

Class-specific tests of spatial segregation based on nearest neighbor contingency tables

Elvan Ceyhan*

Department of Mathematics, Koç University, 34450 Sarıyer, Istanbul, Turkey

The spatial interaction between two or more classes might cause multivariate clustering patterns such as segregation or association, which can be tested using a nearest neighbor contingency table (NNCT). The null hypothesis is randomness in the nearest neighbor structure, which may result from *random labeling* (RL) or *complete spatial randomness* of points from two or more classes (which is henceforth called *CSR independence*). We consider Dixon's class-specific segregation test and introduce a new class-specific test, which is a new decomposition of Dixon's overall chi-squared segregation statistic. We analyze the distributional properties and compare the empirical significant levels and power estimates of the tests using extensive Monte Carlo simulations. We demonstrate that the new class-specific tests have comparable performance with the currently available tests based on NNCTs. For illustrative purposes, we use three example data sets and provide guidelines for using these tests.

Keywords and Phrases: association, clustering, completely mapped data, complete spatial randomness, random labeling, spatial point pattern.

1 Introduction

Spatial patterns have important implications in various fields, such as epidemiology, population biology and ecology. A *spatial point pattern* includes the locations of some measurements, such as the coordinates of trees in a region of interest, while a *marked spatial point pattern* defines the distribution of 'marks' or 'class labels' (such as species of the trees) to the locations of the points and perhaps are the most common type of spatial point patterns (GAVRIKOV and STOYAN, 1995; DIGGLE, 2003; SCHLATHER, RIBEIRO and DIGGLE 2004). These locations are referred to as *events* by some authors, in order to distinguish them from arbitrary points in the region of interest (DIGGLE, 2003). However, in this article, *points* will refer to the locations of events, as we only consider the locations of events. It is of practical interest to investigate the patterns of one type of points with respect to other types (see, for

*elceyhan@ku.edu.tr

example, PIELOU, 1961; WHIPPLE, 1980, DIXON, 1994, 2002a). For convenience and generality, we call the different types of points as ‘classes’, but the class can stand for any (qualitative) characteristic of an observation at a particular location. For example, the spatial segregation pattern has been investigated for *plant species* (DIGGLE, 2003), *age classes* of plants (HAMILL and WRIGHT 1986), *fish species* (HERLER and PATZNER 2005) and *sexes* of dioecious plants (NANAMI, KAWAGUCHI and YAMAKURA 1999). Many of the epidemiologic applications are for a two-class system of case and control labels (WALLER and GOTWAY 2004).

In the analysis of multivariate point patterns, the null pattern is usually one of the two (random) pattern types: *random labeling* (RL) or *complete spatial randomness* (CSR) of two or more classes (i.e. *CSR independence*). We consider two major types of spatial patterns as alternatives: *association* and *segregation*. *Association* occurs if the nearest neighbor (NN) of an individual is more likely to be from another class than to be from the same class. *Segregation* occurs if the NN of an individual is more likely to be of the same class as the individual than to be of another class (see, e.g. PIELOU, 1961).

Many tests of spatial segregation have been proposed in the literature (KULLDORFF 2006). These include comparison of Ripley’s K - or L -functions (RIPLEY, 2004), comparison of NN distances (CUZICK and EDWARDS 1990; DIGGLE 2003) and analysis of nearest neighbor contingency tables (NNCTs) which are constructed using the NN frequencies of classes (PIELOU 1961, MEACHER and BURDICK 1980). (PIELOU 1961) proposed tests (of segregation, symmetry, niche specificity, etc.) and later DIXON (1994) introduced an overall test of segregation, cell- and class-specific tests based on NNCTs for the two-class case and DIXON (2002a) extended his tests to multiclass case. PIELOU (1961) used the usual Pearson’s chi-squared test of independence for testing spatial segregation. Due to the ease of computation and interpretation, Pielou’s test of segregation has been used frequently (MEAGHER and BURDICK 1980) for both completely mapped or sparsely sampled data, although it is not appropriate for such data (MEAGHER and BURDICK 1980, DIXON (1994)). For example, Pielou’s test is used for testing the segregation between males and females in dioecious species (see, e.g. HERRERA 1988 and ARMSTRONG and IRVINE 1989) and between different species (GOOD and WHIPPLE 1982). DIXON (1994, 2002a) derived the correct sampling distribution of the cell counts in the NNCTs under RL and designed overall, cell- and class-specific tests of segregation. CEYHAN (2008c) proposed new cell-specific and overall segregation tests which are more robust to the differences in the relative abundance of classes and have better performance in terms of size and power. Dixon’s overall segregation test is a compound summary statistic and applicable to test (small-scale) multivariate spatial interaction between classes in a given study area; (base)class-specific test provides information on the NN distribution of a base class (if (X, Y) is a pair of points in which Y is the closest point to X , then X is the base point and Y is the NN point), while NN-class-specific test (which is introduced in this article) provides the (base) distribution of classes to which members of a class serve as NNs. Hence, the NNCT tests answer different questions; i.e.

they provide information about different aspects of the spatial interaction (at small scales) compared with each other and the other tests in literature.

In this article, we discuss Dixon's overall and class-specific test of segregation and introduce a new class-specific test. We only consider *completely mapped data*; i.e. for our data sets, the locations of all events in a defined space are observed. We provide the null and alternative patterns in section 2, describe the NNCTs in section 3, provide the base- and NN-class-specific tests in section 4, empirical significance levels of the tests under CSR independence in section 5, empirical power comparisons under the segregation and association alternatives in sections 6 and 7, respectively, examples in section 8, and our conclusions and guidelines for using the tests in section 9.

2 Null and alternative patterns

In the univariate (i.e. one-class) spatial point pattern analysis, the null hypothesis is usually *CSR* (DIGGLE 2003). A spatial pattern exhibits CSR if, given n events in its domain D , the events are an independent random sample from the uniform distribution on D . This implies that there is no spatial interaction; i.e. the locations of these points have no influence on one another.

To investigate the spatial interaction between two or more classes in a bivariate or multivariate process, usually there are two benchmark hypotheses: (i) *independence*, which implies two classes of points are generated by a pair of independent univariate processes and (ii) *random labeling (RL)*, which implies that the class labels are randomly assigned to a given set of locations in the region of interest (DIGGLE, 2003). Under CSR independence, points from each of the two classes satisfy the CSR pattern in the region of interest. On the other hand, RL is the pattern in which, given a fixed set of points in a region, class labels are assigned to these fixed points randomly so that the labels are independent of the locations. That is, CSR independence is a process defining the spatial distribution of event locations, while RL is a process, conditioned on locations, defining the distribution of labels on these locations.

The null hypothesis in this article is 'randomness in NN structure', which may result from – among other patterns – CSR independence or RL. That is, when the points are assumed to be uniformly distributed over the region of interest, then the null hypothesis we consider is

$$H_0 : \text{CSR independence}$$

and when only the labeling (marking) of a set of fixed points – the allocation of the points might be regular, aggregated, clustered, or of lattice type – is considered, our null hypothesis is

$$H_0 : \text{RL.}$$

The distinction between CSR independence and RL is very important when defining the appropriate null model for each empirical case; i.e. the null model depends on the particular ecological context. GOREAUD and PÉLISSIER (2003) assert that CSR (independence) implies that the two classes are *a priori* the result of different processes (e.g. individuals of different species or age cohorts), whereas RL implies that some processes affect *a posteriori* the individuals of a single population (e.g. diseased vs non-diseased individuals of a single species). Although CSR independence and RL are not the same, they lead to the same null model (i.e. randomness in NN structure) in tests using NNCT, which does not require spatially explicit information. We provide the differences in the proposed tests under either null hypotheses.

As spatial clustering alternatives, we consider two major types of spatial patterns: *association* and *segregation*. *Association* occurs if the NN of an individual is more likely to be from another class than to be of the same class as the individual. For example, in plant biology, the points from the two classes might represent the coordinates of two mutualistic plant species, so the species depend on each other to survive. As another example, the points from one class might be the geometric coordinates of parasitic plants exploiting the other plant whose coordinates are the points from the other class. *Segregation* occurs if the NN of an individual is more likely to be of the same class as the individual than to be from a different class; i.e. the members of the same class tend to be clumped or clustered (see, e.g. PIELOU, 1961). For instance, one type of plant might not grow well around another type of plant and vice versa. In plant biology, points from one class might represent the coordinates of trees from a species with large canopy, so that other plants (whose coordinates are the points from the other class) that need light cannot grow (well or at all) around these trees (see, for instance, DIXON, 1994; COOMES, REES and TURNBULL, 1999). As many different forms of segregation (and association) are possible, it is not possible to list all segregation types, but their existence can be tested by an analysis of the NN relationships between the classes (PIELOU, 1961).

3 Nearest neighbor contingency tables

Nearest neighbour contingency tables are constructed using the NN frequencies of classes. We describe the construction of NNCTs for two classes from a binomial spatial process; extension to multiclass case is straightforward. Consider two classes with labels $\{1, 2\}$. Let N_i be the number of points from class i for $i \in \{1, 2\}$ and n be the total sample size; so, $n = N_1 + N_2$. If we record the class of each point and the class of its NN, the NN relationships fall into four distinct categories: (1, 1), (1, 2), (2, 1) and (2, 2), where, in cell (i, j) , class i is the *base class*, while class j is the class of NN of a point from class i . That is, the n points constitute n (base, NN) pairs, which can be categorized with respect to the base and NN labels (row and column categories). Denoting N_{ij} as the frequency of cell (i, j) for $i, j \in \{1, 2\}$, we obtain the NNCT in Table 1 where C_j is the sum of column j ; i.e. number of times

Table 1. NNCT for two classes.

Base class	NN class		Sum
	Class 1	Class 2	
Class 1	N_{11}	N_{12}	n_1
Class 2	N_{21}	N_{22}	n_2
Sum	C_1	C_2	n

class j points serve as NNs for $j \in \{1, 2\}$. Furthermore, we adopt the convention that variables denoted by upper (lower) case letters are random (fixed) quantities. Hence, column sums and cell counts are random, while row sums and overall sum are fixed quantities in an NNCT. However, if we want to have the overall sum to be random also, we might consider that all points are from a Poisson spatial process.

Note that, under segregation, the diagonal entries, N_{ii} for $i = 1, 2$, tend to be larger than expected; while, under association, the off-diagonals tend to be larger than expected. The general two-sided alternative is that some cell counts are different than expected under the null hypothesis (i.e. CSR independence or RL).

To test for segregation, PIELOU (1961) used chi-squared test of independence which requires the independence between cell counts (and rows and columns also), which is violated under CSR independence or RL. This problem was first noted by MEAGHER and BURDICK (1980) who identify the main source of it to be reflexivity of (base, NN) pairs. A (base, NN) pair (X, Y) is *reflexive* if (Y, X) is also a (base, NN) pair. As an alternative, they suggest using Monte Carlo simulations for Pielou's test. DIXON (1994, 2002a) derived the appropriate asymptotic sampling distribution of cell counts using Moran join count statistics (MORAN 1948) and hence the appropriate test which also has a chi-squared distribution asymptotically. Furthermore, CEYHAN (2008b) demonstrated that Pielou's test is only appropriate when the NNCT is constructed from a random sample of (base, NN) pairs, otherwise it is liberal under CSR independence and RL.

4 Class-specific tests of spatial segregation

4.1 Dixon's class-specific test of spatial segregation

Dixon's overall test of segregation tests the hypothesis that all cell counts in the NNCT are equal to their expected values (DIXON, 1994). Under RL, these expected cell counts are as

$$E[N_{ij}] = \begin{cases} n_i(n_i - 1)/(n - 1), & \text{if } i = j, \\ n_i n_j / (n - 1), & \text{if } i \neq j, \end{cases} \tag{1}$$

where n_i is a realization of N_i , i.e. is the fixed sample size for class i for $i = 1, 2, \dots, q$. Observe that the expected cell counts depend only on the size of each class (i.e. row sums) but not on column sums. In the multiclass case with q classes, DIXON (2002a) suggests the quadratic form

$$C = (\mathbf{N} - E[\mathbf{N}])' \Sigma^{-} (\mathbf{N} - E[\mathbf{N}]) \tag{2}$$

where \mathbf{N} is the $q^2 \times 1$ vector of q rows of cell counts concatenated row-wise, $E[\mathbf{N}]$ is the vector of $E[N_{ii}]$ which are as in Eq. (1), Σ is the $q^2 \times q^2$ variance–covariance matrix for the cell count vector \mathbf{N} with diagonal entries being equal to $\text{var}[N_{ii}]$ and off-diagonal entries being $\text{cov}[N_{ij}, N_{kl}]$ for $(i, j) \neq (k, l)$. The explicit forms of the variance and covariance terms are provided in DIXON (2002a). Furthermore, Σ^{-} is a generalized inverse of Σ (SEARLE, 2006) and the prime symbol stands for the transpose of a vector or matrix. The overall test statistic C can be partitioned into q class-specific (or species-specific) test statistics. Consider the NNs of type i base points, i.e. row i in the NNCT denoted by $\mathbf{N}_i = (N_{i1}, N_{i2}, \dots, N_{iq})'$. Then

$$C_B(i) = (\mathbf{N}_i - E[\mathbf{N}_i])' \Sigma_i^{-} (\mathbf{N}_i - E[\mathbf{N}_i]) \tag{3}$$

where $E[\mathbf{N}_i]$ is the vector of expected cell counts for row i and Σ_i^{-} is the generalized inverse of variance–covariance matrix of \mathbf{N}_i . Asymptotically $C_B(i)$ has $\chi^2_{(q-1)}$ distribution whose degrees of freedom is $q - 1$ as Σ_i has rank $(q - 1)$. Furthermore, $C_B(i)$ are not independent, hence do not sum to C . Each $C_B(i)$ tests whether each row i is similar to the one under RL; that is, the frequencies of the NNs for *base class* i is as expected under RL:

$$H_0 : E[\mathbf{N}_i] = (E[N_{i1}], E[N_{i2}], \dots, E[N_{iq}])' \tag{4}$$

Therefore, we call this type of class-specific test as *base class-specific test of segregation* for class i , and hence the notation $C_B(i)$.

