Journal of Vegetation Science 6: 575-582, 1995

Spatial pattern analysis in ecology based on Ripley's *K*-function: Introduction and methods of edge correction

Haase, Peter

Department of Pure and Applied Biology, University of Leeds, Leeds LS2 9JT, UK;
Tel. +44 113 233 2816; Fax +44 113 233 2835; E-mail PABPH@LEEDS.AC.UK

Abstract. Spatial pattern analysis based on Ripley's *K*-function is a second-order analysis of point patterns in a two-dimensional space. The method is increasingly used in studies of spatial distribution patterns of plant communities, but the statistical methods involved are sometimes poorly understood or have been modified without evaluating the effects on results. The procedures of field data acquisition, statistical analysis, and the test for the null hypothesis of complete spatial randomness are described and the presentation of results is discussed. Different methods of edge correction were tested on a computer-generated random pattern and a mapped distribution of a Mediterranean shrubland. The inclusion of buffer zones around mapped plots describes the spatial pattern most accurately, but may not warrant the additional labour involved. Three variations of the weighted edge correction yielded comparable results for the distribution patterns tested. The toroidal edge correction may give biased results for non-random patterns. Recommendations for standardisation of the statistical procedures and data presentation are given.

Keywords: Boundary correction; Buffer zone; Edge correction; Monte Carlo simulation; Second-order statistics.

Introduction

Spatial pattern in plant communities is of particular interest to ecologists because it can reveal information on stand history, population dynamics and competition. Simple statistical procedures like the nearest neighbour method (Greig-Smith 1983) have long been employed in forest ecology with the main purpose of estimating stand density and standing biomass from subsamples of tree populations. With the rapidly increasing importance of statistics in biology, more elaborate analytical methods were developed, such as the test of Clark & Evans (1954; Simberloff 1979), Pielou's (1960) index of non-randomness, the contiguous quadrat technique (Greig-Smith 1983), or Morisita's (1959) index. The primary question to be answered is whether an analysed distribution is random or non-random. Non-random patterns may be either clumped (aggregated) or regular (dispersed). Regular patterns do not always conform with ecological theory, e.g. the apparently regular spacing of certain desert shrubs (e.g. Ebert & McMaster 1981; King & Woodell 1984; Cox 1987).

In particular, the development of density functions using the second moment, i.e. the variance of all point-to-point distances, has supplied a powerful analytical tool for the study of distribution patterns. Spatial point pattern analysis based on Ripley's (1976) *K*-function is now widely used in plant ecology; it has been applied to distribution patterns of herbs (Kenkel 1993), desert shrubs (Prentice & Werger 1985; Skarpe 1991) and tropical forest trees (Sterner et al. 1986). The method can also be used to study sedentary animals such as barnacles, or stationary constructions such as ant hills or birds' nests (Ripley 1981). Some researchers have developed their own computer programs (ter Braak 1980; Getis & Franklin 1987; Duncan 1991), others have used or modified programs made available by colleagues (Prentice & Werger 1985; Stewart & Rose 1990; Leemans 1991; Skarpe 1991). However, most authors provide insufficient information regarding mathematical procedures, and some have presented formulas which deviate slightly from the original version or contain errors. The statistical literature referred to is often hard to understand for non-mathematicians. In view of the expected rise in popularity of spatial pattern analysis among ecologists, it is desirable to introduce some standardization of the statistical methods involved in order to allow for a proper evaluation of the results. This should in particular include the effect of different procedures for edge correction proposed in the literature, which have been very little discussed (Ripley 1979).

This paper describes univariate (single-species) spatial pattern analysis based on Ripley's *K*-function, focusing on the test for spatial randomness, and providing detailed information on the statistical background. Methods for edge correction are evaluated in comparative tests, using as examples a computer-generated random pattern and a real distribution pattern of Mediterranean shrubland.

Methods

Acquisition of field data

Sample plots are typically rectangular, because they are more easily marked and recorded in the field than circular plots, but spatial point pattern analysis can be applied to either shape. Plot size depends on the density of the species investigated; often only specimens above a certain size class are considered. In the publications reviewed here, plot sizes range from 25 m^2 (Kenkel 1993) to 0.8 ha (Szwagrzyk & Czerwczak 1993); Getis & Franklin (1987) used aerial photographs to map dominant conifers in 1.44-ha plots. Plots may be placed randomly or semi-randomly but should contain typical examples of the vegetation. Obvious environmental variation should also be avoided, because this could mask underlying patterns resulting from biotic interactions. The location of all plants in the plot with respect to the plot sides is determined accurately. In small plots distances can be measured to at least two contiguous borders, but with increasing plot size, say > 100 m^2, this method is unreliable, unless the plot is subdivided. The alternative is to measure inter-plant distances, followed by conversion to coordinates using trigonometric functions (Rohlf & Archie 1978); this appears to give accurate data. Besides the pair of coordinates it is useful to record for each plant other variables like canopy height, stem and/or crown diameter, etc., because this allows stratification of the data (e.g. Gibson & Menges 1994). The example of a mapped distribution pattern discussed in this paper was derived from a 10 m × 10 m plot (Fig. 1), which was subdivided into four 5 m × 5 m plots for measuring plant coordinates to the nearest 0.1 m.

