3

Points, lines and graphs

Introduction

In ecological studies, we often wish to perform analyses of the spatial characteristics of objects or structural units by reducing their natural complexity, where practical, representing them with simplified forms such as dimensionless points, one-dimensional lines, whether curving or straight, or simple polygons. Obvious examples would be tree stems in a forest represented by points in a two-dimensional plane, or aquatic stream systems in a landscape, represented by branching linear structures. These points and lines can be studied in separate analyses, or they may be considered together, as we might in a study of a clonally growing plant, using points to represent erect stems and lines to represent the rhizomes or stolons of the system. In some specialized cases, we will still use points to represent the structural units, but the lines between them will now indicate relationships, not physical elements of the system but functional or relational interactions between the structural units. The best approach to describe and analyse that kind of system is to use graph theory. Graph theory provides a body of mathematical knowledge based on simple concepts, in which structural units are represented by dimensionless points and relationships between the units are represented by lines joining the points in pairs. For clarity, we will refer to the points when they are in a **graph** as **nodes**, and the lines between them as **edges** (see Glossary; Figure 3.1*a*). In this chapter, the Glossary provides definitions of graph terminology needed to understand the key terms (in bold in the text) related to graph theory. The nodes may have additional qualitative and quantitative characteristics associated with them, like species identification or stem size, and the edges may have directions (Figure 3.1*b*) and signs (plus or minus) as in Figure 3.1*c*, or other properties such as rates of flow. The concepts of graph theory have been applied to a range of ecological phenomena including interspecific associations in communities, spatial structure and dispersal in landscapes, as well as the relationships among metapopulations and metacommunities. The open and flexible conceptual model that it provides can contribute to our understanding of the relationship between structure and processes, including the mechanisms of configuration effects and compositional differences. This chapter shows how objects, which can be represented as points or lines, can be analysed separately or together, including within the formal structure of a graph. We also offer a review of the concepts of graph theory and how they can be applied in ecological analysis, but with an explicit emphasis on spatial graphs: those in which the nodes and edges that depict structural units and their interactions have spatial locations that are important to the characteristics of the system.

In ecological studies, individual organisms, populations and habitat patches are commonly objects of interest, while behavioural, physical and dispersal activities are some of the main processes that link these objects. Such links vary in weights according to the intensity of ecological processes and as such allow the persistence of metapopulation dynamics (Rozenfeld *et al.* 2008) and the maintenance of interactions within metacommunity structures (Holyoak *et al.* 2005). In the related fields of evolutionary

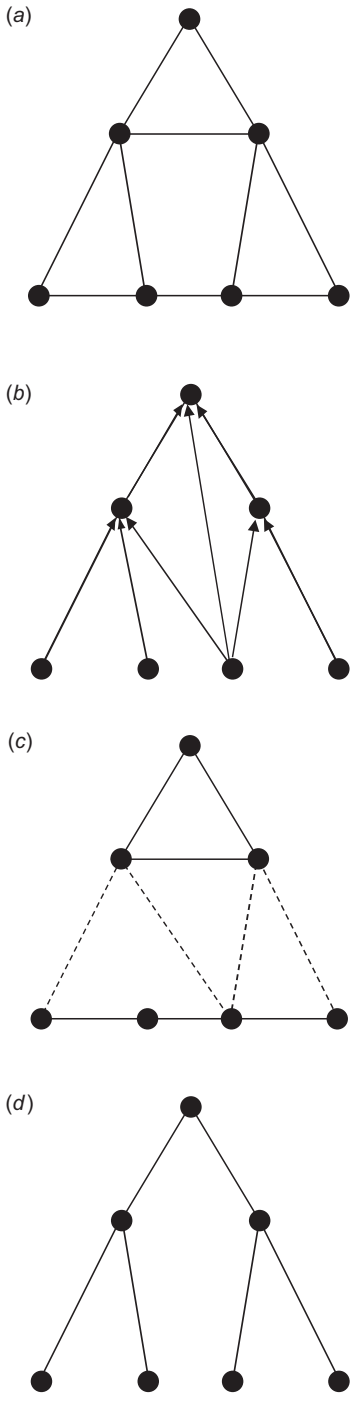


Figure 3.1 How points and lines form graphs. (*a*) A graph consisting of nodes, representing objects, joined by edges, representing relationships between them. (*b*) A digraph with edges that have directions. (*c*) A graph in which the edges have signs. A solid line for positive and a dashed line for negative relations. (*d*) A tree: a graph which has no cycles.

biology, population genetics, and spatial epidemiology, organisms and units such as taxa, traits, genes, molecular markers and so on (sometimes known as Operational Taxonomic Units) are the objects, and the links are the relationships of descent, functional pathways, or observed similarity. In graph theoretical approaches to all these systems, the objects are the nodes of a graph and their interactions or relationships are the edges between pairs of nodes (Harary 1969; West 2001; Bang-Jensen & Gutin 2002; Lesne 2006; Kolaczyk 2009). In these areas of study, organisms and interactions have been modelled and analysed for decades using **aspatial graphs** and **networks** (Dale 1977; Proulx *et al.* 2005; Lesne 2006; Mason & Verwoerd 2007; Dale & Fortin 2010). In this context, *aspatial* designates graphs in which the locations of the nodes and edges that form the graph are determined by convenience and ease of comprehension. They therefore convey no real information and cannot be interpreted, irrespective of the information from which the graph is derived.

Graph theory is not at all novel in these kinds of applications, nor only recently discovered. In the eighteenth century, Euler laid the foundation of graph theory in resolving the "Seven Bridges of Königsberg" problem (see Biggs *et al.* 1976). A century later, Cauchy and L'Huillier formalized the basic concepts of topology, as an approach to "the mathematics of position" (Biggs *et al.* 1976). Subsequently, particular types of graph were proposed such the **tree** (a graph with no **cycles**, i.e. no closed polygons formed by edges, see Figure 3.1*d*) for applications in differential calculus and isomer chemistry. Then in electrical circuits, spanning trees (graphs with no cycles, but with all nodes joined up by **paths** of edges to form a single connected structure) are often used. In the 1950s and 1960s, Erdös and Rényi introduced probabilistic graph theory and the formalizations needed to study random graphs (Erdös & Rényi 1960). The work on probabilistic graph theory has been extended to a

range of different models beyond simple independence models to consider various kinds of dependence in the generation of random graphs, such as small world and scale-free graphs (Barabási & Bonabeau 2003; Barabási 2009) or graphs in a spatial context with an upper threshold to edge **length** (**geometric graphs**; Penrose 2003).

In recent decades, graph theory has seen further important applications in food web studies (Pascual & Dunne 2006), conservation ecology (Keitt *et al.* 1997) and epidemiology (Meyers 2007). Nowadays, graphs are commonly used in ecology for a number of applications depicting physical or functional connectivity among organisms (predation, pollination, competition and other forms of interactions; Bascompte 2009) or within spatially structured groupings of local populations (metapopulations; Fagan 2002; Grant *et al.* 2007), and they can obviously be used for many more applications. It is interesting that although many of the phenomena that provided inspiration for the development of graph theory were themselves spatial (paths, maps, electric circuits, etc.), the graphs originally were not, but simplified the problems by removing the explicit spatial context. It seems to be the case that only latterly were the graphs themselves spatial, as well as the problems that inspired this more abstract approach.

Whether spatial or not, a graph's composition can be summarized as:

$$\text{graph} = \{\text{nodes}\} + \{\text{edges joining pairs of nodes}\}.$$

In formal terms, a **graph** is a mathematical object made up of two sets: **nodes** (the points, also called vertices) and **edges** (the lines, also called arcs or links) that join pairs of nodes indicating connection of the pair (Harary 1969; West 2001).

We will first discuss the analysis of point patterns with the concept of a point event's nearest neighbours. Event x is the first nearest neighbour of event y if no other point event is closer to y than x. If all pairs nearest neighbour events in a map of their locations are joined by lines, a picture of the network of nearest neighbours is produced. Technically, this network is a *graph*, with the point events as nodes, v_i, joined by edges, $e_k = (v_i, v_j)$. A graph is **connected** if there is a sequence of nodes joined by edges (a **path**) between any two nodes in the graph. (Imagine tracing a route from one node to another, following the edges.) A graph of first nearest neighbours is almost never connected, as can be seen in Figure 3.2.

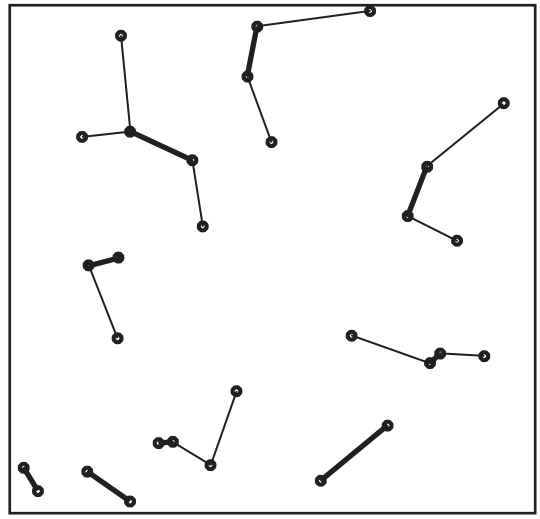


Figure 3.2 Network of first nearest neighbours for an artificial example of point pattern (NN). Mutual nearest neighbour pairs have bold lines, and form a subgraph (MNN) of the example.

In its purest form, a graph is a combinatorial entity, depicting only structure. Therefore, the nodes do not have positions and the edges do not have lengths (or weights, or directions, or shape), except for the purposes of depiction. (Our example of a nearest neighbour graph uses spatial information to determine the pairs that are joined, however, it could be redrawn with the same edges but with the locations of the points changed to reflect other criteria for their arrangement.) A graph as a geometric entity, drawn on a surface with node positions and edge lengths (that is to say, it is *embedded* in the plane), is sometimes referred to as a *topological graph* (Harary 1967), but once again, this need not be a truly spatial graph because the positions of the nodes in the depiction do not need to be the spatial location of objects.

The most basic form of graph has only these two entities as simple structural elements with no other characteristics (Figure 3.1*a*). If there are n nodes, there

are $n(n-1)/2$ possible edges. It is possible to have a graph with nodes that have no edges at all, but each edge needs two nodes to exist, and so the nodes are really the primary entities. At its most essential, a graph represents structure in an abstract way and the positions of the nodes have no meaning in the diagram we use to represent it, only convenience and clarity of presentation; the only real information is which pairs of nodes are linked.

The edges in a graph are usually simple and symmetric but they can have directions, resulting in a *digraph* (a **directed graph**, Figure 3.1*b*; see Bang-Jensen & Gutin 2002). If the edges have directions, there can actually be two edges between any pair of nodes, one in either direction, so that the number of possible edges in a digraph is $n(n-1)$. An edge may also have a sign (positive or negative) as in Figure 3.1*c*, or a magnitude, such as a weight or capacity, associated with it. A network is often understood to be a digraph in which each edge has both direction and a non-negative value related to capacity for flow (West 2001) or rules or equations governing its dynamics (Strogatz 2001). One familiar example is the trophic structure of a community, which is frequently represented by a digraph in which the nodes are species (e.g. arctic willow, snowshoe hare, lynx, etc.), with directed edges indicating predation or consumption. This digraph becomes a network when the edges have quantitative flow rates of energy or the transfer of materials associated with them, or even dynamic equations describing the feedback outcomes of each trophic link (see Glossary).

For studies in ecology and related fields, graph theory can quantify many functional and relational characteristics between the units represented by nodes. Those characteristics include structural complexity, the diversity of the nodes, the diversity of the edges, the dynamic complexity of the whole network, and the interactions of these characteristics (Strogatz 2001). This flexible conceptual model contributes to understanding the relationships between structure and function in ecological systems, including the effects of configuration and compositional differences, and the mechanisms of these relationships. In evolutionary studies, graph theory is used in the quantification of evolutionary traits, the evaluation of interactions between organisms, and in descriptions of speciation. We cannot describe all these applications, but it is useful for us to provide a few examples to illustrate the range of possibilities.

For studies of species associations in plant communities, physical and functional attributes can be combined in a graph depicting the relative frequencies of each species in the proximity of plants of each other species in the community (Hopkins 1957). In such cases, the edges joining species' nodes have signs, with solid edges by convention representing significant positive spatial associations and dashed or dotted edges indicating significantly negative spatial associations (Dale 1977; Figure 3.1*c* is an example of a signed graph). These associations are important spatial and functional characteristics of the community because only plants that tend to be in proximity are likely to interact with each other or to be responding to the same environmental conditions. Therefore, the associations determined by physical location are interpreted in relation to dynamic processes such as dispersal, establishment, growth, competition, mortality, and so on.

In the field of evolutionary biology, the phylogenetic tree is a graph and it aims to depict evolution in a group of related taxa from the ancestral taxon by a series of bifurcations that represent points of evolutionary divergence. This produces a structure with a **degree of the node** of 1 at the end of each branch (a *leaf*), representing a defined taxon, and other nodes, usually of degree 3, representing divergences (Figure 3.1*d*). These graphs are trees because they have no **cycles** in them (no closed polygons of edges) so that branches that have diverged do not rejoin. Different information bases for producing phylogenetic trees for the same group of organisms (such as protein sequences, nuclear genes or mitochondrial sequences) can give rise to somewhat different trees. Penny *et al.* (1982) used graph theory to compare phylogenetic trees of 11 mammalian species, based on five different proteins. The trees were not identical, but they were found to be significantly similar, 'consistent with the theory of evolution' (Penny *et al.* 1982). Phylogenies can also be depicted by phylogenetic networks, which allow cycles in the structure, created by the re-coalescence of

branches that once diverged, or structures that remain unresolved (Huson & Bryant 2006). In some ways, a similar net-like structure can represent a number of possible binary phylogenetic trees taken together (Gauthier & Lapointe 2007). The result is a reticulate phylogenetic structure for which a tree is not an appropriate depiction.

The focus of this chapter is **spatial graphs** for spatial analysis and these can be introduced either by starting with separate spatial sets of point events and of lines and then combining them according to a set of rules to create spatial graphs, or by beginning with graphs as abstract structural depictions with nodes and edges, and then embedding them in a spatial context. Either approach can be helpful, but we will begin with the spatial sets of points and then of lines, as these may be most familiar to the reader and have intuitive appeal.

3.1 Points: spatial patterns of point events

We can start by considering a spatial set of dimensionless points alone in one (on a line), two (in a plane), or three (in a volume) dimensions. These spatial sets of point events are the subject of 'point pattern analysis' (Diggle 1983; Dale 1999; Illian *et al*. 2008; Chapter 4 of this book), and the locations of the events result from a 'point process'. One such process is complete spatial randomness (CSR) and other processes result in the clumping of events (aggregation), or in overdispersion (mutual repulsion). One basic approach calls for testing whether the positions indicate clustering, randomness or repulsion by comparison with the results of CSR. The basic test of CSR is too simplistic for most applications, and analysis is usually to determine the scales of clustering or repulsion, not just to test complete spatial randomness at each scale. For this purpose, we use methods such as 'Ripley's K' (for details see Chapter 4 of this book, Section 4.1). Examining a range of scales is like hierarchical hypothesis testing, but the results for different scales are not all independent, and so determining the statistical significance of the results is not simple, and may be best evaluated using randomization procedures (see Chapter 1 of this book).

If the point events occur in the plane (two dimensions), and we associate with each event those regions of the plane that are closer to it than to any other, the result is a set of polygons (variously known as Voronoi, Dirichlet or Thiessen polygons; Okabe *et al*. 2000) each associated with an individual point. In this case, with each polygon belonging to one point event, it is easy to define the neighbours of any particular point as those events with which the associated polygons are contiguous with a side in common. Treating these events as the nodes of a graph, and creating edges between pairs of nodes that share a side of their polygons produces a graph of neighbours known as the Delaunay triangulation. Technically, this triangulation is the dual of the polygon structure already described, once again defining a set of spatial neighbours for each point event in the pattern. This comment leads to the following discussion of determining the 'neighbours' of point events in a point pattern, which can be done in several different ways, each producing a particular spatial graph depicting the relationships of first-order neighbours.

3.1.1 Topological neighbours

There are several ways to join points in pairs when they are in a two-dimensional plane, thus giving us several algorithms for producing spatial graphs of neighbours in a point process. Probably the simplest way is to determine the nearest neighbour of each point and then join these into pairs by lines, producing a graph of nodes and edges depicting the topology of the network of nearest neighbours. This graph of first nearest neighbours is almost never connected, in the sense that there will rarely be a continuous path of nodes and edges between any two nodes in the graph. The structure usually consists of several components that are disconnected from each other, as shown in Figure 3.2.

The nearest neighbour definition can be narrowed to include only those pairs of point events that are mutually nearest neighbours, producing a graph with fewer edges (the heavy lines in Figure 3.2). Some events will have no neighbour at all based on this definition. The graph of mutually nearest neighbour

Table 3.1 Hierarchy of neighbour networks[a] as spatial graphs

(1) Mutually nearest neighbour pairs (≈0.62)
(2) All nearest neighbours (≈1.4)
(3) Minimum Spanning Tree (≈2.0)
(4) Relative neighbourhood graph (≈2.4)
(5) Gabriel graph[b] (≈4.0)
(6) Delaunay triangulation (≈6.0)

[a] With approximate average number of neighbours for realizations of CSR.

[b] The Gabriel graph is part of another hierarchy of graphs, the β-skeleton series, with a broad range of neighbour numbers.

pairs (MNN) is a **subgraph** of the graph of all first nearest neighbours (NN) because the set of nodes is the same in the two structures and all edges in MNN are also included in NN. Under complete spatial randomness, about 62% of the point events are members of reciprocal nearest pairs and so the expected number of edges per node is about 0.62. This means, of course, that an expected 38% of the nodes have no edge at all under CSR, and are unconnected 'singleton' nodes, although the actual proportion will vary from realization to realization of the same model. For other models of spatial point processes, the expected numbers will change with the model, as well as varying from one realization to another.

The first nearest neighbour graph can be extended in several ways: for example, to include the first and second nearest neighbours, or first, second and third, and so on. With a slightly different approach, we can describe a neighbour hierarchy that begins with the graph of mutually nearest neighbours and ends with the Delaunay triangulation (Table 3.1). The advantage of using several networks for analysis is that a range of numbers of neighbours and a range of average distance to neighbours are available. This approach should provide more insight, for example, into the characteristics of segregation or aggregation of a bivariate pattern.

