

Comparing the Performances of Diggle's Tests of Spatial Randomness for Small Samples with and without Edge-Effect Correction: Application to Ecological Data

Jacques Gignoux,^{1,*} Camille Duby,² and Sébastien Barot¹

¹Ecole Normale Supérieure, Laboratoire d'Écologie (CNRS-URA 258),
46, Rue d'Ulm, 75230 Paris Cedex 05, France

²Institut National Agronomique, Chaire de Mathématiques,
16, Rue Claude Bernard, 75231 Paris Cedex 05, France

* email: gignoux@biologie.ens.fr

SUMMARY. Diggle's tests of spatial randomness based on empirical distributions of interpoint distances can be performed with and without edge-effect correction. We present here numerical results illustrating that tests without the edge-effect correction proposed by Diggle (1979, *Biometrics* **35**, 87–101) have a higher power for small sample sizes than those with correction. Ignoring the correction enables detection of departure from spatial randomness with smaller samples (down to 10 points vs. 30 points for the tests with correction). These results are confirmed by an example with ecological data consisting of maps of two species of trees in a West African savanna. Tree numbers per species per map were often less than 20. For one of the species, for which maps strongly suggest an aggregated pattern, tests without edge-effect correction enabled rejection of the null hypothesis on three plots out of five vs. on only one for the tests with correction.

KEY WORDS: Bias; Distance methods; Monte Carlo; Nearest neighbor; Point process; Power; Spatial pattern.

1. Introduction

Tests of spatial randomness and spatial association based on interpoint distances have been developed to analyse spatial pattern data consisting in maps of points (Ripley, 1981; Diggle, 1983). This kind of data is frequent in ecology (e.g., Fisher, 1993; Haase, 1995; Haase et al., 1996; Moeur, 1993; Skarpe, 1991; Steiner, Ribic, and Schatz, 1986; Szwagrzyk, 1990), where maps of e.g. plant locations are drawn to study interactions between individuals. The most popular tests among ecologists are based on Diggle's F and G functions and on Ripley's K function. These tests enable to assess whether observed spatial patterns can be considered as regular, random or aggregated, the null hypothesis H_0 being that of complete spatial randomness (synonymous for a homogeneous Poisson process in this paper, as in Cressie, 1991, p. 586).

Diggle's tests perform well for relatively small samples (down to 30 mapped points). However, one often wishes to identify the type of spatial pattern with smaller samples such as in the following:

- (1) In a hierarchical pattern with clumps of points, it might be valuable to analyse the pattern of points within clumps, which may be of small size, depending on the particular pattern.
- (2) When a map is drawn in the field, there will almost certainly be classes of points of ecological interest with few individuals, e.g., adult trees are often much less

numerous than younger ones although their ecological role is prominent (Barot, Gignoux, and Menaut, 1998b; Steiner et al., 1986), some species may be rare and still justify a study (Forman and Hahn, 1980), size classes defined for objective reasons (commercial size in a forest) might be rare, etc.

- (3) Field plots are not always designed with spatial analysis as a unique goal—they are often designed for a factorial experiment or a demographic survey—and as such may not be of a sufficiently comfortable size for the spatial analysis of each plot. Given the high cost of field surveys, this kind of information will often be the only that is available to study the spatial pattern.

Thus, there is a need to improve the sensitivity of the most standard existing tests for small samples of 10 to 30 points, a sample size where spatial patterns are still recognizable but not by most of the existing statistical methods, except Ripley's K function (e.g., Ripley, 1981, pp. 182–183). Although the K -based test seems more efficient than the F and G tests, it should not be used alone because the three tests have slightly different meanings and different sensitivities to different types of spatial distributions: the typical hierarchy of powers for samples of 100 points against aggregation is $\text{power}(F) > \text{power}(K) > \text{power}(G)$ and $\text{power}(K) > \text{power}(G) > \text{power}(F)$ against regularity (Diggle, 1979). These powers have been estimated for a single sample size and

for a limited set of alternative hypotheses but constitute, at the moment, the only available information on the sensitivity of these tests since no systematic study of the power depending on the sample size has been performed so far. Working on a large data set of real data (full maps of all trees in five plots of a West African savanna), we found that the G and F tests almost never reject the null hypothesis below approximately 30 points, while tests ignoring the correction for edge effects initially proposed by Diggle (1979) were able to reject the null hypothesis down to sample sizes of 8–10. The goal of the present study is to explain this better sensitivity of uncorrected tests through a simulation study of their power and to document it with an example on real data.

2. Diggle's Tests of Spatial Randomness

The tests rely on the comparison of an observed distribution of nearest-neighbor distances to the theoretical distribution expected under the null hypothesis of complete spatial randomness. They involve two cumulative distribution functions (c.d.f.s) of nearest-neighbor distances based on (using Diggle's notations)

- (i) the distance w between each point of the observed spatial pattern and its nearest neighbor (Diggle's G function);
- (ii) the distance x between each point of a predefined sampling grid and the nearest point of the observed spatial pattern (Diggle's F function).