Statistical background

Spatial point pattern analysis uses all point-to-point (plant-to-plant) distances to describe two-dimensional distribution patterns. Rather than the first-order statistic, i.e. the mean of the distances as in nearest-neighbour methods, the variance of the distances, the second-order statistic, is of interest; hence the method has been called second-order analysis. In spatial point pattern analysis a circle of radius t is centred in each point and the number of neighbours within the circle is counted. For n individual points distributed in an area A, the density ($\lambda = n/A$) gives the mean number of points per unit area. The function $\lambda K(t)$ gives the expected number of further points within radius t of an arbitrary point. If the points are randomly (Poisson) distributed, the expected value of $K(t)$ equals πt^2, i.e. the area of a circle of radius t, and a plot of $\sqrt{K(t)}$ versus t should therefore be linear.

In a first analytical step, the function $K(t)$ is calculated from the data and then tested against the null hypothesis of complete spatial randomness (*CSR* of Diggle 1983), i.e. it is assumed that all points are distributed independently. If the null hypothesis of spatial randomness has to be rejected for the field data, other hypothetical models for particular non-random patterns may be tested (e.g. Diggle et al. 1976; Ripley 1979; Sterner et al. 1986; Kenkel 1993), but these are not discussed in the present paper. Ripley (1976, 1981) gave an approximately unbiased estimator for $K(t)$ as

$$\hat{K}(t) = n^{-2} A \sum_{i \neq j} \sum w_{ij}^{-1} I_t(u_{ij}) \qquad (1)$$

where n is the number of events (plants) in the analysed plot, A is the area of the plot in m^2, I_t is a counter variable, u_{ij} is the distance between events i and j, and w_{ij} is a weighting factor to correct for edge effects. Since distribution patterns may vary depending on the spatial scale of investigation – for instance, tree seedlings may show random or regular patterns at scales of a few m, but a clumped distribution at larger scales – $K(t)$ has to be calculated separately for each distance t. The intervals for t are not prescribed and have to be determined by the investigator depending on the resolution required.

Calculation of the interpoint distance u_{ij}

Since all points are defined by a pair of coordinates x and y, the distance between two points i and j can be calculated by:

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

The calculated interpoint distance is then compared with the current value of t. If $u_{ij} \leq t$, the counter variable $I_t(u_{ij})$ in Eq. (1) is set to 1; otherwise it is set to 0. This procedure is repeated for all point-to-point distances and $I_t(u_{ij})$ is summed for each distance t. This means that each pair of points is considered twice, because although the distance i - j is equal to the distance $j - i$, point

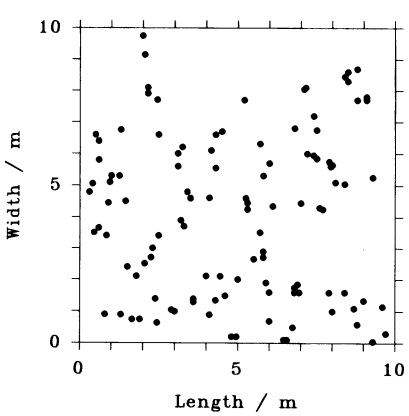

Fig. 1. Mapped distribution pattern of the Mediterranean subshrub *Anthyllis cytisoides* L. (n = 111).

j may have a different spatial neighbourhood in form of a plot boundary, and $I_t(u_{ij})$ then requires a different weighting factor w_{ij} (see below). The term $i \neq j$ in Eq. (1) denotes that self-comparisons are excluded.

The problem of edge effects

If a calculated interpoint distance u_{ij} is greater than the distance between point i and the nearest plot boundary, part of the spatial neighbourhood of point i lies outside the plot and cannot be evaluated without a certain bias. Methods to account for edge effects include the addition of a buffer zone around the plot, considering the plot area as a torus, i.e. a three-dimensional surface without boundaries, and correcting $I_t(u_{ij})$ by the weighting factor w_{ij} [Eq. (1)].

Buffer zones (Fig. 2) were used by Sterner et al. (1986) for rectangular plots and by Szwagrzyk & Czerwczak (1993) for circular plots. Typically, buffer zones must have a width equal to the largest value of t used in the analysis. One shortcoming of this method is that only the plants within the inner plot can be analysed, which means that all plants in an area of up to 4 × the size of the analysed plot have to be recorded. On the other hand, this method is the most realistic one, since only distances to real neighbour plants are used in the calculations. A necessary presumption is that the distribution pattern in the buffer zone is the same as in the inner plot. In the computational procedure, only points in the inner plot are considered as points i, while points j include all points both in the inner plot and the buffer zone.