The third network in the hierarchy of network is the Minimum Spanning Tree (Figure 3.3). A spanning tree is a graph with no cycles that includes all nodes; the Minimum Spanning Tree is the one such tree with the smallest total length of edges. In spatial graphs, the length being minimized is just physical or geographic distance, but in other contexts, different measures may be more appropriate. The Minimum Spanning Tree (MST) contains all the nearest neighbour edges and so the nearest neighbour network is a subgraph. One way to visualize the relationship between the two graphs is to consider that the Minimum Spanning Tree is formed by connecting the disconnected **components** of the nearest neighbour graph, using the shortest edges available. A spanning tree on n nodes must have $n - 1$ edges, giving an average number of neighbouring nodes of $2 - 2/n$, or approximately two neighbours for each event.

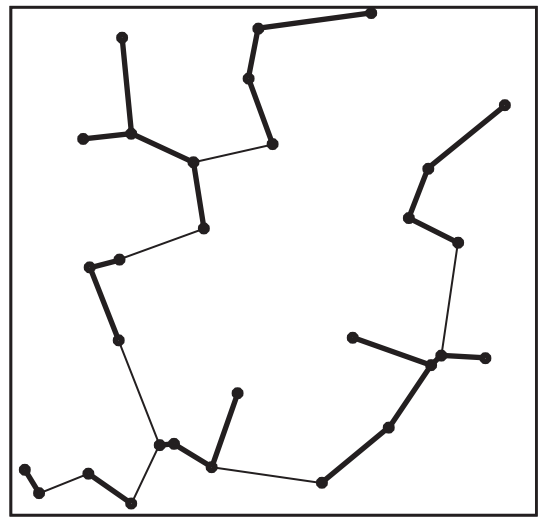


Figure 3.3 Minimum Spanning Tree for the same pattern as shown in Figure 3.2. The bold lines are those in the nearest neighbour graph.

The fourth network is the relative neighbourhood graph (RNG; Toussaint 1980), formed by joining all pairs of nodes, A and B, for which the lens formed by the radii of the two circles AB, centred on A and B, contains no other node (Figure 3.4*a*). This fourth network contains the Minimum Spanning Tree as a subgraph, which is illustrated in Figure 3.4*b*. The average number of neighbours in this network seems to be about 2.4 under CSR.

Next in the hierarchy is the Gabriel graph (GG; Gabriel & Sokal 1969). It is formed by joining all pairs

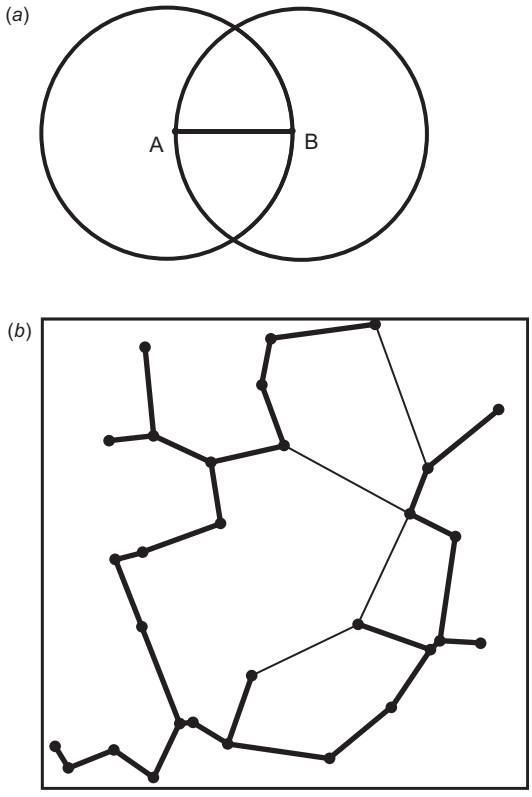


Figure 3.4 The relative neighbourhood graph for the same pattern as shown in Figure 3.2. (*a*) The radii of the two circles AB, centred on A and B, contains no other node. (*b*) Relative neighbourhood graph.

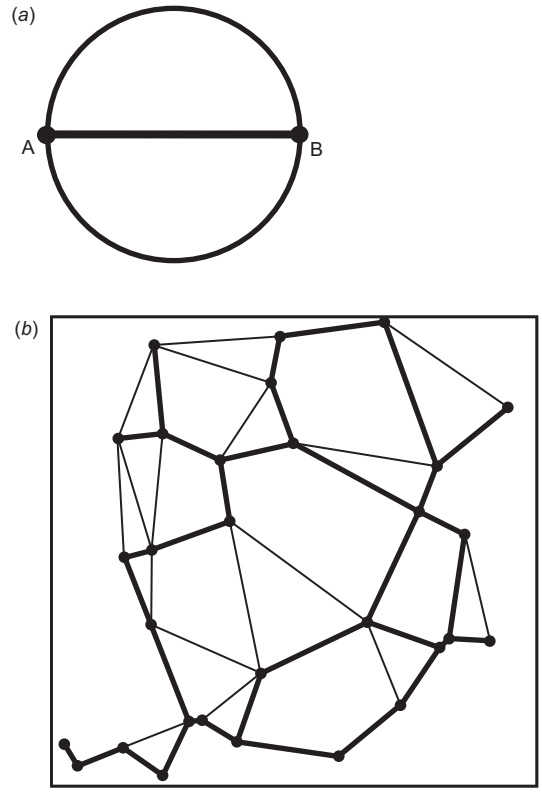


Figure 3.5 Network of a Gabriel graph for the same pattern as shown in Figure 3.2. (*a*) The Gabriel graph is formed by joining all pairs of nodes, A and B, for which the circle on diameter AB is empty radii of the two circles AB, centred on A and B, contains no other node. (*b*) Gabriel graph. The bold lines are those edges of the Relative Neighbourhood Graph.

of nodes, A and B, for which the circle on diameter AB is empty (Figure 3.5*a*). The relative neighbourhood graph is a subgraph of the Gabriel graph (Figure 3.5*b*). Matula & Sokal (1980) showed that the average number of neighbours in the Gabriel network approaches four, under CSR. The Gabriel graph is part of another hierarchy of graphs, the 'β-skeleton' series, with a broad range of neighbour numbers, as will be described below (see Kirkpatrick & Radke 1985).

The most complex neighbour network in this hierarchy is the Delaunay triangulation (DT; Okabe *et al.* 2000). It is formed by joining, with triangles of edges, all triplets of nodes A, B, and C, for which the circumcircle of the triangle ABC contains no other node (Figure 3.6*a*). The Gabriel graph is a subgraph of this triangulation (Figure 3.6*b*). It can be shown that the average number of neighbours in the Delaunay network approaches six, no matter what the spatial arrangement of the events (Upton & Fingleton 1985). One disadvantage of DT graph for ecological applications is that it often has very long edges around the periphery; these are sometimes removed based on an upper distance threshold, producing a reduced Delaunay graph (see Dos Santos *et al.* 2008). The Delaunay is not the only triangulation for neighbours in the literature (there are a number of variants), but it certainly is

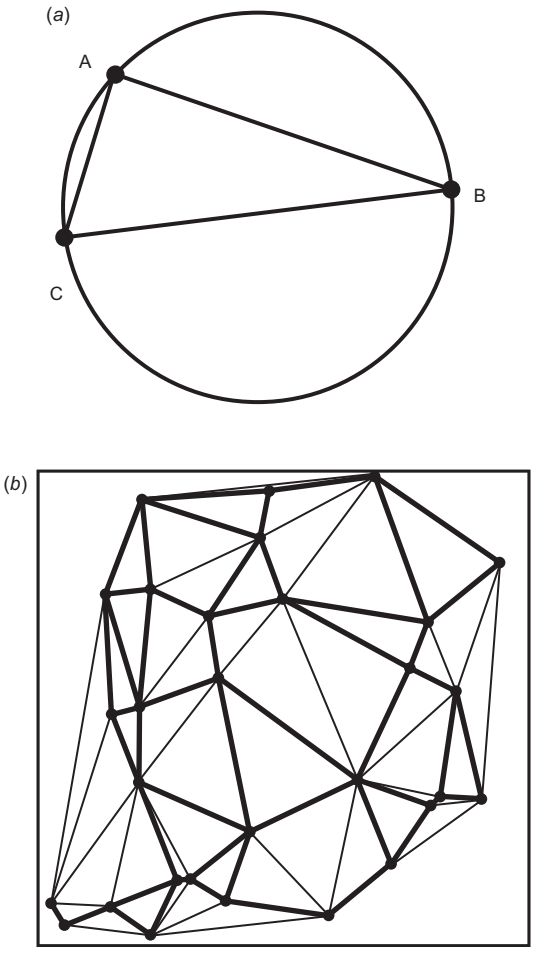


Figure 3.6 Network of a Delaunay triangulation (DT) for the same pattern as shown in Figure 3.2. (*a*) Delaunay triangle formed by joining, with triangles of edges, all triplets of nodes A, B and C, for which the circumcircle of the triangle ABC contains no other node. (*b*) Delaunay graph. The bold lines are those edges of the Gabriel graph.

the most common in spatial applications. For a network of more than six neighbours on average, the Delaunay graph can be augmented by edges to second-order neighbours in the original triangulation, restricted to those where the new edge intersects the opposite side of the triangle.

Delaunay triangulation is closely related to a familiar tessellation of polygons variously known as Dirichlet

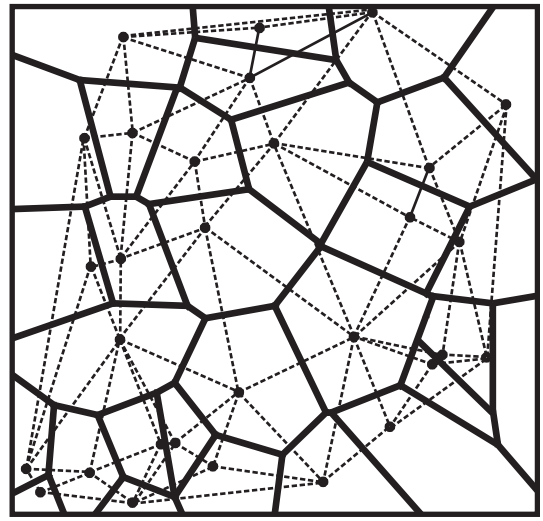


Figure 3.7 The Delaunay network (fine dashed lines) as Dirichlet domains, Thiessen polygons, or Voronoi polygons (bold lines).

domains, Thiessen polygons or Voronoi polygons (Figure 3.7). Any pair of events, the polygons of which have a common boundary in the tessellation, are joined by a line in the triangulation, indicating that they are first-order neighbours in that structure. The Voronoi polygons based on point events have Y-junctions and are usually five- or six-sided polygons. The ecological application of the tessellation comes from the fact that each polygon contains all parts of the plane closest to its own event than to any other. Where the events are plants, the polygon associated with each plant determines the resources it can pre-empt and its success may depend on the size of its polygon (Mithen *et al.* 1984). The Voronoi tessellation algorithm weights each point equally such that the determination of the Voronoi polygons is based only on the topological location of the points. In some cases, however, additional information about points (e.g. tree height, fire intensity) can be added in the delimitation of weighted Voronoi polygons to reflect the greater influence of some points relative to others (Okabe *et al.* 2000).

The hierarchy of networks given in Table 3.1 is unified by the fact that each graph of the neighbour

network is a subgraph of the next most complicated, so that in going up through the hierarchy, edges are only added, never removed or replaced. In ecological terms, this means that more and more events are considered to be neighbours, but no event that was counted as a neighbour is removed from consideration.

To illustrate how these various types of graphs can be informative of point data, the positions of plants of *Solidago canadensis* L. growing at the edge of a hay field near Edmonton, Alberta (quadrat 2; Dale & Powell 1994) were used (Figure 3.8). They are of two types: those that have obvious evidence of attack by insects (stem galls) and those that do not. We used the hierarchy of networks to define which pairs of plants are neighbours, and used a randomization procedure to determine whether there was a significantly large number of neighbour pairs that were of the same type. The results were not significant for the first member of the hierarchy (MNN), but they were for the next two (NN and MST); the results were not significant for RNG and GG, but they were for the DT.

This hierarchy of neighbour networks can be used in a number of other ways. It can be the basis for multivariate point pattern analysis and for the analysis of marked point patterns (Chapter 4), for example by counting the number of like–like joins (for multivariate classification) or correlation coefficients (for quantitative marks) in the range of networks available. In most applications, a 're-labelling' randomization will be an obvious technique to evaluate the significance of any result (Chapter 1), such as the number of like joins compared to the expected, or the value of the correlation coefficient.

The neighbour networks can also be used in many instances to transfer a method of analysis that was originally proposed for a lattice system of observations to a set of observations that are irregularly arranged in space. For example, the general approach to multivariate analysis known as clustering creates groupings of observations based on greater similarity (however measured) among objects within a group than between objects in different groups. One important variant of this technique is 'spatially constrained clustering' where objects are united into groups based on greatest similarity, but conditional on being adjacent

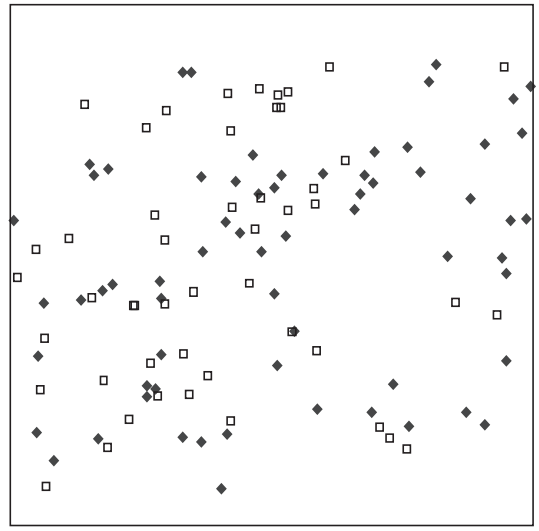


Figure 3.8 Map of *Solidago canadensis* in a 2 × 2 m study plot (quadrat 2) near Edmonton, Alberta. Closed diamonds are 'clean' plants, and open squares are those with galls or other forms of insect attack.

in space. Although that spatial adjacency is usually defined by a lattice or grid, the same approach can also be used when adjacency is defined by one of the neighbour networks described in this section. Spatially constrained clustering has a number of important applications in ecological research and is obviously closely related to techniques for determining spatial boundaries between samples or observations that are different, for example in the delimitation of landscape habitat patches (see Chapter 9 of this book).

This description of the hierarchy is not intended to provide advice on which is the 'best' network for ecological analysis. The suggestion is to use the entire hierarchy of six networks because the differences among them can provide valuable insights. Even more useful to ecologists would be comparison of the results of this hierarchical analysis with the results from artificial data that are designed as realizations of the hypothesized underlying ecological processes, such as dispersal and mortality. Other applications of this hierarchy, as well as extensions or modifications, will be found useful by ecologists for the spatial analysis of various kinds of point patterns (Chapter 4).

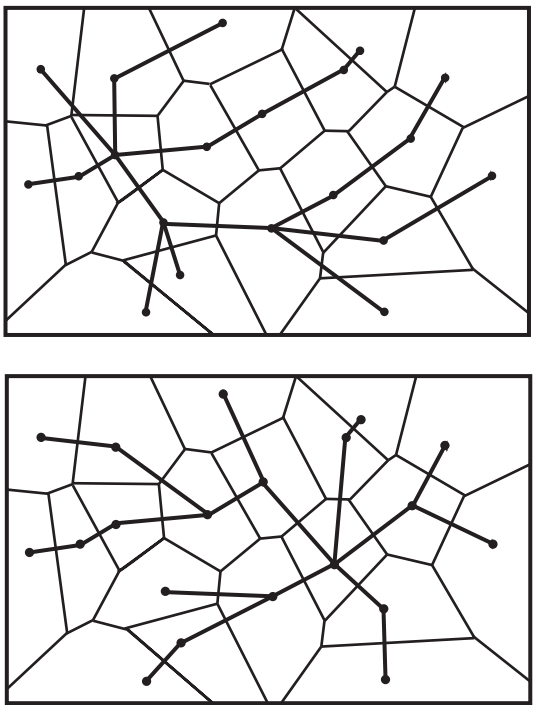


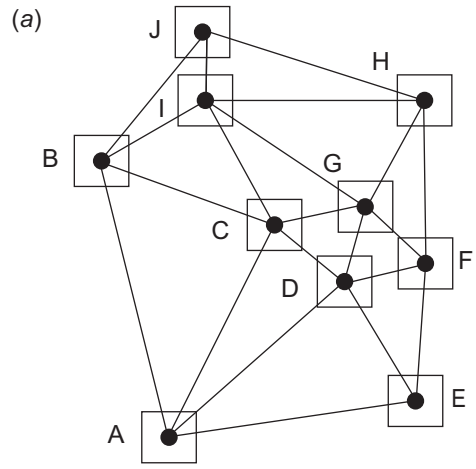
Figure 3.9 Examples of two Ulam trees starting with different nodes.

In addition to those in the hierarchy described here, there are many other neighbour networks possible, including the Ulam tree and the set of networks defined by the 'β-skeleton' rule. The Ulam tree (Figure 3.9; Ulam 1962) is created using the Voronoi tessellation as the basic structure. Then starting at a location of interest on the tessellation; then neighbouring nodes are connected iteratively where nodes become linked to the tree by a new edge only if no node exists between two nodes with an edge of the previous iteration. Another way of explaining this tree is to say that it is a subgraph of the Delaunay triangulation that is based on a focal event (and Voronoi cell), and includes all its first-order neighbours, and then all second-order, third order, and so on, provided their connection creates no cycles in the resulting graph (see Figure 3.9). Order takes precedence over edge length, but edge length is the deciding factor if the orders are the same. This graph, the Ulam tree, is a **rooted tree**, with the initial location as the root node, and different roots give different connections and different trees. This graph has a web-like structure, and may have future applications for spatial analysis in ecological research.

The β-skeleton rule creates its own hierarchy of neighbour graphs, which includes the Gabriel graph of the hierarchy described previously in this section. We said that the Gabriel graph can be defined by the fact that nodes A and B are joined with an edge whenever the circle with diameter AB contains no other nodes. A β-skeleton graph is defined by joining nodes A and B when the intersection or union of two circles, with a common chord AB and equal and opposite tangent angles determined by parameter β, are empty of other nodes (see Kirkpatrick & Radke 1985). The case of $\beta = 1$, is the Gabriel graph and the two circles coincide with AB as the diameter. When $\beta < 1$, the area of intersection is the defining region and it is smaller than the circles (forming a lens) and so there can be more edges in this graph than in the Gabriel graph; when $\beta > 1$, the area of the circles' union is the defining region, and it is larger than the individual circles (the union of two circles of equal size that overlap partially) and so there are fewer edges. Many other rules, and even other kinds of rules, for neighbour networks remain to be explored in detail, but the general principles for application will be similar to those discussed for the first hierarchy defined in this section.

