ANALYSIS OF SPATIAL POINT PATTERNS BY KERNEL IDENTIFICATION[*]

M. Q. VAHIDI-ASL[**] AND M. R. FAGHIHI

Department of Statistics, Shahid Beheshti University, Evin, Tehran, I. R. of Iran 19839

Abstract – In the analysis of spatial point patterns, complete spatial randomness (CSR) hypothesis, which is a restriction of a homogenous Poisson process to study region A, operates as a dividing hypothesis between "regular" and "aggregated" patterns. Meanwhile, many alternatives to CSR in aggregated patterns are extensions of homogenous Poisson processes themselves. Therefore, when the CSR hypothesis is rejected, results related to Poisson processes may be used to formulate plausible alternatives to CSR. In this paper, we propose a new statistic for testing CSR and then by applying it in conjunction with a notion of kernels of a point pattern, we determine the "parents" of a Poisson cluster process when the CSR hypothesis is rejected and a Neyman-Scott process is assumed for the point pattern under alternative hypothesis. We have made power studies for our test statistic by simulation, and have also surveyed the performance of our method on a certain point pattern. Finally, the whole method is carried on certain real life data.

Keywords – Spatial point patterns, complete spatial randomness, poisson processes, Neyman-Scott processes, cluster analysis

1. INTRODUCTION

A spatial point pattern is a set of locations, irregularly distributed within a region of interest, which have been generated by some unknown mechanism. These locations are usually referred to as *events* to distinguish them from arbitrary points of the region in question. A main approach in the analysis of spatial point patterns is the formulation of an explicit model of the underlying mechanism.

Most available methods of analysis are "space-domain" techniques that involve the examination of inter-event distances, but the potential of applying spectral techniques in the study of spatial point patterns is also gaining recognition, at least in relation to the exploratory stages of analysis, [1].

Most analysis begins with a test of *complete spatial randomness* (hereafter CSR). The hypothesis of CSR for a spatial point pattern asserts that: (i) the number of events in any planar region A with area $|A|$ ($|A|$, the Lebesgue measure of the Borel set A), follows a Poisson distribution with mean $|A|$; (ii) given n points x_i in a region A, the x_i's are an independent sample from the uniform distribution on A.

There are at least three reasons for beginning an analysis with a test of CSR: 1. rejection of CSR is a minimal prerequisite for any serious attempt to model an observed pattern; 2. tests are used to explore a set of data to assist in the formulation of plausible alternatives to CSR; 3. CSR operates as a dividing hypothesis between regular and aggregated patterns, (Diggle [2]).

In emphasizing more on the role of CSR tests, and in order to introduce the method presented in this paper, we note that any alternative to CSR in aggregated models is an extension of an homogeneous Poisson process to more general processes. A large class of such processes useful in modeling spatial point patterns is the Poisson cluster process, introduced by Neyman and Scott [3]. In such a process, a set of parent "events" are generated by a homogeneous Poisson process, after which each parent produces a random number S of offspring, realized independently and identically for each parent according to a probability distribution $\{p_s, s = 0,1,2, ...\}$. The positions of the offspring relative to their parents being independently and identically distributed according to a bivariate probability distribution function (PDF) $h(.)$.

Here, in this paper, we propose a method for testing CSR against a particular Poisson cluster process by identifying the parent events based on a set of data in a region A of the plane. We are not identifying $h(.)$, the PDF of the positions of offspring relative to their parents explicitly, but only mention that our method assumes that $h(.)$ is of any general type, and that the probability of an offspring being at a distance d is proportional to d.

Our method rests solely on tests of CSR, and since these tests have a central role in the method, we offer a new procedure for CSR tests based on the projection of higher dimensional data on certain lines and forming a one-dimensional data set, which under CSR hypothesis is a restriction of a one-dimensional homogeneous Poisson process to an interval. A short account of this idea for testing CSR is reported by one of the authors (Vahidi-Asl [4]), where Theorem 2 in subsection 2.1 below is applied to the three data sets given by Diggle [2] as real life prototypes of regular, random, and aggregated patterns, and is led to the same conclusions obtained by Diggle using Monte Carlo tests, (Barnard [5]). In the present work, the proof of Theorem 1 of subsection 2.1 is improved and an explicit formula is given for the distribution of respective statistic.

The paper is organized as follows: in Section 2 we introduce a statistic that is used throughout this paper for our CSR tests. Section 3 is devoted to a power study by simulation in which the power of our test statistic is compared to the power of a test statistic known to be powerful. In Section 4 we introduce a concept called "average concentration level", which is used to identify the parent events in some alternative hypothesis to CSR. In the final section we apply the methods introduced in this paper to a real life data, that of longleaf pine data described in [6].