For the *toroidal edge correction*, the area of the plot is considered to be wrapped around a torus. Points at opposite sides of the plot are now close to each other; the boundary does not exist. In practice, the actual plot is replicated eight times around itself (Ripley 1979, 1981; Upton & Fingleton 1985) (Fig. 3). If t is kept smaller than half the shorter side of the rectangular plot, only the shortest of the possible distances of a given point i to one of the nine replicates of point j is accepted as the interpoint distance u_{ij}. The distances between a given point i and the nine replicates of point j are calculated as follows:

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

$$u_{ij} = \sqrt{(x_i - x_j)^2 + [y_i - (y_j \pm b)]^2}$$

$$u_{ij} = \sqrt{[x_i - (x_j \pm a)]^2 + (y_i - y_j)^2}$$

$$u_{ij} = \sqrt{[x_i - (x_j \pm a)]^2 + [y_i - (y_j \pm b)]^2}$$

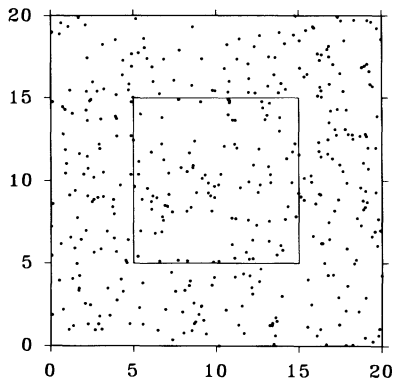

Fig. 2. Simulated random distribution pattern of 400 points in a square plot. The inner plot, containing 100 points, is surrounded by a buffer zone with the same point density.

where a and b are the lengths of the two plot axes. It is not necessary to determine the minimum of the nine values of u_{ij}. $I_t(u_{ij})$ can be conveniently summed for all nine calculations since eight of these will give zero. Both the toroidal edge correction and the buffer zone method do not require the weighting factor w_{ij} in Eq. (1) for the calculation of $K(t)$.

Edge corrections by weighting $I_t(u_{ij})$ were employed by Getis & Franklin (1987) and Andersen (1992). The weighting factor w_{ij} is equal to the proportion of the circumference of the circle with radius u_{ij} and centred on point i and passing through j, that lies within the plot boundaries (Fig. 4). Although this method of edge correction gives approximately unbiased results for values of t of up to $\sqrt{2}/2$ × the side of the (square) plot (Diggle 1983), the formulas given below only apply to values of t of up to 1/2 the shorter side of the rectangular plot. If larger values of t are considered, circles centred in points near the middle of the plot may intersect three boundaries. The boundary correction is based on the assumption that the region surrounding the study plot has a point density and distribution pattern similar to the nearby areas within the boundary (Getis & Franklin 1987). Getis & Franklin's (1987) weighted edge correction for rectangular plots is as follows: if the distance u_{ij} between two points i and j is greater than the distance between i and the nearest boundary (e) (Fig. 4A):

$$w_{ij} = 1 - \cos^{-1}(e_1/u_{ij})/\pi \qquad (3)$$

The inverse cosine function returns the value of α in radians; division by π yields the proportion of this angle with respect to one half of the circumference of the circle (the full circumference is considered in Eqs. 4 and 5), and its subtraction from 1 gives the proportion of the

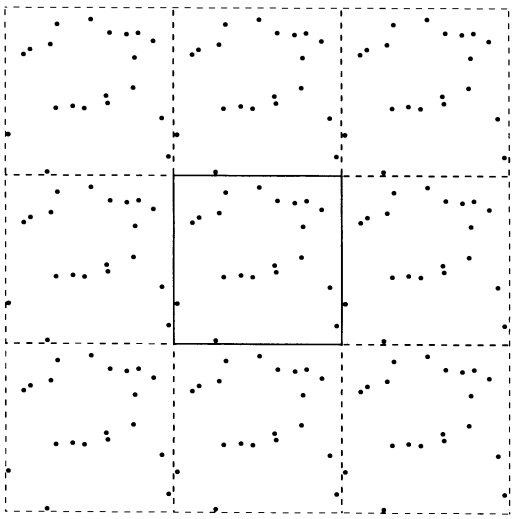

Fig. 3. Realisation of the toroidal edge correction. A plot with 20 points is replicated eight times around itself.

circumference which is inside the plot. Note that w_{ij} is entered as a denominator in Eq. (1), so the value of $I_t(u_{ij})$ increases to > 1. If the distance between i and j is greater than either distance between i and the two nearest boundaries (e_1, e_2) (Fig. 4B), then

$$w_{ij} = 1 - \left[\cos^{-1}\left(e_1/u_{ij}\right) + \cos^{-1}\left(e_2/u_{ij}\right) + \pi/2\right]2\pi \quad (4)$$

The values in radians of α_1 and α_2 are added, the addition of the term $\pi/2$ corresponds to the 90° angle covering the enclosed corner of the plot. Division by 2π gives the proportion of the circumference of the circle outside the plot which is again subtracted from 1.

