

Contiguous units analysis

Introduction

We will begin this chapter with a discussion of spatial analysis in one dimension; the basic form of data is a series of values, x_1 to x_m , representing the density or stem counts of a particular species in a transect of n contiguous sampling units, usually quadrats which represent the most common sampling method in plant ecology. The density will often be measured in per cent or may be expressed as a proportion, running from 0 to 1. In presence : absence form, the data take only the values 0 for absence and 1 for presence. The spatial pattern of a single species is often studied using one of several methods that examine the effects of distance or block size on a calculated variance, with low variance indicating similarity and high variance indicating dissimilarity. The distances that produce high and low values of the variance are closely related to the scales of pattern and the sizes of patches and gaps in the original data.

The methods we describe in the sections that follow can also apply to density data, presence : absence, or to 'count' data such as the number of stems found in each quadrat; these include methods for studying the spatial relationship of pairs of species and for assessing multispecies pattern. Having begun with a single species in one dimension, we also look at methods for two or more spatial dimensions. To compare with variance-based approaches, we describe analysis techniques based on wave forms such as spectral analysis and wavelet analysis. It turns out these approaches are actually similar to the quadrat variance methods, and we conclude with a simplified summary comparison of the methods presented in Chapters 4 and 5.

5.1 Quadrat variance methods

These methods have their origins in the examination of counts of events, such as plant stems, in sampling units, like quadrats, for evidence of nonrandomness in their distribution among units. The underlying argument is that if the events occur randomly, then the counts per unit should follow a Poisson distribution, and that distribution has the convenient property that the variance and the mean are equal. This led to the recommendation that researchers should calculate the variance-to-mean ratio of such count data as an index of dispersion:

$$I_D = s_c^2 / m_c, \quad (5.1)$$

where I_D is the index of dispersion, s_c^2 is the sample variance of the counts per sample unit, and m_c is the sample mean of those counts. The suggested interpretation was that values around 1 indicate randomness, values considerably greater than 1 indicate clumping, and values less than 1 indicate a regular pattern or overdispersion. Of course, the logic of this recommendation is imperfect (see Dale 1999, figure 7.16), but the concept of the variance among quadrat counts as the foundation of pattern analysis became well-established.

If the simple variance-to-mean ratio of counts in sampling units is an imperfect index of spatial pattern, how can it be improved? A number of different suggestions have been made to improve response and interpretability, but we will mention two developments that lead to the methods we will describe in detail. The first improvement is to use sample units of different sizes, in order to determine how scale (in the sense of

grain) affects how the spatial pattern is perceived. It is well-known, for example, that point events such as the positions of tree stems can be overdispersed at very small scales (inter-stem competition or self-thinning of the closest neighbours), clumped at intermediate scales (clonal clusters or patches of seed deposition), and then overdispersed (spacing of clumps) at the largest scales. However, there are practical limitations to the range of sizes we might sample, and when we use summaries from scattered quadrats, even if they represent a range of areas, we lose any spatial information that could be derived from the relative positions of those sample units. A further improvement then is to sample in many small contiguous units which can then be combined into larger samples, as blocks, for which indices of spatial pattern could be calculated for a range of block sizes corresponding to different scales of investigation.

Two 'blocked quadrat' methods of pattern analysis were published by Hill (1973) in which the contiguous quadrats from a string or transect are combined into blocks in a range of sizes. Both are based on the calculation of a variance between or among the blocks, rather than a variance-to-mean ratio within blocks. Both use a continuous range of block sizes and an average over all possible starting positions for the blocking. The first of these is 'two-term local quadrat variance', TTLQV (Hill 1973), which can be thought of as using a two-part window, each part containing b units, as a template for its calculations. The block size affects both the length of the window (the number of quadrats included within each part) and the distance between the two parts of the template.

The variance calculated in TTLQV is

$$V_2(b) = \frac{\sum_{i=1}^{n+1-2b} \left(\sum_{j=i}^{i+b-1} x_j - \sum_{j=i+b}^{i+2b-1} x_j \right)^2}{2b(n+1-2b)}. \quad (5.2)$$

The two sums within the brackets are the totals for each block of the two-part window, and the outer sum creates the average over all positions of the window. This variance is calculated for a range of block sizes and, when plotted, peaks in the variance are interpreted as

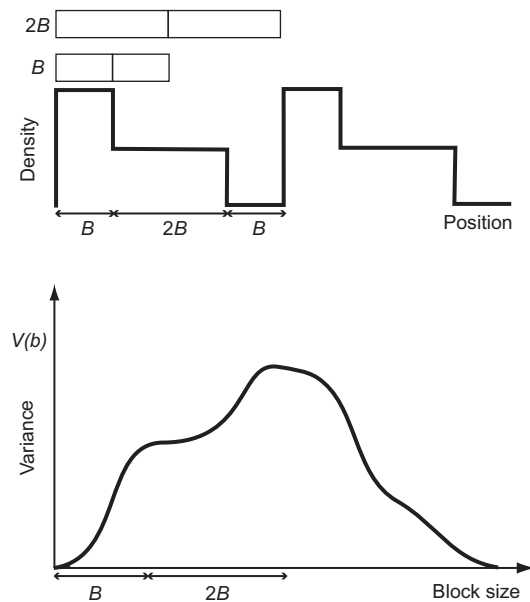


Figure 5.1 Top: the test data are created by adding together two square wave patterns, with scales of B and $2B$. Bottom: the variance (TTLQV) is plotted as a function of block size; peaks (at scales B and $2B$) are interpreted as corresponding to the scales of pattern in the data.

being indicative of scales of pattern in the data (Hill 1973; Dale 1999), as illustrated in Figure 5.1. This method is most often used with density data or estimated cover, but it can be used for presence : absence data, or for counts such as tree stems as in the dispersion index methods on which it is ultimately based.

An alternative is to stay with single quadrats rather than blocks for analysis, and to have a two-part template with the window of each half containing only a single sample unit. This method is known as 'paired quadrat variance', PQV (Ludwig & Goodall 1978), and only the spacing between the halves of the template change, not their size:

$$V_p(d) = \frac{\sum_{i=1}^{n-d} (x_i - x_{i+d})^2}{2(n-d)}. \quad (5.3)$$

The terms within the bracket are now the values for single quadrats and the outer sum creates the average over all possible positions of the window for a given spacing d . As for TTLQV, peaks in the plot of V_p as a

function of d are interpreted as representing scales of pattern in the data (cf. Ludwig & Reynolds 1988). TTLQV and PQV are similar in that they both use a two-part window, but in PQV, the size of the window does not change, only the spacing; in TTLQV, both the size and distance between the centres of the two parts change.

Both TTLQV and PQV can be extended to three-part forms, 'three-term local quadrat variance', 3TLQV (Hill 1973), and 'triplet quadrat variance', tQV (Dale 1999).