The variance–covariance matrix for the tests in Eqs (2) and (3) depend on the quantities Q and R (DIXON 1994), where Q is the number of points with shared NNs, which occurs when two or more points share an NN and R is twice the number of reflexive pairs. Then $Q = 2(Q_2 + 3Q_3 + 6Q_4 + 10Q_5 + 15Q_6)$ where Q_i is the number of points that serve as an NN to other points i times. Notice that, under RL, Q and R are fixed quantities, as they depend only on the location of the points, not the types of NNs.

4.2 NN class-specific tests of spatial segregation

Another way to partition C in Eq. (2) is by columns instead of rows. Consider the frequency of type j points serving as NN to all other types (including class j), i.e. *column* j , namely, $\mathbf{C}_j = (N_{1j}, N_{2j}, \dots, N_{qj})'$.

Under RL, for each column j , we have

$$C_{NN}(j) = (\mathbf{C}_j - E[\mathbf{C}_j])' \Sigma_j^{-1} (\mathbf{C}_j - E[\mathbf{C}_j]), \tag{5}$$

where $E[\mathbf{C}_j]$ is the vector of expected cell counts for column j and Σ_j^{-1} is the inverse of variance–covariance matrix of \mathbf{C}_j . The test statistics $C_{NN}(j)$ are asymptotically distributed as χ^2_q , as Σ_j has rank q . The test statistics $C_{NN}(j)$ are dependent, hence do not sum to C .

If significant, these class-specific tests imply that class j serves more (or less) frequently as NN to other classes (including class j itself) than expected under RL:

$$H_0 : E[C_j] = (E[N_{1j}], E[N_{2j}], \dots, E[N_{qj}])' . \tag{6}$$

So, we call this type of class-specific test as *NN class-specific test of segregation* for class j , and hence the notation $C_{NN}(j)$.

REMARK 1. The status of Q and R under CSR independence and RL: Under CSR independence, the above tests and discussion are as in the RL case, except that under RL the quantities Q and R are fixed, while under CSR independence they are random. That is, under CSR independence, $C_B(i)$ asymptotically has $\chi^2_{(q-1)}$ distribution and $C_{NN}(j)$ asymptotically has χ^2_q distribution conditional on Q and R , because the variances and covariances used in Eqs (3) and (5), and all the quantities depending on these expectations also depend on Q and R . The unconditional variances and covariances can be obtained by using the conditional expectation formula incorporating Q and R . Letting $W = (Q, R)'$, we have

$$\begin{aligned} \text{var}[N_{ij}] &= E_W[\text{var}[N_{ij} | W]] + \text{var}_W[E[N_{ij} | W]] \\ &= E_W[\text{var}[N_{ij} | W]] \end{aligned}$$

as $E[N_{ij} | W]$ is independent of Q and R . Similarly,

$$\begin{aligned} \text{cov}[N_{ij}, N_{kl}] &= E_W[\text{cov}[N_{ij}, N_{kl} | W]] + \text{cov}_W[E[(N_{ij}, N_{kl}) | W]] \\ &= E_W[\text{cov}[N_{ij}, N_{kl} | W]] \end{aligned}$$

as $E[(N_{ij}, N_{kl}) | W]$ is independent of Q and R . Furthermore, $\text{var}[N_{ij} | W]$ and $\text{cov}[N_{ij}, N_{kl} | W]$ are linear in Q and R (DIXON, 2002a). Hence, the unconditional variances and covariances can be found by replacing Q and R with their expectations.

Unfortunately, given the difficulty of calculating the expectation of Q under CSR independence, it is reasonable and convenient to use test statistics employing the conditional variances and covariances even when assessing their behavior under the CSR independence model. Cox (1981) calculated analytically that $E[R | N] = 0.6215N$ for a planar Poisson process. Alternatively, we can estimate the expected values of Q and R empirically. For example, for homogeneous planar Poisson pattern, we have $E[Q | N] \approx 0.63N$ and $E[R | N] \approx 0.62N$ (estimated empirically by 1,000,000 Monte Carlo simulations for various values of N on unit square). Notice that $E[R | N]$ agrees with the analytical result of Cox (1981).

REMARK 2. Comparison of base class- and NN Class- specific tests: Notice that the base-class-specific test for class i tests whether the NN frequencies of base class i follows the expected frequencies under CSR independence (conditional on Q and R) or RL. That is, base class-specific test for class (or species) i is concerned with how other classes interact as NNs with base class i in the NN structure. For example, if

$(q - 1)$ plant species are introduced to a region with species i being already there, the base class-specific test for species i is more appropriate and informative as it provides information on how the newly introduced species interact with the already existent base species i . On the other hand, the NN-class-specific test for class j tests whether the base frequencies for which class j serves as NN follows the expected frequencies under CSR independence (conditional on Q and R) or RL. So, NN class-specific test for class j is concerned with how class j points interact as NNs with other base classes in the NN structure. For example, if species j was introduced to a region with $(q - 1)$ species already present, NN class-specific test for species j is more appropriate and informative, as it provides information on how the newly introduced species interact with the already existent $(q - 1)$ base species.

4.3 The two-class case

In the two-class case, Dixon (1994) calculates $Z_{ii} = (N_{ii} - E[N_{ii}]) / \sqrt{\text{var}[N_{ii}]}$ for both $i \in \{1, 2\}$ and then combines these test statistics into a statistic that is asymptotically distributed as χ^2_2 . The suggested test statistic is given by

$$C = \mathbf{Y}'\Sigma^{-1}\mathbf{Y} = \begin{bmatrix} N_{11} - E[N_{11}] \\ N_{22} - E[N_{22}] \end{bmatrix}' \begin{bmatrix} \text{var}[N_{11}] & \text{cov}[N_{11}, N_{22}] \\ \text{cov}[N_{11}, N_{22}] & \text{var}[N_{22}] \end{bmatrix}^{-1} \begin{bmatrix} N_{11} - E[N_{11}] \\ N_{22} - E[N_{22}] \end{bmatrix}.$$

Notice that this is equivalent to

$$C = \frac{Z_{AA}^2 + Z_{BB}^2 - 2rZ_{AA}Z_{BB}}{1 - r^2}$$

where

$$Z_{AA} = \frac{N_{11} - E[N_{11}]}{\sqrt{\text{var}[N_{11}]}} \quad , \quad Z_{BB} = \frac{N_{22} - E[N_{22}]}{\sqrt{\text{var}[N_{22}]}} \quad , \quad \text{and}$$

$$r = \text{cov}[N_{11}, N_{22}] / \sqrt{\text{var}[N_{11}]\text{var}[N_{22}]}.$$

The overall test statistic C can be partitioned into two base class-specific test statistics $C_B(i)$ which asymptotically have χ^2_1 .

When C is partitioned by columns \mathbf{C}_j , we get the NN class-specific test $C_{\text{NN}}(j)$ which are asymptotically distributed as χ^2_2 . However, in the two-class case, $C_{\text{NN}}(1) = C_{\text{NN}}(2) = C$, as each column \mathbf{C}_j contains all the information about the NNCT.

REMARK 3. The multiclass case with $q > 2$ classes: In the multiclass case with more than two classes, the overall test statistic C can be partitioned into q base class-specific test statistics $C_B(i)$ which asymptotically have χ^2_{q-1} . When C is partitioned by columns \mathbf{C}_j , we get q NN class-specific tests $C_{\text{NN}}(j)$ which are asymptotically distributed as χ^2_q . In the case of $q > 2$, $C_{\text{NN}}(j)$ and $C_{\text{NN}}(l)$ are very likely to be different for $j \neq l$ and $C_{\text{NN}}(j)$ and C are also very likely to be different for each j .

4.4 Asymptotic Structures for the NNCT-Tests

There are two major types of asymptotic structures for spatial data (LAHIRI, 1996). In the first, any two observations are required to be at least a fixed distance apart; hence, as the number of observations increase, the region on which the process is observed eventually becomes unbounded. This type of sampling structure is called *increasing domain asymptotics*. In the second type, the region of interest is a fixed bounded region and more and more points are observed in this region. Hence, the minimum distance between data points tends to zero as the sample size tends to infinity. This type of structure is called *infill asymptotics*, due to CRESSIE (1993).

Under RL, the sampling structure in our asymptotic sampling distribution could be either one of increasing domain or infill asymptotics, as we only consider the class sizes and hence the total sample size tending to infinity regardless of the size of the study region. That is, under RL, the locations on which labels are assigned randomly could be realizations from a process whose asymptotic structure follows either infill or increasing domain asymptotics. On the other hand, under CSR independence, we cannot have the increasing domain asymptotics in the usual sense, as, for any bounded subset of the region, the points will not necessarily follow the uniform distribution as there is a restriction on the interpoint distances. However, if we let the area of the region go to infinity while uniformness is preserved for any bounded subset of the region, this type of ‘modified’ increasing domain asymptotics is applicable for the CSR independence pattern. Furthermore, under CSR independence, the number of points from uniform distribution on a bounded region going to infinity will give rise to the infill asymptotics. However, the values and distributions of Q and R are affected by infill and increasing domain asymptotics. In both asymptotic structures, under CSR independence, the tests will still be conditional on Q and R , whose asymptotic distribution will also depend on the type of the asymptotics. For the actual conditional result, the only possible asymptotics for which the current arguments might be appropriate is what might be called *pattern stamp asymptotics*. In this form of asymptotics, the study region (assumed to be rectangular or at least something that will tile the plane) and its locations are repeated (stamped). This multiple stamping of the same pattern of locations maintains Q and R at fixed constants. However, this is not either of the usual asymptotics for the spatial data nor a realistic case in practice.

Theorem 1. (Consistency) *Under RL, the overall test in Eq. (2) which rejects for $C > \chi_{q(q-1)}^2(1-\alpha)$ where $\chi_{q(q-1)}^2(1-\alpha)$ is the 100(1- α)th percentile of $\chi_{q(q-1)}^2$ distribution, the test against H_0 in Eq. (4) which rejects for $C_B(i) > \chi_{(q-1)}^2(1-\alpha)$ and the test against H_0 in Eq. (6) which rejects for $C_{NN}(j) > \chi_q^2(1-\alpha)$ are consistent. Under CSR independence, consistency follows conditional on Q and R .*

PROOF. The consistency of the overall test is proved in (CEYHAN, 2008b). Under RL, the null hypothesis in Eq. (4) is equivalent to $H_0: \lambda = 0$ where λ is the non-centrality parameter for the $\chi^2_{(q-1)}$ distribution. Any deviation from the null case corresponds to $H_a: \lambda > 0$, then under any specific alternative the power tends to unity by the consistency of usual Chi-squared tests. The consistency of $C_{NN}(j)$ under RL and of both tests under CSR independence follow similarly.

5 Empirical significance levels of the tests under CSR independence

5.1 The two-class case

First, we consider the two-class case with classes X and Y which stand for classes 1 and 2 respectively. We generate n_1 points from class X and n_2 points from class Y both of which are independently uniformly distributed on the unit square, $(0, 1) \times (0, 1)$, for some combinations of $n_1, n_2 \in \{10, 30, 50, 100\}$. Hence, all X points are independent of each other and so are Y points; and X and Y are independent data sets. This will imply randomness in the NN structure, which is the null hypothesis for our NNCT tests. Thus, we simulate the CSR independence as the null case. We repeat the sample generation $N_{mc} = 10,000$ times for each sample size combination for a reasonable precision in the results within a reasonable time period. For each Monte Carlo replication, we construct the NNCT for classes X and Y , then compute the base class- and NN class-specific tests for each class and Dixon's overall segregation test. The empirical sizes are calculated as the ratio of significant results to the number of Monte Carlo replicates. The nominal significance level used in all these tests is $\alpha = 0.05$.