One special case was not considered by Getis & Franklin (1987). If both distances between point i and the two nearest boundaries are smaller than the distance between i and the nearest corner of the plot, part of the circumference included in the term $\pi/2$ in Eq. (4) actually lies within the plot (Fig. 4C). While the resulting error may be small, it is possible to calculate the exact proportion by changing Eq. (4) to:

$$w_{ij} = 1 - \left[2\cos^{-1}\left(e_1/u_{ij}\right) + 2\cos^{-1}\left(e_2/u_{ij}\right)\right]2\pi \quad (5)$$

As an additional computation step it is necessary to calculate the distances of point i to all four corners of the plot. If all these distances are smaller than the distances between i and both nearest boundaries e_1 and e_2, Eq. (4) applies, otherwise Eq. (5).

Diggle (1983) gave formulae for a weighted boundary correction for both rectangular and circular plots. Diggle's equations for the correction in rectangular plots differ slightly from those given by Getis & Franklin (1987) but both versions give identical results. However, Diggle (1983; p. 72) provided different conditions for their application by prescribing a comparison of the squared values of u_{ij} and e_1 and e_2.

Statistical significance and confidence intervals

After the data have been analysed by the method described above, a plot of $K(t)$ versus t may reveal deviations from πt^2, expected under *CSR*. The deviation must now be tested for statistical significance. One method employs the calculation of constant approximate confidence intervals around *CSR*, defined by $\pm$ 1.42 $\sqrt{A/(n-1)}$ and $\pm$ 1.68 $\sqrt{A/(n-1)}$ as reasonable approximations of the 0.05 and 0.01 point level of significance, respectively (Getis & Franklin 1987; Szwagrzyk & Czerwczak 1993). The latter authors give a formula which contains an error, but apparently used the correct version for their calculations.

Although computationally much more time-consuming, most authors, however, employ Monte Carlo methods to determine the statistical significance of their results. The term 'Monte Carlo' has been rarely defined in papers that say they have used it (Judson 1994); here I follow the definition given by Manly (1991): "Essentially the idea is to use computer-generated data to determine the amount of variation to be expected in sample statistics". In the context of spatial pattern analysis, Monte Carlo methods simulate randomly generated plots of the same dimensions as the observed plot. In practice, random sampling is replaced by pseudo-random sampling and it is assumed that the random number generator employed is satisfactory (Besag & Diggle 1977). The simulated plot is then analysed as before. This procedure is repeated 19 times and the lowest and highest value of $K(t)$ for each t is used to define the lower and upper bound of a 95% confidence envelope. A 99% confidence envelope requires 99 simulations. The confidence interval for the null hypothesis is defined by the number of simulations (n) and equals $n/(n+1) \times 100\%$ (Leemans 1991).

Comparison of different methods of edge correction

All five methods of edge correction described above, i.e. buffer zone, toroidal edge correction, and the three versions of the weighted edge correction described by Diggle (1983), Getis & Franklin (1987), and the modification by Haase (this paper), were tested in two experimental set-ups. A square plot of 100 randomly placed points was simulated (inner plot in Fig. 2). This plot was surrounded by a buffer zone with a width of 1/2 the length of the plot side, which contained a further 300 random points. Only the pattern of the inner plot was

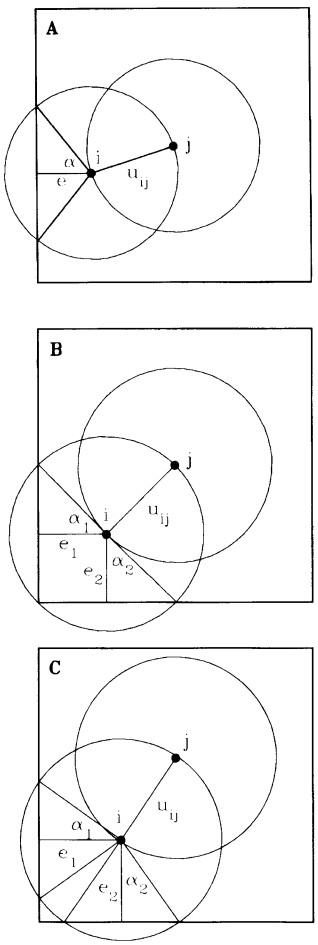

Fig. 4. Examples of three different cases to be considered for the weighted edge correction. The calculation of the parameter $I_t(u_{ij})$ in Ripley's estimator for $K(t)$ requires weighting by the proportion of the circumference of the circle with radius u_{ij} centred on point i and passing through j which lies within the plot. A, case 1 of the weighted edge correction, Eq. (3) has to be applied; B, case 2 of the weighted edge correction, equation (4) applies; C, case 3 (Haase, this paper) of the weighted edge correction, Eq. (5) applies.

analysed. For the generation of 95% confidence envelopes, 19 sets of 400 pairs of coordinates which were distributed between the inner plot and the buffer zone as before, were generated with a random number generator. To avoid variation in the simulated data, these same sets of random coordinates were used in all five methods of edge correction. A mapped 10 m × 10 m plot containing 111 individuals of the Mediterranean sub-shrub *Anthyllis cytisoides* (Fig. 1), was analysed with the same procedure, except that the analysis with the buffer zone method could not be performed because of insufficient field data.