The choice of the sampling grid (point pattern and number of points) has been discussed widely in Ripley (1981) and Diggle (1983). We used a pragmatic approach, as proposed by Upton and Fingleton (1985): We based our tests on a regular grid of $\text{sup}(N, 400)$ points, where N is sample size.

The theoretical c.d.f.s under H_0 are (Diggle, 1983)

$$G_0(w) = 1 - e^{-\pi\lambda w^2} \quad (1)$$

$$F_0(x) = 1 - e^{-\pi\lambda x^2}, \quad (2)$$

where λ is the intensity of the homogeneous Poisson process and the index refers to H_0 .

These expressions are compared to empirical c.d.f.s calculated from a sample of limited spatial extent. Points located near the border of the mapped plot might have their nearest neighbor outside the plot. This edge effect causes a systematic error in the estimation of the empirical c.d.f.s. A classical edge-effect correction (EEC), derived from minus sampling (Miles, 1974), proposed by Diggle (1979) and called the border method by Ripley (1988), relies on the elimination of two groups of points, specifically,

- (1) all points that lie closer to the border than to their nearest neighbor inside the plot are discarded,
- (2) when estimating the empirical c.d.f.s at distance d , all points lying within the strip of width d from the border of the plot are discarded (Figure 1); increasing d leads to discarding a larger number of points.

In this method, the mean proportion of discarded points increases when sample size decreases. This effect is also amplified by increasing dimension [see the three-dimensional version of the method by Baddeley et al. (1993) and Baddeley and Gill (1997)]. Three-dimensional data sets, although rare,

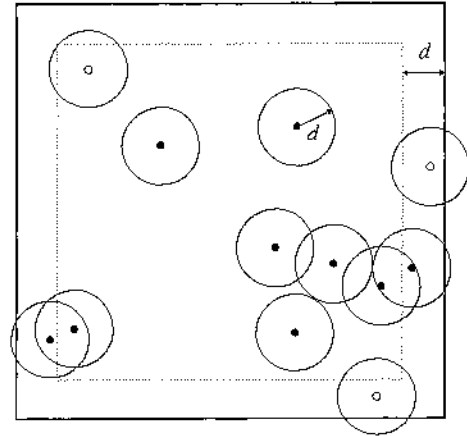


Figure 1. The correction for edge effects in Diggle's F and G tests. Points \circ are discarded because they are closer to the border than to their nearest neighbor. As distance d increases, points \bullet are discarded when they lie within the strip of width d . At the distance shown, only 7 points out of 12 are used to compute $\hat{G}(d)$.

exist in ecology but are costly to obtain and require very efficient tests.

With this correction, the empirical c.d.f.s are (Diggle, 1983)

$$\hat{G}(w) = \#\{w_i \leq w, b_i > w\} / \#\{b_i > w\} \quad (3)$$

$$\hat{F}(x) = \#\{x_i \leq x, b_i > x\} / \#\{b_i > x\}, \quad (4)$$

where w_i (respectively, x_i) is the distance of point i to its nearest neighbor and b_i the distance of point i to the nearest border of the sample plot.

These estimators do not have the same properties (Ripley, 1988). $\hat{F}(x)$ is unbiased (since its denominator is a fixed number, $E[\#\{x_i \leq x, b_i > x\}] / \#\{b_i > x\} = E[\hat{F}(x)] = F(x)$), while $\hat{G}(w)$ is only ratio-unbiased (since its denominator is a random number, $E[\#\{w_i \leq w, b_i > w\}] / E[\#\{b_i > w\}] = G(w)$; Ripley, 1988, p. 27; Baddeley and Gill, 1997). Both are non-monotone (see Figure 6 for an example), particularly at large distances.

The most widely used test statistics are based on the difference between the empirical and theoretical c.d.f.s, with expected value under H_0 zero at all distances (Diggle, 1983), i.e.,

$$dw = \sup_w |\hat{G}(w) - G_0(w)| \quad (5)$$

$$dx = \sup_x |\hat{F}(x) - F_0(x)|, \quad (6)$$

where the sup are taken on all distances smaller than the half-length of the smallest side for a rectangular plot, given the edge-effect correction used. We note $w_{\text{sup}} = \arg(\sup_w |\hat{G}(w) - G_0(w)|)$ and $x_{\text{sup}} = \arg(\sup_x |\hat{F}(x) - F_0(x)|)$, the distances corresponding to dw and dx .