2. TESTING FOR COMPLETE SPATIAL RANDOMNESS

The standard against which spatial point patterns are often compared is (a realization from) a completely spatially random point process. In this paper, following [6], CSR is synonymous with a homogeneous Poisson process in R^d with the definition given in the introduction.

Several different approaches are usually taken to quantify different types of spatial point patterns. One type of descriptive statistic is based on quadrats. Another type of statistic is based on distances between events, or between randomly sampled points (of the study region A, not of the point pattern) and events.

Many of these statistics, based on distances, have explicit, exact or asymptotic distributions, a summary of which is presented in Table 8.6 of [6].

We are taking a new approach here and introduce a test statistic whose exact and asymptotic distribution is readily obtained in the next section.

2.1 Theoretical results

Since the statistics to be introduced are based upon points distributed along a line with exponential distribution for the distance between two consecutive points, we consider the one-dimensional case first.

Let the points $X_1, \ldots, X_{n+1}$ be distributed randomly along the stretch of a line so that the random variables

$$T_i = X_{i+1} - X_i, \qquad i = 1,2,\ldots,n$$

are independent and exponentially distributed random variables with mean λ^{-1}. We place line segments of length c on every point X_i along the supporting line of the X_i's so that X_i is the midpoint of this line segment. These lines either overlap or there is a gap between two consecutive line segments. If we denote the gap between line segments centered on X_i and X_{i+1} by Y_i, then we have

$$Y_i = (T_i - c)^+, \qquad i = 1,2,\ldots,n,$$

in which $x^+ = max(x,0)$.

Let $U_n = \sum_{i=1}^{n} Y_i$. We discuss the distribution of U_n first.

Theorem 1. *Let* $U_n = \sum_{i=1}^{n} (T_i - c)^+$, *where* $T_1, T_2, \ldots, T_n$ *are iid with common exponential distribution of mean* $1/\lambda$. *Then*

$$F_n(u) = P(U_n \leq u)$$

$$= 1 - e^{-\lambda u} \sum_{j=0}^{n-1} \sum_{m=0}^{j} \binom{n}{j+1} p^{j+1} q^{n-j-1} \frac{(\lambda u)^m}{m!}, \qquad n \geq 1, \qquad u \geq 0,$$

where $p = e^{-\lambda c}$ *and* $q = 1 - p$.

Proof. Denoting by N the number of nonzero $(T_i - c)^+$'s, we have

$$F_n(u) = P(U_n \leq u)$$

$$= \sum_{j=1}^{n} P(\sum_{i=1}^{n} (T_i - c)^+ \leq u \mid N = j) P(N = j) + P(0 \leq u, N = 0)$$

$$= \sum_{j=1}^{n} P(\sum_{k=1}^{j} (T_{i_k} - c) \leq u \mid N = j) P(N = j) + 1_{[0, \infty)}(u)[P(T_1 < c)]^n$$

where we have assumed that $T_{i_1}, \ldots, T_{i_j}$ are those random variables among $T_1, \ldots, T_n$ that exceed c.

Now the random variable N has a binomial distribution with parameters $p = P(T_1 > c) = e^{-\lambda c}$ and n, and given $N = j$, and therefore $T_{i_k} > c$, the conditional distribution of $T_{i_k} - c$ is exponential with mean $1/\lambda$. Hence, given $N = j$, the random variable $\sum_{k=1}^{j} (T_{i_k} - c)$ has a gamma distribution with parameters $\alpha = j$ and $\beta = \lambda$.

Therefore, with $u \geq 0$,

$$F_n(u) = \int_0^u \sum_{j=1}^n \frac{\lambda^j}{(j-1)!} e^{-\lambda x}.x^{j-1}\binom{n}{j}p^j q^{n-j}\,dx + q^n,$$

from which we conclude that U_n has a mixed distribution with a discrete part, which is a point mass at 0, and an absolutely continuous part with density

$$f_n(u) = \sum_{j=0}^{n-1} \frac{\binom{n}{j+1}}{(j+1)!}(p\lambda)^{j+1} q^{n-j-1} \int_0^u e^{-\lambda x}x^j\,dx.$$