We present the empirical significance levels for the NNCT tests in Table 2, where $\hat{\alpha}_i^B$ is the empirical significance level for (Dixon's) base class-specific test and $\hat{\alpha}_i^{NN}$ is for the NN class-specific test for $i \in \{1, 2\}$ and $\hat{\alpha}_D$ is for Dixon's overall test of segregation. The empirical sizes significantly smaller (larger) than 0.05 are marked with c (l), which indicate that the corresponding test is conservative (liberal). The asymptotic normal approximation to proportions are used in determining the significance of the deviations of the empirical sizes from the nominal level of 0.05. For the tests regarding the proportions, we also use $\alpha = 0.05$ to test against empirical size being equal to 0.05. Notice that in the two-class case $\hat{\alpha}_1^B \neq \hat{\alpha}_2^B$ but $\hat{\alpha}_1^{NN} = \hat{\alpha}_2^{NN} = \hat{\alpha}_D$; so, only $\hat{\alpha}_1^{NN}$ is presented. The empirical sizes are also plotted in Figure 1 where the horizontal lines are the nominal level of 0.05 and upper and lower limits for the empirical size (i.e. 0.0464 and 0.0536).

Observe that (Dixon's) base class-specific test for class Y (i.e. the larger class in unequal sample size combinations) yields about the desired significance levels in rejecting H_0 : CSR independence for all sample size combinations. However, when one of the samples is small the tests are conservative, but base class-specific test for class X (i.e. the class with the smaller sample size) is more conservative than others.

Table 2. The empirical size estimates for the NNCT tests in the two-class case under H_0 : CSR independence with $N_{mc} = 10,000$, n_1, n_2 in $\{10, 30, 50, 100\}$ at the nominal level of $\alpha = 0.05$.

Sizes (n_1, n_2)	Base		NN*
	$\hat{\alpha}_1^B$	$\hat{\alpha}_2^B$	$\hat{\alpha}_1^{NN}$
(10,10)	0.0454 ^c	0.0465	0.0432 ^c
(10,30)	0.0306 ^c	0.0485	0.0440 ^c
(10,50)	0.0270 ^c	0.0464	0.0482
(30,30)	0.0507	0.0505	0.0464
(30,50)	0.0590 ^l	0.0522	0.0443 ^c
(50,50)	0.0465	0.0469	0.0508
(50,100)	0.0601 ^l	0.0533	0.0560 ^l
(100,100)	0.0493	0.0463 ^c	0.0504

$\hat{\alpha}_i^B$ stands for the empirical significance level for (Dixon's) base class-specific test for class $i=1, 2$ and $\hat{\alpha}_1^{NN}$ is for the NN class-specific test for class 1. (c: empirical size significantly less than 0.05; i.e. the test is conservative. l: empirical size significantly larger than 0.05; i.e. the test is liberal.

*Only $\hat{\alpha}_1^{NN}$ is presented as $\hat{\alpha}_1^{NN} = \hat{\alpha}_2^{NN} = \hat{\alpha}_D$. base, base class-specific test. NN, NN class-specific test.

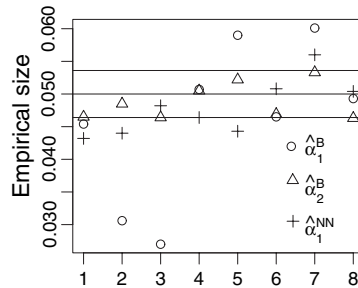


Fig. 1. The empirical size estimates of Dixon's base class-specific test for class 1 (circles, \circ), for class 2 (triangles, \triangle) and NN class-specific test (pluses, $+$) under CSR independence. Classes 1 and 2 represent classes X and Y respectively. The horizontal lines are located at 0.0464 (upper threshold for conservativeness), 0.0500 (nominal level) and 0.0536 (lower threshold for liberalness). The horizontal axis labels are: 1=(10, 10), 2=(10, 30), 3=(10, 50), 4=(30, 30), 5=(30, 50), 6=(50, 50), 7=(50, 100), 8=(100, 100).

For large classes with unequal sizes (i.e. with different relative abundances), the base class-specific test tends to be slightly liberal for the smaller class.

5.2 The three-class case

In the two-class case, NN class-specific tests are equal to the overall test of segregation. However, for the q -class case with $q > 2$, neither $C_{NN}(j)$ are likely to be equal to each other nor any of them are likely to be equal to C for $j=1, 2, \dots, q$. Therefore, in order to better compare the performance of NN class-specific tests with base class-specific tests, we consider the three-class case with classes X, Y and Z (which

stand for classes 1, 2 and 3 respectively) under CSR independence. We generate n_1, n_2 and n_3 points distributed independently uniformly on the unit square $(0, 1) \times (0, 1)$ from classes X, Y and Z respectively, as in section 5.1. We generate data points for some combinations of $n_1, n_2, n_3 \in \{10, 30, 50, 100\}$; and for each sample size combination, the empirical sizes and the significance of their deviation from 0.05 are also calculated as in section 5.1.

The empirical significance levels for the three-class case are presented in Figure 2, where horizontal lines are as in Figure 1. Notice that when at least one class is small (i.e. $n_i \leq 10$) tests are usually conservative, especially the base class-specific tests for the smaller classes. When all classes are large, the tests are about the desired level, and conservative for a few sample size combinations. For the three-class case, we conclude that the NN class-specific tests exhibit better performance than the base class-specific tests in terms of empirical size.

REMARK 4. *Empirical Significance Levels under RL:* Recall that the segregation tests we consider are conditional under the CSR independence pattern. To evaluate their empirical size performance better, Monte Carlo simulations under various RL cases for two- and three-class cases are also performed in CEYHAN (2008a). It is shown that the empirical sizes are about the same under the CSR independence and RL patterns. However, the tests are conservative when at least one sample is small, regardless of whether the null case is CSR independence or RL. Hence, the conservativeness or liberalness of the tests for small sample sizes does not result from

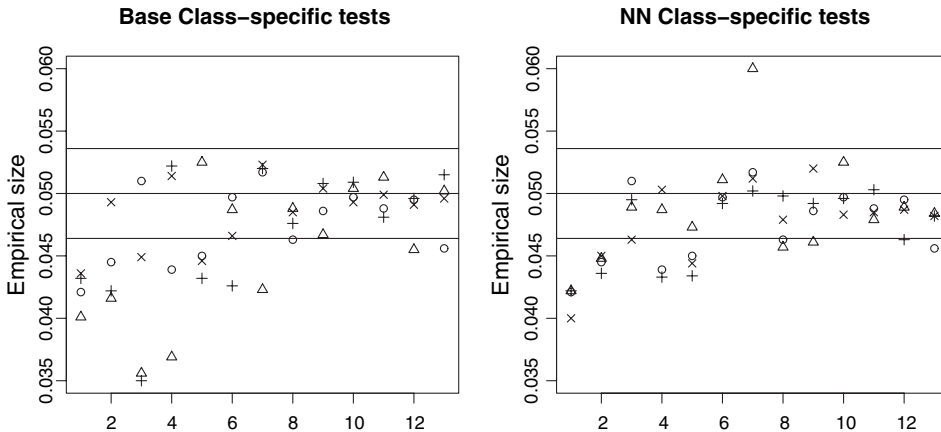


Fig. 2. The empirical size estimates of Dixon’s overall test (circles, ○) and base class-specific tests (left), and the NN class-specific tests (right) under the CSR independence pattern in the three-class case. The triangles (△), pluses (+) and crosses (×) are for classes 1, 2 and 3 (i.e. X, Y and Z), respectively. The horizontal lines are located at 0.0464 (upper threshold for conservativeness), 0.0500 (nominal level) and 0.0536 (lower threshold for liberalness). The horizontal axis labels are 1=(10, 10, 10), 2=(10, 10, 30), 3=(10, 10, 50), 4=(10, 30, 30), 5=(10, 30, 50), 6=(30, 30, 30), 7=(10, 50, 50), 8=(30, 30, 50), 9=(30, 50, 50), 10=(50, 50, 50), 11=(50, 50, 100), 12=(50, 100, 100), 13=(100, 100, 100).

conditioning on Q and R under the CSR independence pattern, as similar results are obtained under the RL pattern also.

REMARK 5. *Monte Carlo critical values:* When sample sizes are small so that some cell count(s) are expected to be ≤ 5 with a high probability, then it will not be appropriate to use the asymptotic approximation (hence the asymptotic critical values) for the overall and class-specific tests of segregation (DIXON, 1994). That is, when some cell counts are small, the tests tend to be either liberal or conservative. In order to better evaluate the empirical power performance of the tests, for each sample size combination, the test statistics at each Monte Carlo simulation under the CSR independence cases are recorded and the 95th percentiles of these test statistics at each sample size combination are found and used as ‘Monte Carlo critical values’ for the power estimation in CEYHAN (2008a). It is seen that the results based on Monte Carlo critical values are very similar to the ones based on the asymptotic critical values, hence are not presented here.

6. Empirical power analysis under segregation alternatives

6.1 The two-class case

For the segregation alternatives, we generate

$$X_i \stackrel{\text{i.i.d.}}{\sim} \mathcal{U}((0, 1 - s) \times (0, 1 - s)) \quad \text{and} \quad Y_j \stackrel{\text{i.i.d.}}{\sim} \mathcal{U}((s, 1) \times (s, 1))$$

for $i = 1, \dots, n_1$ and $j = 1, \dots, n_2$. Notice that the level of segregation is determined by the magnitude of $s \in (0, 1)$. We consider the following three segregation alternatives:

$$H_S^I : s = 1/6, \quad H_S^{II} : s = 1/4, \quad \text{and} \quad H_S^{III} : s = 1/3.$$

Observe that, from H_S^I to H_S^{III} (i.e. as s increases), the segregation between X and Y gets stronger in the sense that X and Y points tend to form one-class clumps or clusters. We calculate the power estimates using the asymptotic critical values based on the corresponding chi-squared distributions and present them in Figure 3, where $\hat{\beta}_i^B$ are for (Dixon’s) base class-specific tests for classes $i = 1, 2$, and $\hat{\beta}_1^{NN}$ is for NN class-specific test for class 1, $\hat{\beta}^D$ is for Dixon’s overall segregation test. We present only $\hat{\beta}_1^{NN}$ as $\hat{\beta}_1^{NN} = \hat{\beta}_2^{NN} = \hat{\beta}^D$. Observe that, for both class-specific tests, as $n = (n_1 + n_2)$ gets larger, the power estimates get larger; for similar $n = (n_1 + n_2)$ values, the power estimate is larger for classes with similar relative abundance (i.e. for similar sample sizes $n_1 \approx n_2$); and as the segregation gets stronger, the power estimates get larger for each sample size combination. When sample sizes are similar, NN class-specific tests tend to have higher power. On the other hand, when relative abundances (i.e. sample sizes) are very different, the base class-specific test for classes X and Y have the highest and lowest power estimates respectively.

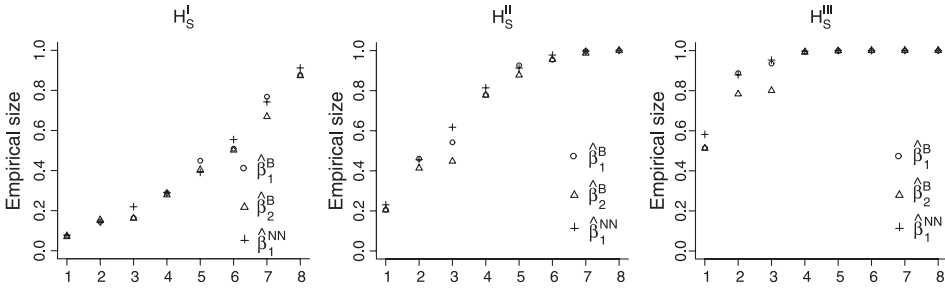


Fig. 3. The empirical power estimates for base class-specific test for class 1 (circles, \circ), for class 2 (triangles, \triangle) and for NN class-specific test (pluses, $+$) under the segregation alternatives in the two-class case. $\hat{\beta}_i^B$ stand for (Dixon’s) base class-specific tests for classes $i=1, 2$, and $\hat{\beta}_1^{NN}$ stands for NN class-specific test for class 1. Classes 1 and 2 represent classes X and Y , respectively. The horizontal axis labels are $1=(10, 10)$, $2=(10, 30)$, $3=(10, 50)$, $4=(30, 30)$, $5=(30, 50)$, $6=(50, 50)$, $7=(50, 100)$, $8=(100, 100)$.