3.1.2 Distance-based spatial neighbours

In the previous section, we introduced the creation of spatial graphs by joining points based on different topological algorithms. There are several circumstances, however, where the actual Euclidean distances between the points or different spatial statistics are employed (Chapter 6), and distance-based algorithms are required to determine neighbours for point events. Indeed, in order to perform spatial analysis with sampled data (Chapter 6), spatial statistics measure the degree of spatial structure based on pairs of sampling units. As the distance-based algorithms are the same for point locations or for sampling units, we will illustrate their use with sampling units where the centroids are treated as point locations.



(b) Connectivity matrix

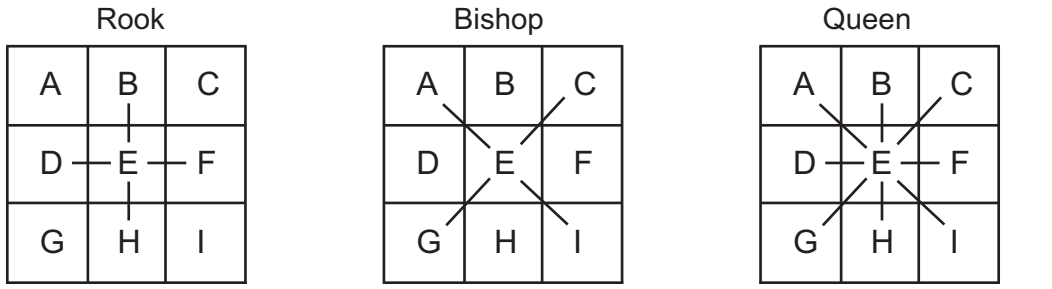
	A	B	C	D	E	F	G	H	I	J
A	-									
B	1	-								
C	1	1	-							
D	1	0	1	-						
E	1	0	0	1	-					
F	0	0	0	1	1	-				
G	0	0	1	1	0	1	-			
H	0	0	0	0	0	1	1	-		
I	0	1	1	0	0	0	1	1	-	
J	0	1	0	0	0	0	0	1	1	-

(c) Distance class connectivity matrix

	A	B	C	D	E	F	G	H	I	J
A	-									
B	1	-								
C	1	1	-							
D	1	2	1	-						
E	1	2	2	1	-					
F	2	3	2	1	1	-				
G	2	2	1	1	2	1	-			
H	3	2	2	2	2	1	1	-		
I	2	1	1	2	3	2	1	1	-	
J	2	1	2	3	3	2	2	1	1	-

Figure 3.10 (*a*) Connectivity based on the topology among 12 sampling locations (A–J). Here, the *x*–*y* coordinates of the centroid of each sampling unit are used to determine a tessellation network linking all the sampling locations. From this series of links, a binary connectivity matrix, (*b*) binary connectivity matrix of first nearest neighbours (1, connected; 0, not connected), can be defined as well as a distance class connectivity matrix (*c*) of *k* nearest neighbours (1, first nearest neighbour; 2, second nearest neighbour; and so on).

The determination of spatial neighbourhoods combines two different representations of space: topological and Cartesian. In topological space, sampling units can be reduced to simple points in the plane, using the *x*–*y* coordinates of their centroids (Figure 3.10), and then any of the neighbour network algorithms presented in Section 3.1.1 can be used to establish the **connectivity** matrix, which lists the links between sampling units. For example, the first-order neighbours of sampling unit A are B, C, D and E, which



	A	B	C	D	E	F	G	H	I
A	-								
B	0	-							
C	0	0	-						
D	0	0	0	-					
E	0	1	0	1	-				
F	0	0	0	0	1	-			
G	0	0	0	0	0	0	-		
H	0	0	0	0	1	0	0	-	
I	0	0	0	0	0	0	0	0	-

	A	B	C	D	E	F	G	H	I
A	-								
B	0	-							
C	0	0	-						
D	0	0	0	-					
E	1	0	1	0	-				
F	0	0	0	0	0	-			
G	0	0	0	0	1	0	-		
H	0	0	0	0	0	0	0	-	
I	0	0	0	0	1	0	0	0	-

	A	B	C	D	E	F	G	H	I
A	-								
B	0	-							
C	0	0	-						
D	0	0	0	-					
E	1	1	1	1	-				
F	0	0	0	0	1	-			
G	0	0	0	0	1	0	-		
H	0	0	0	0	1	0	0	-	
I	0	0	0	0	1	0	0	0	-

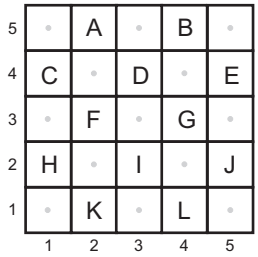
Figure 3.11 Connectivity based on chess moves among nine sampling locations on a lattice (A–I). The targeted sampling location E has four neighbours while using either the rook's or bishop's move and eight neighbours with the queen's move, as shown in the plots and translated in the binary connectivity matrices of first nearest neighbours (1, connected; 0, not connected).

is indicated by code 1 in the connectivity matrix (Figure 3.10). Such a connectivity (or adjacency) matrix indicates which pairs of sampling units have first-order connections by an entry of 1, and those that do not by 0. The connectivity matrix can be extended to include higher order neighbours so that spatial connectivity among the sampling units can be divided into d neighbour orders, giving a 'distance class connectivity matrix' (Figure 3.10*c*), where $d = 1, 2, \ldots, m$, and where m is the smallest number of neighbour steps (the shortest graph theoretical path) needed to link the two most distant sampling units. The distance-based spatial neighbours are usually determined in all directions and can be referred to as isotropic.

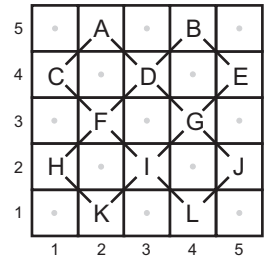
There may be cases in which the sampling units are contiguous squares, or form a lattice with regular spacing. In those circumstances, the tessellation links between sampling units can be described as chess piece moves (Figure 3.11): the rook (castle) makes links in four cardinal directions, the bishop makes only diagonal links, and the queen makes links in all eight directions. By using topological space to establish the spatial relationship among sampling units, we are assuming implicitly that Euclidean distances among sampling units are less important, and that we are interested only in their relative positions. When the absolute distances among sampling units or their positions in some frame of reference (e.g. the study area limits) are important, then Euclidean space should be used.

Euclidean distances among sampling units can be computed using the x–y coordinates of the centroid of the sampling units, resulting in a matrix of straight-line distances (Figure 3.12*a*). Based on these Euclidean distances, a distance class connectivity matrix can be created for a number of distance classes (Figure 3.12*a*).

(*a*)



(*b*)



Euclidean distance matrix

	A	B	C	D	E	F	G	H	I	J	K	L
A	0											
B	2.00	0										
C	1.41	3.16	0									
D	1.41	1.41	2.00	0								
E	3.16	1.41	4.00	2.00	0							
F	2.00	2.83	1.41	1.41	3.16	0						
G	2.83	2.00	3.16	1.41	1.41	2.00	0					
H	3.16	4.24	2.00	2.83	4.47	1.41	3.16	0				
I	3.16	3.16	2.83	2.00	2.83	1.41	1.41	2.00	0			
J	4.24	3.16	4.47	2.83	2.00	3.16	1.41	4.00	2.00	0		
K	4.00	4.47	3.16	3.16	4.24	2.00	2.83	1.41	1.41	3.16	0	
L	4.47	4.00	4.24	3.16	3.16	2.83	2.00	3.16	1.41	1.41	2.00	0

Connectivity matrix

	A	B	C	D	E	F	G	H	I	J	K	L
A	0											
B	0	0										
C	1	0	0									
D	1	1	0	0								
E	0	1	0	0	0							
F	0	0	1	1	0	0						
G	0	0	0	1	1	0	0					
H	0	0	0	0	0	1	0	0				
I	0	0	0	0	0	1	1	0	0			
J	0	0	0	0	0	0	1	0	0	0		
K	0	0	0	0	0	0	0	1	1	0	0	
L	0	0	0	0	0	0	0	0	1	1	0	0

Distance class matrix

	A	B	C	D	E	F	G	H	I	J	K	L
A	0											
B	2	0										
C	1	3	0									
D	1	1	2	0								
E	3	1	4	2	0							
F	2	2	1	1	3	0						
G	2	2	3	1	1	2	0					
H	3	4	2	2	4	1	3	0				
I	3	3	2	2	2	1	1	2	0			
J	4	3	4	2	2	3	1	4	2	0		
K	4	4	3	3	4	2	2	1	1	3	0	
L	3	3	4	3	3	2	2	3	1	1	2	0

Weighted matrix

	A	B	C	D	E	F	G	H	I	J	K	L
A	0											
B	0	0										
C	0.7	0	0									
D	0.7	0.7	0	0								
E	0	0.7	0	0	0							
F	0	0	0.7	0.7	0	0						
G	0	0	0	0.7	0.7	0	0					
H	0	0	0	0	0	0.7	0	0				
I	0	0	0	0	0	0.7	0.7	0	0			
J	0	0	0	0	0	0	0.7	0	0	0		
K	0	0	0	0	0	0	0	0.7	0.7	0	0	
L	0	0	0	0	0	0	0	0	0.7	0.7	0	0

Figure 3.12 Connectivity based on Euclidean distance among 12 sampling locations (A–L). (*a*) The Euclidean distance matrix among the 12 sampling locations and the corresponding distance class matrix when using a distance interval class of 1.5 units. (*b*) The binary connectivity matrix based on first nearest neighbours and the corresponding weight matrix as a function of distance ($1/d = 1/1.41$ units, which gives a weight of 0.7).

When other kinds of a priori knowledge of the point or surface pattern process are available, one may not just want to use a connectivity matrix (Figure 3.12*b*) that indicates in a simple binary fashion which sampling units are neighbours (1, neighbours; 0, otherwise), but a more sophisticated weighted matrix where real values (usually between 0 and 1) stress that nearby neighbours are more important than ones located further away in the estimation of spatial structure (Figure 3.12*b*). The most common weight used is the inverse distance ($1/d$) or the inverse of distance squared ($1/d^2$).

There are some rules of thumb for choosing the number of distance classes, based either on sample size or spatial lag. For example, Legendre & Legendre (2012) suggested using Sturge's rule, based on sample size, to determine the number of histogram classes: for n samples the recommended number of classes is:

$$D = 1 + 3.3 \times \log_{10}\left[\frac{n(n-1)}{2}\right]. \quad (3.1)$$

Once the number of distance classes is selected, there are two ways of distributing the Euclidean distances among them. The first view is that they should be divided by equal distance intervals (also referred to as spatial lag or the spatial interval, Figure 3.13*a*) and the second is that distance classes should have equal numbers of pairs of sampling units in them (as proposed by Sokal & Wartenberg 1983; Figure 3.13*b*). The equidistant approach is the most frequently used and is more intuitive, allowing comparison of results among different studies at a given distance interval. However, the number of pairs per distance class varies, which in turn affects the reliability of the estimation of spatial structure (see Chapter 6). The equal frequency approach provides a better estimation of spatial autocorrelation at each distance class because each distance class must contain the same number of pairs, thus minimizing the edge-effect problem presented in Chapter 1. This implies, however, that the values estimated to describe the spatial structure are not at regular distance intervals. The equal frequency distance class approach is less common, either because it is not widely known or because the results are not as easy to compare from one study area to another when the distance intervals differ.

When using equidistant classes, one can select the distance interval instead of the number of distance classes. The minimum distance interval must be equal to or larger than the sampling unit length (or the longest side length in the case of a rectangular sampling unit). When the sampling units form a regular lattice, this minimum distance interval (i.e. the length of a sample unit, say 1 m) will make links equivalent to the rook's move, so that each sample unit has only four neighbours (Figure 3.13). In order to include all eight neighbours (equivalent to the queen's move), the minimum distance interval needs to include the diagonal, that is $\sqrt{2} = 1.414$ (Figure 3.13). There are some trade-offs to consider in selecting a distance interval (Fortin 1999a): too small a distance interval will result in a very large number of distance classes and there will be fewer pairs of observations in each distance class, making the estimation of spatial structure less reliable. However, too large a distance interval will lead to too few distance classes where each distance class will cover more area and the magnitude of spatial structure will be diminished.

3.1.3 Directional angle-based spatial neighbours

Spatial arrangement of geomorphological features (topography, hydrology) and directionality of processes (wind, stream flow) influences the spatial pattern of ecological data creating elongated (anisotropic) patches rather than circular (isotropic) ones. To analyse such pattern, the spatial neighbouring search cannot be only by distance as it is an isotropic search approach. Hence it is necessary to add another criterion which includes the angle between the sampling

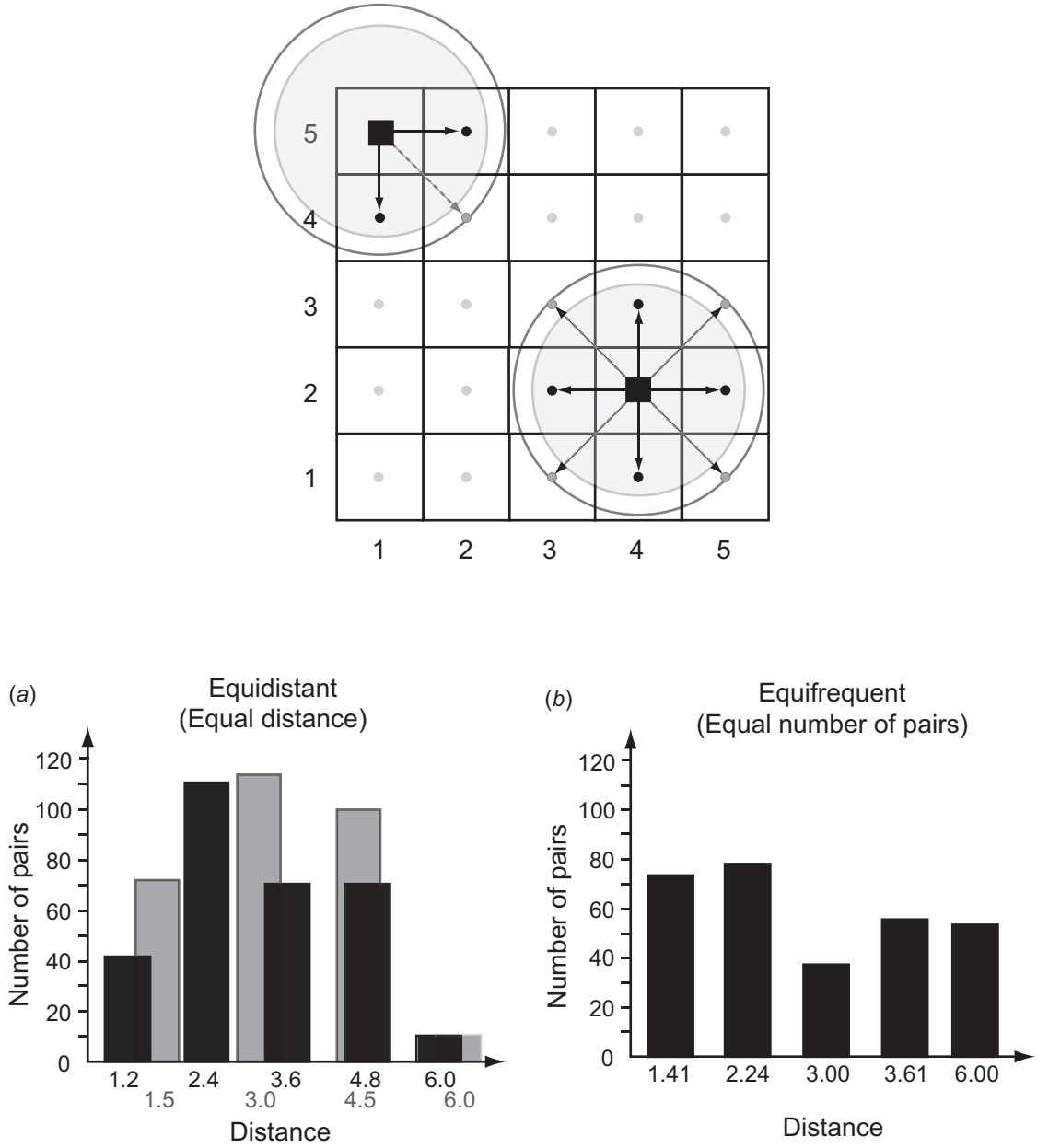


Figure 3.13 Distance class determination based on distance (*a*) or number of pairs (*b*) of sampling locations. (*a*) When defining the distance interval class at 1.2 units (in solid black), there are five equidistant classes (where the diagonal locations are not included in the first distance class – 1.41 units apart); while when the distance interval class is 1.5 units (in solid grey) there are only four equidistant classes (including the diagonal locations in the first distance class – 1.41 units apart) such that there are more pairs of sampling locations in the first distance class. (*b*) With the equifrequent distance class approach, each distance class has more or less 50 pairs of sampling locations, but the Euclidean distance from one distance class to another varies in interval (i.e. 1.41, 0.83, 0.76, 0.61 and 2.39).

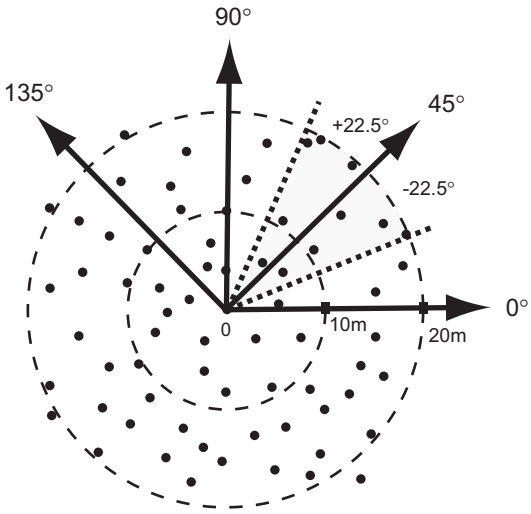


Figure 3.14 Angle-based neighbour determination including both Euclidean distance classes and angle classes for directional neighbour evaluation.

unit locations (Figure 3.14). The spatial neighbours are then determined as follows: (1) use the distance-based search; and (2) include a main search direction, an angle (usually the main cardinal directions: 0°, 45°, 90°, 135°) with a tolerance angle zone on either side of the main direction (usually 22.5°) (Figure 3.14). Because these combined criteria results in fewer pairs of points than just the distance-based approach, it is recommended that larger distance classes be used such that more points can be considered. These implications of this directional-based approach are elaborated in Chapter 6.