Now using the obvious identity of [see e.g. Chung [7], page 202]

$$\int_0^u e^{-\lambda x}x^j\,dx = \frac{j!}{\lambda^{j+1}} \sum_{m=j+1}^\infty \frac{e^{-\lambda u}(\lambda u)^m}{m!},$$

we will finally have

$$F_n(u) = q^n 1_{[\circ,\infty)}(u) + \sum_{j=0}^{n-1} \frac{\binom{n}{j+1}}{j!}(p\lambda)^{j+1} q^{n-j-1} \cdot \sum_{m=j+1}^\infty \frac{e^{-\lambda u}(\lambda u)^m}{m!}$$

$$= q^n 1_{[\circ,\infty)}(u) + \sum_{j=0}^{n-1} \binom{n}{j+1}p^{j+1}q^{n-j-1}[1 - \sum_{m=0}^j \frac{e^{-\lambda u}(\lambda u)^m}{m!}]$$

$$= 1 - e^{-\lambda u}\sum_{j=0}^{n-1} \sum_{m=0}^j \binom{n}{j+1}p^{j+1}q^{n-j-1}\frac{(\lambda u)^m}{m!}, \qquad n \geq 1,\ u \geq 0.$$

The following result is obtained by a straightforward application of the central limit theorem.

Theorem 2. *Let $T_1,...,T_n$ be iid random variables with a common exponential distribution with mean λ^{-1}. Then the random variable $U = \sum_{i=1}^n (T_i - c)^+$ has an asymptotic normal distribution with mean $n.E[(T_n-c)^+] = n\lambda^{-1}e^{-\lambda c}$ and variance $n.Var[(T_n-c)^+] = n\lambda^{-2}e^{-\lambda c}(2 - e^{-\lambda c})$.*

For large n, the computational time in using the exact distribution in later applications is somewhat long, but these computations for asymptotic distribution is immediate. So for large n, using asymptotic distribution is preferred.

To obtain a rough idea of how large n should be in order to use the asymptotic distribution, the quantiles of the exact and asymptotic distributions have been compared for different values of n (5, 20, 30, 40, 60 and 70). The results are shown in Fig. 1. The averaged absolute values of differences between asymptotic and exact quantiles decrease by increasing n, and for $n \geq 40$ this average is less than 0.0099. Hence, it seems that for $n \geq 40$ the asymptotic distribution would be applicable.
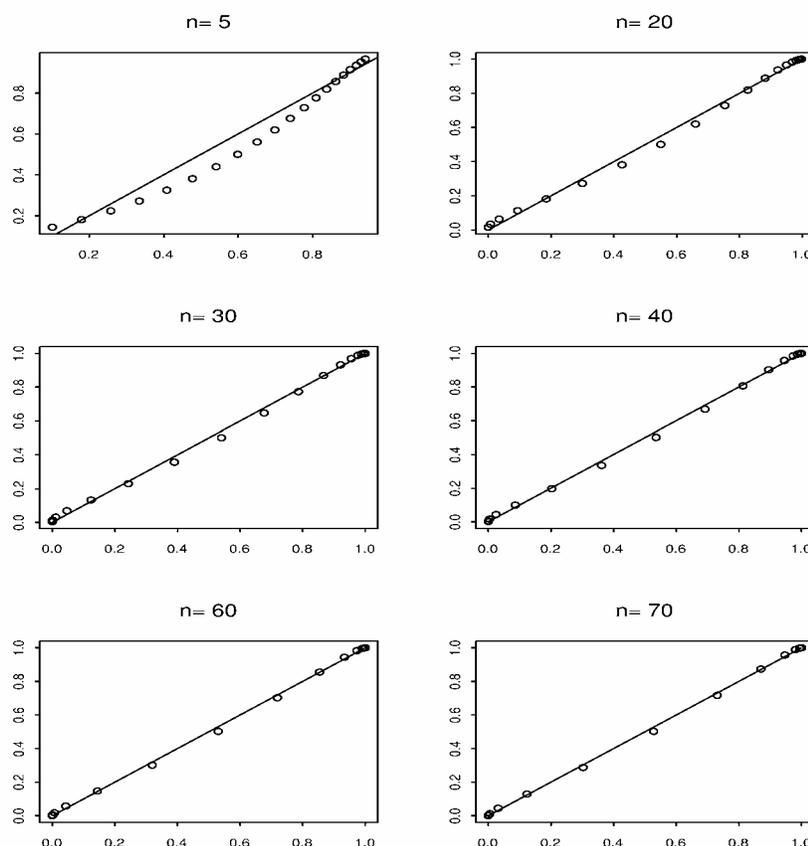
Fig. 1. Comparing quantities of exact and asymptotic distributions for different n.
The x and y axes represent exact and asymptotic quantiles, respectively