The equation for 3TLQV is:

$$V_3(b) = \sum_{i=1}^{n+1-3b} \left(\sum_{j=i}^{i+b-1} x_j - 2 \sum_{j=i+b}^{i+2b-1} x_j + \sum_{j=i+2b}^{i+3b-1} x_j \right)^2 / 8b(n+1-3b). \quad (5.4)$$

For tQV, it is:

$$V_t(d) = \sum_{i=1}^{n-2d} (x_i - 2x_{i+d} + x_{i+2d})^2 / 4(n-2d). \quad (5.5)$$

Again, the contents of the main brackets represent the actions of the window template, and the outer sums generate the averages over all possible positions. In both these methods, peaks in the variance indicate scales of pattern in the data, as in the previous two (cf. the lower part of Figure 5.1). The two-part window methods can filter out the addition of a constant and the three-part window methods can filter out a linear trend. Therefore, 3TLQV and tQV are less sensitive to trends in the data (Dale 1999). Guo & Kelly (2004) studied a full range of these related methods, and concluded that while the three-term version did not always out-perform the two-term methods, the three-term approach did indeed make the results less sensitive to trends in the data.

The concept of lacunarity was introduced in Section 4.4.2 in the discussion of one-dimensional point pattern analysis. The gliding box algorithm for calculating a measure of lacunarity can be applied to the kind of data we are describing here, whether they are densities, counts, or presence : absence in the quadrats (see Dale 2000). In fact, that approach can be seen as something like a one-part window equivalent of the methods we

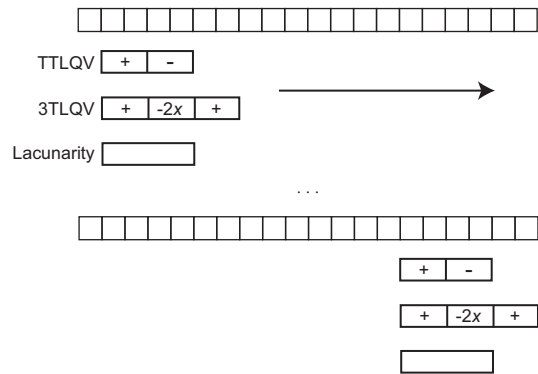


Figure 5.2 Gliding box techniques for one-dimensional data: TTLQV, 3TLQV and lacunarity (lacunarity also shown in Figure 4.15).

have just described (see Figure 5.2, also shown in Figure 4.15). We have already commented, however, that the lacunarity analysis does not give results as precise as those of the quadrat variance approach for patterns of known characteristics (cf. Dale 2000).

Another approach to the analysis of spatial pattern, based on Hill's (1973) quadrat variance methods, is Galiano's (1982) new quadrat variance, which was proposed in both two- and three-term form. In the two-term form it is:

$$V_{N,2} = \sum_{i=1}^{n-2b} |T_2(b, i) - T_2(b, i+1)| / (2b(n-2b)),$$

with

$$T_2(b, i) = \left(\sum_{j=i}^{i+b-1} x_j - \sum_{j=i+b}^{i+2b-1} x_j \right)^2. \quad (5.6)$$

In the three-term form it is:

$$V_{N,3} = \sum_{i=1}^{n-3b} |T_3(b, i) - T_3(b, i+1)| / (8b(n-3b)),$$

with

$$T_3(b, i) = \left(\sum_{j=i}^{i+b-1} x_j - 2 \sum_{j=i+b}^{i+2b-1} x_j + \sum_{j=i+2b}^{i+3b-1} x_j \right)^2. \quad (5.7)$$

Dale (1999) provided a discussion of the properties of these two statistics, but recommended against the use of the three-term version until its properties are better understood. Guo & Kelly (2004) did not find that the three-term version was invariably better than the two-term, although it was found to be less sensitive to trends. In two-phase pattern of patches and gaps, the two-term version gives a variance peak at the average size of the locally smaller phase, and Guo & Kelly (2004) suggested that the three-term version detects the same characteristic. As with many other methods, careful interpretation is necessary if there is non-stationarity in the data.

5.2 Significance tests for quadrat variance methods

In most data analysis, the principal approach to interpretation is a test of statistical significance, whether in evaluating a single data set or in comparing data sets. In spatial pattern analysis, the primary purpose is usually exploratory, with a focus on hypothesis development, rather than a search for evaluating and rejecting or accepting these hypotheses. Nevertheless, statistical significance is definitely of interest (and would be very useful in many cases!), but is made difficult by several forms of lack of independence in the data and in the ways the data are used by the methods we are describing here. First, the values found in adjacent quadrats will tend to be more similar than those at some distance from each other. That fact is part of the underlying logic of spatial pattern analysis, but it also represents spatial autocorrelation in the data that can complicate the evaluation of statistical tests. Positive spatial autocorrelation, in general, tends to make statistical tests too liberal: they give more apparently significant results than the data actually justify (see Chapter 8).

The second source of lack of independence is the fact that each piece of data is used more than once in the analysis, both in calculating the variance at a single block size and in calculating the variances at different block sizes. The overall result of the lack of independence is that it is difficult to provide statistical tests to evaluate the results of a single analysis or to compare results.

Randomization procedures can be used to help evaluate the significance of detected pattern in data, but because that kind of assessment requires both the data and a considerable number of reanalyses, it is not always feasible. Second, complete randomization destroys the spatial structure of the data, so that we can test only the null hypothesis that there is no pattern at all (often not very interesting!) In some cases, restricted randomizations in which the spatial structure is preserved are possible, as we describe in more detail in Chapters 1 and 8. In this section, however, we will show how we might approach evaluating the results of quadrat variance analysis, based on the characteristics of the analysis of random data in which there is no pattern.

Because of its relative mathematical simplicity, we will begin with PQV. Start with a string of data, x_i , $i = 1, 2, 3, \dots, n$, assumed to be independent and random, following some particular probability distribution, say the uniform probability distribution. PQV averages a number of terms of the form $(x_i - x_{i+d})^2$ for a range of distances, d . For independent and random data, $(x_i - x_{i+d})^2$ will have the same mean and variance for any value of d and, therefore, in PQV, $V_p(d)$ will have a constant mean and variance.

For the uniform probability distribution, we know that:

$$E[(x_i - x_j)^2] = 1/6, \text{ and } E[(x_i - x_j)^4] = 1/15, \text{ so that}$$

$$\text{Var}[(x_i - x_j)^2] = 1/15 - (1/6)^2 = 7/180. \quad (5.8)$$

Because $V_p(d)$ is half the average of $n - d$ similar terms, we might be tempted to suggest that:

- (1) $E(V_p(d)) = 1/12 = \mu$ and $\text{Var}(V_p(d)) = 7/720(n - d) = \sigma^2$; and
- (2) being an average of independent terms, $V_p(d)$ should approach the normal distribution, with mean and variance as given.

Then, a significance test of the variance for any distance could be based simply on the interval $\mu \pm 1.645\sigma$. Unfortunately, points (1) and (2) are not quite true.

Point (1) is wrong because the terms contributing to $V_p(d)$ are not completely independent. For example, for $d = 3$, the calculation includes both

$(x_1 - x_4)^2$ and $(x_4 - x_7)^2$, which share the variate x_4 , and therefore have a non-zero covariance. For uniform distribution, those non-zero covariances increase the variance from $7/720(n - d)$ to approximately $9/720(n - d)$.