3.2 Lines: fibre pattern analysis

In addition to reducing spatial objects to dimensionless points as helpful representations of locations, spatial objects can also be treated as lines or as polygons, whichever is more appropriate to their characteristics. Many different kinds of linear events are of interest in ecology, both objects that are one-dimensional and abstractions that have similar properties. The list includes plant stems or roots, rivers, hedge-rows, and roads as physical objects; and migration routes or dispersal trajectories,

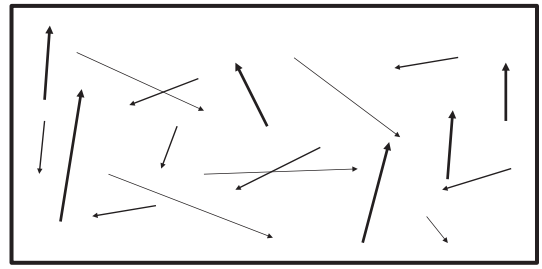


Figure 3.15 Fibres with length, direction and width. The three characteristics seem not to be independent.

and information exchange routes as derived or constructed lines. In these cases, these linear structures are the primary entities of interest, unlike the situation in a graph, where the edges require nodes to join in order to have existence at all. Any such linear events can be analysed using fibre pattern analysis (Stoyan *et al.* 1995), which is the analysis of a set of one-dimensional events resulting from a 'fibre process', the equivalent of a 'point process' previously described.

As a starting point, we can consider fibres that are straight lines of equal length (Figure 3.15), for which we can determine scales of clustering, randomness and repulsion. These fibres have angles of orientation and so we can also answer questions about relative angles as a function of distance between fibres or about angles as a function of location. More generally, we can consider fibres that are not identical, and we can test hypotheses about length, thickness and crossings, using restricted randomization for each characteristic (position, angle, length, width) and then for combinations of two and three characteristics. Similarly, if the fibres belong to several classes, such as species, we may be interested in the equivalent of join count statistics and proceed by evaluating the number of crossings of fibres that belong to each of the possible categories or to summary classes of like–like versus unlike pairs of fibres crossing. Again, randomization techniques seem to be the most likely methods to evaluate these count statistics.

We have alluded to the extensive body of literature on the analysis of dimensionless events located in two or three dimensions, 'point pattern analysis' (e.g. Diggle 1983, 2003) where the events are tree

stems in a forest, retail outlets, or stars in galaxies. Now, we extend that approach to analysing a fibre process or fibre pattern, which is a system of finite linear (possibly curved, but smooth and differentiable) structures located in the plane or in three dimensions (Stoyan *et al.* 1995; Benes & Rataj 2004). A large number of ecologically interesting phenomena can be treated as fibre processes: roots in the soil, the rhizomes or stolons of clonal plants, invertebrate burrows in sediment, animal migration routes, and so on. Because of the potentially very complex and computationally intensive analyses that we might consider for such systems of linear features, we will begin by exploring analyses of simple cases: two dimensions rather than three, and finite straight line segments. There are a variety of different characteristics that might be of ecological interest in such systems: the aggregation or overdispersion of the fibres, anisotropy and the tendency for the fibres to be parallel, spatial autocorrelation of segment lengths, and so on.

3.2.1 Aggregation and overdispersion of fibres

We can base the analysis of fibres on methods that already exist for the analysis of point patterns, a set of dimensionless events arranged in the plane. One of the most commonly used methods for point pattern analysis is Ripley's K (Ripley 1976; see also Chapter 4). This method is based on the concept that, if λ is the density of events per unit area, the expected number of events in a circle radius t centred on a randomly chosen event is $\lambda K(t)$, where $K(t)$ is a function that depends on the pattern of the events. For example, if the events are overdispersed, $K(t)$ will be close to 0 for short distances and increasing for larger distances.

For a given distance, t, the statistic is based on counting all pairs of events separated by distances less than t. The statistic $\hat{K}(t)$ is an estimate of $K(t)$:

$$\hat{K}(t) = A\sum_{i=1}^{n}\sum_{j\neq i}^{n} w_{ij} I_t(i,j)/n^2. \qquad (3.2a)$$

Here, A is the area of the study plot, d_{ij} is taken as the distance between points i and j, $I_t(i,j)$ is an indicator function, taking the value 1 if $d_{ij} \leq t$ and 0 otherwise, and w_{ij} is a weight to correct for edge effects. If the circle centred on i with radius d_{ij} is completely within the study plot $w_{ij} = 1$, otherwise it is the reciprocal of the proportion of that circle's circumference that is in the plot (Diggle 1983). The weight w_{ij} can be replaced with a weight that depends on i and t, rather than on i and j: weight $h_i(t) = 1$ if the circle centred on i with radius t is completely within the study plot, otherwise the reciprocal of the proportion of that circle's area within the plot:

$$\hat{K}(t) = A\sum_{i=1}^{n}\sum_{j\neq i}^{n} h_i(t) I_t(i,j)/n^2. \qquad (3.2b)$$

This approach reduces the number of calculations that need to be made: one weight for each event and each radius, rather than one for each pair of points and each radius. If the events are located completely at random, the expected number of events in a circle radius t is $n\pi t^2/A$ and the function $K(t)$ is just πt^2. The observed statistic $\hat{K}(t)$ is compared with this expected value by subtracting the observed from the expected (or vice versa in some examples, which can cause confusion) after a square root transformation:

$$\hat{L}(t) = t - \sqrt{\hat{K}(t)/\pi}. \qquad (3.3)$$

$\hat{L}(t)$ is then plotted as a function of t, with negative values indicating clumping and positive values indicating overdispersion. (Remember that some analyses use the expected minus the observed so that the signs are reversed.)

In order to modify this approach to deal with linear features rather than dimensionless events, a number of adjustments must be made. To begin with, because lines are not dimensionless, it is more appropriate to base any statistic on the lengths of lines found within a particular area, rather than on counts of lines. Because the point pattern analysis used circles as the templates for event counts, the most direct modification for linear features would be to base any statistic on the template of an envelope of all points of the plane at distance less than or equal to a given distance, t. We will begin by considering only patterns of finite straight line segments, for which this concept produces an envelope

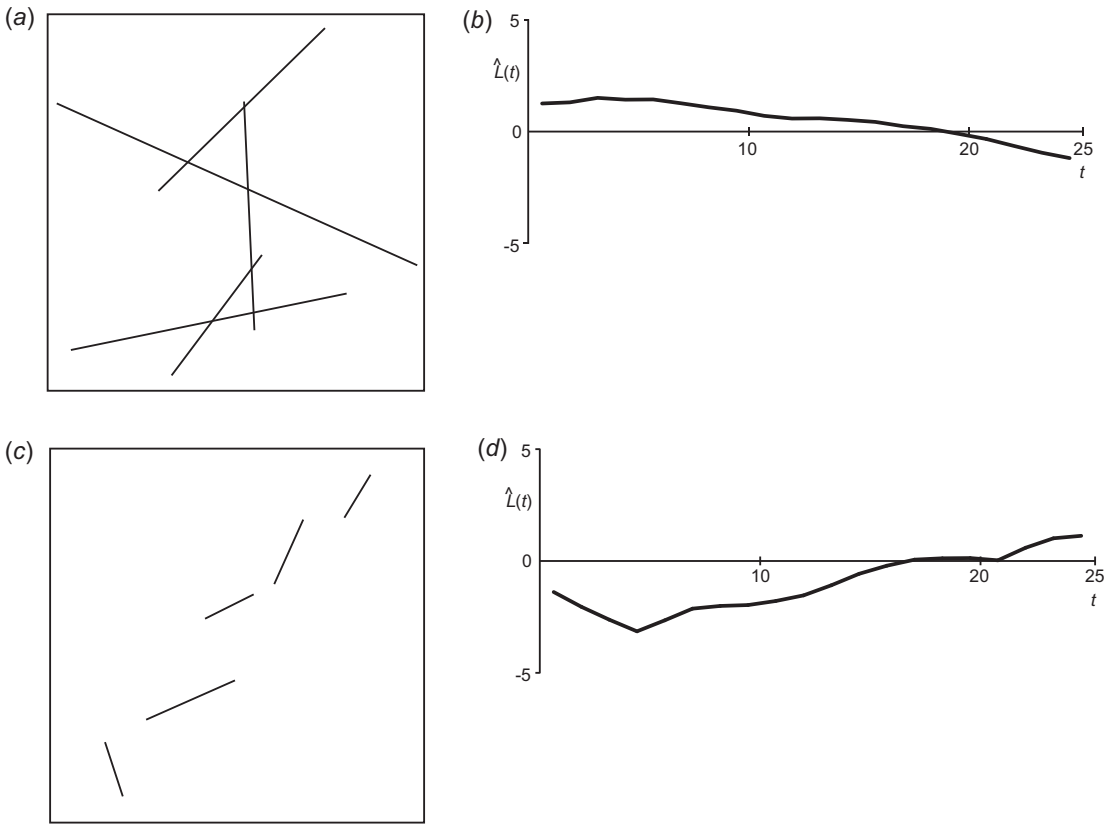


Figure 3.16 Examples of fibres [(*a*), (*c*), (*e*), (*g*)] and corresponding Ripley's *K* analysis [(*b*) (*d*), (*f*), (*h*)].

that is a rectangle with semicircular ends. The area of the envelope for line i and distance t is $a_i(t)$ and the length of line j that lies within it is $l_t(i,j)$. Where L is the total length of lines in the entire area, by analogy with Ripley's point pattern function, the expected length of line within an envelope of radius t centred on a randomly chosen line is $\lambda_f K_f(t)$, where $K_f(t)$ is a function that depends on the pattern of the fibres and λ_f is the length of line per unit area. Then $K_f(t)$ is estimated by:

$$\hat{K}_f(t) = A\sum_{i=1}^{n}\sum_{j\neq i}^{n} h_i(t)l_t(i,j)/n^2L. \tag{3.4}$$

The expected value of $K_f(t)$ under complete spatial randomness of the fibre process is just the average of the areas of the envelopes for radius t: $a_\bullet(t)$. For easy comparison, again by analogy with the point pattern case, we suggest the statistic:

$$\hat{L}_f(t) = \sqrt{a_\bullet(t)} - \sqrt{\hat{K}_f(t)}. \tag{3.5}$$

As in the analogous point pattern analysis, negative values indicate clumping and positive values indicate overdispersion. For the template delimiting area around a fibre, we are using a rectangle with semicircular ends:

$$a_i(t) = \pi t^2 + 2td_i. \tag{3.6}$$

Examples of this analysis are given in the several parts of Figure 3.16, which covers some of the range of possible outcomes.

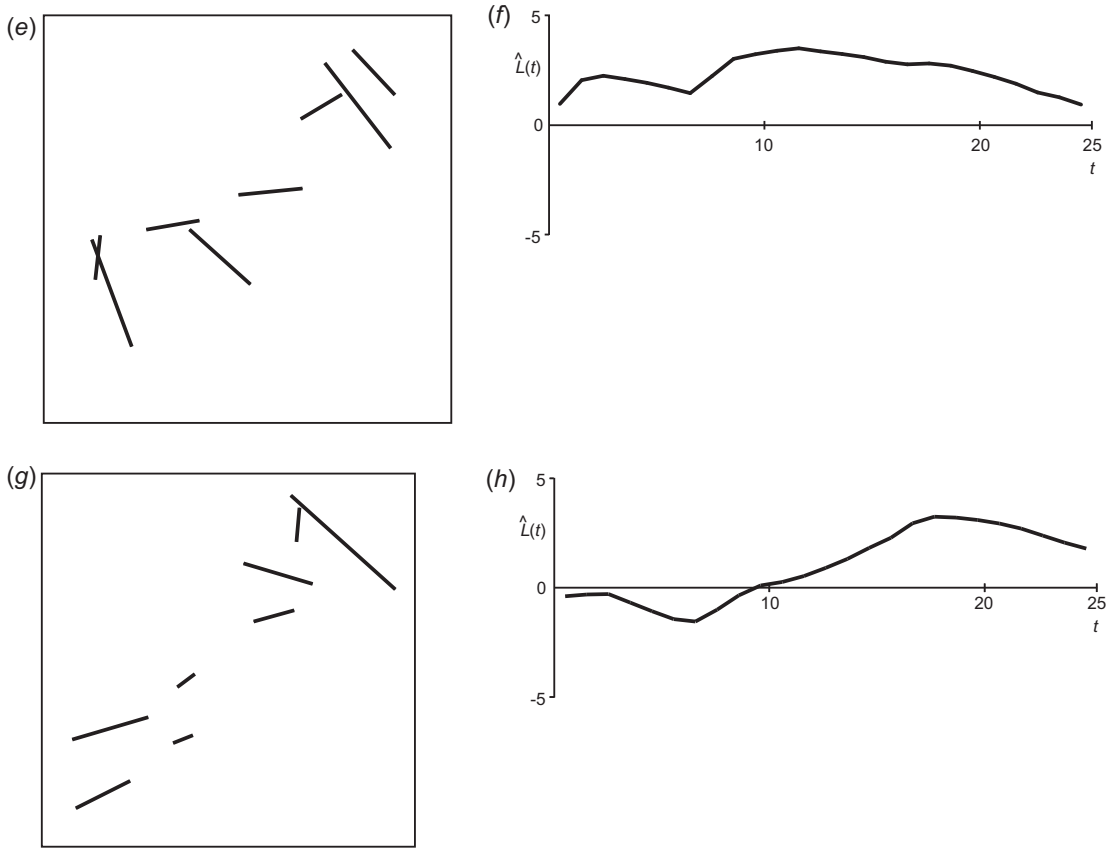


Figure 3.16 (*cont.*)

There has been some discussion in the literature on point pattern analysis about the advantages and disadvantages of a 'decumulated' form of Ripley's *K*, where the evaluation of the neighbourhoods of point events are based on the contents of rings rather than circles, producing a statistic usually called Ω (Condit *et al.* 2000; see Section 4.1.5 of this book). The advantage is attributed to looking at a narrow range of distances for the statistic, say from $t - d$ to t, rather than a full range of distances from 0 to some limit t. While such a modification is possible for fibre pattern analysis statistics, one factor to be considered is that while there is only one value for the distance between any pair of point events, there is a range of values for the distance between any pair of fibre events, and so the effects of changing to a decumulated version of our fibre analysis *K* function are not clear.

Comments that are somewhat similar can be made with reference to another modification of Ripley's *K* analysis for point events which is to produce a map of localized scores, to evaluate place-to-place variation in deviations from the expected rather than just the global overall evaluation (Getis & Franklin 1987). Because a fibre event does not have a single location in the way that a point event has only one location, the scores for individual fibres would have to be spread out along its length, producing either a lack of clarity or genuine ambiguity about what such local scores might represent. This is a topic for further investigation and consideration that may well be worth pursuing.

3.2.2 Fibres with properties

A fibre process may give rise to a fibre pattern that includes a qualitative (e.g. living root versus dead) or quantitative (e.g. thickness or age) property for each fibre, and these characteristics can be included in the analysis by modifications of the method just described. This is done in the same way as the simple univariate Ripley's K can be modified to deal with bivariate or multivariate data (Chapter 4) or to include measured values as a mark correlation analysis (Chapter 4).

For categorical variables, we can explain the bivariate case in some detail, with the multivariate case to be created by extension. For two kinds of fibres, in order find the length of Type 2 fibres in Type1 t-neighbourhoods and the length of Type 1 fibres in Type 2 t-neighbourhoods, we calculate:

$$\hat{K}_f(t)_{12} = A\sum_{i}^{n_1}\sum_{j}^{n_2} h_i(t)l_t(i,j)/n_1 n_2 L_2. \tag{3.7}$$

$$\hat{K}_f(t)_{21} = A\sum_{i}^{n_1}\sum_{j}^{n_2} h_i(t)l_t(j,i)/n_1 n_2 L_1. \tag{3.8}$$

Then

$$\hat{L}_f(t)_{12} = \sqrt{a\bullet(t)} - \sqrt{[n_2\hat{K}_f(t)_{12} + n_1\hat{K}_f(t)_{21}]/(n_1+n_2)}. \tag{3.9}$$

Negative values indicate aggregation and positive values indicate segregation. For multivariate data, we can consider a variety of 'interspecific' statistics of the form $K_f(t)_{I,\sim I}$ and so on which can be evaluated on their own or compared with 'intraspecific' measures of the form $K_f(t)_{I,I}$ as described in Chapter 4 for point pattern analysis.

If the fibres belong to C categories, we can evaluate the number of crossings of fibres that belong to each of the $C \times C - C$ categories created by the vertical order of the crossings; consider a log of Species A crossing on top of a log of Species B, different from B crossing on top of A. Where the order cannot be recovered, depending on the system being studied, there are only half as many categories for the fibre crossings, and in some cases, it may be desirable to use summary classes of like-with-like versus unlike pairs of fibres crossings. In any of these cases, we suggest that (restricted) randomization techniques seem to be the best approach to evaluating these count statistics.

For circumstances in which the fibres have quantitative attributes, we can calculate

$$\hat{K}_f(t)_m = A\sum_{i=1}^{n}\sum_{j\neq i}^{n} h_i(t)l_t(i,j)m_i m_j/n^2 L. \tag{3.10}$$

Then, where μ is the overall mean of the quantitative attribute, plot

$$\hat{L}_f(t)_m = \sqrt{a\bullet(t)} - \sqrt{\hat{K}_f(t)_m/\mu^2}. \tag{3.11}$$

Positive values indicate overdispersion of the marks, and negative values indicate their aggregation. This approach is often referred to as *marked fibre process analysis* (Stoyan *et al.* 1995).

3.2.3 Curving fibres

In this discussion, so far, we have been concentrating on fibres that are finite straight line segments, but some applications will require curving fibres. There are a number of ways in which curving finite-length fibres can be produced, but the most practical approach is probably to use Bézier curves, sometimes referred to as 'French curves'. This technique cannot produce all possible curves, but it is very flexible, and very complex curves can be built up from Bézier subunits by a spline technique. A curve can be created of any order, but the most commonly used is the cubic form, based on four points at positions q_1 to q_4. The parametric equation of the curve is then

$$B(z) = (1-z)^3 q_1 + 3z(1-z)^2 q_2 + 3z^2(1-z)q_3 + z^3 q_4. \tag{3.12}$$

An example is shown in Figure 3.17*a*.