2.2 Adapting to planar data

Our method for testing for CSR rests solely on data distributed on a line segment so that the distances between consecutive points are iid random variables with common exponential distribution. This is true, for example in case *n* points belonging to a one-dimensional homogeneous Poisson process with "edge effects" due to confinement to a line segment completely ignored. The idea of exploiting the statistic U_n above is very simple. If the *n*+1 events are almost regularly distributed along their supporting line, and *c* is chosen close to the average distance between all of the events, then U_n will be small. For events randomly distributed under CSR hypothesis, the value of U_n should be moderate and for aggregated data or events somehow showing clustering, the value of U_n should be large. In the sequel, the value of *c* is taken as the range of values of x_i divided by *n*-1.

We now try to adapt it to planar data. Considering a rectangle *A* with sides *a* and *b* so that the hypothesis of CSR is valid in this rectangular region, we divide the sides of lengths *a* and *b* into *l* and *m* equal parts, respectively, and draw lines parallel to the sides so that the original rectangle is

subdivided into m rectangles with sides a and b/m and l rectangles with sides b and a/l. Now consider the horizontal rectangles first. If we denote these rectangles by $A_1, A_2, \ldots, A_m$, we project all the events inside $A_j; j=1,2,\ldots,m$ on the base of the rectangle, i.e., the side with length a. If we denote two consecutive projections in the rectangle A_j by X_{ij} and $X_{i+1,j}$ and the distance between the events X_{ij} and $X_{i+1,j}$ by T_{ij}, then under CSR

$$P(T_{ij} > t) = e^{-\lambda \frac{b}{m} t}$$

and therefore, T_{ij}'s are independent and each T_{ij} has an exponential distribution with mean $m(\lambda b)^{-1}$, and hence we are back to the one-dimensional case again. Of course the edge effects are completely ignored again. Doing the same on all the strips $A_1, \ldots, A_m$, we may consider the statistic

$$U = \sum_{i,j} (T_{ij} - c)^+$$

which is the sum of $n-m$ iid random variables with a common exponential distribution with mean $m(\lambda b)^{-1}$. Using the same procedure for vertical strips, we obtain a statistic

$$V = \sum_{i,j} (T'_{ij} - c')^+$$

which is the sum of $n-l$ iid random variables with a common exponential distribution with mean $l(\lambda a)^{-1}$. Let $c = \dfrac{ma}{n-m}$ and $c' = \dfrac{lb}{n-l}$. As for the "linear" data, "small" values of U and V simultaneously correspond to the regular data, and large values of U and V correspond to aggregated data. Therefore we will reject the CSR hypothesis unless

$$(u_1 \leq U \leq u_2 \text{ and } v_1 \leq V \leq v_2).$$

Rephrasing, if we denote the two-dimensional statistic (U, V) by W, then we reject the CSR hypothesis if W is not in the rectangle $(u_1, u_2) \times (v_1, v_2)$.

Remark. There remains the choice of m and l in applications. We have chosen $m = [\sqrt{nb/a}]$ and $l = [n/m]$, where $[x]$ is the integer part of the real number x. In practice we have used the y-range for b and x-range for a where (x, y) is the coordinate of a typical event.

In the sequel, ignoring the edge effects, we have pieced together the horizontal strips in consecution and made up a single strip. The same has been done for vertical strips too. Necessary modifications are made in the values of m and l in applying the above theory. We also mention that, due to the rotational invariance of the homogenous Poisson process, we may project the events on any two sets of equidistant parallel lines intersecting the study region A, but in that case we need to discard some non-rectangular subregions at both ends of the strips between two parallel lines before piecing together the consecutive strips. Therefore, because of the rectangular shape of the study region A, using lines parallel to the sides of A will result in no loss of data.

Finally, we have used the maximum likelihood estimator of λ, that is $\hat{\lambda} = (n-1)/\sum_{i=1}^{n-1} (x_{i+1} - x_i)$, wherever needed (for more details, see Karr [8]).

2.3 An application

We now apply our method for testing CSR to a real life data, that is to longleaf-pine data (Cressie [6], page 579). These data consist of coordinates of all longleaf pine trees at least 2 cm in diameter at breast height in 4 ha of forest in 1979, making a total of 584 such trees.

Using the *W* statistic of subsection 2.2, we obtain a p-value of 0.0003. Thus rejecting the CSR hypothesis [to be compared with table 8.7 of Cressie [6] (page 608)].