Results

Presentation and interpretation

Results of spatial pattern analysis using Ripley's *K*-function are normally presented as graphs with the sample statistic $K(t)$ or a derived variable plotted against the independent variable t. The original sample statistic $K(t)$ is rarely plotted, however. Andersen (1992) preferred $\sqrt{K(t)}$ while Getis & Franklin (1987) used their own equivalent of $\sqrt{[K(t)/\pi]}$. Since $K(t) = \pi t^2$, the transformation $t = \sqrt{[K(t)/\pi]}$ yields a linear plot of the sample statistic against t (Fig. 5). Following Besag (1977), some authors preferred the derived variable $L(t) = \sqrt{[K(t)/\pi]}$ (Sterner et al. 1986; Szwagrzyk 1990; Duncan 1991; Szwagrzyk & Czerwczak 1993). Note that plots of $\sqrt{[K(t)/\pi]}$ as shown in Fig. 5 may have a poor resolution when the values of the sample statistic are close to spatial randomness and the confidence envelope is narrow.

Most authors are mainly interested in the deviation of the sample statistic from complete spatial randomness and plot $\sqrt{[K(t)/\pi]} - t)$ or $(L(t) - t)$ against t. The advantage of this transformation is that, under the null hypothesis of complete spatial randomness, the derived function has an expectation of 0 for all values of t (Skarpe 1991). The resulting plots are more informative and also yield a much higher resolution (Figs. 6, 7). Prentice & Werger (1985), Leemans (1991), and Skarpe (1991) named the corresponding derived variable $K^*(t) = \sqrt{[K(t)/\pi]} - t$, but the latter two authors give an erroneous formula for its calculation.

If the deviation of the sample statistic from zero expectation is positive, and above the upper limit of the confidence envelope, a clumped distribution of the sampled points can be assumed, while negative deviation indicates a dispersed or regular pattern. If the sample statistic remains within the bounds of the confidence envelope for any given t, the null hypothesis of complete spatial randomness cannot be rejected. Occasionally, $t - \sqrt{[K(t)/\pi]}$ has been plotted against t (Upton & Fingleton 1985; Kenkel 1988, 1993; Rebertus et al. 1989). This preference, however, reverses the sign of the sample statistic, so that values of $\sqrt{[K(t)/\pi]} > t$ now appear on the negative scale. The interpretation of clumped versus regular distribution pattern is thus reversed, which may result in some confusion.

Edge corrections and their effects on results

Results of spatial pattern analysis of the simulated plot by the five different methods of edge correction discussed above show basically the same shape for the curve of the derived sample statistic $\sqrt{[K(t)/\pi]} - t$ (Fig. 6). All plots show a distribution pattern close to com-

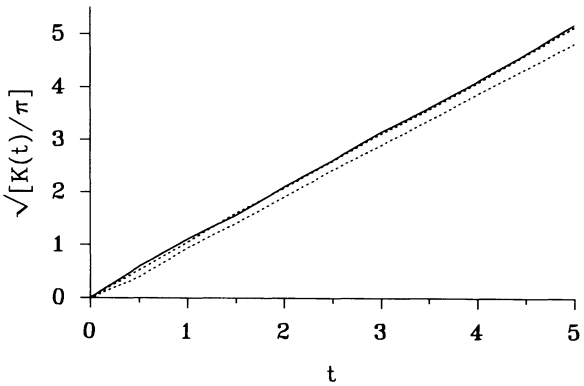

Fig. 5. Result of spatial pattern analysis of the distribution of 111 *Anthyllis cytisoides* L. shrubs using the weighted edge correction method of Getis & Franklin (1987). The function $K(t)$ was calculated for each 0.5-m interval and the derived sample statistic $\sqrt{[K(t)/\pi]}$ is plotted against t (solid line). The dotted lines give a 95% confidence envelope for complete spatial randomness.