One can obtain test significance for dw or dx analytically for simple cases only (homogeneous Poisson processes in a plot of infinite surface) but, as observed samples are usually of finite area, edge effects lead to untractable distribution theory

(Diggle, 1983). A solution is to use Monte Carlo simulations to estimate a confidence interval for $G(w)$ under H_0 at each distance (the same procedure holds for F) using the following procedure:

- (i) First, compute the empirical c.d.f. $\hat{G}(w)$ from the observed data and the test statistic according to (5).
- (ii) Generate s random spatial patterns (i.e., patterns under H_0) of the same intensity as the observed pattern.
- (iii) Compute $\hat{G}_{0j}(w)$ and dw_j for each simulated pattern ($j = 1, \dots, s$) according to formulas (3) to (6) to get an estimate of the distribution of the test statistic under H_0 .
- (iv) Estimate a rejection limit or a P -value (e.g., if a test at the 5% level is performed with 100 simulations, the rejection limit will be given by the fifth greatest value of simulated dw 's; the P -value will be estimated as the proportion of simulated samples for which that dw is greater than the observed dw). The greater the number of simulations, the better the estimated P -value or rejection limit (Marriott, 1979).

When H_0 is rejected, the sign of the difference between the empirical and theoretical c.d.f. with the greatest absolute value enables deciding whether there is a tendency toward aggregation or regularity: if $\hat{G}(w_{\text{sup}}) - G(w_{\text{sup}}) > 0$, the pattern is aggregated, otherwise it is regular; the reverse holds for the F test. One-sided tests could also be constructed, but they would probably be of little use in ecology since many ecological processes can act simultaneously in a system (e.g., the spatial pattern of a population of plants results from the initial pattern of seeds, usually very clumpy, and from competition, leading to regular distribution; hence, regularity as well as aggregation can be simultaneously expected, depending on which ecological factor is prominent).

3. How to Build Tests Without Edge-Effect Correction

To build a consistent test, we need to treat edge effects in the same way for the theoretical and the observed c.d.f.s. Two methods are possible:

- (1) the classical method described in Section 2: if we use an EEC for the observed c.d.f., we can use the theoretical c.d.f. in its analytical form given by expressions (1) and (2);
- (2) another method: get an estimate or an analytical expression of the theoretical and observed c.d.f.s for the same sampling window, both computed without EEC.

Method (2) has the advantage over method (1) of making profit of all the information of the data set, which is particularly interesting for small samples. The only problem with method (2) is that the analytical expression of the theoretical c.d.f. affected by a sampling window is usually unknown. Diggle (1983) proposed using the mean of the Monte Carlo simulated distributions as an estimate of the theoretical c.d.f. when the analytical expression of the theoretical c.d.f. is unknown. We can use this method to estimate the theoretical c.d.f. for a given sampling window for a homogeneous Poisson process.

Without EEC, the estimations of the observed and theoretical c.d.f.s, and the test statistics, respectively, become

$$\hat{G}^*(w) = \#\{w_i \leq w\}/N \quad (7)$$

$$G_0^*(w) = \frac{1}{s} \sum_{j=1}^s \#\{w_{ij} \leq w\}/N \quad (8)$$

$$dw^* = \sup_w |\hat{G}^*(w) - G_0^*(w)|, \quad (9)$$

where N is the sample size (number of observed points), s is the number of Monte Carlo simulations used to compute rejection limits or P -values, and w_{ij} is the distance of the i th point of simulation j to its nearest neighbor, and with similar expressions for the F function. The bias due to edge effects of the c.d.f.s computed from expressions (7) and (8) are noted $B_0(w)$ for $G_0^*(w)$ and $B_\bullet(w)$ for $\hat{G}^*(w)$. Under H_0 , these two biases are equal and do not affect dw^* and the type I error of the test is unaffected by the bias. Under H_1 , the biases are different, dw^* is affected by the bias and, as a result, the power of the tests based on dw^* could be affected. An unbiased test statistic would be given by

$$dw^{**} = \sup_w |[\hat{G}^*(w) - \hat{B}_\bullet(w)] - [G_0^*(w) - \hat{B}_0(w)]|, \quad (10)$$

where $\hat{B}_\bullet(w)$ and $\hat{B}_0(w)$ are the estimates of $B_\bullet(w)$ and $B_0(w)$. The test based on dw^{**} can only be performed with simulation data since H_1 has to be known to estimate the bias. We used it to evaluate the effect of the bias on the power of the tests based on dw^* through a simulation study.