3. A POWER STUDY BY SIMULATION

To give an idea of the performance of this method, we make a comparison of the relative frequencies of rejections by two test statistics on a certain point pattern to be described below. One of the test statistics, denoted as *K* in Cressie [6] (page 608), is due to Diggle [9] whose value is

$$48\left[n\log\left\{\sum\left(2D_i^2 + Z_i^2\right)/n\right\} - \sum\log\left(2D_i^2 + Z_i^2\right)\right]/(13n+1),$$

where D_i is the distance between the ith sample point to the nearest event and Z_i is the distance between that nearest event to the nearest event in half plane not containing the sample point i (Cressie [6], page 602). The other test statistic is, of course, the *W* statistic described at the end of section 2.2.

As for our point pattern, using some of Matern's idea [10], we generate each data set for simulation in two stages: in the first stage, we generate 40 points (events) in a square of side one under CSR hypothesis. In other words, these 40 events are distributed uniformly and independently in the unit square, and these are called the parent events. In the second stage, corresponding to each parent event, 7 additional points called daughter events are generated uniformly and independently inside a circle of a radius of 0.05 drawn around parent events. Parent and daughter events are both present in our point pattern, although we distinguish between them in our records for later uses.

A typical realization of such a process is shown in Fig. 2(a). Fig. 2(b) represents the set of parent events generated by S-plus.
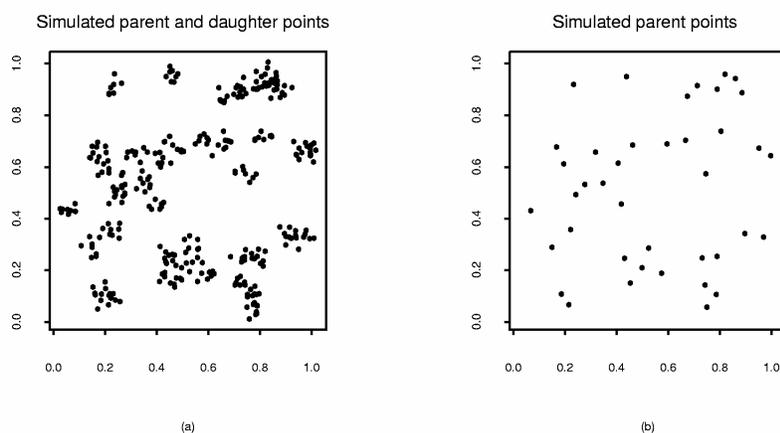

Fig. 2. Examples of simulated events

At each simulation, we generate the above point pattern, and compute the *p*-values corresponding to the values of *K* and *W* statistics. Then we remove one of the daughter events relating to each parent event at random, and compute the *p*-values corresponding to the values of *K* and *W* statistics again for the new set of events with one daughter event removed ($r=1$, where *r* denotes the

number of removed daughter events). In the next step, we remove another daughter event relating to each parent event at random ($r=2$), and repeat the calculation for the p-values. We continue these steps until all daughter events are removed and only the parent events remain.

We repeat the simulation 500 times, and using these p-values, we obtain the relative frequencies of rejection of H_0 by K and W statistics compared to the values of $\alpha=0.01, 0.05, 0.1$ as significance levels. So for a specific α, these relative frequencies can be plotted versus r for each statistic.

The result is plotted in Fig. 3 and indicates that the test statistic W is more powerful than the K statistic, at least for the point pattern described above. We have singled out the K test statistic from Table 8.7 of Cressie [6] because of its better performance compared to many others listed in this table.
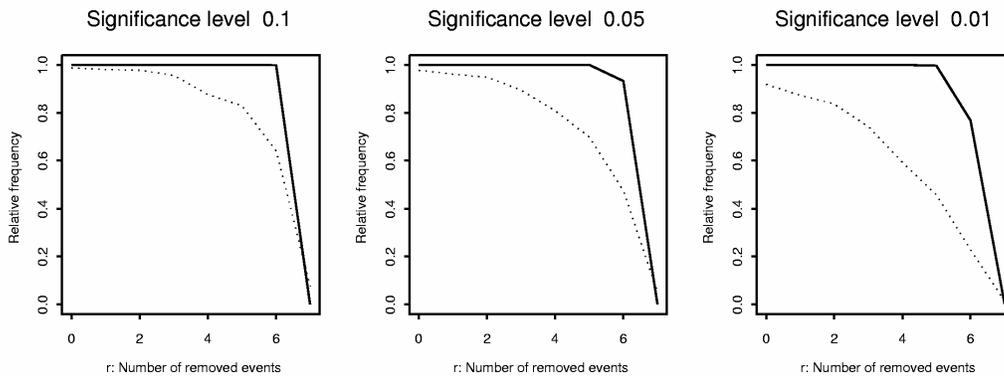