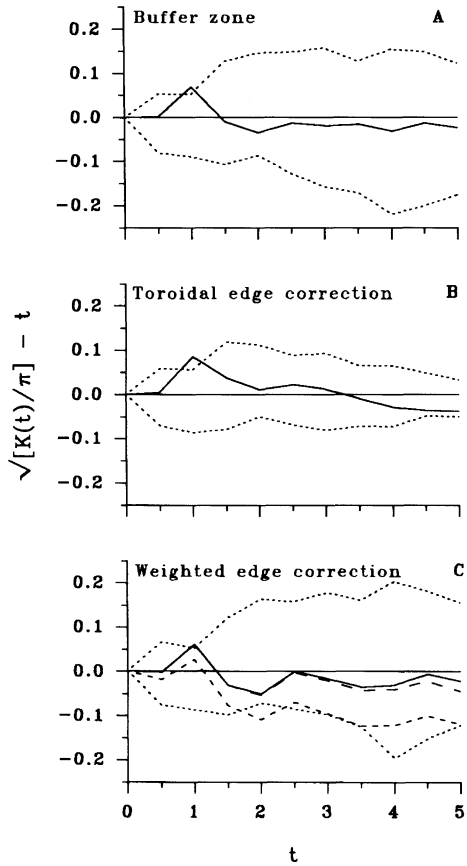

Fig. 6. Analyses of a simulated random pattern (central plot in Fig. 2) by different methods of edge correction. The derived sample statistic $\sqrt{[K(t)/\pi]} - t$ (solid lines) has been plotted against t for each 0.5 m interval. The dotted lines give a 95% confidence envelope for complete spatial randomness. A, buffer zone method; B, toroidal edge correction; C, weighted edge corrections of Getis & Franklin (solid line), Haase (long dash) and Diggle (short dash). Only the confidence envelope derived by the method of Getis & Franklin is shown.

plete spatial randomness, except for the values for $t = 1$, which are just outside the confidence envelope (Fig. 6). These significant deviations from randomness are picked up by all five methods. The method considering a buffer zone gives almost identical results to those of Getis & Franklin, and Haase. The plot for the toroidal edge correction, which basically creates its own buffer zone by replicating its distribution pattern around itself, gives somewhat different values for the sample statistic and shows narrowing confidence envelopes with increasing values of t. The three methods for the weighted edge correction produce plots of exactly the same shape, but varying in their deviation from πt^2. The method of Haase (this paper), which additionally considers case 3 for the edge correction procedure, gives slightly lower values at $t > 1.5$ m. This is to be expected, because the computations following Getis & Franklin ignore in some cases a small part of circumference actually within the plot (see Fig. 4C) thus giving overestimates of the sample statistic. The differences between the two results obtained are small for the values of t considered in this paper, but become proportionally larger as t (plot size) increases. The weighted edge correction of Diggle (1983) gives underestimates of the values of $K(t)$ because of the particular criteria set for the application of the boundary correction. Diggle uses the squares of the interpoint distance u_{ij} and of the distance from point i to the nearest boundary, while Getis & Franklin compare simple distances. These two conditions are not equivalent. In some cases Diggle's conditions ask for a single boundary correction when a double one is necessary, or for none when a single correction is needed. The result is a considerable underestimate of the value of $K(t)$.

The natural stand of *Anthyllis cytisoides* reveals a significantly clumped pattern at distances of up to 0.8 m and again at 3–5 m when they are analysed with any one of the three weighted edge correction methods (Fig. 7). The toroidal edge correction, however, yields a plot which does not significantly differ from random at distances > 1 m. The weighted edge correction of Getis & Franklin shows again small overestimates of the sample statistic at values of $t > 1.5$ m compared to the correction of Haase. The sample statistic and confidence envelopes derived by Diggle's weighted edge correction show a considerable shift to lower values, although still allowing the same interpretation of the observed pattern.

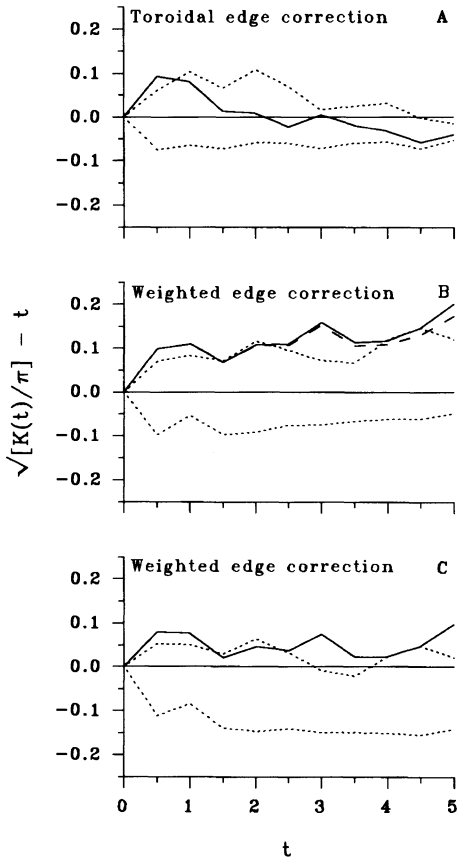

Fig. 7. Analyses of a mapped non-random pattern (shown in Fig. 1) using different methods of edge correction. The derived sample statistic $\sqrt{[K(t)/\pi]} - t$ (solid lines) has been plotted against t for each 0.5 m interval. The dotted lines give a 95% confidence envelope for complete spatial randomness. A, toroidal edge correction; B, weighted edge correction of Getis & Franklin (solid line), Haase (long dash) and confidence envelope by Getis & Franklin (dotted lines); C, weighted edge correction of Diggle.