4. Simulation of 'Observed' Spatial Patterns

To study the performances of the tests with and without EEC, we generated 'observed' spatial patterns through simulation. We chose two contrasted alternative hypotheses, a highly aggregated spatial pattern H_{1a} and a highly regular pattern H_{1b} (Figure 2), generated, respectively, from a Poisson cluster process and a sequential inhibition process (Diggle, 1979; Ripley, 1977). Those hypotheses are as follows:

- (1) The Poisson cluster process is based on (i) a homogeneous Poisson process with intensity ρ of parent points, (ii) a probability distribution to describe the number of offspring points per parent, and (iii) a bivariate distribution describing the position of offspring relative to parents. In the present case, we used a Poisson distribution for (ii) and a symmetric radial normal distribution for (iii), as in Diggle (1979). Parameters were the intensity of the whole process $\lambda = 10$, the mean number of offspring per parent $\mu = 1.5$ (thus $\rho = \lambda/(1 + \mu) = 6.667$), and the average distance of offspring to parents $\sigma = 0.08$. Parent points were kept in the final pattern.
- (2) The sequential inhibition process assumes that (i) all points are distributed uniformly over the whole plot, with the constraint that (ii) the i th point must be distant at least δ from all the previously located points. We chose as parameters an intensity $\lambda = 10$ and a minimum separation distance between points $\delta = 0.2$.

For H_0 , a homogeneous Poisson process with intensity $\lambda = 10$ was used.

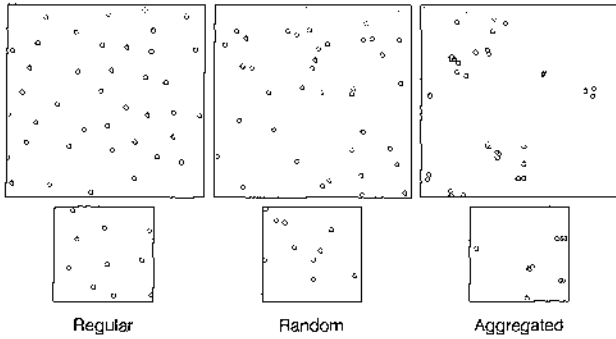


Figure 2. Examples of realizations conditioned by sample size of the three types of processes used in all simulations: a regular pattern (H_{1a}) simulated from a sequential inhibition process with intensity $\lambda = 10$ and $\delta = 0.2$; a random pattern (H_0) simulated from a homogeneous Poisson process with intensity $\lambda = 10$; an aggregated pattern (H_{1b}) simulated from a Poisson cluster process with $\lambda = 10$, $\mu = 1.5$, and $\sigma = 0.08$. Top, samples of size 40; bottom, samples of size 10.

We generated processes for seven samples sizes on a square plot: 10, 15, 20, 25, 30, 40, and 50 points. Conditioning by sample size is straightforward for the homogeneous Poisson and sequential inhibition processes; for the Poisson cluster process, conditioning was based on a rejection method. Process parameters were chosen to produce rather distinct patterns since our main focus was on the effect of sample size rather than on the alternative hypotheses. Differences are conspicuous on large samples and sometimes still visible on small samples (Figure 2).

5. Study of the Bias and Variance of the Test Statistics

The power of the tests depends on both the bias and the variance of the test statistics. We estimated the variance of the test statistics dx , dw , dx^* , and dw^* from 1000 Monte Carlo simulations of a homogeneous Poisson process. The estimates of the theoretical distributions for the tests without EEC were computed from the same 1000 simulations. Computations were done at each sample size. The statistics of the tests without EEC have a much lower variance than those of the tests with EEC (Figure 3). A direct consequence of this is that the rejection limits for a sample of size 10 are very close to the maximal possible value of the test statistics of 1 in the tests with EEC: e.g., the rejection threshold at 5% for the dw -based test is 0.94 but drops to 0.47 in the dw^* -based test.

To estimate $B_1(w)$ (bias of $\hat{G}^*(w)$ under H_1), we need to estimate $E[\hat{G}^*(w)]$, which is easy (expression (7)), and estimate $G(w)$, which is unknown for H_{1a} and H_{1b} . Since edge effects are due to the possibility that the nearest neighbor of a point of the sample lies outside the mapped area, we simulated a large sample, of which only the central points were used to build the estimate. The sample was large enough to guarantee that the nearest neighbor of each point of the central area laid within the larger sample. The resulting unbiased estimated $G(w)$ was computed as

$$G^{**}(w) = \frac{1}{s} \sum_{j=1}^s \# \{w'_{ij} \leq w\} / N, \quad (11)$$

where w'_{ij} is the distance of the i th point of simulation j to its nearest neighbor, with neighbors being searched eventually