Fig. 3. The power functions for W and K test statistics for 3 different
significance levels, solid and dotted lines, respectively

4. AVERAGE CONCENTRATION LEVEL AND KERNEL IDENTIFICATION

As mentioned in the introduction, many alternatives to the CSR hypothesis in aggregated models of point patterns are extensions of homogeneous Poisson processes. A notable case is the Neyman-Scott [3] process in which for each event (parent event), a random number of offspring (daughter events), are distributed independently of other parents, with locations according to some specified bivariate PDF $h(.)$. The question is whether the parent events, or events assuming such a role, could be recaptured by some mechanism. Of course due to the interactions by daughter events, we do not expect to obtain the "exact parents", but events that "dominate" others in some neighborhood. We will call such events "kernels" for reasons to be explained below.

As a tool for distinguishing these kernels, we define the "average concentration level" of each event in A in the following manner: By assuming as before that there are n events in A, labeled by $i=1,2,...,n$, let $N_i(s)$ denote the number of events in a circle of radius s with center at i, $i=1,2,...,n$. Let $\{t_k, k=1,2,...,\binom{n}{2}\}$ denote the sequence of distances between all pairs of events in A arranged in an increasing order. To take into account the edge effects, we denote by $A_i(s)$ the circle with center at the event labeled by i, $i=1,2,...,n$, and radius s. Define $w_i(s)$ as the weight in some s-neighborhood as

$$w_i(s) = \frac{|A_i(s)|}{|A_i(s)\cap A|} = \frac{\pi s^2}{|A_i(s)\cap A|},$$

where $|B|$ denotes the Lebesgue measure of the set B in R^2.

Now define the "average concentration level" at the event i with parameter $p \in N$ as

$$a_i(p) = \frac{1}{p} \sum_{k=1}^{p} w_i(t_k) N_i(t_k), \quad p \leq \binom{n}{2}.$$

In order to use $a_i(p)$ in finding the kernels, we compute $a_i(p)$ for all i and for some p, so that a maximum value is attained by $a_i(p)$ for some i. We call such an event a "kernel". In theory, p should be taken equal to $\binom{n}{2}$, but for values of n that are not very small, this requires a large number of computations, let alone its usefulness in practice. Therefore we may use low to moderate "values" for p. In the present work, we have started with $p=5$ and have increased it by steps of 1 if no unique maximum values for $a_i(p)$'s have been achieved. Obtaining values of $a_i(p)$ in this manner, we select the event with highest concentration level, i.e. with a maximum value of $a_i(p)$'s. Specifically, the method is carried out according to the following algorithm:

1. Take $j=1$ with j the stage number and form the kernel set S which is the whole set of events at the start.

2. Perform CSR test and stop the algorithm if it is not rejected. Otherwise go to the next step to find the kernels for the jth stage.

3. Compute $c_j = \dfrac{|A|}{n_j - 1}$ where n_j is the number of events in A, with $n_1 = n$.

4. Take $p=5$.

5. Compute $a_i(p)$ for all $i = 1, 2, ..., n_j$.

6. If there is no unique i with a maximum value of $a_i(p)$'s, then increase p by one and go to step 5.

7. Take point i for which $a_i(p)$ is a maximum as kernel k_j.

8. Form a set, say R_j, of events other than k_j which are inside a square with side $\sqrt{c_j}$ and the kernel k_j as its center.

9. Call all points in R_j the relatives of k_j.

10. Set $S = S - R_j$ and then $j = j+1$.

11. Start the next stage by going to step 2 and using the set of remaining events as the data set.

After conducting this algorithm to the end, we will end up with the set of kernels S from which we can form the set of "relatives" to each "parent" by adjoining to each kernel the set of events eliminated on behalf of its kernel. Note that if any of the kernels $k_1, ..., k_{j-1}$ at stage j is included in R_j, that kernel and its relatives will be included in R_j. Also some kernels in S may have no relatives. We call this algorithm the *K*-algorithm for later references. Also, we call the kernels found by using the *K*-algorithm the kernel set.

Remark. The *K*-algorithm above may be regarded as an *ad hoc* method in cluster analysis. In other words, the kernel set may be regarded as the "centers" of the clusters, to be called α-clusters to emphasize the role of significance of level α, and the set of daughters corresponding to each kernel as the data set belonging to that cluster.