Point (2), also, is not quite right, because of the reuse of the same data. Because calculations of $V_p(d - 1)$ and $V_p(d)$ use the same variates, they are not independent and with some effort (quite a lot actually), we can show that

$$\text{Cov}[V_p(d-1), V_p(d)] = \frac{(2n-3d+1)(1/90)}{4(n-d)(n-d+1)} \approx 1/180n. \quad (5.9)$$

This value may seem small, and it is; but not only will the variances of successive distances have non-zero covariance, such as $d = 4$ and $d = 5$, but so will $d = 4$ and $d = 6$, $d = 4$ and $d = 7$, and so on. This lack of independence is an important obstacle to the development of statistical tests.

All that being said, however, we can use our understanding to provide some guidance to the interpretation of PQV plots, specifically to do with the position of the first variance peak. Because $E(V_p(d))$ is constant and the distribution of $V_p(d)$ is independent of d , $V_p(d_j) > V_p(d_k)$ with probability of about 0.5 for all values of d_j and d_k . For $d = 1$ to be the first peak, all that is required is $V_p(1) > V_p(2)$ and, with random data, that occurs with a probability of approximately 0.5. For $d = 2$ to be the first peak, we need $V_p(2) > V_p(1)$ and $V_p(2) > V_p(3)$, which occurs with a probability of $0.5 \times 0.5 = 0.25$. In general, then, the first peak for random data occurs at distance d , with probability $(0.5)^d$. The probabilities are not exact because of non-zero covariances, but this provides a direct mathematical explanation of the observation by Campbell *et al.* (1998) that, for random data, variance peaks occur at block sizes 1, 2 and 3 with frequencies of approximately 50%, 25% and 12% (their figure 1). A second comment is that there is an important distinction to be made between testing whether the data as a set are nonrandom and have significant spatial pattern in them and providing significance tests for the variance at a particular block size or spacing.

If we turn our attention from PQV to TTLQV, it is obvious that the problems for any derivation will be greater. The blocking of data introduces more dependence and more covariance to the calculations for any single block size and more dependence and more covariance to the calculations at successive block sizes. It could be done, but is it worth the effort?

NO!

It is not worth the effort for TTLQV; and it is not worth the effort and ingenuity invested for PQV either. The reason is that the values derived for the mean and variance depend strongly on the underlying distribution of the data. The real problem for deriving statistical tests of the significance of results from pattern analysis is that we never know for certain the true underlying distribution or whether the distribution is stationary for the length of the transect. For that reason alone, notwithstanding analytical complexities, these methods will remain techniques for data exploration, not for statistical testing unless some meaningful restricted randomization approach is available... and perhaps not even then.

5.3 Adaptations for two or more species

An obvious extension of single species analysis is to examine scales of association for pairs of species, by looking at the effect of scale on covariance (Greig-Smith 1961, 1983; Dale & Blundon 1991). Adapting the quadrat variance methods to evaluate covariance proceeds on the basis that the covariance of two variables can be derived from their individual variances and the variance of their sum (cf. Dale 1999):

$$\text{Cov}(A, B) = [\text{Var}(A + B) - \text{Var}(A) - \text{Var}(B)]/2. \quad (5.10)$$

This formula can be applied to any of the methods just described for variances to produce the corresponding covariances TTLQC, 3TLQC, PQC, tQC, and so on. For example, Figure 5.3 shows the application of 3TLQC to presence : absence data from a study of sedge meadows on Ellesmere Island (Young *et al.* 1999).

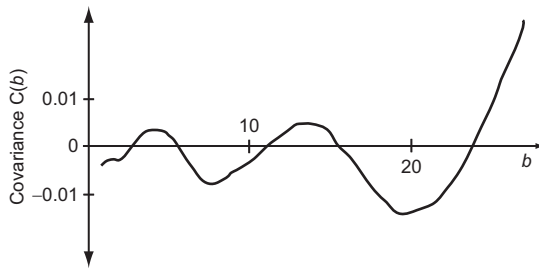


Figure 5.3 Covariance analysis (3TLQC) for the Ellesmere Island sedge meadow data, where b is block size. The association of *Carex aquatilis* and *Eriophorum triste* cycles between positive and negative.

The relationship between *Eriophorum triste* and *Carex aquatilis* changes with block size between negative and positive. The negative association at small scales can be attributed to the divergent ecological ‘preferences’ of the two species, one more mesic and one semi-aquatic, coexisting in a wet hummock–hollow habitat, but not in close proximity.

In many cases, it will be of interest to evaluate the spatial pattern of a whole community, and the analysis of all possible pairs of species (and there are many such pairs!) may be very difficult to interpret. Therefore, researchers have proposed several methods to examine multispecies pattern, evaluating the spatial pattern of all species together.

The concept of multispecies pattern in vegetation may have arisen because of a perception that a plant community is like a mosaic of distinguishable phases or of vegetation types, defined by the combinations of species densities or occurrence. Each phase need not be homogeneous, but there must be greater similarity within a phase than among phases, even if the boundaries between phases are not sharp. Local similarity, however, is not sufficient to give true multiple species pattern with a detectable scale; it requires the repetition of similar species combinations (see Chapter 10). Different phases of a mosaic can have different scales of pattern, and so the scale of multiple species pattern for the whole community should be defined as half the distance that maximizes the probability of finding the most similar combination of species’ densities (Dale & Zbigniewicz 1995).

The recommended method for the evaluation of multispecies pattern is the one introduced by Noy-Meir & Anderson (1971) and modified by Ver Hoef & Glenn-Lewin (1989) and by Dale & Zbigniewicz (1995). The method is based on a combination of the quadrat (co)variance calculations described above and the ordination method of principal components analysis (PCA), and it is usually called multiscale ordination (MSO; Ver Hoef & Glenn-Lewin 1989; Wagner 2004).

For an analysis of k species, we need the $k \times k$ variance–covariance matrix for each block size, b , from 1 to some maximum, B : call the matrices $C(1)$, $C(2)$, \dots , $C(B)$. The variances and covariances can be from any of the methods described above, and Ver Hoef & Glenn-Lewin (1989) used TTLQV and its covariance TTLQC. We recommend 3TLQV and 3TLQC because they are less affected by trends in the data, as previously suggested. The matrices, $C(b)$, for all block sizes 1 to B , are summed and that sum matrix is eigen-analysed as in principal components analysis. Eigenanalysis creates linear combinations of the original variables, the species densities, x_1 to x_k , that are mutually orthogonal and the linear combinations, y_1 to y_k in order, must explain as much of the total variance in the data as they can. For each y_i , its eigenvalue, usually λ_i , is the proportion of the total variance that it explains and its eigenvector is the vector of the weights in the linear combination of the x s producing y_i .

In this method, the largest eigenvalues are each partitioned into the amounts of variance contributed by each block size, using the weights in the eigenvectors (see Ver Hoef & Glenn-Lewin 1989 or Dale & Zbigniewicz 1995), and these are plotted as a function of scale. Peaks or plateaux in each of these variance plots are interpreted as indicating scales of pattern in the vegetation (Ver Hoef & Glenn-Lewin 1989). Because larger block sizes tend to dominate the analysis, the covariance matrices should be weighted prior to summing and Dale & Zbigniewicz (1995) suggested weighting of the variance–covariance matrices calculated for each block size, b , by the factor $6b^2/(b + 2)$. Dale (1999) provided a detailed example of this analysis, using artificial data, which illustrates the ability of the method to recover the important properties of the data.