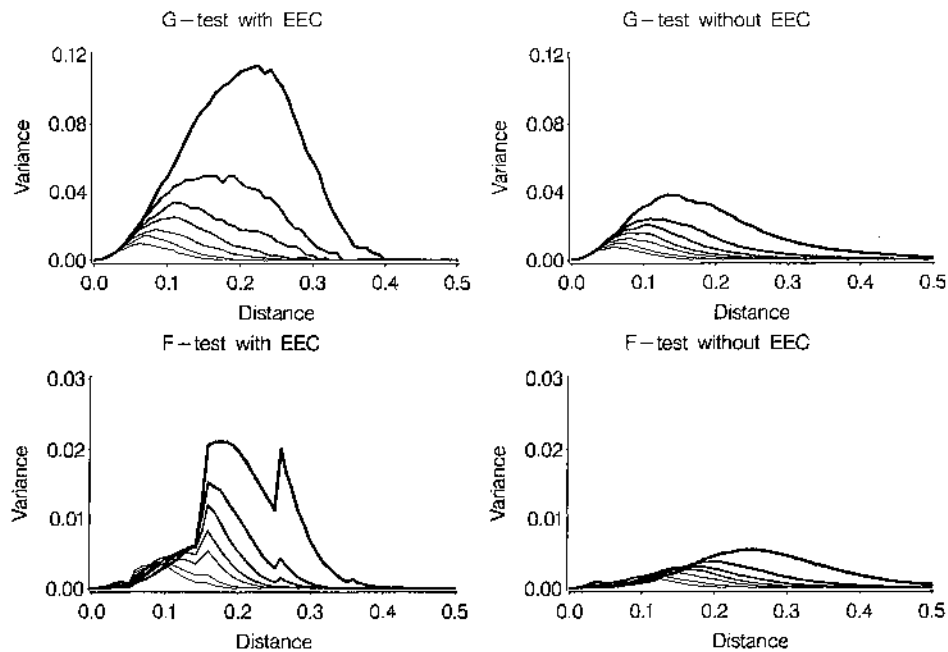


Figure 3. Variance of the test statistics dx , dw , dx^* , dw^* for sample sizes 10, 15, 20, 25, 30, 40, 50 (thickest line: $N = 10$; thinnest line: $N = 50$). Estimates based on 1000 simulations. Distances normalized to a square plot of unit side.

Table 3

Summary of the analysis of the spatial patterns of Figure 4 with the *F* and *G* tests with and without edge-effect corrections and with Ripley's *K* test. Significant results at the 5% level in bold. Signed test statistics used. *P*-values estimated with 500 Monte Carlo simulations of H_0 .

Species	Plot	<i>N</i>	<i>F</i> test				<i>G</i> test				<i>K</i> test	
			<i>dx</i>	<i>P</i>	<i>dx</i> *	<i>P</i>	<i>dw</i>	<i>P</i>	<i>dw</i> *	<i>P</i>	<i>L_m</i>	<i>P</i>
<i>Bridelia ferruginea</i>	A	7	0.339	0.092	0.027	0.972	0.277	0.980	-0.222	0.942	-3.634	0.976
	C	12	-0.124	0.508	-0.054	0.650	-0.925	0.078	0.225	0.710	2.633	0.884
	G	22	-0.031	0.964	0.032	0.802	-0.155	0.886	0.175	0.564	-1.827	0.670
	H	39	0.024	0.982	0.024	0.876	-0.214	0.238	-0.244	0.052	-1.003	0.718
	I	18	0.094	0.522	-0.065	0.366	-0.350	0.314	-0.176	0.666	2.548	0.454
<i>Cussonia arborea</i>	A	11	0.133	0.470	-0.103	0.184	0.494	0.312	0.396	0.128	5.258	0.180
	C	7	-0.294	0.148	0.041	0.848	-0.585	0.508	-0.430	0.278	-7.664	0.230
	G	14	-0.152	0.316	-0.253	0.000	0.583	0.124	0.514	0.006	6.441	0.004
	H	60	-0.114	0.020	-0.143	0.000	0.320	0.000	0.311	0.000	2.494	0.000
	I	10	-0.165	0.360	-0.087	0.330	-0.733	0.228	0.657	0.002	6.722	0.080

Power was estimated for H_{1a} and H_{1b} . Five hundred realizations of both processes for each of the seven sample sizes constituted our observed samples. The six tests (based on *dw*, *dx*, *dw**, *dx**, *dw****, *dx****) were completed for each of these. Rejection thresholds were computed with 1000 other Monte Carlo simulations of a homogeneous Poisson process with the same sample size as the (simulated) observed process for each sample size. Tests were performed at the 0.01 and 0.05 significance level.

For both alternative hypotheses and for small samples, *dw*-based tests have a much lower power than *dw**-based tests, which in turn have a slightly lower power than the *dw****-based tests (Table 2). Rejection of H_0 can occur for samples as small as 10 points for tests without EEC, while the power of the tests with EEC drops below 25–30 points. The same conclusion holds for the *dx*-, *dx**-, and *dx****-based tests. For large samples (*N* = 50), the tests with and without EEC tend to have the same powers.

We found a maximum absolute difference of 0.014 between the powers of *dw**- and *dw****-based tests and of 0.066 between the powers of *dx**- and *dx****-based tests (Table 2). These differences are rather small compared to the gain in power obtained from using tests without EEC instead of tests with EEC and justify the use of *dw** and *dx** as suitable statistics.

The *G*- and *F*-based tests have different power curves: against regularity, we find, like Diggle (1979), that the *dw**-based test has a higher power than the *dx**-based test. Against aggregation, both tests have rather similar powers, contrary to what Diggle found. Maybe a systematic study of power for more alternative hypotheses would help in assessing this point.