4.1 An example

To demonstrate and examine the performance of the above method, we consider a Matern [10] point process with (M, λ, r) as parameters where M is the number of parent events uniformly and independently generated in the unit square, r is the radius of circles with centers at each parent event,

and λ is the mean of the Poisson number of daughter events in each circle. Evidently this is a special case of a Neyman-Scott point process.

For our example we have chosen *M*=10, λ=6 and *r*=0.05. The original Matern process and the result after applying the *K*-algorithm is shown in Fig. 4 and Fig. 5. Unlike the common practice, here we have included the parent events in our point patterns everywhere.
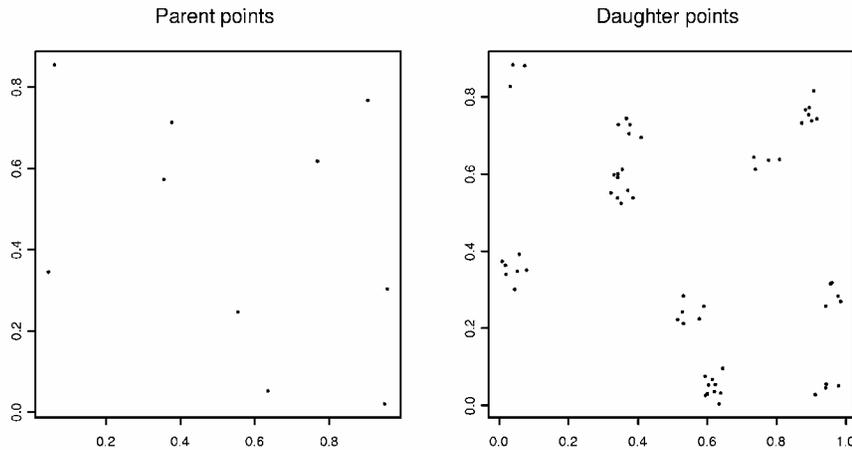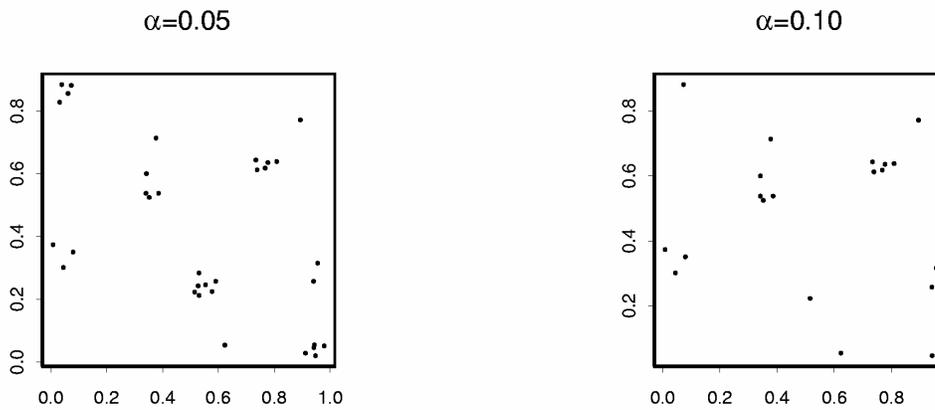

Fig. 4. A simulated Matern point process

Fig. 5. The result of applying the K-algorithm to a simulated Matern point process

Now the question remains about the "proximity" of the real parent events to the kernels found by our *K*-algorithm. This is especially crucial since because of the sequential nature of the tests leading to the kernel set-that assume the role of parent events-it is very difficult to give the overall size of the type 1 error. Therefore, for examining the performance of the *K*-algorithm over this particular point process, i.e. Matern's, we find the Voronoi tessellation corresponding to the real parent events (Gordon [11]). We recall that each such parent event belongs to a region that is the set of all points of the region in question, here the unit square, which is closer to that event than to any other one. We call this region a "face" of the Voronoi tessellation corresponding to the kernel set.