Discussion

Most methods for analysing spatial pattern in plant communities fall into two categories (Goodall & West 1979): those based on distances, e.g. the Clark-Evans test (Clark & Evans 1954), and those based on area (quadrats), e.g. the contiguous quadrat method of Greig-Smith (1983). Both lines of pattern analysis have been continuously modified, but problems remain (e.g. Ludwig & Goodall 1978; Goodall & West 1979) and an ever increasing variety of such methods, giving different results for the same patterns (Ludwig & Goodall 1978; Goodall & West 1979), is now in use. The most powerful of the area-based methods appears to be the Two Term Local Quadrat Variance (TTLQV, Hill 1973) and its modifications (Ludwig & Goodall 1978; Dale & Blundon 1990; Ver Hoef et al. 1993).

As an alternative, combined count-distance methods such as Ripley's *K*-function (Diggle 1983) appear to offer solutions for some of the problems encountered by the methods mentioned above. Spatial point pattern analysis by Ripley's *K*-function has proved to be a useful and informative tool for the study of plant distribution patterns and its popularity among ecologists will probably rise. This calls for a certain level of standardisation of the statistical methods, so that published results can be more easily evaluated. While one understands that some researchers wish to modify the basic method to suit a particular problem, the 'standard user' should adhere to the original version, i.e. use the unbiased estimator for $K(t)$ proposed by Ripley (1976, 1977) given in Eq. (1) and plot the derived sample statistic $\sqrt{[K(t)/\pi]} - t$ or $L(t) - t$ (Ripley 1979, 1981) against t.

There is some scope for choice of the method of edge correction, but possible effects of a particular method on the results should be discussed. The method of Getis & Franklin (1987) as modified in this paper gave similar results for the patterns tested, but the discrepancy increased towards larger values of t. It is therefore recommended to incorporate case 3 of the edge correction into present or future programs. Although the formulae for calculations of the case 1 and case 2 edge corrections given by Diggle (1983) give identical results to those of Getis & Franklin, underestimates of $K(t)$ result because of the inadequate criteria set for their application. While the buffer zone method most realistically reflects the existing distribution pattern, it may not warrant the additional field work involved. If several adjacent plots are established, however, it may be convenient to analyse the inner parts by this method for comparison. The toroidal edge correction may give results that differ from the other methods tested because the addition of replicate areas to each side of the plot is a potential source of error. When non-random and bivariate patterns are analysed, clusters at opposite sides of the plot will be replicated at close proximity to each other. The estimation of the distribution of point-to-point distances then becomes biased to an unknown extent and the ecological interpretation of the spatial relationships may be flawed.

The null hypothesis is not limited to the test of spatial randomness of distribution patterns, although this is normally the first pattern tested for. If a clumped or dispersed pattern is analysed, appropriate models which generate the corresponding pattern, are used with Monte Carlo methods to create a confidence envelope for this pattern (Ripley 1977; Diggle 1983; Upton & Fingleton 1985). A further strength of the technique is that it also can be employed to investigate bivariate distribution patterns and give information on the spatial relationships between two species, two size classes, two life stages, etc. (e.g. Cox 1987; Duncan 1991).

Acknowledgements. The research leading to this paper was carried out as part of the MEDALUS II (Mediterranean Desertification and Land Use) collaborative research project, funded by the EC under its Environment Programme, contract number EV5V CT92-0164. I thank the Consejo Superior de Investigaciones Científicas and particularly Dr. J. Puigdefábregas for use of the facilities of the Estación Experimental de Zonas Áridas and the Rambla Honda field site. I also thank the referee who commented on toroidal edge correction, and Rainer Haase who critically read the manuscript.

References

Andersen, M. 1992. Spatial analysis of two-species interactions. *Oecologia* 91: 134-140.

Besag, J. 1977. Contribution to the discussion of Dr Ripley's paper. *J. R. Stat. Soc. B* 39: 193-195.

Besag, J. & Diggle, P.J. 1977. Simple Monte Carlo tests for spatial pattern. *Appl. Stat.* 26: 327-333.

Clark, P.J. & Evans, F. C. 1954. Distance to nearest neighbor as a measure of spatial relationships in populations. *Ecology* 35: 23-30.

Cox, G.W. 1987. Nearest-neighbour relationships of overlapping circles and the dispersion pattern of desert shrubs. *J. Ecol.* 75: 193-199.

Dale, M.R.T. & Blundon, D. J. 1990. Quadrat variance analysis and pattern development during primary succession. *J. Veg. Sci.* 1: 153-164.

Diggle, P.J. 1983. *Statistical Analysis of Spatial Point Patterns*. Academic Press, London.

Diggle, P.J., Besag, J. & Gleaves, J.T. 1976. Statistical analysis of spatial point patterns by means of distance methods. *Biometrics* 32: 659-667.

Duncan, R.P. 1991. Competition and the coexistence of species in a mixed podocarp stand. *J. Ecol.* 79: 1073-1084.