7. Example with Real Data

As a real case study, we analysed the spatial pattern of two species of trees on five 50 by 50 m plots of a West African savanna (Lamto, Côte d'Ivoire; see Menaut and César (1979) for a complete description of the site and study species). All trees ≥ 2 m, considered adults, were mapped (Figure 5). The two species selected here (*Bridelia ferruginea* and *Cussonia arborea*) belong to the group of the four dominant species

in this savanna type. Although the plots all share the same soil type and main grass species, they have very different tree densities—a characteristic of these savannas. The spatial pattern of trees is of great importance in savanna ecosystems since it conditions most of the physical and biological processes driving the ecosystem, such as competition for light and water between trees and grass and young tree mortality induced by the annual fires that burn the grass layer (Menaut et al., 1990).

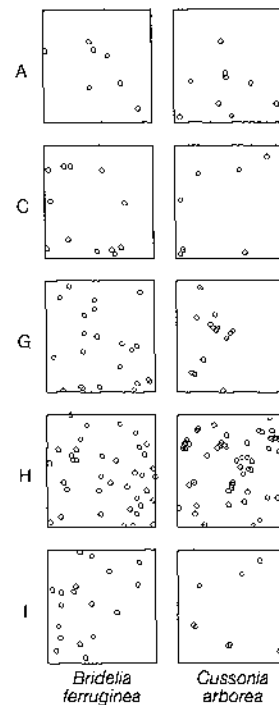


Figure 5. Positions of all adult (≥ 2 m in height) trees of two different species on five 50 × 50 m plots in a humid savanna of West Africa (Lamto, Côte d'Ivoire).

