be offered here are some 'starters' that will give some feel for what is available.

## 7.2 General Geography and Geographic Information Science projects of relevance

Although it might be assumed that spatial concepts would be addressed by instructors working over the entire range of academic geography, in practice most concern seems to have been shown by the geographical information science community and/or those concerned with computer based learning. Over the years there have been a number of projects concerned to advance teaching in the geographical sciences that have developed resources of use in teaching and learning spatial concepts. Examples include:

a) In USA Ken Foote's *Geographers Craft* and *Virtual Geography Department* materials remain as an example of some very forward looking ideas and can be accessed at:

<http://www.colorado.edu/geography/gcraft/contents.html>

and

<http://www.colorado.edu/geography/virtdept/resources/educatio/courses/courses.htm>

b) Also in USA the original 1990 *National Center for Geographic Information and Analysis (NCGIA) Core Curriculum in GIS* and the materials that were submitted in 2000 for what was intended to be an update can be found either at

<http://www.ncgia.ucsb.edu/giscc/> (2000 version) or as benchmarked by Brian Klinkenberg of the University of British Columbia at:  
<http://www.geog.ubc.ca/courses/klink/gis.notes/ncgia/toc.html> (1990 version)

Both contain relevant curriculum materials.

c) In UK during the 1990s a series of discipline-specific projects were centrally funded with a variety of pedagogic missions related to developing and/or making available resources for teaching in the universities. Of these a few survive or have been benchmarked at various websites. Probably the most useful are those assembled by the *Geography Discipline Network (GDN)* at the University of Gloucester, see:

<http://resources.glos.ac.uk/ceal/gdn/index.cfm> and  
<http://www2.glos.ac.uk/gdn/index.htm>

These materials include the complete text of the book by Gold *et al.* (1992) *Teaching Geography in Higher Education*, which remains perhaps the most comprehensive account of pedagogic issues in the field (see <http://www2.glos.ac.uk/gdn/gold/index.htm>) and to which some exercises in the workbook make reference. The same site allows keyword searches of a data base of resources at <http://www2.glos.ac.uk/gdn/keywords/keywords.htm>

d) For the past decade much of the work done by GDN and its predecessors have been collected under the umbrella of the (UK) *Higher Education Academy Subject Centre for Geography, Earth and Environmental Sciences (GEES)*, see:

<http://www.gees.ac.uk/db/search.php>

e) Most users of this workbook will be working in institutions that have an *Environmental Systems Research Institute Inc (ESRI)* site license and thus will have access to exercises that illustrate GIS concepts, keyed to sections of Longley *et al.* (2011), see <http://training.esri.com/gateway/index.cfm>

f) Some GIS textbooks list student exercises of the sort included in this workbook or have accompanying materials that do the same thing. Of these the instructor's teaching manuals that accompany *Geographic Information Systems and Science*, the well-known text by Longley, Goodchild, Maguire and Rhind will prove a useful source of ideas and projects, see

<http://www.wiley.com/college/longley>. The teacher's manual accompanying the Third Edition is Singleton, A.D., Unwin, D., and Kemp, K. (2010) *Geographic Information Systems and Science: Teaching Manual* (New York: Wiley) .A second text that contains short 'thought exercises' to illustrate basic concepts is O'Sullivan D and Unwin D J (2010) *Geographic Information Analysis* (Second edition, Hoboken, NJ, Wiley)

If your interest is primarily in analytical concepts, then there is a very good resource available as print or via WWW as:

deSmith, M. Goodchild M.F. and P.A. Longley (2006-2009) *Geospatial Analysis: a Comprehensive Guide to Principles, Techniques and Software Tools*. The three editions of this book (published by Troubador, Leicester, UK) are hard copy versions of an extensive and detailed technical website, <http://www.spatialanalysisonline.com>.

#### g) Intended learning outcomes

An examination of this workbook will show that drafting and listing intended learning outcomes (ILO) appropriate for each of the spatial concepts it illustrates is neither trivial nor an activity to be addressed after design of the resource. To the contrary it is central to the entire exercise. Relevant guidance on how to write them can be found at:

<http://ncgia.ucsb.edu/giscc/units/format/outcomes.html>

The 2006 American GIS&T Body of Knowledge is articulated through use of ILO and serves as a source for many relevant to spatial concepts:

DiBiase, D. , DeMers, M. , Johnson, A. B. , Kemp, K. K. , Plewe, B. P. and Wentz, E. A. (eds.) (2006) *The Geographic Information Science and Technology Body of Knowledge* Association of American Geographers , Washington, DC

## 7.3 Spatial concepts

Narrowing the focus to projects concerned solely with spatial literacy, there are at least three projects that attempt to address spatial literacy and that have created resources of use in teaching the associated spatial concepts.

#### a) *Spatial Literacy in Teaching* (SPLINT)

Although no longer running, the project that sponsored this workbook has websites at each of the collaborating institutions at:

<http://www.le.ac.uk/gg/splint/> (Leicester University)  
<http://www.spatial-literacy.org> (University College London)  
<http://www.nottingham.ac.uk/splint/> (Nottingham University)

#### b) *Teach Spatial*

The team at University of California, Santa Barbara (UCSB) has a number of other relevant, nationally funded projects that can be accessed from <http://www.spatial.ucsb.edu>

These projects include Don Janelle's *Spatial Perspectives on Analysis for Curriculum Enhancement* (SPACE, see <http://www.csiss.org/space>), Michael Goodchild's *Center for Spatially Integrated Social Science* (CSISS, see <http://www.csiss.org>), and the late Reginald Golledge's *Spatial Thinking* and they have done a great deal to isolate the issues and develop resources that illustrate basic spatial concepts. The most recent is *teachspatial* that has a website at <http://teachspatial.org> with a series of useful papers and resources relating to spatial concepts including a series of lessons based on spatial thinking contributed by Joseph Kerski and some outlines of concept based spatial thinking learning modules from Josh Bader. All are very worth looking at. The same website has some useful GIS-based tools that would enable simple exercise to be created using standard GI software.