6.2 The three-class case

For the segregation alternatives, we generate

$$X_i \stackrel{i.i.d.}{\sim} \mathcal{U}((0, 1 - 2s) \times (0, 1 - 2s)), \quad Y_j \stackrel{i.i.d.}{\sim} \mathcal{U}((2s, 1) \times (2s, 1)) \quad \text{and} \\ Z_l \stackrel{i.i.d.}{\sim} \mathcal{U}((s, 1 - s) \times (s, 1 - s))$$

for $i=1, \dots, n_1, j=1, \dots, n_2$ and $l=1, \dots, n_3$. Notice that the level of segregation is determined by the magnitude of $s \in (0, 1/2)$. We consider the following three segregation alternatives:

$$H_S^1 : s = 1/12, \quad H_S^2 : s = 1/8, \quad \text{and} \quad H_S^3 : s = 1/6.$$

Observe that, from H_S^1 to H_S^3 (i.e. as s increases), the segregation gets stronger in the sense that X , Y and Z points tend to form one-class clumps or clusters. Furthermore, under each segregation alternative, X and Y are more segregated compared with X and Z or Y and Z .

The empirical power estimates using the asymptotic critical values are presented in Figure 3. Observe that, at each sample size combination, as the segregation gets stronger, the power estimates get larger. Furthermore, the power estimates for the similar sample sizes tend to be larger compared with the different sample sizes when the total sizes $n=(n_1 + n_2 + n_3)$ are similar. The power estimates also confirm that the segregation between classes X and Y is stronger compared with the segregation between X and Z or between Y and Z .

For large samples, Dixon’s overall test tends to have the highest power estimates, but class-specific tests provide more information about the pattern. For all (small or large) similar relative abundance values, NN class-specific tests have slightly better power performance, while, for very different relative abundances, base class-specific tests have slightly better power performance. However, overall, we can conclude that

in the three-class case, NNCT tests have about the same performance in terms of power.

7 Empirical power analysis under association alternatives

7.1 The two-class case

For the association alternatives, we consider three cases. First, we generate $X_i \stackrel{\text{i.i.d.}}{\sim} \mathcal{U}((0, 1) \times (0, 1))$ for $i = 1, 2, \dots, n_1$. Then we generate Y_j for $j = 1, 2, \dots, n_2$ as follows. For each j , we pick an i randomly, then generate Y_j as $X_i + R_j(\cos T_j, \sin T_j)'$ where $R_j \stackrel{\text{i.i.d.}}{\sim} \mathcal{U}(0, r_y)$ with $r_y \in (0, 1)$ and $T_j \stackrel{\text{i.i.d.}}{\sim} \mathcal{U}(0, 2\pi)$. In the pattern generated, appropriate choices of r_y will imply association between classes X and Y . That is, it will be more likely to have (X, Y) or (Y, X) NN pairs than the same class NN pairs (i.e., (X, X) or (Y, Y)). The three values of r_y we consider constitute the following three association alternatives;

$$H_A^I : r_y = 1/4, \quad H_A^{II} : r_y = 1/7 \quad \text{and} \quad H_A^{III} : r_y = 1/10.$$

Observe that, from H_A^I to H_A^{III} (i.e. as r decreases), the association between X and Y gets stronger in the sense that X and Y points tend to occur together more and more frequently. By construction, for similar sample sizes the association between X and Y are at about the same degree as association between Y and X . For very different samples, smaller sample is associated with the larger, but the abundance of the larger, sample confounds its association with the smaller.

The empirical power estimates, based on the asymptotic critical values are plotted in Table 5. Observe that when $n_1 \approx n_2$, for both class-specific tests and Dixon's overall segregation test, as n_i gets larger, the power estimates get larger; and, as the association gets stronger, the power estimates get larger for each sample size combination.

When at least one class has a small size (i.e. $n_i \leq 10$), the base class-specific test for the smaller class, say class i , fails to detect any deviation from CSR independence for row i . However, the base class-specific test for the larger class is robust to the differences in relative abundance (i.e. sample sizes) and behaves as expected. Likewise, the NN class-specific test is not affected by the differences in relative abundance as long as the smaller class is associated with the larger (the corresponding simulation results are not presented). In our set-up, this corresponds to the case that class Y is much larger than class X . If, on the contrary, the smaller class is associated with the larger (i.e. class X is much larger than class Y), then NN class-specific test has very poor power performance.

When sample sizes are similar, the base class-specific tests for class Y (the class less associated with the other) have higher power under weak association, while NN class-specific tests have higher power under strong association. When sample sizes are very different, the base class-specific tests for classes Y and X have highest

and lowest power estimates respectively. Base class-specific test for class i measures the association of other classes with class i , while NN class-specific test for class j measures the association of class j with other classes.

7.2 The three-class case

For the association alternatives, we also consider three cases for which we generate X_i for $i=1, 2, \dots, n_1$ and Y_j for $j=1, 2, \dots, n_2$ as in section 7.1 and Z_l for $l=1, 2, \dots, n_3$ as follows. For each l , we pick an i randomly, then generate $R_l^Z \stackrel{i.i.d.}{\sim} \mathcal{U}(0, r_z)$ with $r_z \in (0, 1)$ and $U_l \stackrel{i.i.d.}{\sim} \mathcal{U}(0, 2\pi)$ and set $Z_l := X_i + R_l^Z(\cos U_l, \sin U_l)'$.

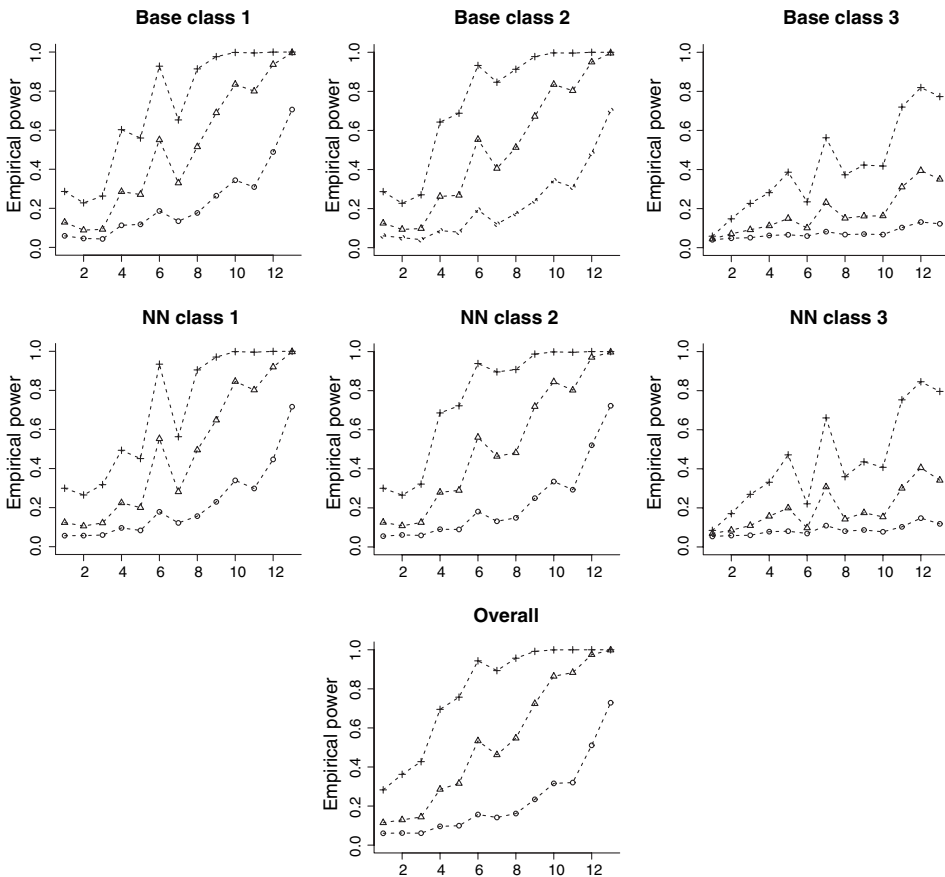


Fig. 4. The empirical power estimates based on asymptotic critical values for Dixon's overall test, base class specific tests and NN class-specific tests under the segregation alternatives H_S^1 (circles, \circ), H_S^2 (triangles, Δ) and H_S^3 (pluses, $+$) in the three-class case. Classes 1, 2 and 3 represent classes X , Y , and Z , respectively. The horizontal axis labels are 1=(10, 10, 10), 2=(10, 10, 30), 3=(10, 10, 50), 4=(10, 30, 30), 5=(10, 30, 50), 6=(30, 30, 30), 7=(10, 50, 50), 8=(30, 30, 50), 9=(30, 50, 50), 10=(50, 50, 50), 11=(50, 50, 100), 12=(50, 100, 100), 13=(100, 100, 100).

In the pattern generated, appropriate choices of r_y (and r_z) values will imply association between classes X and Y (and X and Z). The three association alternatives are as

$$H_A^1 : r_y = 1/7, r_z = 1/10, \quad H_A^2 : r_y = 1/10, r_z = 1/20, \quad H_A^3 : r_y = 1/13, r_z = 1/30.$$

Observe that, from H_A^1 to H_A^3 (i.e. as r_y decreases), the association between X and Y gets stronger in the sense that Y points tend to be found more and more frequently around the X points. The same holds for X and Z points as r_z decreases. Furthermore, by construction, classes X and Z are more associated compared with classes X and Y . On the other hand, classes Y and Z are not associated, but perhaps are mildly segregated (Figures 4 and 5).

We present the empirical power estimates using the asymptotic critical values in Figure 6. Observe that for each sample size combination, as the association gets stronger, the power estimates tend to be higher. For all sample size combinations, NN class-specific tests have better performance in terms of power in the sense that the highest power estimate for the three NN class-specific tests tends to be larger than the highest power estimate for the three base class-specific test at each sample size combination. Moreover, among the base class-specific tests, the one for class 3 (i.e. class Z) has the highest power, while among the NN class-specific tests, the one for class 1 (i.e. class X) has the highest power estimates.

Remark 6. Edge correction for the NNCT tests. The CSR independence pattern assumes that the study region is unbounded for the analyzed pattern, which is not the case in practice. So, the edge (or boundary) effects might confound the test results in the analysis of empirical (i.e. bounded) data sets if the null pattern is CSR independence and much effort has gone into the development of edge correction methods (YAMADA and ROGENSEN, 2003). Two correction methods for the edge effects on NNCT tests, namely *buffer zone correction* and *toroidal correction*, are investigated in CEYHAN (2007, 2008b) where it is shown that the empirical sizes of the NNCT tests are not affected by the toroidal edge correction under CSR

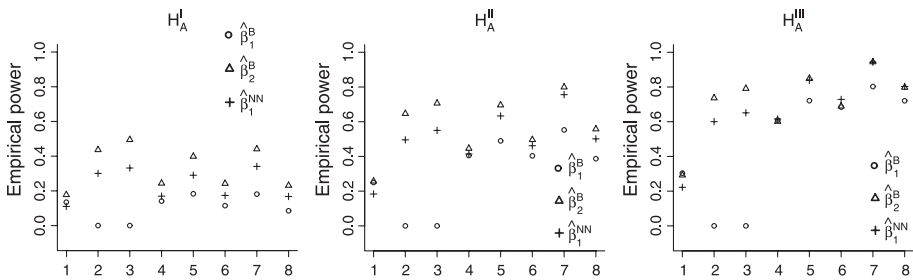


Fig. 5. The empirical power estimates for Dixon's tests (solid lines) and new tests (dashed lines) under the association alternatives in the two-class case. The power notation is as in Figure 3. The horizontal axis labels: 1=(10, 10), 2=(10, 30), 3=(10, 50), 4=(30, 30), 5=(30, 50), 6=(50, 50), 7=(50, 100), 8=(100, 100).