One feature of the multispecies pattern that may be of interest is the relative strength of the contributions of the various species to the pattern. If one species overly dominates an eigenvector, the pattern detected is not truly multispecies. A measure of species' contributions can be created, based on the fact that each new variable produced by the eigenanalysis is a linear combination of the original species densities, with weights for that linear combination given by the eigenvector:

$$y_i = \sum_{j=1}^k u_{ij}x_j, \text{ with } \sum_{j=1}^k u_{ij}^2 = 1. \quad (5.11)$$

We can propose a measure of how evenly the species contribute to pattern using the variance of the absolute values of the weights, u_{ij} . If all species have equal weights, the variance will be 0; and if one weight is 1.0 and the rest are 0, then the variance is $(k-1)/k^2$. Let C be the coefficient of variation, the square root of the variance divided by the mean. Its maximum value is the square root of k , and a measure of the evenness of the weights in the i th eigenvector is therefore

$$E_i = 1 - C_i/\sqrt{k}. \quad (5.12)$$

The application of this approach may be clarified by an example. We established six 50 m transects at a site near Fort Assiniboine, Alberta, which was dominated by Jack Pine and Aspen, and we sampled the understorey in each transect with 200 contiguous 25×25 cm quadrats. The species list is typical of the boreal forest including vascular plants, such as *Linnaea borealis*, *Maianthemum canadense*, and *Aralia nudicaulis*, species of *Vaccinium*, and feather mosses, such as *Ptilium crista-castrensis*, *Pleurozium schreberi* and *Hylocomium splendens*. We analysed the data with multiscale ordination based on 3TLQV/3TLQC, as described above. We have chosen one transect, the southern east-west transect, as an example to discuss. Based on the 12 most common species, the first three axes explained 22.6%, 16.8% and 15.2% of the variance (54.6% in total, a high proportion for only three axes). The evenness of the eigenvector weights was also high, 0.868, 0.801 and 0.819, indicating true multispecies pattern, not domination by any single species. Figure 5.4a shows the data for this transect and

Figure 5.4b shows the partitioned variances as a function of block size for the first three eigenvalues. The three axes provide clear evidence of pattern in the range of 17–27 units (8.5–13.5 m). This is in agreement with an informal evaluation of the data, which seem to have a pattern of about 10 m.

5.4 Two or more dimensions

The basic concepts of the quadrat variance methods, described for a single dimension, can also be extended to two-dimensional data collected on a grid of contiguous units. Again, any of the methods can be adapted for this purpose.

For example, in one dimension, tQV had the formula:

$$V_t(d) = \frac{\sum_{i=1}^{n-2d} (x_i - 2x_{i+d} + x_{i+2d})^2}{4(n-2d)}. \quad (5.13)$$

Expanded to two dimensions, it could become:

$$V_5(d) = \frac{\sum_{i=1}^{n-2d} \sum_{j=1}^{m-2d} (x_{i,j} + x_{i+d,j} - 4x_{i+d,j+d} + x_{i+2d,j} + x_{i,j+2d})^2}{20(n-2d)(m-2d)}. \quad (5.14)$$

The other extensions to two dimensions are similar and simple in concept, but the equivalents of TTLQV and 3TLQV require somewhat long and complicated equations (see Dale 1999). They are understood intuitively from the templates for their calculation. Figure 5.5a shows the template for $V_5(d)$ and compares it with the template for $V_9(b)$, which is the two-dimensional equivalent of 3TLQV, Figure 5.5b. Expansion of the concept of $V_5(d)$ to compare any sample unit with all units that are very approximately a distance d away from it, as in Figure 5.5c, comes very close to the estimation of the omni-directional variogram, as described in Chapter 6.

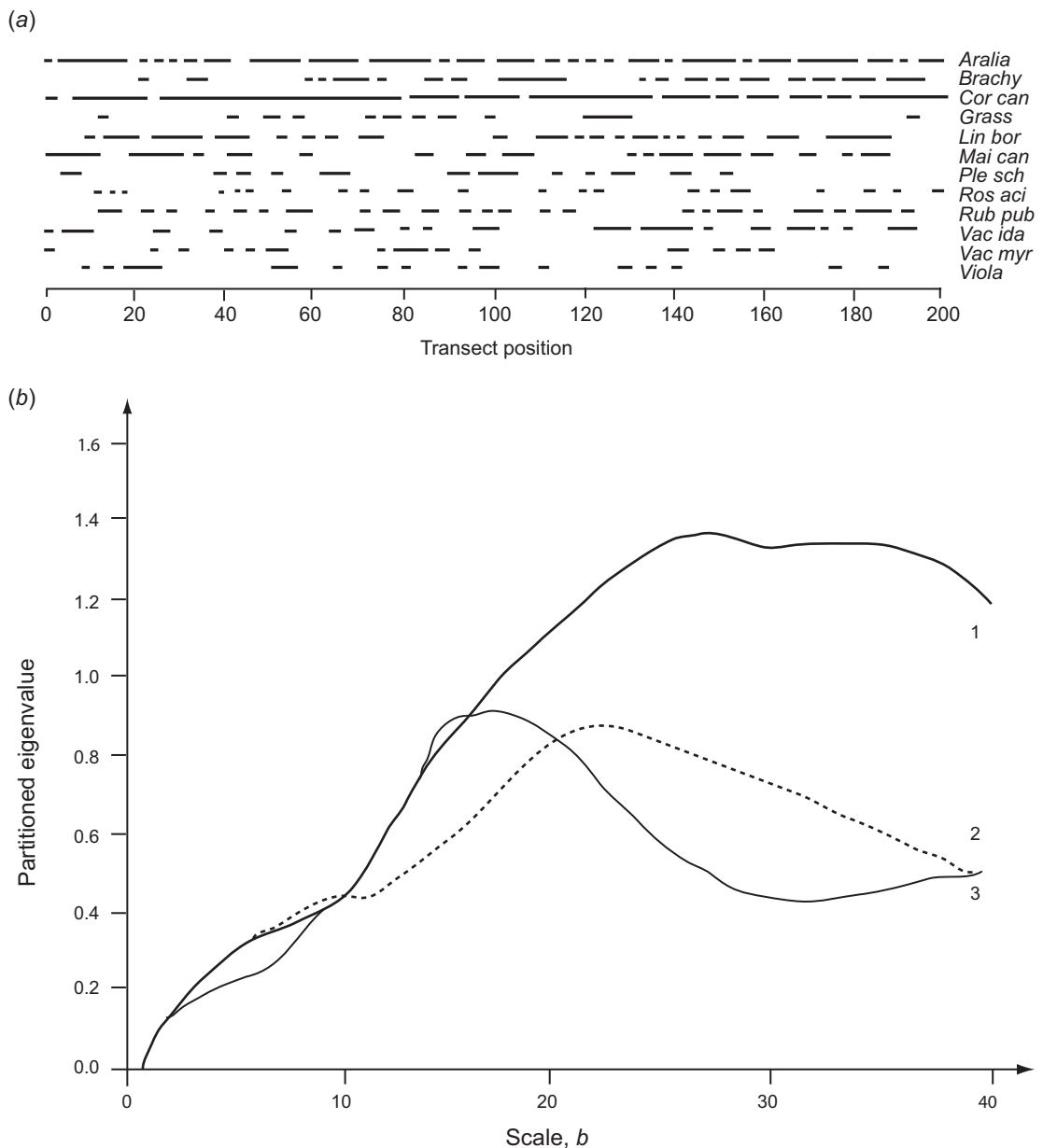
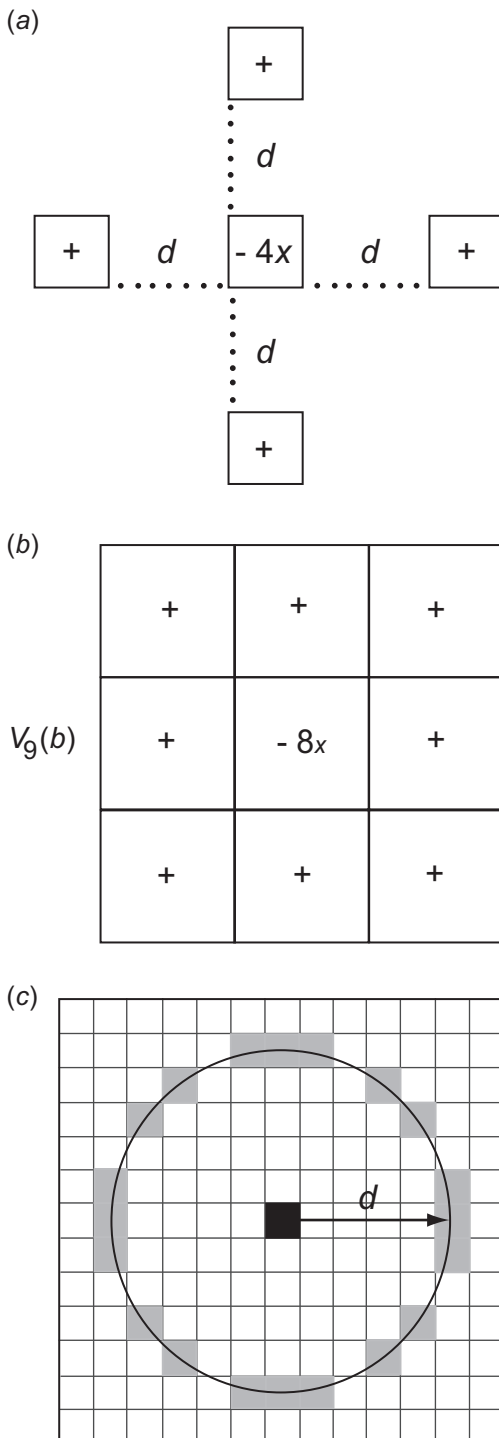