Elsewhere on the website of the UCSB Geography Department will be found papers that develop appropriate ideas and theory relating to spatial concepts, such as, for example papers submitted by a number of contributors to a workshop on spatial concepts in GIS and design held in 2008 at:

<http://ncgia.ucsb.edu/projects/scdg/participants-scdg.php>

#### c) *Learning Spatially*

Diana Sinton at the University of Redlands (USA) has an interesting project (LENS) that uses spatial concepts for learning, see:

<http://www.spatial.redlands.edu/lens/about.aspx>

The section under 'learning pathways' has a number of useful links to data and mapping resources of relevance.

## 7.4 Software used in the Workbook

In addition to standard software the workbook uses three free-of-charge software systems of which two have dedicated websites with their own workbooks and related resources.

a) *GeoDa*<sup>TM</sup>

The first complete version of the *GeoDa*<sup>TM</sup> workbook contains 244 pages with 25 chapters of step by step guidelines and exercises to learn all the features of *GeoDa*<sup>TM</sup>, including spatial regression analysis and can be downloaded from:

<http://www.csiss.org/clearinghouse/GeoDa/>

or:

<http://geodacenter.asu.edu>

In addition the second of these sites has a revised version of the same workbook, a short teaching demonstration for spatial autocorrelation, an introduction to exploratory data analysis using *GeoDa*<sup>TM</sup> and an introduction to spatial autocorrelation analysis.

b) *CrimeStat III*

The website at:

<http://www.icpsr.umich.edu/CrimeStat>

has both a comprehensive *User Manual* and *Workbook* available for download. The workbook shows how to prepare data for *CrimeStat*, produce results and import them into ArcGIS<sup>TM</sup> for further analysis or presentation. It also covers entering data into *CrimeStat*, basic descriptive spatial statistics, measures of clustering, several 'Hot Spot' detection techniques, and using both single and dual Kernel Density Interpolation. The same site has a link to the US National Archive of Criminal Justice Data which could be source of further data if (and when) BOOK, BANK and SNOW become too familiar.

Luc Anselin at <http://geodacenter.asu.edu> provides a short *Introduction to Point Pattern Analysis using CrimeStat* .

c) *3Dfield*

As noted in Chapter 6, *3Dfield* is available from:



<http://field.hypermart.net/>

where there is an on-line manual/help facility. The system does not have facilities for estimating and modelling semi-variograms but if you are prepared to be patient and take things step by step, the freeware version of *VarioWin* should be available from

<http://www.sst.unil.ch/research/variowin/index.html> or from <http://www.ai-geostats.org> which often will form the starting point for a search for anything geostatistical.

There is an excellent guide to its use from Luc Anselin *An Introduction to Variography using Variowin* available, once again, from <http://geodacenter.asu.edu>

Doubtless these resources will date quite rapidly, and their web URLs change, but use of a search engine will often allow them to be re-discovered.

## 7.5 Linked Resources: Data Files

The workbook is accompanied by six small data files, all in simple text format:

| File                               | Comment                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
|------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Alta_temperatures.txt              | Transcribed from paper sources many years ago, see O'Sullivan and Unwin (2003) page 222                                                                                                                                                                                                                                                                                                                                                                             |
| BANK.txt                           | This is a famous data set that has been analyzed many times, notably by the statistician Brian Ripley. These data were taken from the website associated with the text by Davis (2002). $n = 47$                                                                                                                                                                                                                                                                    |
| BOOK.txt                           | These are the 12 sample data taken from Table 5.2 page 131 of O'Sullivan and Unwin (2010) $n = 12$                                                                                                                                                                                                                                                                                                                                                                  |
| SNOW.txt                           | This is probably the most famous point data set ever to be analyzed. It consists of the locations of 578 deaths from cholera recorded by Dr. John Snow in the Soho area of London during an outbreak of cholera in 1854. These data were digitized at the request of Professor Waldo Tobler (UCSB) by Rusty Dodson of the US National Center for Geographic Information Analysis from a reprint of Snow's book <i>On Cholera</i> (Oxford University Press, London). |
| Striae_in_Finland.txt<br>(and xls) | Davis (2002, page 317) presents these data for the recorded directions of 51 glacial striae in a 35km <sup>2</sup> area                                                                                                                                                                                                                                                                                                                                             |

|               |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
|---------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|               | of southern Finland. These are recorded as angular directions in degrees clockwise from north.                                                                                                                                                                                                                                                                                                                                                                                                       |
| Topo_data.txt | This has $(x, y, z)$ earth surface relief data taken directly from Davis (2003, page 373) made available on the website at <a href="http://www.kgs.ku.edu/Mathgeo/Books/Stat/index.html">http://www.kgs.ku.edu/Mathgeo/Books/Stat/index.html</a> as the text file NOTREDAM.TXT. Interestingly, John Davis tells me that they were originally used by him at Notre Dame University to introduce a student exercise on contouring. $n = 50$ , reduced by 2 from the original to get them into 3Dfield. |