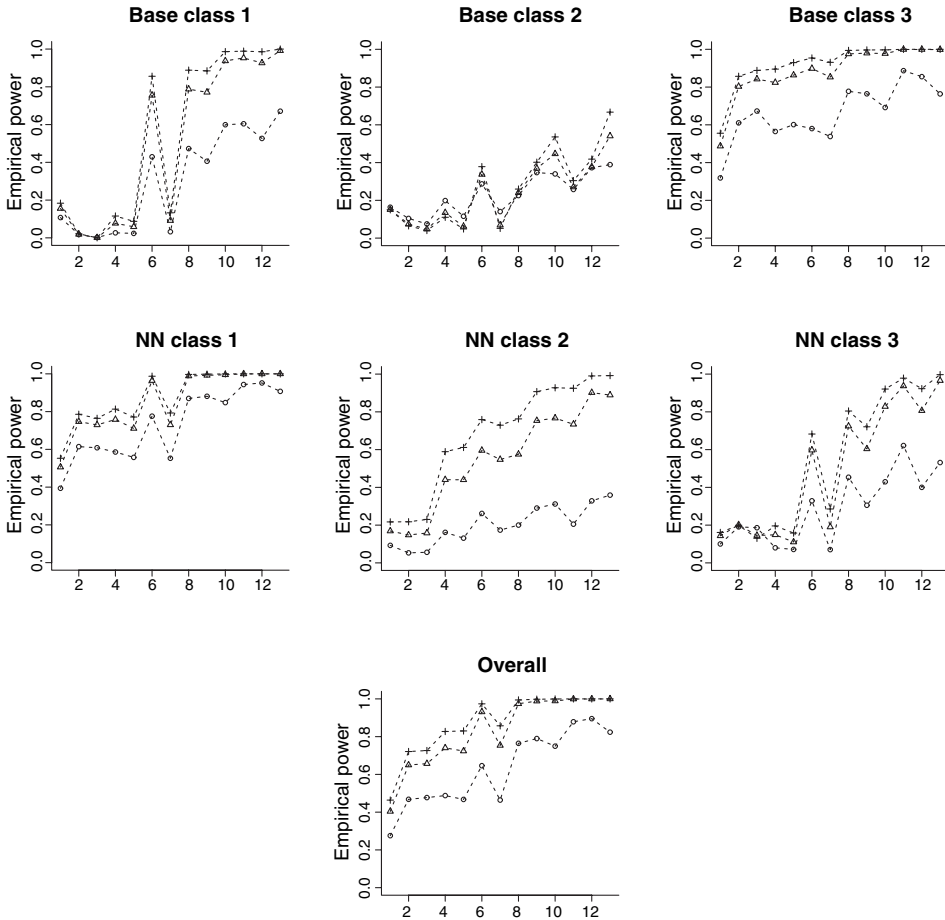


Fig. 6. The empirical power estimates based on asymptotic critical values for Dixon's overall test, base class-specific tests and NN class-specific tests under the association alternatives H_A^1 (circles, \circ), H_A^2 (triangles, \triangle) and H_A^3 (pluses, $+$) in the three-class case. The horizontal axis labels are as in Figure 4. Classes 1, 2, and 3 represent classes X , Y and Z respectively.

independence. However, toroidal correction is biased for non-CSR patterns. In particular if the pattern outside the plot (which is often unknown) is not the same as that inside it, it yields questionable results (HAASE 1995 and YAMADA and ROGERSEN 2003). The bias is more severe especially when there are clusters around the edges. Under CSR independence, the (outer) buffer zone edge correction method seems to have a slightly stronger influence on the tests compared with toroidal correction. However, for these tests, buffer zone correction does not change the sizes significantly for most sample size combinations. This is in agreement with the findings of BAROT, GIGNOUX and MENANT (1999) who say that NN methods only require a small buffer area around the study region. A large buffer area does not help much as one only needs to be able to see far enough away from an event to find its NN.

Once the buffer area extends past the likely NN distances (i.e. about the average NN distances), it is not adding much helpful information for NNCTs. Hence, we recommend inner or outer buffer zone correction for NNCT tests with the width of the buffer area being about the average NN distance. We do not recommend larger buffer areas, as they are wasteful with little additional gain.

8 Examples

We illustrate the tests on three example data sets: an ecological data set (swamp tree data), an epidemiological data set (leukemia data) and an MRI data set (pyramidal neuron data).

8.1 Swamp tree data

GOOD and WHIPPLE (1982) considered the spatial patterns of tree species along the Savannah River, South Carolina, USA. From this data, DIXON (2002b) used a single $50 \text{ m} \times 200 \text{ m}$ rectangular plot to illustrate NNCT methods. All live or dead trees with 4.5 cm or more dbh (diameter at breast height) were recorded together with their species. Hence, it is an example of a realization of a marked multivariate point pattern. The plot contains 13 different tree species, four of which comprising over 90% of the 734 tree stems. The remaining tree stems were categorized as ‘other trees’. The plot consists of 215 water tupelo (*Nyssa aquatica*), 205 black gum (*Nyssa sylvatica*), 156 Carolina ash (*Fraxinus caroliniana*), 98 bald cypress (*Taxodium distichum*) and 60 stems of eight additional species (i.e. other species). A 5×5 NNCT analysis is conducted for this data set. If segregation among the less frequent species were important, a more detailed 13×13 NNCT analysis should be performed. The locations of these trees in the study region are plotted in Figure 7 and the corresponding 5×5 NNCT together with percentages based on row and column sums are provided in Table 3. For example, for black gum as the base species and Carolina ash as the NN species, the cell count is 26 which is 13% of the 205 black gums (which is 28% of all trees) and 15% of the 171 times Carolina ashes serves as NN (which is 23% of all trees). Observe that the percentages and Figure 7 are suggestive of segregation for all tree species, especially for Carolina ashes, water tupelos, black gums and the other trees, as the observed percentage of species with themselves as the NN is much larger than the marginal (row or column) percentages.

The locations of the tree species can be viewed *a priori* resulting from different processes; so, the more appropriate null hypothesis is the CSR independence pattern. Hence, our inference will be a conditional one (see Remark 1). We calculate $Q=472$ and $R=454$ for this data set. We present the overall test of segregation, class-specific test statistics and the associated P -values in Table 4, where p_{asy} stands for the P -value based on the asymptotic approximation (i.e. the corresponding chi-squared distribution), p_{mc} is the P -value based on 10,000 Monte Carlo replication of the CSR independence pattern in the same plot and p_{rand} is based on Monte Carlo

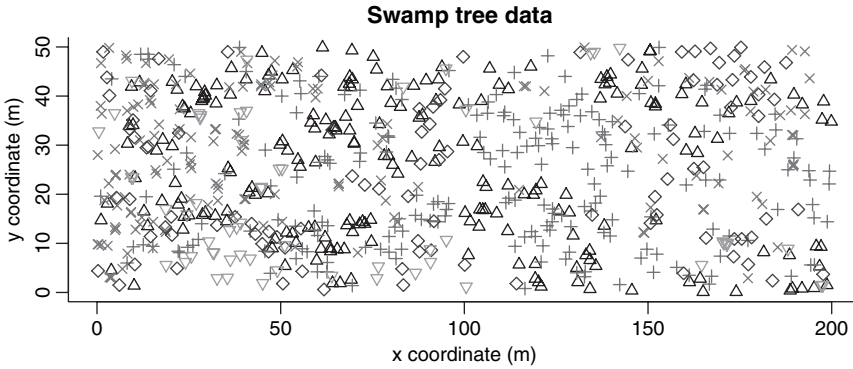


Fig. 7. The scatter plot of the locations of water tupelos (triangles \triangle), black gum trees (pluses +), Carolina ashes (crosses \times), bald cypress trees (diamonds \diamond) and other trees (inverse triangles ∇) in the swamp tree data.

Table 3. The NNCT for swamp tree data (top), the percentages of each row, i.e. the percentages of tree species serving as NN to each tree species (middle) and the percentages of each column, i.e. the percentages of tree species serving as base to each tree species (bottom).

		NN					
Base	WT	BG	CA	BC	OT	sum	
WT	112	40	29	20	14	215	
BG	38	117	26	16	8	205	
CA	23	23	82	22	6	156	
BC	19	29	29	14	7	98	
OT	7	8	5	7	33	60	
sum	199	217	171	79	68	734	
WT	52%	19%	13%	9%	7%	29%	
BG	19%	57%	13%	8%	4%	28%	
CA	15%	15%	53%	14%	4%	21%	
BC	19%	30%	30%	14%	7%	13%	
OT	12%	13%	8%	12%	55%	8%	
WT	56%	18%	17%	25%	21%		
BG	19%	54%	15%	20%	12%		
CA	12%	11%	48%	28%	9%		
BC	10%	13%	17%	18%	10%		
OT	4%	4%	3%	9%	49%		
sum	27%	30%	23%	11%	9%		

WT, water tupelo; BG, black gum; CA, Carolina ash; BC, bald cypress; OT, other tree species.

randomization of the labels on the given locations of the trees 10,000 times. Observe that p_{asy} , p_{mc} and p_{rand} are similar for each test. Overall test of segregation is significant implying significant deviation from the CSR independence pattern for at least one of the tree species. Base class-specific tests are all significant for all species but bald cypresses, implying significant deviation in rows than expected under CSR independence, except for bald cypress trees. These findings are in agreement with the results of DIXON (2002b). NN class-specific tests are significant for all species but bald cypresses, implying significant deviation in columns than expected under

Table 4. The chi-squared statistics for the overall and base class- and NN class-specific segregation tests and the corresponding P -values for the swamp tree data.

	Chi-squared statistic	p_{asy}	p_{mc}	p_{rand}
Overall	275.64	<0.0001	<0.0001	<0.0001
Base				
WT	42.27	<0.0001	<0.0001	<0.0001
BG	65.13	<0.0001	<0.0001	<0.0001
CA	70.99	<0.0001	<0.0001	<0.0001
BC	70.09	0.1313	0.1315	0.1291
OT	117.48	<0.0001	<0.0001	<0.0001
NN				
WT	61.37	<0.0001	<0.0001	<0.0001
BG	75.96	<0.0001	<0.0001	<0.0001
CA	810.06	<0.0001	<0.0001	<0.0001
BC	10.73	0.0571	0.0611	0.0518
OT	118.23	<0.0001	<0.0001	<0.0001

p_{asy} stands for the P -value based on the asymptotic approximation, p_{mc} is the P -value based on 10,000 Monte Carlo replication of the CSR independence pattern in the same region and p_{rand} is based on Monte Carlo randomization of the labels on the given locations of the trees 10,000 times. WT, water tupelo; BG, black gum; CA, Carolina ash; BC, bald cypress; and OT, other tree species.

CSR independence for all species except for bald cypress trees. Hence, except for bald cypresses, each tree species seems to result from a (perhaps) different first-order inhomogeneous Poisson process.

On the basis of the NNCT tests above, we conclude that tree species exhibit significant deviation from the CSR independence pattern (except for bald cypresses) toward the segregation of the species. Then, we might also be interested in the causes of the segregation and the type and level of interaction between the tree species at different scales (i.e. distances between the trees). To answer such questions, we also present the second-order analysis of the swamp tree data. We calculate Ripley’s (univariate) L -function, $L(t)$, under CSR, $L(t) - t = 0$ holds. If the univariate pattern exhibits aggregation, then $L(t) - t$ tends to be positive, if it exhibits regularity then $L(t) - t$ tends to be negative. For a rectangular region, it is recommended to use t values up to one-fourth of the length of the smaller side of the rectangle to reduce the bias (see DIGGLE, 2003 for more detail). So, we take the values $t \in [0, 12.5]$ in our analysis. In Figure 8, we present the plots of $\hat{L}_{ii}(t) - t$ functions together with the upper and lower (pointwise) 95% confidence bounds for each species as well as the plot of all trees combined.

We also calculate Ripley’s bivariate L -function, $L_{ij}(t)$, under CSR independence, $L_{ij}(t) - t = 0$ holds. If the bivariate pattern is segregation, then $L_{ij}(t) - t$ tends to be negative; if it is association then $L_{ij}(t) - t$ tends to be positive (see DIGGLE, 2003 for more detail). In Figure 9, we present the bivariate plots of $\hat{L}_{ij}(t) - t$ functions together with the upper and lower (pointwise) 95% confidence bounds for each pair of species (due to the symmetry of $L_{ij}(t)$, we plot only the 10 different pairs).