Figure 5.4 (a) Presence : absence of 12 understorey species along a 50 m transect at Fort Assiniboine, Alberta. (b) Partitioned variances as a function of block size for the first three eigenvalues in multiscale ordination of the data in (a). Key to species abbreviations: Aralia, *Aralia nudicaulis* L.; Brachy, *Brachythecium* sp.; Cor can, *Cornus canadensis* L.; Grass, Poaceae; Lin bor, *Linnaea borealis* L.; Mai can, *Maianthemum canadense* Desf.; Ple sch, *Pleurozium schreberi* (Brid) Mitt.; Ros aci, *Rosa acicularis* Lindl.; Rub pub, *Rubus pubescens* Raf.; Vac ida, *Vaccinium vitis-idaea* L.; Vac myr, *Vaccinium myrtilloides* Michx.; Viola, *Viola* sp.



Dale (1995, 1999) described a ‘random paired quadrat frequency’ method for the analysis of two-dimensional mosaics such as communities of crustose saxicolous lichens. As the name suggests, it compares the frequency of particular species combinations in randomly chosen pairs of quadrats, as a function of the x and y displacements, with the expected value based on occurrence in the entire data set. This approach provides an easy assessment of anisotropy; for details see Dale (1995) or (1999).

The next approach we describe is the set of SADIE techniques, from ‘spatial analysis by distance indices’, developed by Perry and co-workers (Perry 1994, 1995, 1996, 1999). The basic concept begins with the counts of events (e.g. insects) in a grid of contiguous sampling units (e.g. a field divided into quadrats). The approach then uses either the total distance that individuals would have to move in order for all to occupy a single quadrat (distance to crowding) or the total distance they would have to move in order to have equal numbers in all quadrats (distance to regularity). The two different versions can be seen as maximizing or minimizing the variance-to-mean ratio of the counts (see Section 5.1, above), but the spatial arrangement of the counts is used to evaluate the distance characteristics of those re-arrangements.

Consider the ‘distance to regularity’ approach in greater detail: let m be the mean of the counts of N events in a grid of n quadrats. Each unit that has a count of more than m events, call it c_i , must lose $c_i - m$ events, and each unit that has a count of fewer than m events, call it c_j , must gain $m - c_j$ events. If there are p units with counts greater than m , and q units with counts less than m , there are pq pairs of units between which a movement of numbers might take place from a source unit, i , to a sink unit, j , with magnitude v_{ij} (which is not necessarily an integer). Although there are many different ways in which the flow of numbers

Figure 5.5 (a) Template used for the calculation of $V_5(d)$. (b) Template used for the calculation of $V_9(b)$. (c) An illustration of extending the $V_5(d)$ concept to compare a sample unit to all others at a distance of approximately d .