There are many advantages to this approach, one of the most obvious being that transformations such

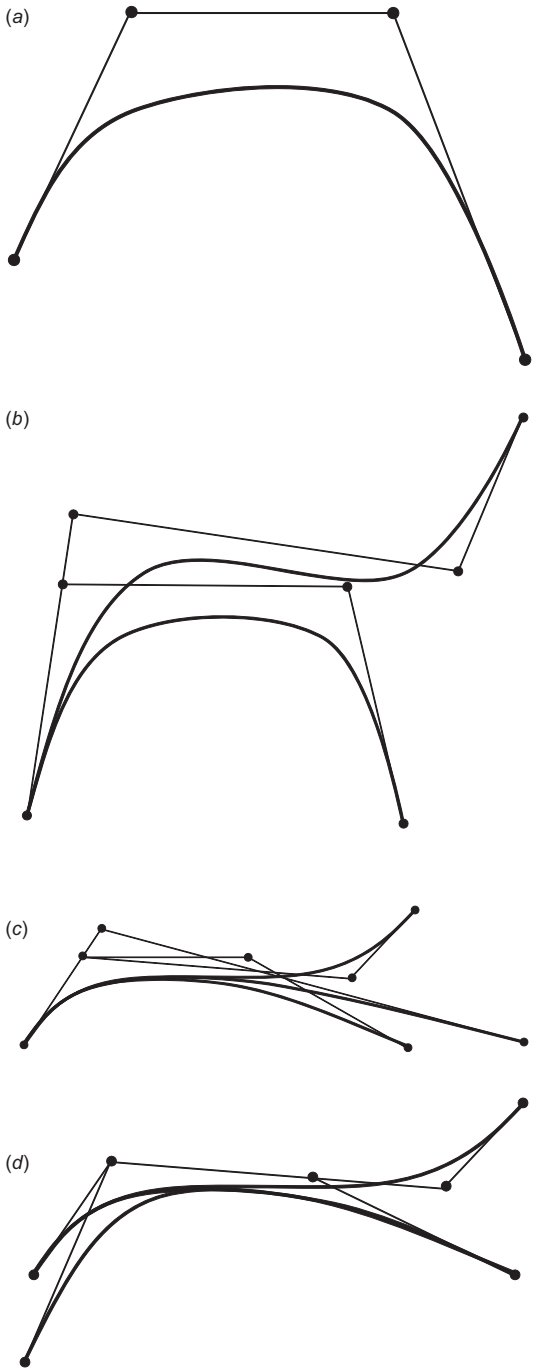


Figure 3.17 Bézier curve example. (*a*) simple (*b*) one branch (*c*) two branches (*d*) double bifurcation.

as change of scale, position, or angle can be generated easily for the curve by applying the transformation to the four 'control' points (Bézier 1977). A second advantage is that for any two such curves, one based on points at q_1 to q_4 and the other on points g_1 to g_4, the midline between the two curves is easily based on $m_i = (q_i + g_i)/2$:

$$M(z) = (1-z)^3 m_1 + 3z(1-z)^2 m_2 + 3z^2(1-z)m_3 + z^3 m_4. \quad (3.13)$$

We can then use the midline as the basis for comparing properties of two curving fibres, as will be shown below.

3.2.4 Branching curved fibres

Although Bézier curves have been used in applications ranging from automotive design to the creation of type fonts, we have found no discussion of their use to create branching curves. This is easily done for cubic curves of this family, by using one additional control point, q_5. Then the two branches of the curve, which begin identically, diverge, following the equations

$$B_1(z) = (1-z)^3 q_1 + 3z(1-z)^2 q_2 + 3z^2(1-z)q_3 + z^3 q_4 \text{ and} \quad (3.14)$$

$$B_2(z) = (1-z)^3 q_1 + 3z(1-z)^2 q_2 + 3z^2(1-z)q_3 + z^3 q_5, \quad (3.15)$$

(see Figure 3.17*b*). A second branching can be introduced using a sixth control point, q_0:

$$B_3(z) = (1-z)^3 q_0 + 3z(1-z)^2 q_2 + 3z^2(1-z)q_3 + z^3 q_4. \quad (3.16)$$

This is illustrated in Figure 3.17*c*.

Further complexities can be introduced using a second set of four control points, q_6 to q_9 and forcing q_6 to be collinear with the line q_2 – q_1 then

$$B_3(z) = (1-z)^3 q_8 + 3z(1-z)^2 q_7 + 3z^2(1-z)q_6 + z^3 q_1 \text{ and} \quad (3.17)$$

$$B_4(z) = (1-z)^3 q_9 + 3z(1-z)^2 q_7 + 3z^2(1-z)q_6 + z^3 q_1. \quad (3.18)$$

This combination produces a doubly bifurcating fibre (Figure 3.17*d*). Combining such substructures allows the creation of multibranched structures resembling trees, or anastomosing structures of curved fibres forming networks. Again, this is an area that shows promise for the spatial analysis of ecologically important objects, but it is yet to be pursued.

3.2.5 Congruence and parallelism of curved fibres

For straight line segments, the concept of parallelism is simple: if two line segments are at the same angle, they are parallel, and any departure from parallelism can be quantified by the difference in their angles. Line segments are all the same shape, and can differ only in length, but curving fibres can differ in shape as well. This fact requires us to introduce some terminology for the discussion. We will follow the vocabulary usually applied to triangles (Gellert *et al.* 1977), and will focus on Bézier curves. Two such curves are *congruent*, if they can be exactly superimposed on each other by translation and rotation. This means that they are the same size with the same shape; if they are exact mirror images, and can be superimposed by reflection followed by rotation and translation, they are *inversely congruent*. If they are the same shape but not the same size, the curves are *similar*, and can be superimposed exactly by rescaling, translation and rotation. The terms apply equally to the Bézier curves and to their polygons of control points. For two curves to be truly parallel, they must be congruent and have a constant distance between corresponding parts of the curves. Of course, our interest is not only in congruent or not, or parallel or not, but in some measure of the degree of congruence or similarity of shape, and some measure of parallelism. Clearly, these thoughts are only the beginnings of what is obviously a very rich field of endeavour for spatial analysis, but it will require a great deal more work to develop it fully, and to realize its potential in applications to ecological questions.

3.3 Points and lines together

We can consider combining point events and lines as objects in several ways.

(1) We can have the two kinds of events resulting from two separate processes (e.g. seedlings versus downed logs on the forest floor) with the null hypothesis of complete independence within and between processes. A good example of examining sets of points and lines for independence or for the form of dependence comes from outside the ecological literature: Berman (1986) looked at the point locations of copper deposits (the point events) in relation to linear geological features (lineaments) in Queensland, Australia. This is an especially good example because the orientation of the lineaments was also considered (see also Foxall & Baddeley 2002). Berman's approach was based on a simple test statistic that uses the nearest point-to-line-neighbour distances, D_i being the distance from the ith point event to the nearest point on any of the lines:

$$S_n = \sum_i^n D_i. \tag{3.19}$$

The smaller the statistic, the more closely associated the two processes will be known to be. The significance of an observed value can be evaluated by a numerical procedure, either by a Monte Carlo process or by a randomization procedure using the lengths (and potentially angles) of the observed fibres and the number of point events.

Foxall & Baddeley (2002) recommended the use of a variant of the function devised by van Lieshout & Baddeley (1999), and described in greater detail for point events alone in the following chapter (Section 4.1.5). Essentially, it compares the 'empty space function' F_Y, which is based on the average shortest distance from a random point in the plane to an element of the set of interest, Y, here the fibres, with the 'nearest neighbour function' G_{XY}, which is based on the average shortest distance from a random event in set X to an element of the set Y. Then, for a radius limit, r, the J-function is defined as:

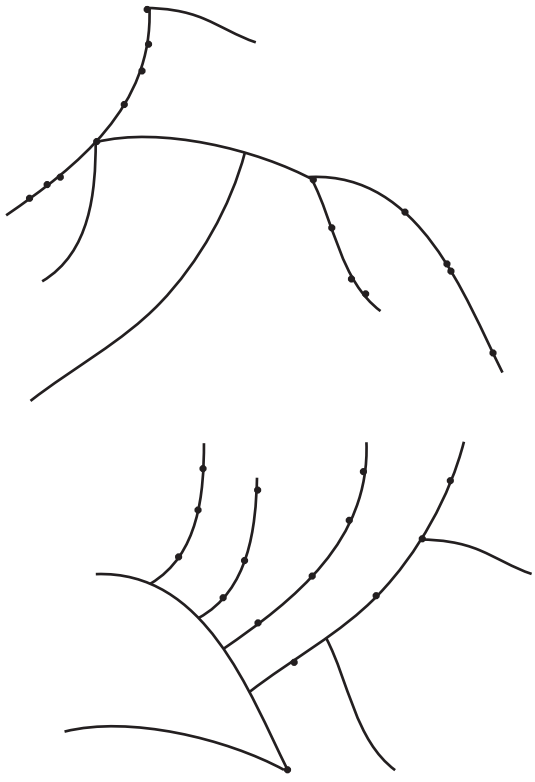


Figure 3.18 Examples of the overdispersion and clumping of point events on fibres.

$$H_{XY}(r) = \frac{1 - G_{xy}(r)}{1 - F_Y(r)}. \tag{3.20}$$

Again, the statistic is most easily evaluated by a Monte Carlo or randomization procedure.

(2) We can have the point events are constrained to be located on the fibres. This occurs in many ecological examples, such as herbaceous stems (the events) arising from rhizomes (the fibres), foraging episodes on an animal's path of movement, or waterfalls on a river system. This constraint gives rise to a hierarchy of hypotheses to be tested, for example, distinguishing between 'Are the events over- or underdispersed compared to CSR?' and 'Given the lengths and positions of the fibres, are the events over- or underdispersed?' This situation is illustrated in Figure 3.18, with the events in the upper part of the figure being clumped on some of the fibre branches, whereas they are overdispersed in the lower fibre system, although appearing to form an open clump in ordinary two-dimensional space. The objects here are constrained asymmetrically, with the point events occurring only on lines (e.g. waterfalls on rivers) which are 'the given' or put in place first; the points on linear structures can be analysed using a modified version of Ripley's K (Okabe & Yamada 2001; Okabe *et al.* 2006). The procedure is to count the number of events within distance t of any event in the network of fibres, where the distance is measured along connected fibres. The average observed count $K(t)$ is compared with the expected, again by Monte Carlo or randomization techniques; for details see Okabe & Yamada (2001), Okabe & Satoh (2009), and their software SANET. Spooner *et al.* (2004) provided an example of ecological application of this technique, in both univariate and bivariate versions, to look at *Acacia* tree populations along a road network in Australia. The analysis showed significant clustering over a range of scales, with the suggestion that road maintenance activities may be more important than environmental factors in determining population dynamics.

One application where the 'point events on a line' analysis seems not to have been fully utilized is in the evaluation of data from line transect sampling, where individual animals are recorded on a sampling line as a way of determining population numbers and spatial dispersion (see, for example, Hedley & Buckland 2004). There seems to be much opportunity for learning more from these data sets by further 'point events on lines' analysis.

(3) In other ecological examples, we have the opposite situation with the lines constrained to occur only at the point events (e.g. roots from plant stems). This situation may be less frequently encountered or less interesting, especially if the linear objects are constrained to be straight line segments, in which case there is little room for variation in the structure. One example of obvious ecological

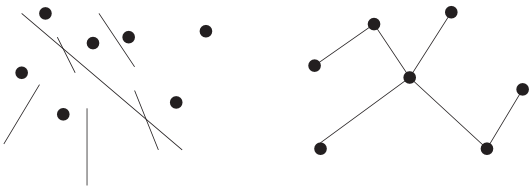


Figure 3.19 Mutual constraint of point events and lines as objects (right panel) versus separate and independent processes (left panel).

importance would be the growth of rhizomorphs as agents of fungal infection, such as *Armillaria*, from one tree stem (or stump) to another.

(4) We also must consider the case of is mutual constraint of point events and lines, possibly with the constraints being sequential (Figure 3.19): first a point event, then some lines constrained by that point, then more points constrained by the lines, and so on, as with plant shoots and their rhizomes. There are a number of obvious examples of this kind of spatial structure in ecology, and the events do not need to be physical objects; consider the linear sections of animal movement, with feeding episodes (or such) being the point events. Clearly there is mutual constraint of the elements of a point-and-line depiction of this spatial phenomenon. We will say more about the analysis of animal movement in Chapter 11 of this book.

3.4 Points and lines: spatial graphs

Moving on from points as events in space, with lines as spatial objects, we can now focus on points and lines as the nodes and edges of a graph in a spatial context, making the graph a fully spatial object (Figure 3.1). In its original form, a graph is purely structural, not spatial, but there is a continuum from the truly **aspatial** to fully and explicitly **spatial graphs**, in which the nodes have location, and possibly size, shape, etc., and the edges are links between nodes with location and extension, and possibly length, width, shape, capacity, etc. An aspatial graph is one in which the positions of the elements are not considered (i.e. arbitrary) and thus space is ignored. In a spatial graph, the nodes have physical positions as well as the edges. Hence edges have lengths determining whether or not the location of the edge is required to be explicit along its length. For example, we can draw an edge between the positions of a tree and one of its seedlings, indicating the relationship but without knowing the real route the seed travelled during dispersal. On the other hand, the edges may actually have physical existence and position, as in the mycorrhizal networks associated with forest trees, such as Douglas Fir (see Beiler *et al.* 2009). The nodes also may have characteristics or weights, such as the size and shape of habitat patches, and so may the edges; for example, the dispersal corridors between habitat patches have characteristics that may enhance or impede dispersal (see Urban *et al.* 2009). As one specific example, Dyer & Nason (2004) used a spatial graph to identify geographic clusters of cactus populations. They depicted two clusters that had significantly few connections between them as determined by a probability calculation based on graph theory. Fortuna *et al.* (2008) studied insect pollination in *Prunus mahaleb* using the connectance and modularity of a bivariate spatial graph in which the nodes were pollen donors or mother trees. They found that most mother trees had comparable numbers of donors except that the modularity analysis showed a significant population substructure of well-defined groups of mother trees as well for their shared pollen donors (Fortuna *et al.* 2008).

In Section 3.1.1, we described a hierarchy of spatial graphs based on the definition of which nodes were to be considered neighbours. A hierarchy of spatial graphs can also be created based on a series of threshold distances, each distance giving a graph that includes an edge for each pair of nodes separated by less than the given threshold. Larger distance thresholds allow more edges, but must include all the edges determined by any smaller threshold. Any evaluation of autocorrelation or similar structure measure that is based on distance between samples implicitly uses a series of distance-defined spatial graphs. Brooks (2006) provided an example of this approach in studying the scale of plant–pathogen association in *Silene latifolia* using a spatial graph with nodes being classified as

diseased and healthy plants. The scale of **connectivity** was determined for a range of distances, using hierarchical analysis. The mean node number and **diameter** of the largest connected subgraphs revealed that while most patches were weakly connected, there were a few patches with high connectivity. This indicates that the disease was disproportionately prevalent where the plants were in close proximity with more neighbours.

In some cases, a more restricted range of distances may be of interest in analysis than the set of all distances from zero to some maximum, D. In that situation, the range can be defined by two thresholds, a lower limit and an upper limit, $D_{\min}$ and $D_{\max}$. A two-threshold distance graph can then be created by including all pairs of nodes that fall within the specified distance range. This may allow the evaluation of more narrowly defined questions about the spatial structure than the use of the single upper threshold approach. For example, in the landscape context, we might be interested in the availability of habitat patches that are nearby, but not right next to, focal patches of habitat; the patches would be the nodes of a spatial graph, and the edges of interest would be defined by a lower and an upper threshold for inter-patch distances (Fall *et al.* 2007). In an evolutionary context, we might be interested in subgraphs in a lineage that include 'close' relatives that are not full siblings. In any case, if the full range of edge lengths defined by the nodes is divided into mutually exclusive distance classes, and two-threshold graphs are accordingly created, the overall effect is to take the nodes' complete graph and partition it into a set of subgraphs, each with the same set of nodes but with only those edges that fall into their particular distance class.

3.4.1 Signed and directed graphs and networks

The edges of a graph that indicate relationships can have associated signs for positive or negative (+ or −, a binary weight) depending on the nature of the relationship, giving a *signed* graph (the signs are often indicated by a solid for positive and dashed for negative, recall Figure 3.1*c*). Graph edges can have numerical weights, with or without signs, which can provide measures of any kind (e.g. similarity or dissimilarity, or the taxonomic distance between species).

For asymmetric relationships, the edges have directions, giving a digraph. One form of hierarchical analysis for digraphs is therefore a comparison of results with and without the inclusion of the edge directions (for an example, see Fortuna *et al.* 2006). We can also have asymmetric relationships that are positive or negative, such as the nurturing effects of nurse plants or the 'interference competition' of chemical allelopathy, producing edges that have both signs and directions. Another obvious application of digraphs is in the description and analysis of movements through time, where the nodes represent subsequent positions of an individual object observed at different occasions, with the edges showing displacement between subsequent observations and the direction of the edges showing their order. The set of edges for any one individual in such a *displacement digraph*, when it is also an explicitly spatial graph, is a graph-theoretical path that sketches the actual path of movement but shows the displacement of each step rather than actual distance of movement.

Often the objects that are represented by the graph nodes fall into two categories, such as hosts and parasites. In a strict sense, in *bipartite graph* the nodes are in two subsets with edges only between subsets, not within. For example, we can have a bipartite graph depicting the relationships between pollinators and plants with edges joining each plant species to its own pollinators. The concept of a bipartite graph applies also to plants and their herbivores and to other parts of food webs (Higashi *et al.* 1991). As an extension, plants, herbivores and carnivores may form a tripartite graph, although the three levels may not be fully separate because of omnivores. Quadripartite graphs are not unusual, at least in metabolic studies, with nodes representing genes, enzymes, reactions and metabolites (Nikoloski *et al.* 2008), but these graphs are not often spatial, mainly providing non-spatial summaries of spatially indexed processes.

A spatial graph where each node has an associated graph representing some other structure (possibly

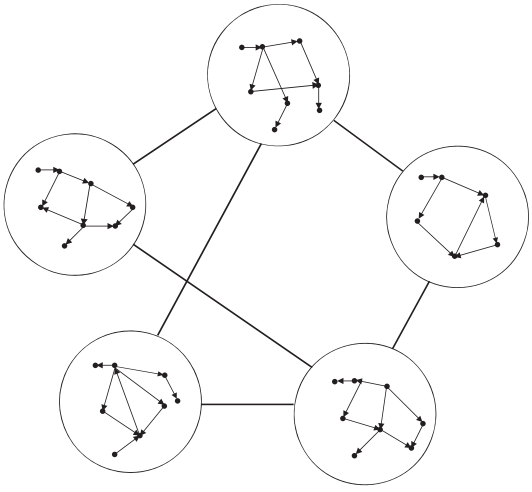


Figure 3.20 A graph-of-graphs: each node of the main graph has a second-order graph associated with it. Usually the main graph will be spatial, showing locations and neighbours, with the inserted node graphs depicting some structure or set of functional relationships for each location.

aspatial; e.g. food web), as shown in Figure 3.20, is called a **graph-of-graphs**. Fortuna & Bascompte (2008) gave an example, showing a graph of a metacommunity consisting of nodes, each of which has its own graph of multispecies interactions. Comparing two complete graphs of five nodes representing habitats, Melián *et al.* (2005) found that each habitat node has its own tri-trophic digraph of a simple food chain or food web that includes omnivory.