However, Ripley’s K -function is cumulative; so, interpreting the spatial interaction at larger distances is problematic (WIEGAND *et al.* 2007) and the results based

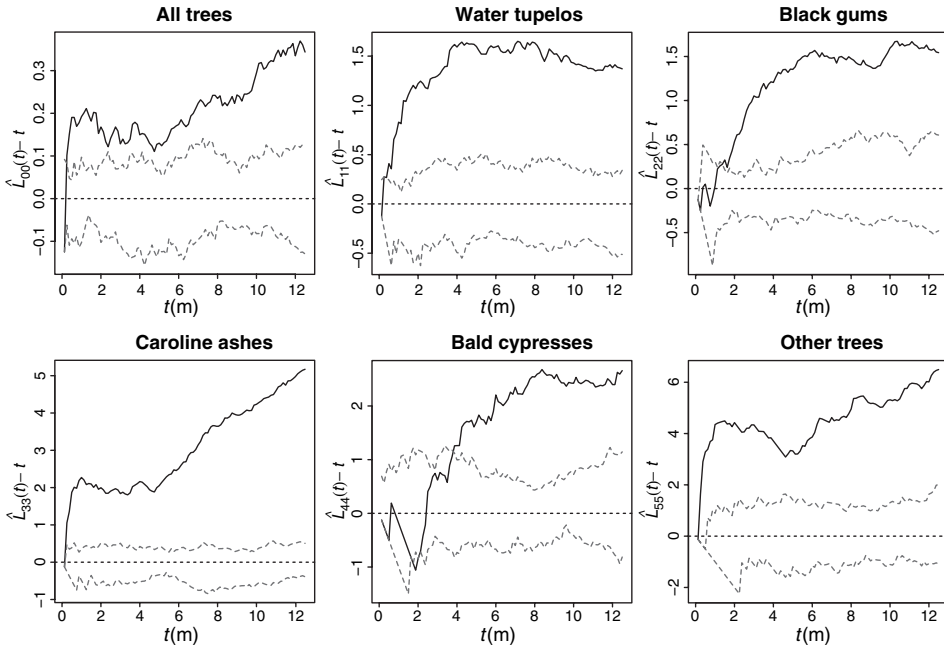


Fig. 8. Second-order properties of swamp tree data. Functions plotted are Ripley's univariate L -functions $\hat{L}_{ii}(t) - t$ for $i=0, 1, \dots, 5$, where $i=0$ stands for all species combined, $i=1$ for water tupelos, $i=2$ for black gums, $i=3$ for Carolina ashes, $i=4$ for bald cypresses, and $i=5$ for other trees. Wide dashed lines around 0 are the upper and lower 95% confidence bounds for the L -functions based on Monte Carlo simulation under CSR independence.

on Ripley's L -function is reliable at about the average NN distances (CEYHAN 2008c). In particular, for the swamp tree data, average NN distance (\pm standard deviation) is about 1.8 (± 1.04) m; so, we consider the distances about 2 m in Figures 8 and 9. In Figure 8, observe that for all trees combined there is significant aggregation of trees (the $L_{00}(t) - t$ curve is above the upper confidence bound). Water tupelos, Carolina ashes, and other trees exhibit significant aggregation for $t \approx 2$ m. Black gums and bald cypresses exhibit no deviation from CSR independence for $t \lesssim 2$ m. Hence, segregation of the species might be due to different levels and types of aggregation of the species in the study region. In Figure 9, observe that water tupelos and black gums exhibit significant segregation ($\hat{L}_{12}(t) - t$ is below the lower confidence bound), water tupelos and Carolina ashes are significantly segregated and black gum and other trees are significantly segregated for $t \approx 2$ m. On the other hand, bald cypresses and other trees are significantly segregated for distances $t \approx 2$ m. All other pairs do not have significant deviation from the CSR independence pattern for distances up to 2 m.

For distances larger than the average NN distance, the (accumulative) pair correlation function $g(t)$ is better for analyzing the spatial interaction (STOYAN and STOYAN, 1994). The pair correlation function of a (univariate) stationary point process is defined as $g(t) = K'(t)/2\pi t$ where $K'(t)$ is the derivative of $K(t)$. The same

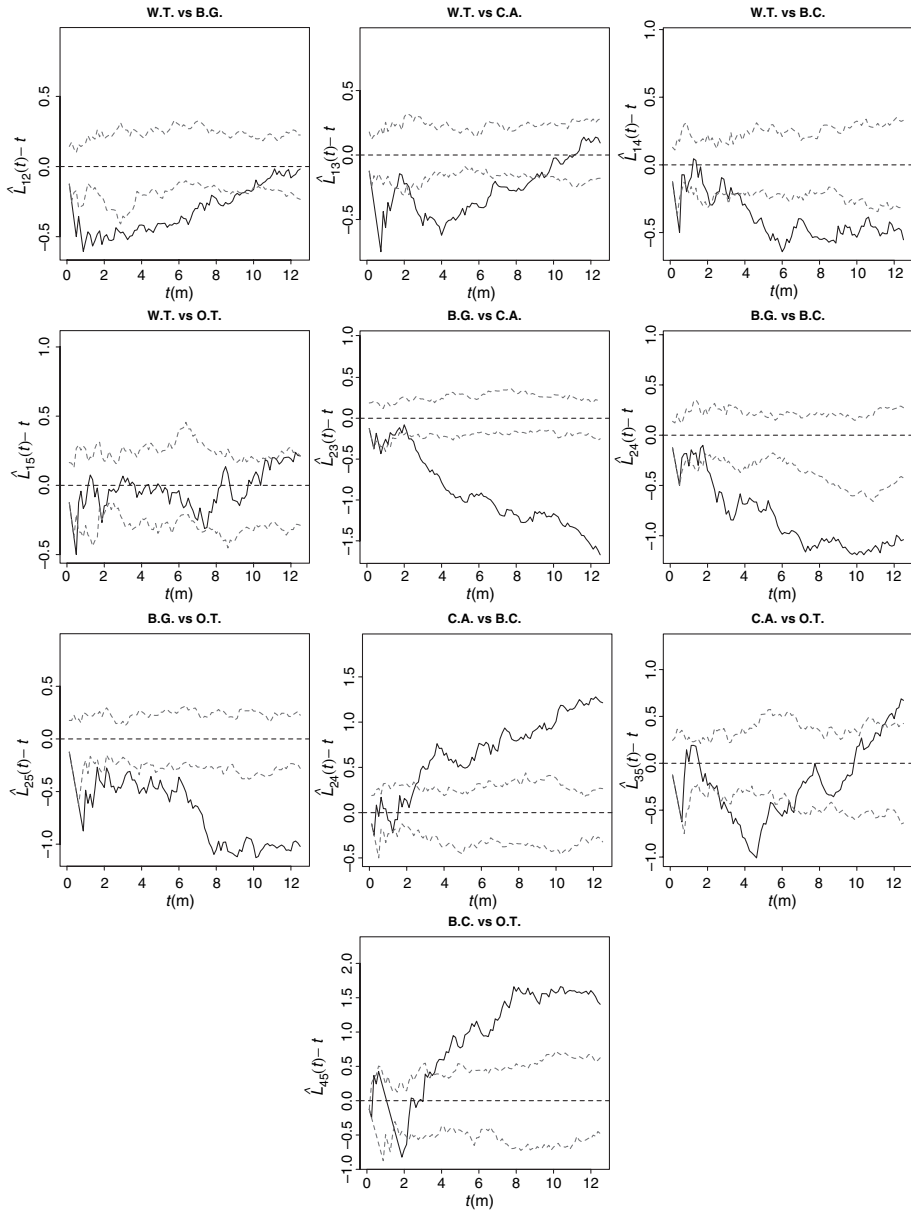


Fig. 9. Second-order properties of swamp tree data. Functions plotted are Ripley's bivariate L -functions $\hat{L}_{ij}(t) - t$ for $i, j = 1, 2, \dots, 5$ and $i \neq j$ where subscript labeling is as in Figure 8. Wide dashed lines around 0 are the upper and lower 95% confidence bounds for the L -functions based on Monte Carlo simulations under the CSR independence pattern. WT, water tupelo; BG, black gum; CA, Carolina ash; BC, bald cypress; OT, other tree species.

definition of the pair correlation function can be applied to Ripley's bivariate K - or L -functions as well. However, the pair correlation function estimates might have critical behavior for small t if $g(t) > 0$, as the estimator variance and hence the bias are considerably large. This problem gets worse especially in cluster processes (STOYAN and STOYAN 1996). So, pair correlation function analysis is more reliable for larger distances and it is safer to use $g(t)$ for distances larger than the average NN distance in the data set. The pair correlation functions for the swamp tree data are presented in CEYHAN (2008a) where we see that Ripley's L and pair correlation functions usually detect the same large-scale pattern but at different ranges of distance values. Ripley's L suggests that the particular pattern is significant for a wider range of distance values compared with $g(t)$, as values of L at small scales confound the values of L at larger scales where $g(t)$ is more reliable to use (LOOSEMORE and FORD (2006).

While second-order analysis (using Ripley's K - and L -functions or pair correlation function) provides information on the univariate and bivariate patterns at all scales (i.e. for all distances), NNCT tests summarize the spatial interaction for the smaller scales (for distances about the average NN distance in the data set). It is interesting to observe that both methods indicate that bald cypresses have a very different pattern compared with that of others.

8.2 Leukemia data

CUZICK and EDWARDS (1990) considered the spatial locations of 62 cases of childhood leukemia in the North Humberside region of the UK between the years 1974 and 1982 (inclusive). A sample of 143 controls are selected using the completely randomized design from the same region. We analyze the spatial distribution of leukemia cases and controls in this data using a 2×2 NNCT. We plot the locations of these points in the study region in Figure 10 and provide the corresponding NNCT together with percentages based on row and column sums in Table 5. Observe that the percentages in the diagonal cells are about the same as the marginal (row or column) percentages of the subjects in the study, which might be interpreted as the lack of any deviation from the null case for both classes. Figure 7 is also supportive of this observation.

We can assume that some processes affect *a posteriori* the population of North Humberside region so that some of the individuals get to be cases, while others continue to be healthy (i.e. they are controls). So, the appropriate null hypothesis is the RL pattern. We calculate $Q = 152$ and $R = 142$ for this data set. In Table 6, we present the overall test of segregation, class-specific test statistics and the associated P -values based on the asymptotic approximation, Monte Carlo simulation and Monte Carlo randomization methods. Observe that p_{asy} , p_{mc} , and p_{rand} are similar for each test and none of the tests yields a significant result.

On the basis of the NNCT tests above, we conclude that the cases and controls do not exhibit significant clustering (i.e. segregation). However, NNCT methods

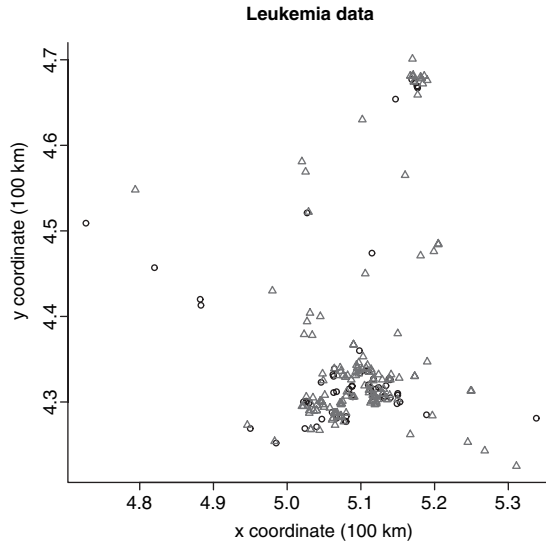


Fig. 10. The scatter plot of the locations of controls (circles \circ) and childhood leukemia cases (triangles \triangle) in North Humberside, UK.

Table 5. The NNCT for the North Humberside childhood leukemia data (top) and the corresponding percentages (middle and bottom).

Base	NN		
	Case	Control	Sum
Case	25	41	66
Control	39	113	152
Sum	64	154	218
Case	38%	62%	30%
Control	26%	74%	70%
Case	39%	27%	
Control	61%	73%	
Sum	29%	71%	

provide information only on spatial interaction for distances about expected NN distance in the data set; so, it might be the case that the type and level of interaction might still be different at larger scales (i.e. distances between the subjects' locations). However, the locations of the subjects in this population (cases and controls together) seem to be from an inhomogeneous Poisson process (see also Figure 10). Hence, Ripley's K - or L -functions in the general form are not appropriate to test for the spatial clustering of the cases (KULLDORFF, 2006). However, DIGGLE (2003) suggests a version based on Ripley's univariate K -functions as $D(t) = K_{11}(t) - K_{22}(t)$ where $K_{ii}(t)$ stands for Ripley's univariate K -function for class i . In this set-up, the case of 'no spatial clustering' is equivalent to RL of cases and controls on the locations in the sample, which implies $D(t) = 0$, as $K_{22}(t)$ measures the degree of spatial aggregation of the controls (i.e. the population at risk), while $K_{11}(t)$ measures this

Table 6. The chi-squared statistics for the overall, base- and NN-class-specific segregation tests and the corresponding P -values for the North Humberside childhood leukemia data.

	Chi-squared statistic	p_{asy}	p_{mc}	p_{rand}
Overall*	2.25	0.3249	0.3303	0.3519
Base				
Case	1.44	0.2293	0.2266	0.2389
Control	1.65	0.1995	0.2041	0.2237

P -value labeling is as in Table 4.

*Only the overall test statistics are provided as NN class-specific tests and overall tests are identical in the two-class case.