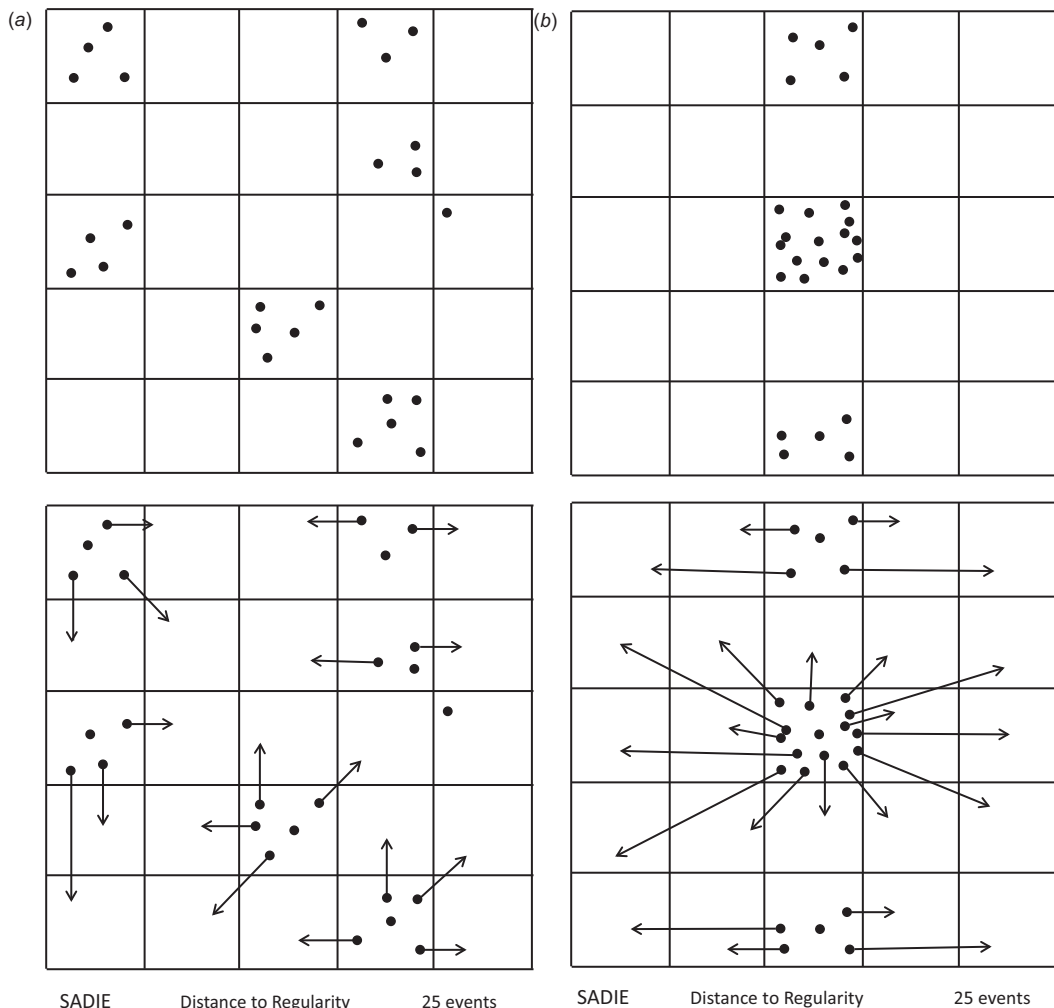


Figure 5.6 SADIE algorithm for distance to regularity. Observed clusters of points are distributed out to maximum regularity with one event per grid cell (here 25 cells). (a) Initially several small clumps (25 events). (b) Initially few larger clumps (25 events).

from sources to sinks could occur, the SADIE algorithm finds the combination that has the smallest total flow distance. That is, where d_{ij} is the distance between the two units, the algorithm minimizes the value of the sum, D :

$$D = \sum_{i=1}^p \sum_{j=1}^q v_{ij} d_{ij}. \quad (5.15)$$

Figure 5.6 illustrates this flow from the observed pattern of point events to maximum regularity with one event per grid cell. The observed value of D from data can be tested for statistical significance using a randomization procedure.

There are several elaborations and modifications of the basic method available, as described in the literature. For example, a spatially explicit result can be obtained by either creating a diagram of the flow

of numbers or by colouring the hot spots of high density one colour (red) and the cold spots of low density another (blue), producing 'red-blue' plots of the data (see Perry 1999). The method can also be modified to deal with point pattern data, rather than the quadrat count data described here. In either version, the full action of the SADIE algorithm can be visualized as a displacement graph (in the sense of a spatial graph as described in Chapter 3) with nodes at the original locations of individuals or the centres of source quadrats, with directed edges joining those original nodes to nodes of the new locations for the events being counted as sinks for the flow dictated by the algorithm (see Figure 5.6).

Another method for the analysis of two-dimensional grid data is the technique of lacunarity analysis which was introduced in Section 4.4.2 in the context of a one-dimensional point pattern, and was then alluded to in Section 5.1 in the discussion of one-dimensional contiguous unit arrays. The technique is most popular for use with two-dimensional arrays of contiguous units, the pixels of satellite images, in particular. For example, there have been several recent publications using the 'gliding box' algorithm on air photographs or Landsat images to calculate measures of lacunarity for the study of fragmentation in tropical landscapes or similar applications (Peralta & Mather 2000; Wu *et al.* 2000; Weishampel *et al.* 2001; Mahli & Román-Cuesta 2008; Dong 2009).

The extension of these methods to three dimensions is certainly possible, although we have not found many examples in the ecological literature. The units in a three-dimensional sample array are usually referred to as 'voxels', to parallel the term 'pixel'. Voxels and voxel analysis are used extensively in the analysis of medical imagery, and also in computer graphics and the development of computer games (consider the term 'doxel' for a dynamic voxel, adding the dimension of time!). Applications of voxels are still rare in ecological studies, but Fukushima *et al.* (1998) counted the leaves of trees in $972\ 1.8 \times 1.8 \times 0.9$ m (vertical) cells in order to test methods of estimating foliage profiles. They did not apply the kinds of methods we have described

here; they looked at foliage height diversity. A closer application is due to Morsdorf *et al.* (2004) who deduced the geometric characteristics of a single tree in the boreal forest from LiDAR data, in order to determine vegetation structure for fire risk assessment and fire behaviour modelling. Wang *et al.* (2008) used LiDAR point clouds as the basis for modelling the canopy layer of a forest or the crowns of individual forest trees. In a somewhat similar approach, Bienert *et al.* (2010) used an analysis of voxel attributes derived from terrestrial laser scans to determine tree density and spatial distance patterns as the basis for meteorological models related to gas exchange.

As we noted above, however, in the discussion of three-dimensional point pattern analysis, as these approaches become better known, ecologists will quickly see their usefulness and take them up more frequently for application in ecological studies, especially when the data for this sort of analysis becomes more and more readily available.

5.5 Spectral analysis and related techniques

Spectral analysis is a technique that detects repeating pattern in spatial density data by fitting sine and cosine functions to the data, thus determining which frequencies or wavelengths best fit the data (Ripley 1978). The data to which this analysis is usually applied are quantitative measures in continuous or evenly spaced series. This approach is most suitable for 'rich' data sets, with large numbers of observations, and in situations in which the assumption of stationarity is justified. One well-known technique for spectral analysis is the Fourier transform, which decomposes the 'signal' (the variability of the data) into sine and cosine waves of various frequencies and starting positions (Figure 5.7a; see Legendre & Legendre 1998). Spectral analysis has been applied to two-dimensional ecological data by Renshaw & Ford (1984) and although originally developed for the analysis of continuous signals, such as time series, it can also be applied to point pattern data (see Muggleston & Renshaw 1996).

For a transect of n values, x_i , the values are expressed as a weighted sum:

$$x_i = \bar{x} + \sum_{p=1}^{n/2-1} c_p \cos(2\pi i p/n) + s_p \sin(2\pi i p/n), \quad (5.16)$$

where

$$c_p = \frac{2}{n} \sum_{i=1}^n x_i \cos(2\pi i p/n) \text{ and} \\ s_p = \frac{2}{n} \sum_{i=1}^n x_i \sin(2\pi i p/n). \quad (5.17)$$

A closely related technique is the Walsh transform which decomposes the signal into square waves instead of sine waves, of various frequencies and positions (Figure 5.7*b*; see Ripley 1978).