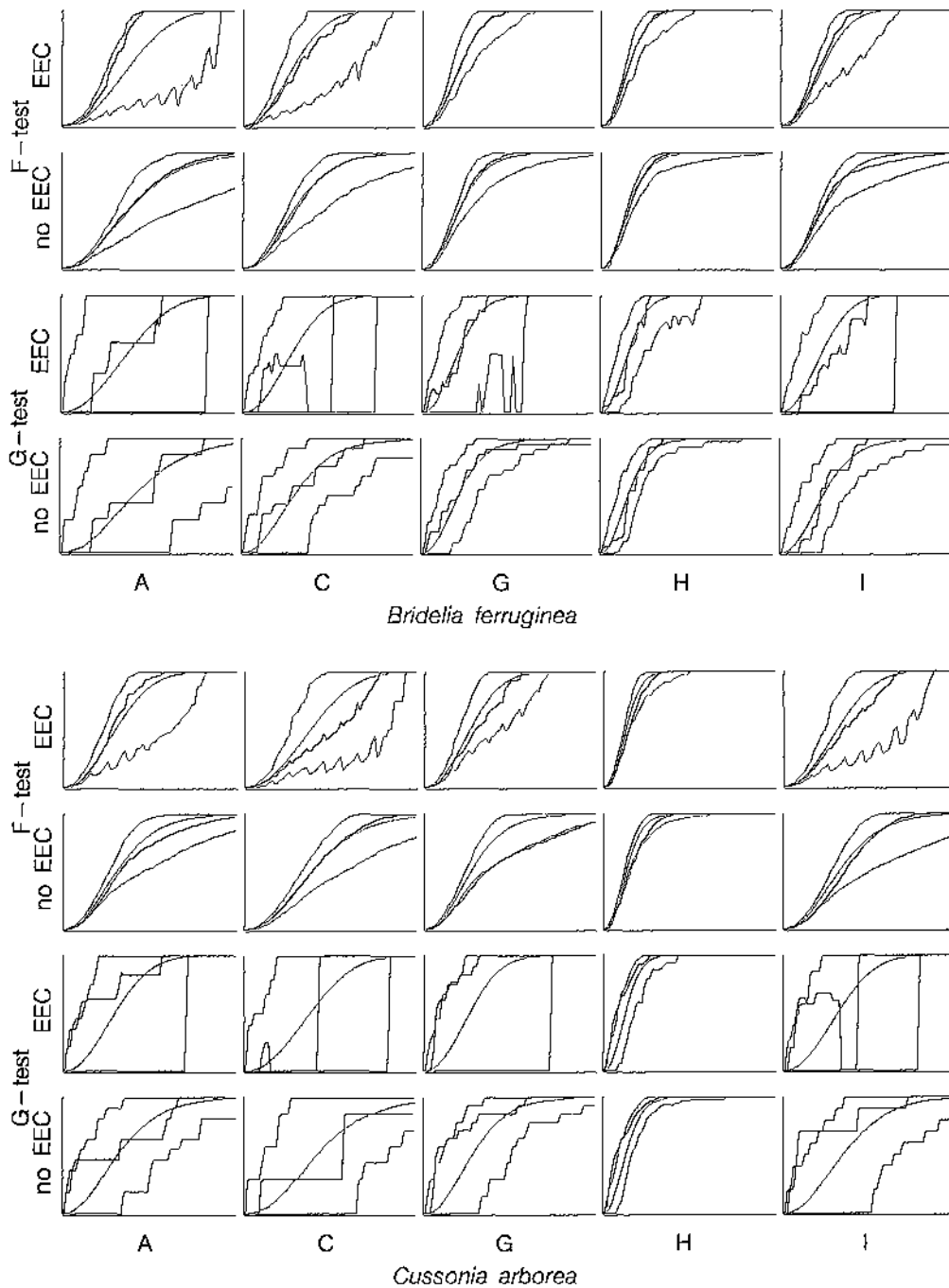


Figure 6. Results of the analysis of the spatial patterns of Figure 4. Abscissa: distance w (respectively, x) on a scale from 0 to 25 m. Ordinates on a scale from 0 to 1: thick line, empirical c.d.f. of the observed sample; thin lines, theoretical c.d.f. against which it is compared and envelope of 500 simulations. The theoretical c.d.f. is equal to $G_0(w)$ (respectively, $F_0(x)$) for the tests with edge-effects correction and to $G_0^*(w)$ (respectively, $F_0^*(x)$) for the tests without EEC. Summary statistics are provided in Table 3.

We analysed the spatial pattern of trees on the five plots with the F and G tests, with and without EEC. We also performed the test based on Ripley's K function as a reference. We used the L_m test statistic as defined by Ripley (1981). The K test is able to reject H_0 for samples as small as 14 points

(e.g., in Ripley, 1981, pp. 182–183) and had the highest power against regularity according to Diggle's (1979) estimates. Results are summarized in Table 3 and Figure 6.

The K test leads to rejection of H_0 for *Cussonia* on plots G and H and indicates aggregation. In all other cases, H_0 is

not rejected. The P -value for *Cussonia* on plot I is low (0.08), suggesting that a more powerful test could detect aggregation on this plot.

With the dw and dx test statistics, the null hypothesis is rejected only once and indicates aggregation for *Cussonia* on the H plot, the plot with the highest tree density.

However, this species seemed to form small clumps of trees on at least three other plots (A, G, and I: see Figure 5), and the K test indicates aggregation on plots G and H and is almost significant ($P = 0.08$) for plot I. Rejection of H_0 is effectively obtained with the dw^* - and dx^* -based tests on plots G and H and with the dw^* -based test on plot I. The sign of the test statistics all indicate a tendency toward aggregation. There is probably such a tendency on plot A also, indicated by a relatively low P -value for the dw^* -based test (0.128), the general shape of the G function (Figure 6), and the map (Figure 5). The species is then probably characterized by an aggregated pattern, not always detectable with samples below size 12 but detectable above this limit. Such a pattern can be due to a variety of ecological causes, the three main ones being the seed dispersal pattern, a high sensitivity to fire (Menaut et al., 1990), and association with nutrient-rich patches (Barot, Gignoux and Menaut, 1998a).

Contrary to *Cussonia*, the use of tests without EEC does not change the results for *Bridelia*. These results are consistent with those of the K test. We therefore conclude that *Bridelia* has a random spatial pattern on all plots. The ecological interpretation of this result is that this species should be highly fire-resistant since it has been demonstrated that fire-sensitive species should have an aggregated pattern in this savanna (Menaut et al., 1990). There is a slight tendency towards regularity (Figure 5, plots H and I), which is almost significant on plot H (dw^* -based test; $P = 0.052$). Such a tendency is usually interpreted as an indication of competition for resources (light, water, or nutrients).

Figure 6 emphasizes the nonmonotonicity of the \hat{F} and \hat{G} estimators, specially at large distances (e.g., empirical distributions of *Bridelia* on plot C and *Cussonia* on plots C and I).

8. Conclusion

The correction for edge effect in Diggle's tests based on the G and F functions with EEC is too prudent and should either be ignored (dw^* - and dx^* -based tests), as in the procedure suggested by Diggle (1983) and studied in detail here, or replaced by more sophisticated corrections leading to a lower loss in data, like those proposed by Ripley for his K function (Ripley, 1981), or others (Haase, 1995; Ripley, 1988; Baddeley and Gill, 1997). Ignoring edge-effect correction leads to a gain in power for small samples sizes, which enables analysis of very small samples of spatial patterns with a sensitivity comparable to or greater than that of the test based on Ripley's K function. Based on this result, it seems that, when analysing real data, the three tests (based on dw^* , dx^* , and L_m) should be used simultaneously since they are sensitive to slightly different types of processes.

ACKNOWLEDGEMENTS

We would like to thank two anonymous referees for very constructive comments on the first version of this paper.