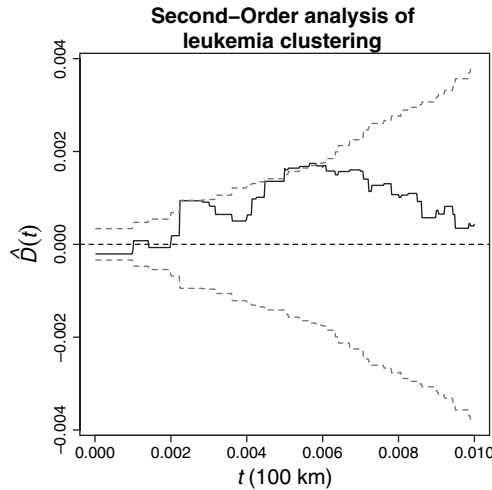


Fig. 11. Second-order analysis of North Humberside childhood leukemia data: Function plotted is Diggle’s modified bivariate K -function $\hat{D}(t) = \hat{K}_{11}(t) - \hat{K}_{22}(t)$ with $i=1$ for leukemia cases and $i=2$ for controls. Wide dashed lines around 0 are plus and minus two standard errors of $\hat{D}(t)$ under RL of cases and controls.

same spatial aggregation plus any additional clustering due to the disease. The test statistic $D(t)$ is estimated by $\hat{D}(t) = \hat{K}_{11}(t) - \hat{K}_{22}(t)$. Figure 11 shows the plot of $\hat{D}(t)$ plus and minus two standard errors under RL. Observe that at distances about 200 and 600 m, there is evidence for mild clustering of diseases (i.e. segregation of cases from controls) as the empirical function $\hat{D}(t)$ gets close to or a little above the upper limit. At smaller scales, plot in Figure 11 is consistent with the results of the NNCT analysis. In particular, average NN distance for leukemia data is 700 (\pm 1400) m, and NNCT analysis summarizes the pattern for about $t = 1000$ m which is depicted in Figure 11. This same data set was also analyzed by DIGGLE (2003, 131–132) and similar plots and results were obtained.

8.3 Pyramidal neuron data

This data set gives the (x,y) coordinates of pyramidal neurons in area 24, layer 2 of the cingulate cortex. The data are taken from a unit square region (unit of measurement unknown) in each of 31 subjects, grouped as follows: controls consists of 12 subjects and correspond to cell numbers 1–655, schizoaffectives consists of nine subjects and correspond to cell numbers 656–1061 and schizophrenics consists of 10 subjects and correspond to cell numbers 1062–1400. Controls are the subjects with no previous history of any mental disorder, schizoaffective disorder is a psychiatric disorder where both the symptoms of mood disorder and psychosis occur and schizophrenia is a psychotic disorder characterized by severely impaired thinking, emotions and behavior. DIGGLE, LANGE and BENES (1991) applied several methods for the analysis of the spatial distributions of pyramidal neurons in the cingulate cortex of human subjects in three diagnostic groupings. With a scaled Poisson analysis, they found significant differences between the groups in the mean numbers of neurons in the sampled region, as well as a high degree of extra-Poisson variation in the distribution of cell counts within these groups. They employed two different functional descriptors of spatial pattern for each subject to investigate departures from completely random patterns, both between subjects and between groups, while adjusting for cell count differences. As the distributions of their main functional pattern descriptor and of their derived test statistic are unknown, they applied a bootstrap procedure to attach P -values to their findings.

As the definition of the rectangular domain for identifying neuron positions is independent of neuronal cell density or the pattern and this sampling domain is almost identical for each subject, DIGGLE, LANGE and BENES (1991) merged (pooled) the data for each group. That is, the pyramidal neuron locations from control subjects were pooled into one group, from schizoaffective into another, and schizophrenic into another. Although the spatial distributions between subjects are not the same, we think pooling the data by group might reveal more than what might be concealed. DIGGLE *et al.* (1991) computed and compared Ripley's univariate K -functions to detect differences between patterns across the three groups. Pattern analysis of the cellular arrangements demonstrated significant departures from CSR in favor of spatial regularity for each group, in particular for schizophrenics. On the pooled data, we also apply multivariate pattern tests. In particular, we apply a 3×3 NNCT analysis and Ripley's bivariate L -functions (in addition to the univariate L -functions).

We plot the locations of these points in the study region in Figure 12 and provide the corresponding NNCT together with percentages based on row and column sums in Table 7. Observe that the pyramidal neuron locations appear to be somewhat regularly spaced, especially for schizophrenics. Moreover, the percentages are slightly smaller for the diagonal cells, especially for the controls, compared with the marginal (row or column) percentages, which might be interpreted as the presence of a deviation from CSR independence in favor of association.

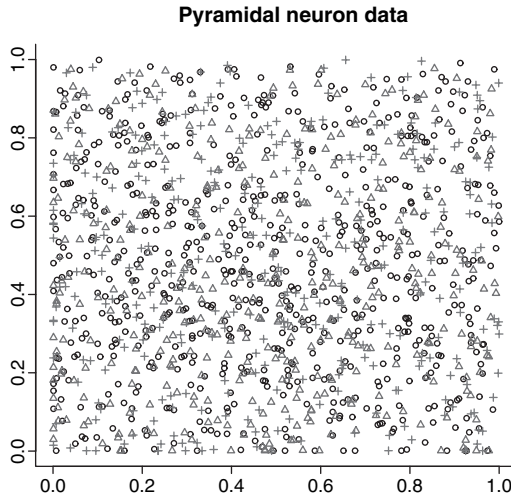


Fig. 12. The scatter plots of the locations of neurons of controls (circles \circ), schizoaffectives (triangles \triangle) and schizophrenics (pluses $+$) in the pyramidal neuron data.

Table 7. The NNCT for the pyramidal neuron data (top) and the corresponding percentages (middle and bottom).

Base	NN			sum
	Ctrl	S.A.	Sch.	
Ctrl	271	216	171	658
S.A.	212	107	89	408
Sch.	175	89	75	339
sum	658	412	335	1405
Ctrl	41%	33%	26%	47%
S.A.	52%	26%	22%	29%
Sch.	52%	26%	22%	24%
Ctrl	41%	52%	51%	
S.A.	32%	26%	27%	
Sch.	27%	22%	22%	
sum	47%	29%	24%	

Ctrl, control; S.A., schizoaffective; sch., schizophrenics.

The locations of the pyramidal neurons can be viewed *a priori* resulting from different processes; so, the more appropriate null hypothesis is the CSR independence pattern. We calculate $Q=888$ and $R=892$ for this data set. In Table 8, we present the overall test of segregation, class-specific test statistics and the associated P -values based on the asymptotic approximation, Monte Carlo simulation and Monte Carlo randomization. Observe that p_{asy} , p_{mc} and p_{rand} are similar for each test. The overall test of segregation is not significant; however, base class- and NN class-specific tests are significant for controls only, implying significant deviation in row 1 and column 1 (i.e. row and column for controls) than expected under CSR independence. On the basis of the NNCT for this data set, we observe that the

Table 8. The chi-squared statistics for the overall, base class- and NN class-specific segregation tests and the corresponding P -values for the pyramidal neuron data.

	Chi-squared statistic	p_{asy}	p_{mc}	p_{rand}
Overall	9.91	0.1283	0.1271	0.1288
Base				
Ctrl	7.43	0.0243	0.0267	0.0230
S.A.	4.19	0.1229	0.1208	0.1233
Sch.	3.12	0.2098	0.2104	0.2101
NN				
Ctrl	9.57	0.0226	0.0229	0.0241
S.A.	6.36	0.0953	0.0968	0.0936
Sch.	2.91	0.4060	0.4052	0.4058

P -value labeling is as in Table 4. Ctrl, control; S.A. schizo-affective; and Sch; Schizophrenics.

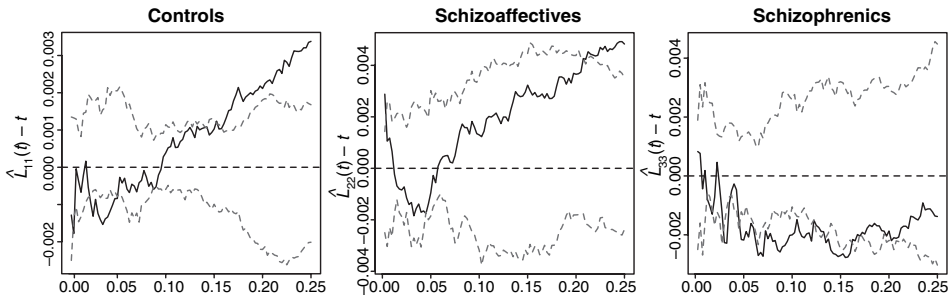


Fig. 13. Second-order properties of pyramidal neuron cells. Functions plotted are Ripley's univariate L -functions $\hat{L}_{ii}(t) - t$ for $i=1, 2, 3$ where $i=1$ for controls, $i=2$ for schizoaffectives and $i=3$ for schizophrenics. Wide dashed lines around 0 are the upper and lower 95% confidence bounds for the L -functions based on Monte Carlo simulations under CSR independence.

deviation is toward association of controls with schizoaffectives and vice versa. As mentioned in DIGGLE *et al.* (1991), the locations of the pyramidal neurons for controls and schizoaffectives are more similar than those of schizophrenics. Thus, it is possible to support the idea that alterations of cortical function may occur in schizophrenics, which might agree with an earlier suggestion that neuronal numbers are lower in the cerebral cortex of schizophrenic patients.

On the basis of the NNCT tests above, we conclude that the pyramidal neurons exhibit significant deviation from the CSR independence pattern for the control subjects toward the association of controls and schizoaffectives. Then, to find out what might be causing the association, and what is the type and level of interaction at different scales, we plot Ripley's (univariate) L -function for all data combined and for each (pooled) group in Figure 13 where the upper and lower 95% confidence bounds are also provided. As Ripley's L is reliable for about the average NN distance [the average NN distance for this data set is 0.014 (± 0.007)], we only consider distances up to 0.02. For small scales ($t \approx 0.02$), controls and schizophrenics exhibit significant regularity, while schizoaffectives do not significantly deviate from the CSR pattern. This is along the lines of the NNCT analysis results, which

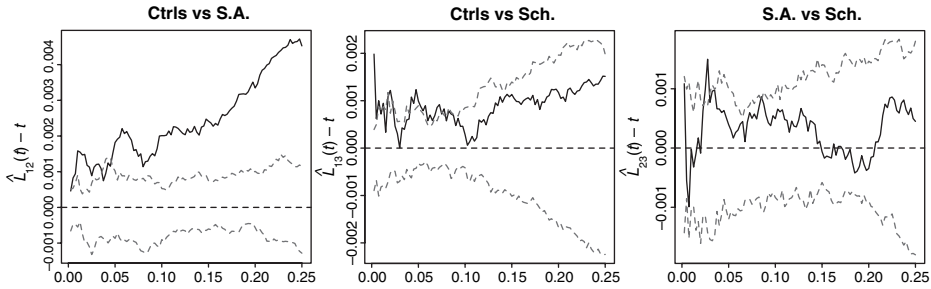


Fig. 14. Second-order properties of pyramidal neuron cells. Functions plotted are Ripley's bivariate L -functions $\hat{L}_{ij}(t) - t$ (right) for $i \neq j$ where $i=1$ for controls, $i=2$ for schizoaffectives and $i=3$ for schizophrenics. Wide dashed lines around 0 are the upper and lower 95% confidence bounds for the L -functions based on Monte Carlo simulations under the CSR independence pattern.

indicate deviation from CSR independence at smaller scales. The significant spatial regularity of the schizophrenics and controls might explain the association of neurons of controls and schizoaffectives. We also plot Ripley's bivariate L -function together with the upper and lower 95% confidence bounds in Figure 14. Observe that for small scales ($t \approx 0.02$) controls and schizoaffectives exhibit significant spatial association, and so do the controls and schizophrenics. So, at smaller scales (i.e. $t \lesssim 0.02$) the univariate and bivariate L -functions seem to be in agreement with the NNCT results which indicate the association of controls and schizophrenics. For larger distances, the pair correlation functions are more reliable and are presented in (CEYHAN 2008a).

9 Discussion and Conclusions

In this article, we consider (Dixon's) overall and base class-specific tests and introduce NN class-specific tests of segregation based on NNCTs. These NNCT tests are used in testing various forms of randomness in the NN structure between two or more classes. The overall test is used for testing any deviation from randomness in the cells of the NNCT; base class-specific tests are used for testing any deviation from randomness in the NN distribution of the base class in question (i.e. any deviation in the corresponding row); NN class-specific tests are used for testing any deviation from randomness in the base distribution of the classes to which the class in question serves as the NN (i.e. any deviation in the corresponding column). The randomness in the NN structure is implied by the CSR independence or RL patterns. However, we demonstrate that under the CSR independence pattern, NNCT tests are conditional on Q and R , while under the RL pattern, they are unconditional tests.