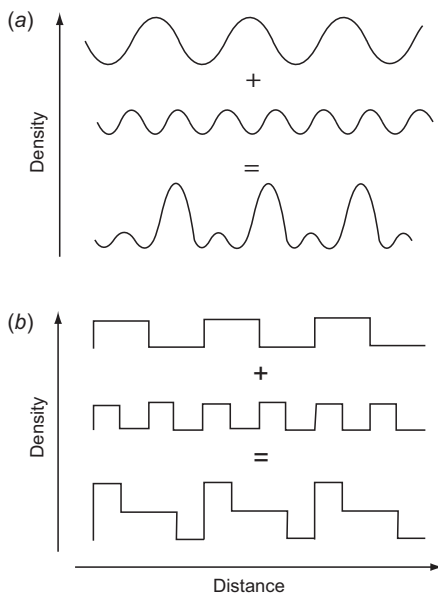


Figure 5.7 Wavelet template shape. (a) The concept of Fourier transform: the combined pattern of sine and cosine waves is resolved back into its original components. (b) A Walsh transform resolves a combination of different frequencies of square waves into its components.

5.6 Wavelets

Wavelet analysis is closely related to spectral analysis, but using a finite wavelet as the template for localized comparison rather than sine or cosine functions applied over the entire data sequence. This method provides both local evaluation of spatial structure and a global measure of fit using the wavelet variance, which gives variance peaks at scales of pattern in the data like those in the quadrat variance methods described above. The analysis evaluates how well the wavelet template, at different sizes, matches the data over a range of positions. The data used for this approach are typically quantitative in a continuous or evenly spaced series. Where g is the wavelet function, its transform, T , is a function of the wavelet size and position:

$$T(b, u_i) = \frac{1}{b} \sum_{j=1}^n y(u_j) g[(u_j - u_i)/b], \quad (5.18)$$

where b is the wavelet's relative width (Figure 5.8); $y(u_j)$ is the data value at position u_j . The transform is essentially calculating the inner product of $y(u)$ with the wavelet function localized in size and position (Daubechies 1993). $T(b, u_i)$ takes large positive values when the match of the wavelet function and the data centred at u_i is very good and large negative values when the match is very poor (see Dale 1999, figure 9.8, or Dale *et al.* 2002, figure 6). As a technical note, the wavelet transform given in Eq. (5.18) is the discrete

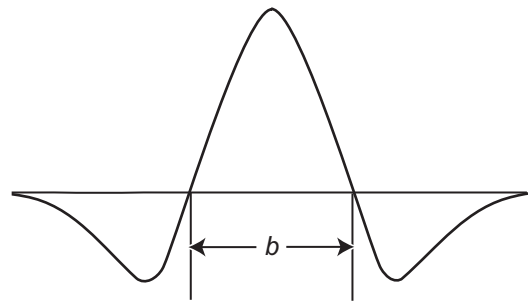


Figure 5.8 Mexican hat wavelet with relative width b .

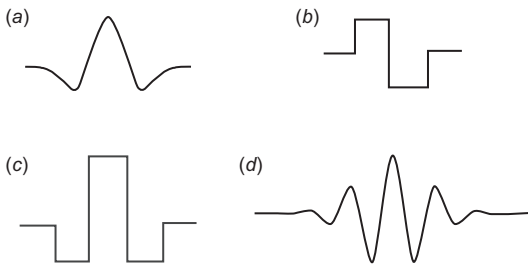


Figure 5.9 Wavelet templates of different shapes: (a) Mexican hat, (b) Haar, (c) French top hat, (d) Morlet.

form, using summation, whereas the continuous wavelet transform uses integration.

Wavelet analysis can employ many different functions as the template, but the 'Mexican hat' template (Figure 5.8) is one of the most common, e.g. Bradshaw & Spies (1992). For $b = 1$, its equation is

$$g_M(u) = \frac{2}{3^{0.5}} \pi^{-0.25} (1 - 4u^2) e^{-2u^2}. \quad (5.19)$$

We can define a wavelet variance, based on the transform, as

$$V_W(b) = \sum_{i=1}^n T^2(b, u_i)/n. \quad (5.20)$$

A great variety of shapes are possible, but they must have an integral of zero; three more are shown in Figure 5.9: the Haar, the French top hat (FTH), and the Morlet. The wavelet variance based on the Haar wavelet is equivalent to TTLQV and that based on the French top hat wavelet is equivalent to 3TLQV (Dale & Mah 1998). Both are also related to spectral analysis using the Walsh transform, which decomposes a signal into combinations of square waves of different frequencies (cf. Ripley 1978). The FTH wavelet for one-dimensional data is also very similar in form and concept to the 'boater wavelet' for point patterns in two dimensions that we introduced in the discussion of circumcircle template for analysis in the previous chapter (Section 4.5). This is quite obvious from the template shapes, and indicates also the close relationship between 3TLQV in one dimension, and circumcircle analysis for point patterns in two.

We can also use the wavelet approach to perform a local equivalent of spectral analysis by using a sine wavelet:

$$g_s(u) = \begin{cases} \sin(\pi u), & \text{if } -1 \leq u \leq 1; \\ 0, & \text{otherwise.} \end{cases} \quad (5.21)$$

If the sine wavelet was expanded into more cycles, indefinitely, the resulting very long wavelet would produce something almost identical to Fourier analysis.

Whatever wavelet is used, wavelet variance analysis can be modified to give a wavelet covariance for bivariate data using Eq. (5.10), and thus to multivariate analysis, but we will not describe this feature in detail. Wavelet analysis can also be applied to two-dimensional data such as densities measured in a plane, for example, vegetation cover in an area of grassland (Csillag & Kabos 1996). One wavelet for such an analysis is that created by rotating the Mexican hat wavelet in Figure 5.8 about its centre, which resembles a true three-dimensional sombrero. In this three-dimensional form the 'brim', the negative part of the template, must be narrower so that the integral of the whole remains zero.

Wavelets have many applications in technology, including image compression, and we will briefly discuss these properties of wavelets as they emerge in a description of boundary detection and spatial clustering in two dimensions (Chapter 9).

5.7 Concluding remarks

In considering the material that has been included in this chapter, we can identify three themes.

The first is the 'relatedness' of the various approaches used, both conceptually and mathematically. The basic concept of wavelets, the comparison of the data with some kind or shape of template, recurs and unifies many of the approaches. Not just wavelets, but Ripley's K -function, TTLQV, and many of the other techniques (almost all!) can be considered to follow this approach. Box 5.1 illustrates this fact. Mathematically this same fact can be expressed as using a cross-product approach, see Getis (1991)

Box 5.1 Comparison of spatial analysis methods with rules for spatial graphs

In this chapter (5), we have described a number of methods for the analysis of pattern in transects or grids of sample units like quadrats. In Chapter 4, we reviewed a number of different methods for point pattern analysis, and prior to that (Chapter 3), we described different rules for the creation of edges in a spatial graph. These included rules based on distance thresholds, topological considerations and functional relationships between nodes. All the methods for point pattern analysis rely on, or are closely related to, one of the spatial graph edge rules. The methods for quadrat transects and grids can also be described by spatial graphs, although their graphs are necessarily more uniform and regular in structure. The purpose of this table is to make explicit those relationships.