We can also consider the relationship between physical and functional connectivity using a graph-of-graphs. For a set of locations, 1 to n, let the local graph of interest (such as a food web like Figure 3.1*b*, or summary structure of plant species associations, like Figure 3.1*c*) be g_i. The locations themselves, as second-order nodes, are joined by a network of connections, depicted by graph G (Figure 3.20). The graph G may be part of a hierarchy of graphs of increasing connectedness, such as the series shown in Figures 3.2 to 3.7. We can then study the differences among the $\{g_i\}$, for example the gains and losses of nodes or edges, as a function of the characteristics of the position in G. The positions of the second-order nodes can be characterized by **cut-points** and **cut-edges** (**bridges**), **connectedness**, **degree of a node**, and so on (see Glossary). For example, we could ask: 'Does more connectivity in G increase the similarity of the $\{g_i\}$?' In addition, the pairwise similarities of the graphs $\{g_i\}$ can be measured using the methods of Dale (1985) and those similarities then mapped onto G for further analysis. As with many of the approaches described in this chapter, this is one that has been used, but rarely, and should be worth further consideration.

3.4.2 How to create subgraphs

The concept of a subgraph was introduced informally in the discussion of the hierarchy of neighbour networks (Section 3.2). A subgraph is derived from a graph by taking a subset of the nodes or a subset of the edges, or some of both (you just cannot add any new nodes or edges). In ecological studies, we are often interested in determining a particular kind of subgraph, in which there is a much higher ratio of edges to nodes than in the graph as a whole. These subgraphs are sometimes referred to as *modules* (connected **subgraphs**). *Modularity* is then the divisibility of a graph into highly connected subgraphs with relatively few edges between them. A closely related concept is that of *blocks* within a graph, where a block is a strongly connected (no bridges or cut-points) component (subgraph) of a graph. The existence of modules or blocks may reflect the community substructure of any spatial graph. Newman (2006) provided an algorithm for dividing a univariate graph into highly connected subgraphs based on the eigenvectors of a characteristic matrix for the graph. The characteristic matrix is not the same as the more familiar *adjacency matrix*, which is an $n \times n$ matrix with 0s indicating pairs of nodes that have no edge and non-zero values, usually 1s, indicating the presence of edges (see Figure 3.10).

This division of the graph into modules is much like spatial partitioning which splits the nodes of networks or similarity graphs into subsets, often based on the boundary detection algorithm called triangulation-wombling (Chapter 9; Fortin 1994; Jacquez *et al.* 2000). The algorithm works by identifying edges between

pairs of nodes that are most different, which will indicate the steepest changes between adjacent locations; these edges are then candidates to be parts of a boundary between regions of higher internal similarity. The technique is also closely related to the converse process of spatially constrained clustering, of which there are many versions, where the most similar nodes are joined into clusters but only if they are also adjacent in the spatial structure (that's the constraint).

These are only a few examples of the use of subgraphs in ecological analysis, but there are many others. As a final comment, we note that a graph that consists of several highly connected subgraphs can be considered as a graph-of-graphs (as described above), because each block (**component**) can be reduced to a single node, and then the structural relationships among the nodes can be depicted in a new reduced graph. Each node of the new graph is a block, which is also itself a graph. Where we have spatial graphs, the reduced version can be used to simplify spatial structure of complicated arrangements to make the essential spatial features more obvious and available for further analysis.

3.4.3 Graph models

We can classify some graphs according to the kind of models that generate them or at least by comparison with the known characteristics of graphs generated by particular models. One such graph is an Erdös & Rényi *random graph*, which is fully random in that it begins with a set of nodes, adding edges, one by one, between randomly chosen pairs of nodes that do not yet have an edge, so that new edges are independent of the positions of existing edges (Erdös & Rényi 1960). Graphs that are *scale-free* are formed by the random but preferential attachment of edges to those nodes that already have more edges, so that a small number of nodes, known as hubs, have many edges but most nodes have only a few (Barabási & Bonabeau 2003; Barabási 2009). The graphs that are *small world* graphs are somewhat similar in having preferential attachment (Watts 1999) although a classic algorithm is to start with a ring graph (a graph that creates a circular polygon) and then randomly move some of the edges to new attachment points (Watts & Strogatz 1998). The most important difference is that these graphs have a number of subgraphs with a high proportion of the possible edges between the nodes, so that they are 'highly connected', in the terminology described below. These three models produce graphs that differ in a number of properties but the distribution of node degrees is an important characteristic that distinguishes them.

(1) In a *random graph*, the distribution is a binomial or Poisson.
(2) In a *scale-free graph*, the distribution is a power law.
(3) In a *small world graph*, the distribution is an exponential.

These three models of graphs differ providing insights about the processes that may have produced them: for example, how vulnerable is its function to the removal of any particular node (think about airline routes in a 'hub and spoke' model and the familiar effects of snowstorms!).

In addition to examining the distribution of the node degrees, it may also be useful to evaluate the relationship between the degree of a node and the distribution of degrees of its first-order neighbours, which are those nodes joined to it by an edge. For example, there can be an important structural difference between two graphs which both have a small number of high-degree nodes and many low-degree nodes, depending on whether the high-degree nodes are directly connected to each other or whether paths between them are longer and depend on nodes with low degrees.

3.5 Network analysis of areal units

There are at least two obvious extensions to our discussion of the analysis of patterns of events in a two-dimensional plane using neighbour networks. The first is to look at areal units, such as habitat units ('patches') in a landscape, as the objects for direct spatial analysis, rather than reducing them to dimensionless points. The second is to look at patterns of events in fewer than the usual two dimensions, such as point events along a linear structure or straight line,

or in more than two dimensions, such as point events in a three-dimensional space.

Our treatment of studying patches in the plane is, in part, prompted by the graph theoretic approach to studying landscape connectivity (Cantwell & Forman 1993; Urban & Keitt 2001; among others). That series of papers brought renewed attention to graph theory as a useful context for the evaluation of ecological structure, and graph theory is obviously a favourite topic of the authors of this book (Dale 1977; Fortin 1994; Dale & Fortin 2010)!

A common underlying assumption in landscape ecology is that an area can be represented by a set of identifiable landscape units, referred to as 'patches'. Studies of fragmentation consider patches of habitat suitable for the focus organisms (e.g. a rare or endangered species) situated within a matrix of unsuitable landscape elements (e.g. roads, pastures, and cultivated fields surrounding woodlots in an agricultural region). In the graph theoretic approach to evaluating this kind of structure, habitat patches are represented by the nodes of a graph and the edges that join them represent connections between the patches, related to ecological processes. That is, the landscape is represented as a functional network, with colonization or dispersal being the process of primary concern in conservation-oriented studies. In the preceding section we provided a nontechnical introduction to some of the terminology used in graph theory. We need a few graph theory terms for this application. A graph is **connected** if there is at least one **path** (a sequence of nodes joined by edges) between any two nodes in the graph. The **degree** of a node is the number of edges that are attached to it. The most highly connected graph is a **complete graph**, in which each node has an edge to every other node, so that all nodes have degree $n - 1$. A **planar graph** is one that can be drawn in the plane without the edges intersecting; for example: the complete graph of order 4 (4 nodes, 3 edges at each) is planar, but the complete graph of order 5 (5 nodes, 4 edges at each) is not (try it!).

In a connected graph, a **cut-point** is any node, the removal of which causes the graph to be no longer connected. Similarly, any edge, the removal of which disconnects the graph, is called a **cut-edge** or a **bridge**. Obviously, a complete graph has no cut-points and no cut-edges. A graph that is not itself connected will be made up of a number of connected subgraphs, called **components**; a connected subgraph is a component if it is maximal: that is, if it is contained in no larger subgraph that is also connected. The *order* of a component is just the number of connected nodes it contains. For a connected graph, there are (at least) two measures of how strongly it is connected: *node* **connectivity** is the minimum number of nodes that must be removed to disconnect the graph. **Edge connectivity** is the minimum number of edges that must be removed to disconnect the graph. Figure 3.21

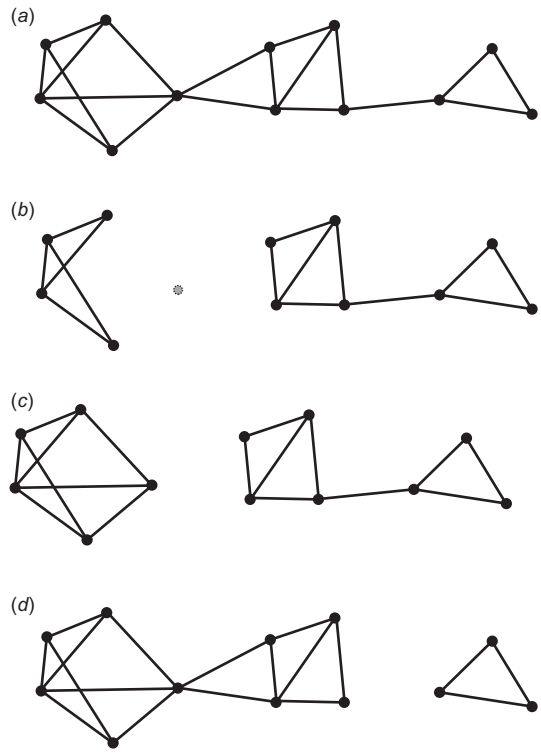


Figure 3.21 A series of graphs to illustrate some terms related to connectivity: (*a*) original graph, (*b*) the point now isolated is a cut-point because its removal disconnects the graph, (*c*) two lines removed disconnect the graph but (*d*) the line removed is a cut-edge because its removal disconnects the graph.
The edge connectivity is 1 and the node connectivity is also 1.

illustrates these various terms. There is an obvious analogy between the connectivity of a spatial landscape graph and the number of dispersal routes, made up of corridors and patches, available in the landscape it represents.

Graph theory is about the structure of connections and, in the most general form, the edges of a graph do not have properties such as length or weight, but in some applications, it is useful for the nodes and edges to have properties of their own. For example, each edge in the graph may have a length associated with it, $d(e_k)$, which could be the physical distance between nodes which have locations, as drawn on a map, or some other property of the network, such as cost of transport or resistance to movement. The graph theoretical **distance** between two nodes in the graph, δ_{ij}, is the minimum path length (smallest sum of edge lengths) of any path between the two nodes. This means that a Minimum Spanning Tree (MST), which we constructed in Section 3.1.1 using the criterion of minimum physical distance, could also use the criterion of minimizing any other measure of edge length. For many applications in ecology, for example in the analysis of landscapes, we can think of a graph of landscape objects as being embedded in the plane, with the nodes having geographic locations or positions, and the edges having the length of the (Euclidean) distances between the nodes they join, thus forming a spatial graph. Whatever the measure of length, the *eccentricity* of any node, $\varepsilon(v_j)$, is the maximum graph theoretical distance to any other node in the graph. Last, the **diameter** of a graph is the maximum eccentricity of any node in the graph.

Urban & Keitt (2001) examined the properties of landscape graphs by considering the effects of removing edges from the graph and of removing nodes, analogous to the loss of dispersal corridors in the first case and of habitat patches in the second. The characteristics they suggested for evaluating edge removal are purely graph theoretic: the number of components that result, the diameter of the largest component, and the order of the largest component. The procedure they suggested is to start with a complete graph (one in which all possible pairs of nodes are joined by edges) and then to use a series of threshold distances and to remove edges, leaving only those shorter than that threshold, thus creating a series of threshold distance networks. The response of the graph theoretical properties of these networks to the threshold distance then provides an evaluation of the landscape patch structure.

As an example, Figure 3.22*a* shows the approximate sizes and locations of 21 lakes in an extensive peatland near the Alberta–Saskatchewan border (55° 45′ N, 110° 45′ W) and their distance-based Minimum Spanning Tree. There are no direct permanent surface-water connections between these lakes, and it is reasonable to take as a working hypothesis that dispersal between them is inhibited by distance. Figure 3.22*b* shows their Delaunay triangulation (DT) graph. Using the Urban & Keitt (2001) approach on this graph, Figures 3.22*c* and *d* show the threshold distance networks for 7 km and 4 km. At 7 km, the graph is a single component of all 21 nodes, with no cut-points or bridges. The diameter of the single component is slightly larger than that of the original complete graph (30.3 rather than 26.5), because (for example) there is no longer an edge between Lakes 5 and 21, and the path between them goes through Lake 4. At 6 km (not shown), the graph is no longer connected, but the largest component includes 20 of the 21 lakes, and its diameter has decreased (27.6) due to the loss of the most isolated lake. The main component contains several cut-points (e.g. Lake 8) and cut-edges. The change from 7 km to 4 km for the network threshold is dramatic. There are now nine components, of which the largest has only five lakes and a diameter of only 6.1. The fairly abrupt transition from a few components, some of which are large, to many small ones is similar to Urban & Keitt's (2001) observations on a hypothetical but realistic landscape of patches.

Distance graphs with the edges determined by threshold distances are known as **geometric graphs**, and there is a growing body of theoretical results on random geometric graphs (Penrose 2003). This work can provide the background needed for the evaluation of observed properties for graphs of ecological phenomena by comparison with random graph characteristics. Some of the properties that can be tested include the distribution of node degrees and the connectivity

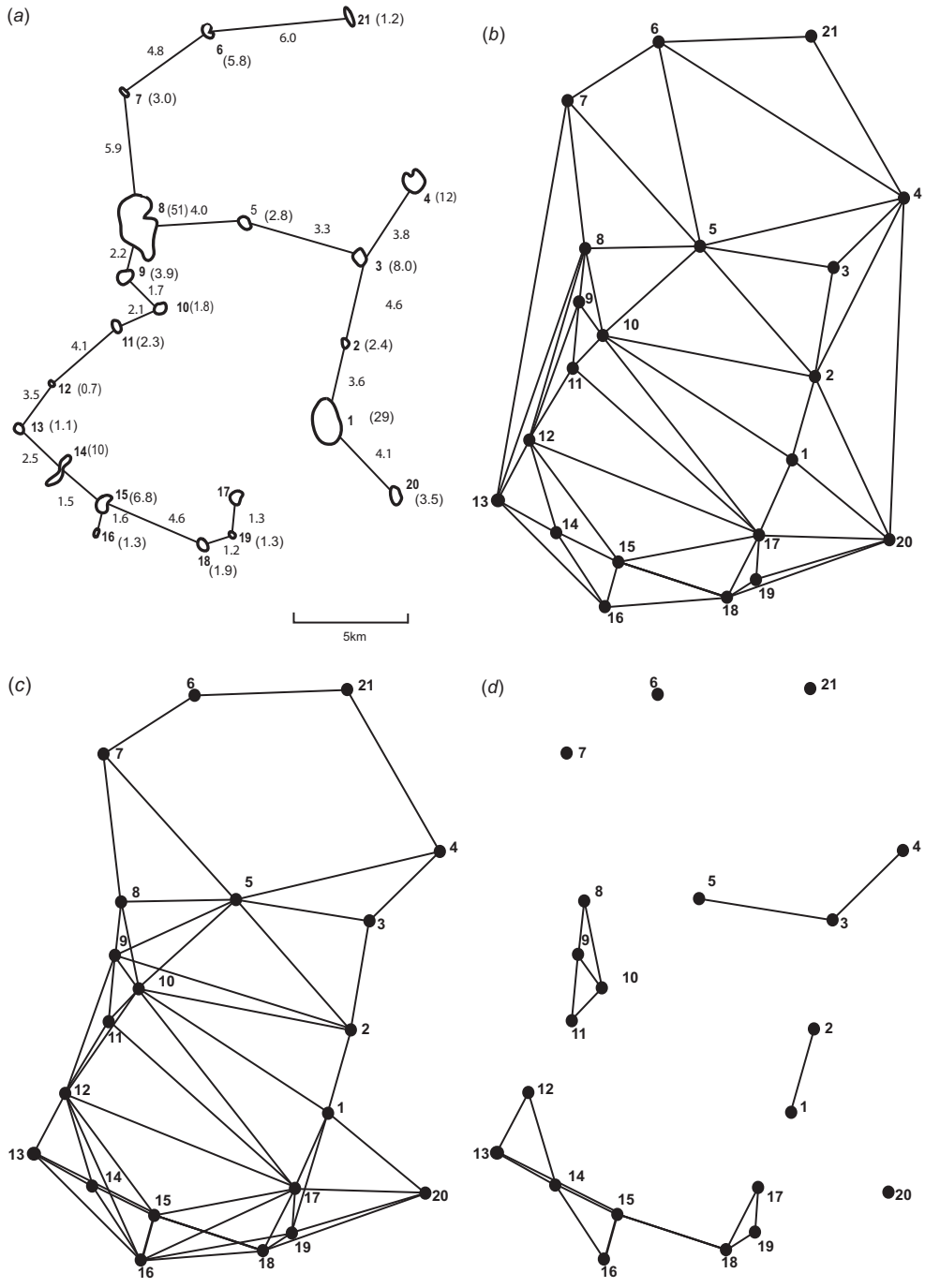


Figure 3.22 Illustration of connectivity of lakes from a peatland near the Alberta–Saskatchewan border. (*a*) The Minimum Spanning Tree for 21 lakes in a peatland, with approximate distances between them (in kilometres) and their relative sizes (arbitrary scale). (*b*) Delaunay triangulation graph of the same set of lakes. (*c*) and (*d*) Network of lake neighbours using thresholds of 7 km and 4 km. Different distance thresholds result in different graphs, but a graph with a smaller threshold is a subgraph of any with a greater threshold.

of a graph and the number of its components (Penrose 2003). We expect to see many applications of this approach in the near future.

If we return to the Delaunay triangulation of the example landscape depicted in Figure 3.22*b*, we can suggest a simple measure of the importance of any edge, e_{ij}, in the graph. That measure is the minimum 'cost' of its removal, either in absolute distance of edges required to replace the connection between v_i and v_j, or as a proportion of its length:

$$c_d(e_{ij}) = l(e_{ik}) \quad + l(e_{kj}) - l(e_{ij}) \tag{3.21}$$

or

$$c_r(e_{ij}) = \frac{l(e_{ik}) + l(e_{kj}) - l(e_{ij})}{l(e_{ij})} \tag{3.22}$$

For example, the shortest replacement path for e_{35} is through v_4, so that

$$c_d(e_{35}) = l(e_{34}) + l(e_{45}) - l(e_{35}) = 3.8 + 6.1 - 3.3 = 6.6;$$
$$c_r(e_{35}) = 6.6/3.3 = 2.0. \tag{3.23}$$

For $e_{9,11}$, the cost of its loss is much less:

$$c_d(e_{9,11}) = l(e_{9,10}) + l(e_{10,11}) - l(e_{9,11})$$
$$= 1.7 + 2.1 - 2.2 = 1.6;$$
$$c_r(e_{9,11}) = 1.6/2.2 = 0.73. \tag{3.24}$$

To evaluate node removal, the criteria that Urban & Keitt (2001) suggested are more explicitly ecological than those for edge removal:

(1) a recruitment index, R (a weighted sum of patch areas, with the weighting being a measure of patch quality);
(2) an index of dispersal flux, F (a weighted sum of patch areas, as for recruitment, but including the probability of dispersal from the focal patch to another); and
(3) a measure of traversability, T (the diameter of the largest component that remains after the node is removed).