Ebert, T.A. & McMaster, G.S. 1981. Regular pattern of desert shrubs: a sampling artefact? *J. Ecol.* 69: 559-564.

Getis, A. & Franklin, J. 1987. Second-order neighborhood analysis of mapped point patterns. *Ecology* 68: 473-477.

Gibson, D.J. & Menges, E.S. 1994. Population structure and spatial pattern in the dioecious shrub *Ceratiola ericoides*. *J. Veg. Sci.* 5: 337-346.

Goodall, D.W. & West, N.E. 1979. A comparison of techniques for assessing dispersion patterns. *Vegetatio* 40: 15-27.

Greig-Smith, P. 1983. *Quantitative Plant Ecology*. Blackwell Scientific Publications, Oxford.

Hill, M.O. 1973. The intensity of spatial pattern in plant communities. *J. Ecol.* 61: 225-235.

Judson, O.P. 1994. The rise of the individual-based model in ecology. *Trends Ecol. Evol.* 9: 9-14.

Kenkel, N.C. 1988. Pattern of self-thinning in jack pine: testing the random mortality hypothesis. *Ecology* 69: 1017-1024.

Kenkel, N.C. 1993. Modeling Markovian dependence in populations of *Aralia nudicaulis*. *Ecology* 74: 1700-1706.

King, T.J. & Woodell, S.R.J. 1984. Are regular patterns in desert shrubs artefacts of sampling? *J. Ecol.* 72: 295-298.

Leemans, R. 1991. Canopy gaps and establishment patterns of spruce (*Picea abies* (L.) Karst.) in two old-growth coniferous forests in central Sweden. *Vegetatio* 93: 157-165.

Ludwig, J.A. & Goodall, D.W. 1978. A comparison of paired- with blocked-quadrat variance methods for the analysis of spatial pattern. *Vegetatio* 38: 49-59.

Manly, B.F.J. 1991. *Randomization and Monte Carlo Methods in Biology*. Chapman and Hall, London and New York.

Morisita, M. 1959. Measuring of the dispersion of individuals and analysis of the distributional patterns. *Mem. Fac. Sci., Kyushu Univ. E, Biology* 2: 215-235.

Pielou, E.C. 1960. A single mechanism to account for regular, random and aggregated populations. *J. Ecol.* 48: 575-584.

Prentice, I.C. & Werger, M.J.A. 1985. Clump spacing in a desert dwarf shrub community. *Vegetatio* 63: 133-139.

Rebertus, A.J., Williamson, G.B. & Moser, E.B. 1989. Fire-induced changes in *Quercus laevis* spatial pattern in Florida sandhills. *J. Ecol.* 77: 638-650.

Ripley, B.D. 1976. The second-order analysis of stationary processes. *J. Appl. Prob.* 13: 255-266.

Ripley, B.D. 1977. Modelling spatial patterns. *J. R. Stat. Soc. B* 39: 172-212.

Ripley, B.D. 1979. Tests of 'randomness' for spatial point patterns. *J. R. Stat. Soc. B* 41: 368-374.

Ripley, B.D. 1981. *Spatial statistics*. J. Wiley, New York, NY.

Rohlf, F.J. & Archie, J.W. 1978. Least-squares mapping using interpoint distances. *Ecology* 59: 126-132.

Simberloff, D. 1979. Nearest neighbor assessment of spatial configurations of circles rather than points. *Ecology* 60: 679-685.

Skarpe, C. 1991. Spatial patterns and dynamics of woody vegetation in an arid savanna. *J. Veg. Sci.* 2: 565-572.

Sterner, R.W., Ribic, C.A. & Schatz, G.E. 1986. Testing for life historical changes in spatial patterns of four tropical tree species. *J. Ecol.* 74: 621-633.

Stewart, G.H. & Rose, A.B. 1990. The significance of life history strategies in the developmental history of mixed beech (*Nothofagus*) forests, New Zealand. *Vegetatio* 87: 101-114.

Szwagrzyk, J. 1990. Natural regeneration of forest related to the spatial structure of trees: A study of two forest communities in Western Carpathians, southern Poland. *Vegetatio* 89: 11-22.

Szwagrzyk, J. & Czerwczak, M. 1993. Spatial patterns of trees in natural forests of East-Central Europe. *J. Veg. Sci.* 4: 469-476.

ter Braak, C.J.F. 1980. *Binary mosaics and point quadrat sampling in ecology*. IWIS-TNO A80 ST 9637, IWIS-TNO, Wageningen.

Upton, G. & Fingleton, B. 1985. *Spatial Data Analysis by Example. Vol. 1. Point Pattern and Quantitative Data*. John Wiley, New York, NY.

Ver Hoef, J.M., Cressie, N.A.C. & Glenn-Lewin, D.C. 1993. Spatial models for spatial statistics: some unification. *J. Veg. Sci.* 4: 441-452.

Received 30 September 1994;
Revision received 7 February 1995;
Final revision received 19 May 1995;
Accepted 12 June 1995.