Point pattern analysis method	Spatial graph edge rule and conditions
<i>Univariate</i>	
1st nearest neighbour distances (1° analysis)	nearest neighbour rule (NN network)
<i>i</i> th nearest neighbours	<i>i</i> th nearest neighbour graph
refined nearest neighbour distance	nearest neighbour and random points
Ripley's <i>K</i> (2° analysis) & Getis map	distance threshold
Condit's Ω	two-threshold distance graph (upper and lower threshold distances)
circumcircle and 3° analysis	Delaunay triangulation
<i>Bivariate (categorical node labels)</i>	
Ripley's <i>K</i> (2° analysis)	distance threshold & node labels
Condit's Ω	two-threshold graph & node labels (upper and lower distance limits)
circumcircle and 3° analysis	Delaunay triangulation & node labels
<i>Multivariate</i>	
Ripley's <i>K</i> & Condit's Ω	threshold distance graphs & node labels
circumcircle and 3° analysis	Delaunay triangulation & node labels
Dixon's segregation index, S_I	neighbour networks (NN, MST...) & labels
Reich's within-group distance, δ	same-label complete subgraphs
<i>Quantitative node labels</i>	
mark correlation	distance thresholds & node labels
<i>SADIE</i>	
distance to uniformity	displacement digraphs: dispersal of events
distance to coalescence	displacement digraphs: convergence of events
<i>Quadrat variance analysis method</i>	
Edge rule and other graph conditions	
PQV, etc.	two nodes at a time: weights +1 and -1 edge length = scale
TTLQV, etc.	2 <i>b</i> nodes at a time: weights <i>b</i> +1s and <i>b</i> -1s edges all unit length
3TLQV	3 <i>b</i> nodes: weights 2 <i>b</i> +1s and <i>b</i> -2s edges all unit length
PQV, TTLQV, etc., in 2D	unit length becomes $\sqrt{2}$ for diagonals
wavelets	as for TTLQV; weights determined by wavelet function

and Dale *et al.* (2002). Some of the methods we describe in this chapter have fallen out of favour, such as TTLQV, and are being replaced by wavelet analysis, now that it is more readily available. We have retained our treatment and discussion of the older techniques, not so much for historical context, but because the concepts and the mechanics of their application are so basic, and provide the conceptual and mathematical basics for the more advanced methods that followed.

Many of the methods can also be explained in the terminology of graph theory as it applies to spatial graphs. Previously, we described the SADIE approach as being closely related to a 'displacement' spatial graph which showed the movement of observed counts from their original locations to the final locations that maximized either crowding or regular spacing in the sample area. In the analysis of transects of contiguous quadrats, each instance of the PQV template is a spatial graph with two nodes, one in the centre of each quadrat of the pair, one with weight +1 and one with weight -1, joined by an edge of length equal to the scale being tested. (This may not be a very interesting graph, we admit!) Similarly, the template for scale b in TTLQV is a spatial graph of $2b$ nodes, the first b with weight +1 and the second b with weight -1, joined in a string by edges all of length 1 (again not very exciting). Other methods in this group can be expressed in this way, including wavelets, for which the weights on the nodes are determined by the wavelet function. Some minor changes are necessary when the analysis moves into two dimensions (grids of sample units) or more, but these are very straightforward. Table 5.1 and Box 5.1 provide summaries of the methods of this and the previous chapter in relation to the spatial graphs that are closely related.

The second theme in this chapter is the concern for using null hypotheses other than randomness for the evaluation of spatial structure. In Chapter 4, the null model for point pattern is CSR, but we noted some examples where the Poisson-Poisson

distribution or a Markov inhibition model seemed more appropriate for comparison. As stated in the discussion of contiguous unit data analysis, the null hypothesis of complete randomness is not really expected to be true and so other hypotheses should be considered. More needs to be said about this theme and more work needs to be done on this topic, especially in areas other than univariate point pattern analysis.

The third theme is the usefulness of Monte Carlo and randomization techniques to solve or circumvent problems with analytical approaches. It is often going to be most useful to test the observed results against something that does not destroy all the spatial structure or assume its absence. Again, more could be said on this and more work followed up, but it is clear that this approach has found its place in this age of fast and easy computation. This comment applies not only to the methods described in this chapter, of course, but is a common thread through the entire book.

In developing recommendations on which methods to use, it probably became clear that for point patterns, the set of methods based on the concept of Ripley's K -function covers a lot of different kinds of data and a range of situations. We showed that set of methods has much to recommend it. On the other hand, this chapter shows that the set of methods using the wavelet approach has flexibility, based on the choice of the wavelet, and a conceptual sophistication that is very appealing. Those methods can also be adapted to deal with a wide range of different data types. The decision on the method to use will depend, of course, on the data and the question being asked, but our recommendation is to use two or more complementary approaches, so as not to miss important features of the data.

By way of a helpful conclusion, Table 5.1 gives a summary of the methods described in the Chapters 3, 4 and 5, emphasizing the similarities and differences of the approaches by focusing on the template or window function used in the calculation of the statistic or measure used.

Table 5.1 Summary of the spatial analysis methods presented in Chapter 5

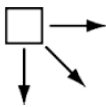








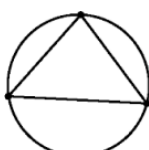
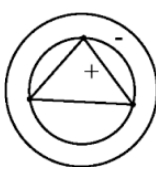
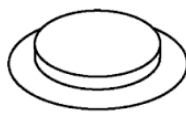
Spatial analysis methods	Template
Contiguous units	
Two-dimensional lacunarity	 Gliding box
Spectral analysis: Fourier	 Sine wave
Walsh transform	 Square wave
Wavelets	
One-dimensional data	 Mexican Hat  Haar  French Top Hat  Morlet  Sine
Two-dimensional data	 Sombrero
Circumcircle	 Simple   Double = boater wavelet

Table 5.1 (cont.)

Spatial analysis methods	Template
Contiguous units	
TTLQV and TTLQC	<div><div><div>b</div><div>$+$</div></div><div><div>b</div><div>$-$</div></div></div>
PQV and PQC	<div><div>$+$</div><div>d</div><div>$-$</div></div>
3TLQV and 3TLCV	<div><div><div>b</div><div>$+$</div></div><div><div>b</div><div>$-2x$</div></div><div><div>b</div><div>$+$</div></div></div>
tQV and tQC	<div><div>$+$</div><div>d</div><div>$-2x$</div><div>d</div><div>$+$</div></div>
Lacunarity	<div><div></div><div>→</div><div>One part window</div></div>
2DtQV = 5QV	<div><div><div>$V_5(d)$</div><div><div>$+$</div><div>d</div><div>$+$</div></div><div><div>$+$</div><div>d</div><div>$-4x$</div><div>d</div><div>$+$</div></div><div><div>$+$</div><div>d</div><div>$+$</div></div></div></div>
2D 3TLQV = 9TLQV	<div><div><div>$V_9(b)$</div><div><div><div>$+$</div><div>$+$</div><div>$+$</div></div><div><div>$+$</div><div>$-8x$</div><div>$+$</div></div><div><div>$+$</div><div>$+$</div><div>$+$</div></div></div></div></div>