They examined the iterative node removal from the complete graph by three different procedures: (1) random choice for removal; (2) removal of the node with smallest patch area; and (3) small-leaf removal where the smallest area node (i.e. a *leaf*) attached only to one other node in the current MST is removed (e.g. Lake 4 in Figure 3.22*a*). Not surprisingly, they found that small-leaf removal degraded the characteristics of the network less rapidly than did random removal. The same three measures, R, F and T, were then used to evaluate the importance of individual nodes according to each criterion. This was done by examining the difference in the measure before and after the node's removal from the entire graph. Because this evaluation was based on the complete graph, which includes all possible edges, evaluation of importance may be very different from one based on the DT or on the MST itself. For example, in a complete graph of our landscape of lakes, removing Lake 8 would have little effect on the graph's diameter, whereas the removal of Lake 21 would make a big difference. In the MST of the same landscape, the removal of Lake 21 reduces the graph's diameter somewhat, but the removal of Lake 8 has a profound effect on the graph's characteristics.

Again, as with edges, other measures of the importance of individual nodes in the landscape can be considered. For example, we could compare the average length of edges attached to a node in the DT graph, with the average length of new edges in a DT graph when the node is removed. Thus, the average length of edges attached to Lake 5 in Figure 3.22*b* is 5.5 km; when Lake 5 is removed, and the DT graph is reformed, the average length of all edges is 8.4, a ratio of 1.51. On the other hand, if Lake 14 is removed, no new edges are required in the DT graph, and the ratio is 0.

If we use a complete graph of the landscape and treat it as a purely combinatorial structure, there is no 'topological' distinction among the points or among the edges. Once the graph is embedded in the plane, as a map, the points are differentiated, with some more central in location and some more peripheral.

More specifically, we can define the *perimeter* nodes of the graph as those that are in the convex hull (Lakes 21, 4, 20, 18, 16, 13, 7, and 6 in our example), with the rest being *interior* nodes. In a complete graph, there are few ways to evaluate the importance of a particular node to the overall connectedness of the graph, because there are no cut-points, cut-edges, and so on. To discuss node importance, we should turn our attention to the Minimum Spanning Tree, threshold distance network graph, or Delaunay triangulation graph. The importance of a node will be related to its position in the graph (perimeter or interior), its degree (number of edges), and the distances to the neighbours to which it is joined. The removal of a perimeter node reduces the network extent (its 'footprint'); the removal of particular perimeter nodes will reduce the graph's diameter, and the removal of interior nodes will reduce the number of alternate paths between other pairs of nodes. In the DT graph, node removal does not result in disconnection, but it often does in the MST. The removal of an interior node with a degree of three or more from the MST will result in the creation of more components of considerably lower order and smaller diameter than the original graph. For example, the removal of Lake 8 in Figure 3.22*a* would have a large potential impact on dispersal or colonization in this group of lakes.

This is an interesting area of research and, clearly, there is more work needed, particularly on evaluating which features of habitat patch networks are most important to the dispersal of particular kinds of organisms.

3.6 Spatial analysis of movement

Although we will provide a more detailed treatment of the analysis of spatio-temporal data, which includes the analysis of movement, in Chapter 11, it seems appropriate to extend this chapter's introduction to the use of graphs and graph theory to include movement as a specific topic.

3.6.1 Transport and gravity models

Transport models look at the effects of distances, costs, and advantages of particular transportation corridors on the flow of materials. The question is usually treated as a linear programming problem focusing on a single product that is available in a number of locations and needed in several other locations; the problem to solve is to move the material from sites of availability to sites of requirement in such a way as to minimize the amount moved, or the distance moved or the costs of transportation. The least-cost path is often used, but it is just one possible choice of approach to solving the general problem of multi-source : multi-sink transportation.

Gravity models (Bossenbroek *et al.* 2001) also evaluate the transportation of materials, people, or information. These models are based on the idea that the strength of a link between two patches will be proportional to the product of their weights (however determined appropriately for the subject) divided by the square of the distance between them. When patches with area and shape are used rather than dimensionless points, edge-to-edge links between the patches are used rather than point-to-point or centre-to-centre. There can be differences between as well as compatibilities between Euclidean distances and least-cost links. There is a relationship between the least-cost links between patches and the Voronoi polygons that can be defined for the set of patches, but that relationship is not as precise as between Voronoi polygons and the Delaunay triangulation, which are duals (Fall *et al.* 2007).

Gravity models have been used in invasive species studies using Euclidean distance between source and sink lakes (Bossenbroek *et al.* 2001; Leung *et al.* 2006) as well as least-cost distances between lakes (Drake & Mandrak 2010). Applications of gravity models have also been used in landscape genetics studies to determine the environmental factors affecting the genetic variability of Columbia spotted frog metapopulation structure (Murphy *et al.* 2010).

3.6.2 Least-cost paths

Structural questions arise when we study the movement of organisms between habitat patches in a

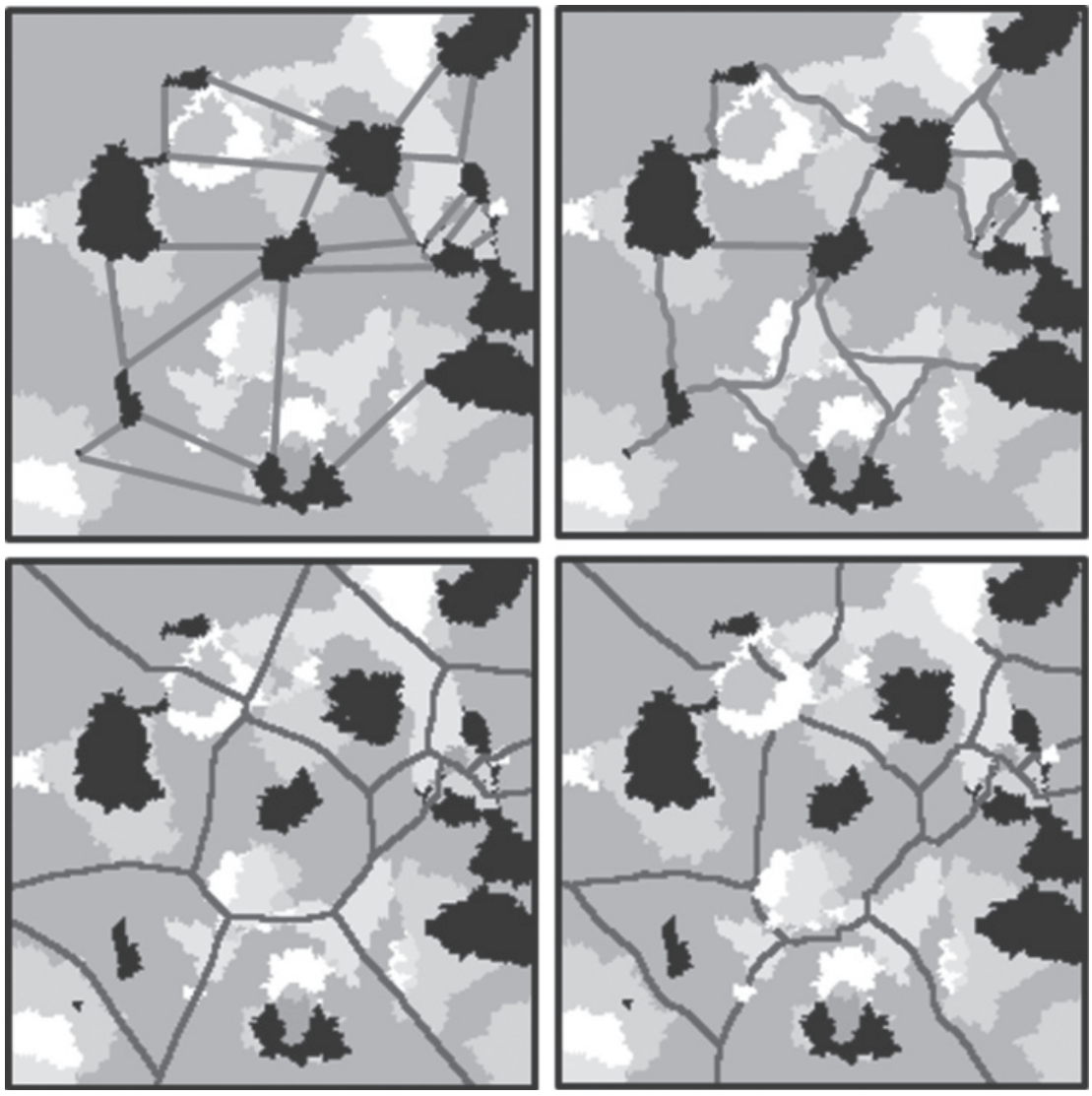


Figure 3.23 Examples of spatial graphs and Voronoi polygons. Left upper panel: spatial based on Euclidean distances and the minimum planar graph (equivalent of Delaunay tessellation algorithm when Euclidean distances are used). Left lower panel: Voronoi based on Euclidean distances but weighted by the size of the patches. Right upper panel: spatial based on least-cost distances and the minimum planar graph (equivalent of Delaunay tessellation algorithm when least-cost distances are used). Right lower panel: Voronoi based on least-cost distances and the patch size weight. Nodes are the black patches. Cost (resistance) values are in the shades of grey: white is the highest cost to dark grey for the lowest cost values.

landscape. One approach is to look at connectivity and 'least-cost' links (Figure 3.23). The costs of movement are weights on the network edges, but there may be several edges or several paths between any pair of nodes or patches (which have shape as well as location). The assumption is that organisms move preferentially among those nodes using the least costly paths. When crossing is forbidden, we get a **minimum**

planar graph (Fall *et al.* 2007). There is also a relationship between the least-cost links and the actual paths between patches. Applications of such spatial graphs include the design and effectiveness of natural reserves as described by Fall *et al.* (2007) among many others. The choice of resistance values used in least-cost path analysis can be subjective however. Rayfield *et al.* (2010) showed how the selection of resistance values influences the position and length of the least-cost paths in fragmented landscapes.

3.6.3 Circuit theory

Another approach is to evaluate all possible paths between any pair of nodes, as in electrical circuits (McRae *et al.* 2008). Thus more than just the single 'easiest' path is considered, and a number of parallel paths between two patches may provide a corridor that allows greater flow than a single line path (Figure 3.24). This leads to the concept of considering the sum of all possible paths between patches. Several ecological studies have used this circuit theory approach to map species' movement probability in fragmented landscapes or to identify areas to be protected as they are used as movement corridors between protected areas (Koen *et al.* 2010).

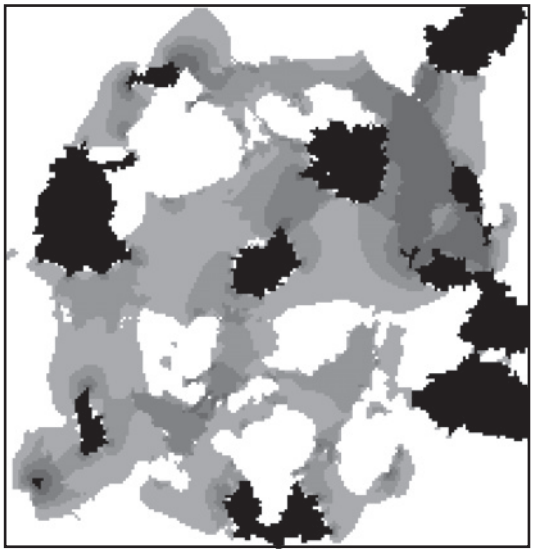


Figure 3.24 Example of circuitscape analysis. Map of the potential probability of movement between the nodes (black patches). The probability of movement between patches is in grey (the highest probability of movement is in dark grey and the lowest in light grey).

3.6.4 Spatial graphs and movement

In this chapter, we have discussed spatial graphs, in which the positions of the nodes represent actual locations in space, and a number of different kinds of rules for determining which pairs of nodes are joined by edges in the graph. For example, we can join all pairs of nodes that are separated by less than a chosen threshold distance (Euclidean or least-cost) or we can join each node to its (single or first) nearest neighbour. One possible edge rule that we have yet to discuss requires an additional dimension in some ecological data sets, that of time. One inspiration for this suggestion begins with data consisting of the physical locations of an animal taken at intervals through time by direct observation, radio-telemetry, or similar technique. Given a time sequence of GPS locations, an obvious rule to create edges between nodes is to join each location to the location of the subsequent observation, thus following the path of the animal's movement, at least approximately, and creating a connected directed graph (also a spanning tree) of the locations. Of course, this is really a displacement digraph since the actual physical path followed between one location node and the next is not accurately shown by the position of the directed edge; it is only indicative of the beginning and the end of the movement made in the time interval and not the distance actually travelled.

3.6.5 Corridors

In considering the movement of animals particularly in the context of migration between landscape patches, it must be kept in mind that the pathway used is not just a simple one-dimensional path, whether straight or curving; often the pathway is a corridor consisting of an elongated two- or three-dimensional spatial structure through which the animal moves. This is most obvious in the case of some easily defined linear structure such as a hedge-row or grown-over ditch in an agricultural

landscape, but it is equally true in less obvious cases, where the path follows routes of less exposure or greater safety. Thus a corridor could consist of the path itself together with a buffer zone around it or a series of least-cost paths that have comparable path length cost (Pinto & Keitt 2009).

3.7 Testing hypotheses with graphs

The basic concept of using graphs to test ecological hypotheses is to compare how well different graphs can explain observed structure in the data, whether we are dealing with one or several single independent variables, or with a set of dependent variables and their presumed predictors. With a single variable, we can determine which graph, out of a number of choices, produces the highest autocorrelation of values of the nodes or of the edges, indicating the best explanatory structure for the pattern of observations. For example, if similarity between sites decays with distance, the first-order neighbour autocorrelation would be significantly lower in a complete graph (in which all nodes are first neighbours) than with edges determined by a low distance threshold (in which only nearby sites are neighbours).

With different kinds of variables, we can evaluate which graphs can best explain the dependent variable characteristics of nodes (e.g. genetic composition or genetic diversity) or of edges (e.g. neighbouring site similarity or observed migration rates), as a function of independent variable characteristics such as patch size (nodes) or inter-patch distances (edges). The graph with the best-supported hypothesis is the one that produces the strongest regression or the largest correlation coefficient for the relationship between the explanatory and dependent variables at adjacent nodes or edges.

The basic approach is to compare the results for a null-hypothesis condition, having fewer distinguishing characteristics for nodes or edges, with those for an alternate-hypothesis condition, having more characteristics considered, as summarized in Tables 3.2–3.4. For example, to test the effects of the definition of neighbours, we can begin with the null hypothesis that all nodes have equal influence on the characteristics of any other node; that null hypothesis is represented by a complete graph of all possible edges. The alternative hypothesis is that the nearest neighbour nodes are the most important explanatory factors, and this alternative can be represented by one of the topological graphs such as that of first-order nearest neighbours. A more

Table 3.2 How to compare different graph properties (modified from Dale & Fortin 2010)

Comparison type	Graph characteristics	Methods
Graph theory properties	Same nodes at the same locations	Network measures (Rayfield *et al.* 2011) Co-inertia analysis
	Same number of nodes but locations not the same	Test against model graphs (e.g. random graph, small world)
Area connected, Length of path	Nodes as patches: edges determined by Euclidean or resistance distances	Amount of patch area connected based on: dispersal limits in Euclidean distance, or least-cost path of resistance values (O'Brien *et al.* 2006)
Strength of relationships between nodes	Between-nodes flow of organisms, information, or energy	Correlate edge weights and directions with ecological or genetic data using Mantel test
Nodes and edges	Same number of nodes not at the same location	Area overlap of Voronoi polygons of the nodes
	Relations among nodes given edge weights	Gravity model (Murphy *et al.* 2010) Incidence function (James *et al.* 2011)

specific example would be the situation where each node is a site with a single value for species richness associated with its node. The null hypothesis is that site richness is the same as the average for all sites (complete graph), and the alternative hypothesis is that site richness is at least partially determined by the richness of the nearest neighbour sites (NN graph).

Table 3.3 Neighbour network graphs[a] for hypothesis testing (bold indicates graphs in the hierarchy given in Table 3.1)

0. Null graph, no edges (0)
1. Mutually nearest neighbour pairs (≈0.62)
2. All nearest neighbours (≈1.4)
  2a. Distance threshold graphs (0 to ≈ *n*?)
  2b. Double threshold graphs (0 to ≈ *n*?)
3. Minimum Spanning Tree (≈2.0)
  3a. Ulam tree (≈2.0)
4. Relative neighbourhood graph (≈2.4)
5. Gabriel graph (≈4.0)
  5a. β-Skeleton graphs (0 to ≈ *n*?)
6. Delaunay triangulation (≈6.0)
  6a. Other triangulations (≈6.0)
7. Augmented Delaunay graph (≈12.0)
8. Complete graph (n – 1)

[a] With approximate average number of neighbours for realizations of CSR.

In schematic form, this is:

$\{H_0 =$ complete graph$\} \equiv$ "All nodes predict the focal node equally well"

versus

$\{H_1 =$ nearest neighbour graph$\} \equiv$ "The closest nodes best predict the focal node".

As described above, the two hypotheses are then compared by an evaluation of autocorrelation, regression, or correlation. The hypotheses evaluated then depend on the rules upon which the structure of the graph is based (Tables 3.2–3.4).

Whatever the nodes represent, there are always a number of different classes of rules for determining the edges of the graph, spatial or not. For example, we can select random pairs of nodes to join with edges such that: node ratio from the complete graphs with all $n(n-1)/2$ edges to the empty graph with none. We can use a distance threshold and join all pairs of nodes with separations up to that threshold to produce a geometric graph or a series of them (Penrose 2003); we can combine distance and position characteristics in a topological rule, such as the rule for a Minimum Spanning Tree; or we can join nodes based on functional connections such as dispersal corridors between landscape patches. Many of these rule classes can each produce several levels of graph, often giving a hierarchy of graphs within the set, for example by physical distance in a threshold rule (1 km, 2 km, 3 km, etc.),

Table 3.4 Hypothesis testing using graphs

Element	H_0	H_A
Nodes	Only the locations (zero-dimension) of samples	Added patch characteristics e.g. size, shape, buffer area
	Only the locations (zero-dimension) of samples	Individuals or groups of individuals
	Unweighted or equal weights	Weights of patch quality, site area, population size, population density, etc.
Edges	Euclidean distance (one-dimension)	Added buffer around straight line (two-dimensional)
	Symmetric or undirected	Asymmetric or directed; e.g. dispersion, species interactions
	Complete (all possible edges)	Distance threshold, topological algorithm, species dispersal ability, barrier to movement, etc.
	Unweighted or equal weights	Weights of flow capacity, landscape resistance (landcover type, barriers), transport costs, etc.
Graph (both)	Unweighted	Weights for nodes and edges as above

or by inclusiveness in the topological series of NN, MST, and so on. In any of the hierarchical series, the graphs with more edges include all the edges of the graphs with fewer edges. (An obvious exception would be the double threshold graph where the edge lengths are constrained between both upper and lower limits. The double threshold distance graphs with upper and lower limits to edge lengths in fact partition the structure of the complete graph into separate graphs defined by classes of edge lengths.)