RÉSUMÉ

Les méthodes d'analyse des répartitions spatiales basées sur les fonctions F et G proposées par Diggle peuvent inclure ou non une correction pour les effets de bords. Nous illustrons par des simulations et sur des données écologiques que les tests ignorant cette correction ont une puissance plus élevée pour les petites tailles d'échantillon que les tests avec correction. Les tests sans correction peuvent détecter l'écart à la distribution Poissonienne pour des échantillons de taille plus petite (jusqu'à 10 points), alors que les tests classiques sont limités à 30. Ces résultats sont illustrés par un exemple utilisant des données provenant de cartes de répartition des arbres dans des savanes d'Afrique de l'Ouest. Les données ont été recueillies sur de petites parcelles (0,25 ha) où les densités d'arbres par espèce étaient en général < 20 . Pour l'une des espèces à répartition nettement agrégative, les tests sans correction ont permis de rejeter l'hypothèse nulle sur 3 parcelles sur 5, alors que les tests avec correction ne l'ont rejeté que sur une parcelle.

REFERENCES

- Baddeley, A. J. and Gill, R. D. (1997). Kaplan–Meier estimators for interpoint distance distributions of spatial point processes. *Annals of Statistics* **25**, 263–292.
- Baddeley, A. J., Moyeed, R. A., Howard, C. V., and Boyde, A. (1993). Analysis of a three-dimensional point pattern with replication. *Applied Statistics* **42**, 641–668.
- Barot, S., Gignoux, J., and Menaut, J.-C. (1998a). Spatial patterns in a humid savanna. I. Use of nutrient rich patches by a palm tree. *Ecology*, submitted.
- Barot, S., Gignoux, J., and Menaut, J.-C. (1998b). Spatial patterns in a humid savanna. II. Dispersal and intraspecific competition in a palm tree. *Ecology*, submitted.
- Cressie, N. A. C. (1991). *Statistics for Spatial Data*. New York: Wiley.
- Diggle, P. J. (1979). On parameter estimation and goodness-of-fit testing for spatial point patterns. *Biometrics* **35**, 87–101.
- Diggle, P. J. (1983). *Statistical Analysis of Spatial Point Patterns*. London: Academic Press.
- Fisher, M. J. (1993). Fine-scale distributions of tropical animal mounds: A revised statistical analysis. *Journal of Tropical Ecology* **9**, 339–348.
- Forman, R. T. T. and Hahn, D. C. (1980). Spatial patterns of trees in a Caribbean semievergreen forest. *Ecology* **61**, 1267–1274.
- Haase, P. (1995). Spatial pattern analysis in ecology based on Ripley's K -function: Introduction and methods of edge correction. *Journal of Vegetation Science* **6**, 575–582.
- Haase, P., Pugnaire, F. I., Clark, S. C., and Incoll, L. D. (1996). Spatial patterns in a two-tiered semi-arid shrubland in southeastern Spain. *Journal of Vegetation Science* **7**, 527–534.
- Marriott, F. H. C. (1979). Barnard's Monte Carlo tests: How many simulations? *Applied statistics* **28**, 75–77.
- Menaut, J. C. and César, J. (1979). Structure and primary productivity of Lamto savannas, Ivory Coast. *Ecology* **60**, 1197–1210.

- Menaut, J. C., Gignoux, J., Prado, C., and Clobert, J. (1990). Tree community dynamics in a humid savanna of the Côte d'Ivoire : Modelling the effects of fire and competition with grass and neighbors. *Journal of Biogeography* **17**, 471-481.
- Miles, R. E. (1974). On the elimination of edge effects in planar sampling. In *Stochastic Geometry*, E. S. Harding and D. G. Kendall (eds). 228-247. London: Wiley.
- Moeur, M. (1993). Characterizing spatial patterns of trees using stem-mapped data. *Forest Science* **39**, 756-775.
- Ripley, B. D. (1977). Modelling spatial patterns. *Journal of the Royal Statistical Society, Series B* **39**, 178-212.
- Ripley, B. D. (1981). *Spatial Statistics*. New York: Wiley.
- Ripley, B. D. (1988). *Statistical Inference for Spatial Processes*. Cambridge: Cambridge University Press.
- Skarpe, C. (1991). Spatial patterns and dynamics of woody vegetation in an arid savanna. *Journal of Vegetation Science* **2**, 565-572.
- Steiner, R. W., Ribic, C. A., and Schatz, G. E. (1986). Testing for life historical changes in spatial patterns of four tropical tree species. *Journal of Ecology* **74**, 621-633.
- Szwagrzyk, J. (1990). A small scale cartographic approach to forest community structure: Problems and preliminary results. *XXXIII Symposium of I.A.V.S.*, Warsaw, 47-50.
- Upton, G. J. G. and Fingleton, B. (1985). *Spatial Data Analysis by Example*. New York: Wiley.

Received October 1996. Revised May and December 1997.

Accepted January 1998.