On the basis of our Monte Carlo simulations, we conclude that the asymptotic approximation for the NNCT tests is appropriate only when each cell count in the NNCT is larger than 5. The tests tend to be conservative when the row or column

they pertain to contain cell(s) whose count(s) is(are) ≤ 5 . DIXON (1994) recommends Monte Carlo randomization when some cell count(s) are ≤ 5 in a NNCT for the overall segregation test, we extend this recommendation for the base class and NN class-specific tests. A cell in the NNCT has a high probability of being ≤ 5 , when at least one sample size is ≤ 10 or when relative abundances (i.e. sample sizes) are very different. For large samples, the NNCT tests have similar empirical significance levels which are about the nominal level. For larger samples, the tests yield empirical sizes that are about the desired nominal level. Under the segregation alternatives, for large samples with similar sizes, we conclude that NN class-specific tests have higher power; when at least one sample is small, base class-specific test for the smaller (larger) class has the highest (lowest) power estimates. Under the association alternatives, we observe that when at least one sample size is small, base class-specific test for the smaller class has extremely poor performance, while for the larger class it has higher power than the NN class-specific tests. On the other hand, for larger samples, base class-specific test for the largest (smallest) class has the highest (lowest) power. However, base-class- and NN class-specific tests answer related but different questions, so it is not appropriate to compare the NNCT tests. Thus, we recommend both of the base-class- and NN class-specific tests to get more information on different aspects of the NN structure.

NNCT tests summarize the pattern in the data set for small scales; more specifically, they provide information on the pattern around the average NN distance between the points. On the other hand, pair correlation function $g(t)$ and Ripley's classical K - or L -functions and other variants provide information on the pattern at various scales. Ripley's classical K - or L -functions can be used for testing (i.e., inference) when the null pattern can be assumed to be CSR independent; i.e., when the null pattern assumes first-order homogeneity for each class. When the null pattern is the RL of points from an inhomogeneous Poisson process, they are not appropriate (KULLDORFF 2006). Diggle's D -function is a modified version of Ripley's K -function (DIGGLE 2003) and adjusts for any inhomogeneity in the locations of, e.g. cases and controls. Furthermore, there are variants of $K(t)$ that explicitly correct for inhomogeneity (see BADDELEY, MØLLER and WAAGEPETERSEN, 2000). Ripley's K -, Diggle's D - and pair correlation functions are designed to analyze univariate or bivariate spatial interaction at various scales (i.e. inter-point distances). Our examples illustrate that for distances around the average NN distance (i.e. for small scales), NNCT tests and Ripley's L - and Diggle's D -functions yield similar but not identical results.

The overall NNCT test and Ripley's L -function provide similar information in the two-class case at small scales. For $q > 2$ classes, overall tests provide information on the (small-scale), while the Ripley's L -function requires performing all bivariate spatial interaction analysis. On the other hand, the class-specific tests can serve as a means of *post hoc* analysis when the overall test is significant. Furthermore, the class-specific tests are testing the spatial interaction of one class, say class i , with all classes (including class i) as part of the multivariate interaction between

all the classes. On the other hand, Ripley's univariate K - or L -functions are restricted to one class and bivariate K - or L -functions are restricted to two classes they pertain to, ignoring the potentially important multivariate interaction between all classes in the study area. However, there are forms of the J -function which is derived from the well-known G and F functions (VAN LIESHOUT and BADDELEY 1999) and deal with this multitype setting (i.e. consider the pattern of type i in the context of the pattern of all other types). VAN LIESHOUT and BADDELEY (1999) define two basic types of J -functions. First is a type- i -to-type- j function which considers the points of type i in the context of the points of type j . The second one is the type- i -to-any-type function which considers the points of type i in the context of points of all types including type i . Other forms can be derived from them by re-defining the types. For example, if we want to consider the points of type i in the context of points of all other types, we collapse all the other types $j \neq i$ into a single type i' and then use the type- i -to-type- i' function. Several authors have written about the bivariate K -function, which is of the type- i -to-type- j form (DIGGLE and HAASE 1995; DIGGLE 2003). Type- i -to-type- j K -function can easily be modified to type- i -to-any-type K -function. Thus, essentially there is only one family of multitype K -functions in the literature. However type- i -to-type- j K -function is comparable with a NNCT analysis based on a 2×2 NNCT restricted to the classes i and j . Similarly, type- i -to-type- i' K -function is comparable with the NNCT analysis based on a 2×2 NNCT with classes i and the rest of the classes labeled as i' . Furthermore, type- i -to-type- i' K -function provides similar information with the base class-specific test for class i at small scales, but the K -function ignores the interaction of class i with itself. On the other hand, by definition none of the variants of the K -function is comparable with the NN class-specific tests, as they provide information on very different aspects of the spatial interaction between classes. As the pair correlation functions are derivatives of Ripley's K -function, most of the above discussion holds for them also, except that $g(t)$ is reliable only for large-scale interaction analysis. Hence, NNCT tests and pair correlation function are not comparable but provide complimentary information about the pattern in question.

For a data set for which CSR independence is the reasonable null pattern, we recommend the overall segregation test if the question of interest is the spatial interaction at small scales (i.e. about the mean NN distance). If it yields a significant result, then to determine which classes have significant spatial interaction with themselves and all other classes, the class-specific tests can be performed. One can also perform Ripley's K - or L -function and only consider distances up to about the average NN distance and compare the results with those of NNCT analysis. If the spatial interaction at higher scales is also of interest, pair correlation function is recommended (LOOSMORE and FORD 2006). On the other hand, when the RL pattern is the reasonable null pattern for the data, we recommend the NNCT tests if the small-scale interaction is of interest and Diggle's D -function if the spatial interaction at higher scales is also of interest.

Acknowledgements

I to thank the associate editor and the referees, whose constructive comments and suggestions greatly improved the presentation and flow of the paper. Most of the Monte Carlo simulations presented in this article were executed on the Hattusas cluster of Koç University High Performance Computing Laboratory.

References

- ARMSTRONG, J. E. and A. K. IRVINE (1989), Flowering, sex ratios, pollen-ovule ratios, fruit set, and reproductive effort of a dioecious tree, *Myristica Insipida* (*Myristicaceae*), in two different rain forest communities, *American Journal of Botany* **76**, 75–85.
- BADDELEY, A., J. MØLLER and R. WAAGEPETERSEN (2000), Non- and semi-parametric estimation of interaction in inhomogeneous point patterns, *Statistica Neerlandica* **54**, 329–350.
- BAROT, S., J. GIGNOUX and J. C. MENAUT (1999), Demography of a savanna palm tree: predictions from comprehensive spatial pattern analyses, *Ecology* **80**, 1987–2005.
- CEYHAN, E. (2007), Edge correction for cell- and class-specific tests of segregation based on nearest neighbor contingency tables, in: H. GÖKÇEKUS (ed), in: *Proceedings of the International Conference on Environment: Survival and Sustainability*, Educational Foundation of Near East University 385
- CEYHAN, E. (2008a), *Class-specific tests of spatial segregation based on nearest neighbor contingency tables*, arXiv:0810.0758v1 [stat.ME], technical Report # KU-EC-08-7.
- CEYHAN, E. (2008b), On the use of nearest neighbor contingency tables for testing spatial segregation, *Environmental and Ecological Statistics* doi: 10.1007/s10651-008-0104-x.
- CEYHAN, E. (2008c), Overall and pairwise segregation tests based on nearest neighbor contingency tables, *Computational Statistics and Data Analysis* doi: 10.1016/j.csda.2008.08.002.
- COOMES, D. A., M. REES and L. TURNBULL (1999), Identifying aggregation and association in fully mapped spatial data, *Ecology* **80**, 554–565.
- COX, T. F. (1981), Reflexive nearest neighbours, *Biometrics* **37**, 367–369.
- CRESSIE, N. A. C. (1993), *Statistics for spatial data*, Wiley, New York.
- CUZICK, J. and R. EDWARDS (1990), Spatial clustering for inhomogeneous populations (with discussion), *Journal of the Royal Statistical Society, Series B* **52**, 73–104.
- DIGGLE, P. J. (2003), *Statistical analysis of spatial point patterns*, Hodder Arnold Publishers, London.
- DIGGLE, P. J. and A. G. CHETWYND (1991), Second-order analysis of spatial clustering for inhomogeneous populations, *Biometrics* **47**, 1155–1163.
- DIGGLE, P. J., N. LANGE and F. BENES (1991), Analysis of variance for replicated spatial point patterns in clinical neuroanatomy, *Journal of American Statistical Association* **86**, 618–625.
- DIXON, P. M. (1994), Testing spatial segregation using a nearest-neighbor contingency table, *Ecology* **75**, 1940–1948.
- DIXON, P. M. (2002a), Nearest-neighbor contingency table analysis of spatial segregation for several species, *Ecoscience* **9**, 142–151.
- DIXON, P. M. (2002b), Nearest neighbor methods, in: A. H. EL-SHAARAWI and W. W. PIEGORSCH (eds), in: *Encyclopedia of Environmetrics*, Vol. 3, New York: John Wiley & Sons Ltd, 1370–1383.
- GAVRIKOV, V. and D. STOYAN (1995), The use of marked point processes in ecological and environmental forest studies, *Environmental and Ecological Statistics* **2**, 331–344.
- GOOD, B. J. and S. A. WHIPPLE (1982), Tree spatial patterns: South Carolina bottomland and swamp forests, *Bulletin of the Torrey Botanical Club* **109**, 529–536.
- GOREAUD, F. and R. PÉLISSIER (2003), Avoiding misinterpretation of biotic interactions with the intertype K_{12} -function: population independence vs. random labelling hypotheses, *Journal of Vegetation Science* **14**, 681–692.

- HAASE, P. (1995), Spatial pattern analysis in ecology based on Ripley's K -function: Introduction and methods of edge correction, *The Journal of Vegetation Science* **6**, 575–582.
- HAMILL, D. M. and S. J. WRIGHT (1986), Testing the dispersion of juveniles relative to adults: a new analytical method, *Ecology* **67**, 952–957.
- HERLER, J. and R. A. PATZNER (2005), Spatial segregation of two common Gobioid species (Teleostei: Gobiidae) in the Northern Adriatic Sea, *Marine Ecology* **26**, 121–129.
- HERRERA, C. M. (1988), Plant size, spacing patterns, and host-plant selection in *Osyris quadripartita*, a hemiparasitic dioecious shrub, *Journal of Ecology* **76**, 995–1006.
- KULLDORFF, M. (2006), Tests for spatial randomness adjusted for an inhomogeneity: a general framework, *Journal of the American Statistical Association* **101**, 1289–1305.
- LAHIRI, S. N. (1996), On consistency of estimators based on spatial data under infill asymptotics, *Sankhya: The Indian Journal of Statistics, Series A* **58**, 403–417.
- VAN LIESHOUT, M. N. M. and A. J. BADDELEY (1999), Indices of dependence between types in multivariate point patterns, *Scandinavian Journal of Statistics* **26**, 511–532.
- LOOSMORE, N. and E. FORD (2006), Statistical inference using the g or k point pattern spatial statistics, *Ecology* **87**, 1925–1931.
- MEAGHER, T. R. and D. S. BURDICK (1980), The use of nearest neighbor frequency analysis in studies of association, *Ecology* **61**, 1253–1255.
- MORAN, P. A. P. (1948), The interpretation of statistical maps, *Journal of the Royal Statistical Society, Series B* **10**, 243–251.
- NANAMI, S. H., H. KAWAGUCHI, and T. YAMAKURA (1999), Dioecy-induced spatial patterns of two codominant tree species, *Podocarpus nagi* and *Neolitsea aciculata*, *Journal of Ecology* **87**, 678–687.
- PIELOU, E. C. (1961), Segregation and symmetry in two-species populations as studied by nearest-neighbor relationships, *Journal of Ecology* **49**, 255–269.
- RIPLEY, B. D. (2004), *Spatial statistics*, Wiley-Interscience, New York.
- SCHLATHER, M., P. RIBEIRO JR, and P. DIGGLE (2004), Detecting dependence between marks and locations of marked point processes, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **66**, 79–93.
- SEARLE, S. R. (2006), *Matrix algebra useful for statistics*, Wiley-Intersciences, New York.
- STOYAN, D. and H. STOYAN (1994), *Fractals, random shapes and point fields: methods of geometrical statistics*, John Wiley and Sons, New York.
- STOYAN, D. and H. STOYAN (1996), Estimating pair correlation functions of planar cluster processes, *Biometrical Journal* **38**, 259–271.
- WALLER, L. A. and C. A. GOTWAY (2004), *Applied spatial statistics for public health data*, Wiley-Interscience, New York.
- WHIPPLE, S. A. (1980), Population dispersion patterns of trees in a Southern Louisiana hardwood forest, *Bulletin of the Torrey Botanical Club* **107**, 71–76.
- WIEGAND, T., S. GUNATILLEKE, and N. GUNATILLEKE (2007), Species associations in a heterogeneous Sri Lankan dipterocarp forest, *The American Naturalist* **170**, 77–95.
- YAMADA, I. and P. A. ROGERSEN (2003), An empirical comparison of edge effect correction methods applied to K -function analysis, *Geographical Analysis* **35**, 97–109.

Received: December 2006. Revised: November 2008.