Within almost any individual graph, there is a series of neighbour node sets: first-order, second-order, and so on, determined by path lengths of 1, 2, 3, and so on edges. This series of neighbour sets within a graph permits the evaluation of characteristics (e.g. spatial autocorrelation of node weights) as a function of path length. In addition, we can compare two different characteristics, such as two different node labels like species richness and patch area, or two different edge labels like community similarity and migration rate, as a function of path length within the graph.

These considerations allow us to distinguish (at least) three levels of comparison.

(A) By path length within a single graph = neighbour order (1^o, 2^o, etc.).
(B) By level within a rule class, such as distance threshold, edge: node ratio, or capacity of a functional connection, for a single set of nodes, producing an inclusive series of graphs.
(C) By class of rule for edge creation, such as physical distance versus functional connection.

These comparisons then allow the test of different hypotheses by comparing the same properties and based on different edge rules, their levels, or neighbour order for the same set of nodes.

Comparing graph structures as explanations of observed node or edge characteristics requires an evaluation of explanatory power. A regression or correlation approach seems the most logical, plotting autocorrelation of neighbouring node or edge weights (or correlation of two different edge weights) as a function of distance threshold or level in graph hierarchy, and as a function of path length. For two variables of interest, this is the process of creating a cross-correlogram with path length as the measure of distance. (A similar investigation of the **complement graph** may also provide some insights! The complement graph of G, call it G', has the same nodes and edges between all pairs of nodes that do not have an edge in G, and no edges between pairs in G that do have an edge between them.) The graph that produces the highest regression scores or correlation coefficient, or the graph that produces the best difference between its own test value and that of its complement is the one with the best-supported hypothesis. Formal tests with results of known probability values are made difficult because of lack of independence among such tests.

These three characteristics of the graphs we use, the different kinds of rules for edges, the hierarchies of graphs within rule classes, and the series of node neighbours determined by path length within each graph, make it possible to develop, evaluate, and compare a range of sophisticated hypotheses about the systems that the graphs represent. As a particular example, the use of spatial graphs as the basis for analysis can provide great insights in studies of ecological communities on virtual islands, or of sites on environmental gradients.

As an example, consider sample sites on two streams that run more or less parallel through two adjacent valleys (Figure 3.25; Peterson *et al.* 2013). We have sampled the community composition at each site, and have chosen to analyse the data using two different edge rules for spatial graphs: closest physical distance and closest stream distances. The similarity of first-order neighbours in the physical distance-based graph will be low, whereas in the stream-distance graph it will be high. For second-order neighbours, there will be high similarity in the physical-distance graph, but probably not quite as high as for first-order neighbours in the stream-based graph.

We can use a range of appropriate pairs of null and alternate hypotheses, each based on characteristics of nodes or edges, or both, to evaluate explanations of observed structures in ecological systems. In this way, graphs, with the great variety of rules available to create them, provide ecological researchers with a rich and sophisticated approach to hypothesis testing.

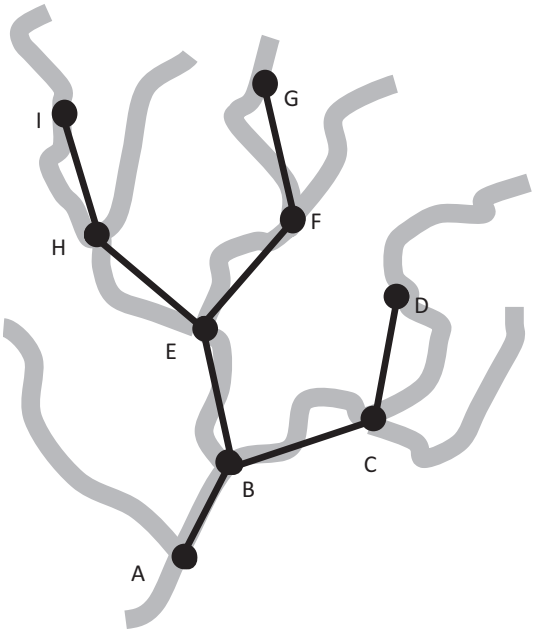


Figure 3.25 Graphs for hypothesis testing: Euclidean distance (black) versus stream distance (grey) between sampling locations (labelled A to I). Locations F and D (or G and H) are very close by Euclidean distance, 'as the crow flies', but more distant if the distance along streams is used.

3.7.1 Comment on spatial graph randomization

Because of their structural complexity, graphs are difficult to evaluate using formal analytical methods, and so the more successful approach is to use permutation and randomization. Let us start with an example: a graph of six nodes and five undirected edges, with all the edges joining one node with all the others, forming a star. If the null hypothesis is that all six nodes are equivalent, we may wish to test whether the observed degree of the 'centre' node is unexpectedly high. We proceed to calculate the probability that, under complete randomness, five of five edges occur at a single node. There are 15 possible positions for undirected edges, and so that probability is 6 (nodes) $\times$ (5/15) $\times$ (4/14) $\times$ (3/13) $\times$ (2/12) $\times$ (1/11) $= 0.002$. It is highly improbable that, if the nodes are somehow equivalent, a single one would end up with so many edges. Changing the example to be a digraph, let us now suppose four of the five edges point outward from the high-degree node. If we ask whether this is highly improbable given the same null hypothesis, the answer is 'yes' because its probability is even less than 0.002 already calculated. If, however, the question is replaced by 'Given the positions of the edges, what is the probability that that four or more of five point outward?', the answer is different. It is then $5 \times (1/2)^5 + 1 \times (1/2)^5 = 0.1875$, adding together the probabilities for four edges and for five pointing inward, and giving a probability that would not lead us to reject the null hypothesis.

Note that in this example, we consider placing exactly the observed number of edges, five, onto a set of six nodes. This is different from assigning an edge (yes or no) to each of 15 possible positions with a probability of 5/15. In the first case, each example created has the same number of edges. In the second case, the number of edges in a set of iterations will have a distribution. There is a finite probability that a graph thus created will have no edges at all (probability $[10/15]^{15} = 0.0023$) or that it will have edges in all 15 locations (probability $[5/15]^{15} \approx 10^{-7}$).

In this example, because the graph is small and simple, we can calculate probabilities directly, but in a real spatial graph, not only will the number of nodes and edges be large, but some edge positions may be forbidden by the physical arrangement of the units under study (e.g. physical barriers that completely prevent dispersal). For that reason, it is only possible to perform the probability calculations needed to evaluate null hypothesis, whether single or in a hierarchical series, using randomization or permutation techniques. In randomization methods, some of the characteristics of the graph of interest remain fixed (e.g. the number of nodes and the number of edges), and others are allowed to vary (e.g. positions of the edges) in creating a reference set of graphs for comparison. Then, for some statistic of interest (e.g. the maximum node degree) the number of times the observed value is matched or exceeded in the reference set generated is used to evaluate the null hypothesis. In permutation tests, a similar process is followed, but usually the actual structure of the graph is maintained as constant throughout the development of the reference set, but the labels

or weights associated with the nodes or edges are redistributed among the available objects. For example, in a bivariate spatial graph of diseased and healthy plants with the edges indicating Delaunay neighbours, the observed numbers of labels for diseased and healthy (N_D and N_H) can be redistributed 1000 times and the proportion of healthy-to-healthy edges counted in each new graph. Redeploying the set numbers of the two labels among the graph's nodes is different from the process of creating new graphs of N_T nodes by labelling each node as diseased with probability N_D/N_T and as healthy with probability N_H/N_T; in the permutation case, the numbers in each class are constant, in randomization, the probabilities are constant, but the actual numbers in each iteration vary.

Because of the complexities of spatial graphs, and because of the possible ranges of hierarchical hypotheses that may be of interest in their application, tests based on permutation and randomization, and in particular randomization with certain restrictions, are the most powerful techniques available.

3.8 Concluding remarks

In this chapter, we have illustrated two different conceptual developments that lead to spatial graphs, which are mathematical objects consisting of points (nodes) joined by lines (edges) with the locations of the nodes and the lengths of the edges having a direct meaning and interpretation in a spatial context. One way is to begin with spatial patterns of point events and spatial patterns of linear objects, and then put them together according to the rules that the lines can only be placed where they join two points. The second approach is to begin with graphs as mathematical objects depicting structure in the abstract and then to embed them in space, usually two dimensions, thus forcing their spatial nature. Both developments are equally logical and have different intuitive appeals. Either leads to a very useful class of mathematical constructs, because spatial graphs have already proven to have a broad range of applications in spatial analysis, and undoubtedly, there are many more yet to be developed or discovered.

Glossary: graph definitions and properties

A **graph** is a mathematical object made up of two sets.

(1) {nodes}. These are the points, which are dimensionless but are usually depicted by dots in a diagram representing the graph. We will call them the nodes of a graph, saving 'points' as a more general term; they are also called vertices.

(2) {edges} = {pairs of nodes}. These are the lines of a graph but we will use "edge" to indicate an element of a graph, with 'line' being more general; they also called arcs or links. They are represented by simple lines in a diagram of the graph, whether they are straight or curving being unimportant because they usually represent relationships not objects in space. The edges are usually allowed to cross with no node occurring at the intersections, which are not themselves structural.

A **complete graph** is one in which there is an edge between every pair of nodes.

The **complement** of graph G is the graph G', with the same set of nodes and edges between all pairs of nodes that do not have an edge in G, and no edge between pairs that do have an edge in G. The edges of graph G and its complement G' together are the edges of the complete graph, being all possible pairs of nodes.

A **cut-point** (also **cut-node** or **articulation point**) is any node the removal of which will disconnect the graph.

A **cut-edge** (also **bridge**) is any edge the removal of which will disconnect the graph.

A **graph-of-graphs** is a graph of which every node either is, or has associated with it, a second level graph with its own nodes and edges.

An **aspatial** graph is one in which the position of the elements is arbitrary, even if the structure represented was originally a spatial one or the relationships depicted are originally spatial; that is, space is ignored.

A **planar graph** is one that can be drawn on the plane with no edges crossing (bold lines in Figure 3.26); each edge must be a continuous curve but need not be a straight line. A planar graph does not have to be a spatial graph, but the definition is essentially spatial

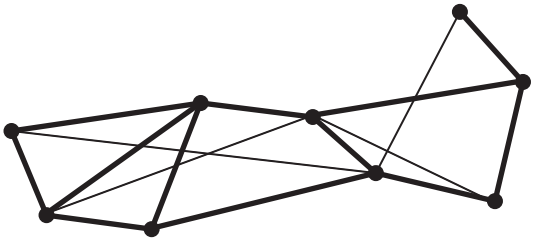


Figure 3.26 Graph as a collection of nodes and edges. Non-planar graph (all the lines). Planar graph (bold lines only).

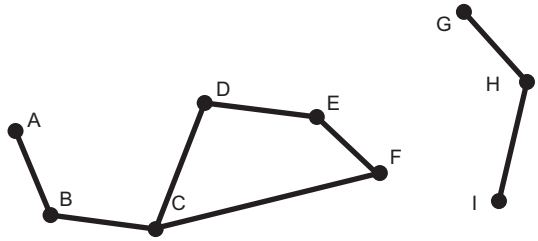


Figure 3.27 Two subgraphs (A, B, C, D, E, F and G, H, I) of graph presented in Figure 3.26.

and planar graphs are used as spatial graphs in the landscape ecology context (Fall *et al.* 2007).

A **subgraph** of any given graph uses nodes that are a subset of the graph's nodes, and edges that are a subset of the graph's edges joining pairs in the chosen subset of nodes.

A **directed graph** consists of a set of nodes with a set of edges that have direction, so that the edge from A to B is different from the edge from B to A. It is possible that two edges, one in each direction, join any pair of nodes.

In a **spatial graph**, the nodes have locations and the edges have their end-points tied to those locations, whether or not the positioning of the edge is explicit along its full length.

A **spatial digraph** is one where the direction of edges adds to or modifies the properties of the graph.

A **distance graph** is a kind of spatial graph, in which the nodes occur at locations in a Euclidean metric space and the edges between them take the values of the Euclidean distances between the points in that space.

A **geometric graph** is a spatial graph in *d*-dimensional space, which uses some measure of distance within the space, possibly, but not necessarily Euclidean. Connections by edges within the graph are determined by an upper threshold on internode distance (see Penrose 2003).

A **minimum planar graph** is essentially a spatial planar graph (no edges crossing) that uses the smallest total length of edges. These are used as spatial graphs in the landscape connectivity context (Fall *et al.* 2007).

The **degree** of a node is the number of edges attached to it. In Figure 3.27, the subgraph nodes A, G, and I have degree of 1; B, D, E, F and H have degree 2; and node C has degree 3.

A **path** is an alternating sequence of nodes and edges where each node in the path is attached to the next node in the sequence by an edge that joins them. Conversely, each edge in the path is joined to the subsequent edge at the node to which they are both attached. No nodes or edge may be used more than once in the same path.

The **length** of a path is the sum of the weights of the edges in the path. In simple cases, it is just the number of edges in the path. The graph theory distance between any two nodes in a graph is just the shortest path between them. In some applications, the shortest or minimum weight path is called a *geodesic* path.

A **cycle** (sometimes called a **loop**) is a closed path: it begins and ends at the same node, usually restricted to having no node or edge used twice. A cycle is a path that is a closed polygon of edges. In the subgraph example (Figure 3.27), CDEF is a cycle; GHI is not.

A **tree** is a connected graph that has no cycles. A typical phylogenetic tree depicting evolutionary history will have nodes of degree 1 (the *leaves* at the top of Figure 3.28) representing the taxa, a *root* of degree 1 (at the bottom of Figure 3.28), representing the common ancestor, and *branch nodes* of degree 3 (the mid levels of the Figure 3.28), indicating points of evolutionary divergence. Any *rooted tree* has an intrinsic orientation created by the root node.

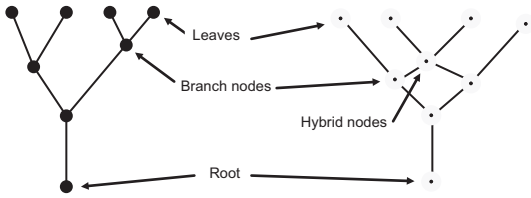


Figure 3.28 Phylogenetic tree (see Figure 3.1*d*) and a phylogenetic net with a hybrid node added.

A **network** is a graph that includes quantitative information, usually representing aspects of a real system: The nodes represent *things* whether they are spatial or not (e.g. species, islands) with weights indicating size, abundance, potential etc. The edges, whether spatial or not, represent *connections* (e.g. interactions like predation, relationships such as similarity, physical links like contiguity, or migration routes) with weights indicating characteristics of the connections such as influence, intensity, rates, capacity, etc.

The edges can also have signs (+ or −, essentially a binary weight) indicating the nature of the relationship portrayed; in a graph's diagram, signs are usually indicated by a solid line for positive and a broken or dashed line for negative; these graphs are called **signed** graphs (Figure 3.1*c*). Edges can also have numerical weights, such as measures of similarity or 'distance' (not necessarily spatial, e.g. taxonomic distance between species).

In most biological applications the terms *network* and *graph* are often used more-or-less interchangeably in a nontechnical way. In phylogenetic studies, the term *network* is sometimes used as in contrast to the graph theoretical *tree*: a phylogenetic tree depicts the division of an evolutionary line going forward through time, but with no rejoining of the evolutionary branches (Figure 3.28); an evolutionary network, unlike a tree, can be reticulate with linkages forming at previously separated branches, for example by hybridization or introgression (Huson & Bryant 2006; Gauthier & Lapointe 2007). In some ways, the netted structure can be interpreted as representing a number of possible binary phylogenetic trees taken together (Gauthier & Lapointe 2007).

As a matter of clarification, while **connectedness** and **connectivity** are used more-or-less interchangeably, **connectance** is a different property. Connectance is the proportion of the positions in the graph where edges may occur that are actually occupied by an edge. For a simple graph of E edges on n nodes, it is $E/[n(n-1)/2]$. Connectance has very little to do with the topological properties of a graph; two graphs with the same numbers of nodes and edges can have very different structural characteristics, including different levels of connectedness.

The **degree of a node** is the number of edges attached to it, and the distribution of degrees observed in a graph can allow us to distinguish among models of graph structure. At one extreme is a **regular graph**, which has nodes all of the same degree, and a **complete graph**, which has all possible edges and in which each node has the maximum degree possible.

The **diameter** of a graph is the maximum distance between any two nodes in the graph, where the distance is defined as the number of edges in the shortest path between the two. For a given number of nodes, the smaller the diameter, the easier it is to form or follow paths between any two nodes in the graph. In most graph structures, having more edges per node will decrease the diameter of the graphs. Once again, this characteristic will be closely related to the ease of transport or communication across the entire system portrayed by the graph. This applies to both spatial and aspatial graphs.

Centrality of a node is a measure of the importance of its position in the shortest paths between other pairs of nodes. One example of the use of this measure is in Fisher (*Martes pennanti*) populations in Ontario, for which node centrality was correlated with the proportion of immigrants, suggesting that the central nodes were high-quality habitat sources of emigrants to other sites. (Garroway *et al.* 2008). A node of high graph theoretical centrality does not need to be near the physical centre of a spatial graph (see example in Bieler *et al.* 2009).

The **betweenness** in a spatial graph is a measure of the frequency with which a node occurs in the shortest paths between other pairs of nodes in the graph. This measure can be used to determine patterns of

sympatry using a reduced Delaunay triangulation (very long edges are removed as not pertinent) and iterative node removal (Dos Santos *et al.* 2008).

A **connected graph** has a path between any two nodes; otherwise it is disconnected. There are several terms related to connectedness characteristic of a graph. **Edge connectedness** is the smallest number of edges that can be removed that will disconnect the graph. **Node connectedness** is the smallest number of nodes that can be removed that will disconnect the graph. A graph is said to be highly connected or to have high **connectivity** if either the edge connectedness or node connectedness is large. If a graph is not connected, it consists of a number of maximally connected subgraphs called **component***s*.