Corresponding to each parent event, we say that a "hit" has occurred if one, and only one, event of the kernel set exists in the face of that event, otherwise we call it a "miss". Now, the number of hits

is an appropriate, but a very strict criterion for examining the proximity of the real parents to the ones obtained by *K*-algorithm, i.e. the kernel set. The reason for its being strict is that when there is a miss, for example when there is no event of the kernel set in one of the faces, it is very likely that this event will appear in another face, thus causing another miss and reducing the number of hits. The results of applying the *K*-algorithm and the above criterion for examining the proximity is given for two different values of α- the significance level-and three different values of *r*-the radius parameter in the Matern process-in Fig. 6. These Figs. show the relative frequency of hits for 200 simulations.
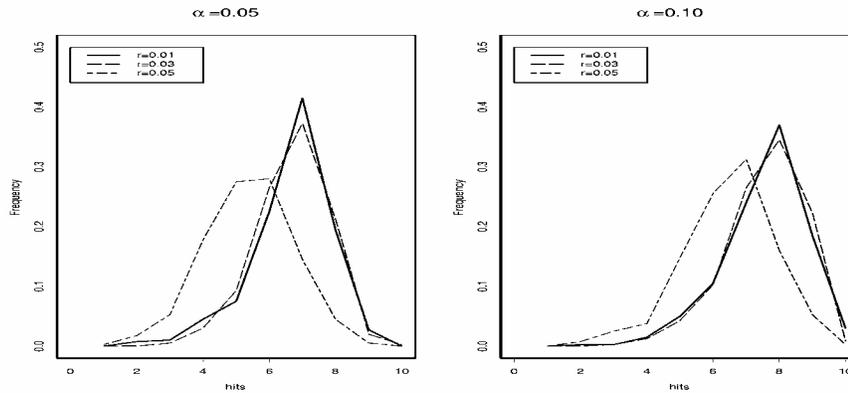
Fig. 6. The frequencies of different simulated hits

5. AN APPLICATION

As a final application, we have conducted the *K*-algorithm to longleaf pine data described in Section 2.3. We have used four different values of α. The kernel set consists of 466, 426, 420, 420, corresponding to the values of 0.05, 0.10, 0.15, and 0.20 for α. Note the relative stability of the kernel set with respect to the increasing value of α. The result is shown in Fig. 7, the y-axis is labeled as frequency in the plots.
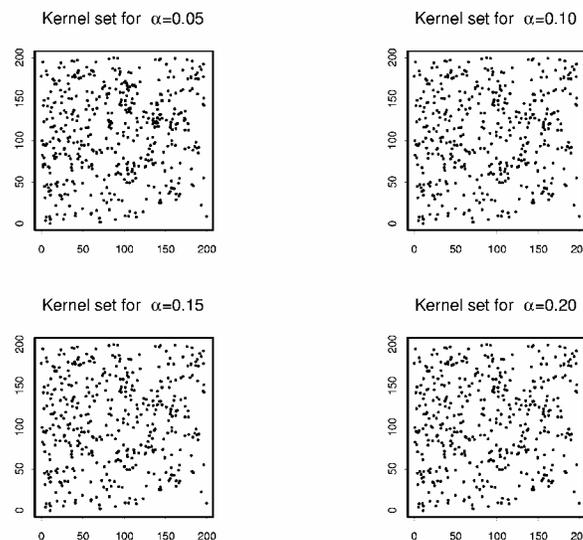
Fig. 7. Kernel set for long-leaf pine data

REFERENCES

1. Mugglestone, M. A. & Renshaw, E. (2001). Spectral tests of randomness for spatial point patterns. *Environmental and Ecological Statistics*, 8, 238–251.

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

3. Neyman, J. & Scott, E. L. (1958). Statistical approach to problems of cosmology (with discussion). *Jr. R. Stat. Soc. B*, 20, 1–43.

4. Vahidi-Asl, M. Q. (2000). Testing for "randomness" in spatial point patterns, using statistics based on one-dimensional inter-event distances. *Journal of Sciences of the Islamic Republic of Iran*, 11, 143–149.

5. Barnard, G. (1963). Contribution to the discussion of Professor Bartlett's paper. *Jr. R. Stat. Soc. B*, 25, 294.

6. Cressie, N. A. C. (1993). *Statistics for Spatial Data*. John Wiley, 2nd edition.

7. Chung, K. L. (1978). *Elementary Probability Theory with Stochastic Processes*. Springer-Verlag, New York, Inc.

8. Karr, A. (1991). *Point processes and their statistical inference.* Second edition. Marcel Dekker, Inc. New York.

9. Diggle, P. J. (1977). The detection of random heterogeneity in plant populations, Biomatrika, 62, 39-48.

10. Matern, B. (1971). *Doubly stochastic Poisson processes in the plane*. In Patil, G. P., Pielou, E. C., & Waters, W. E., editors, *Statistical Ecology*, vol. 1, 195 –213, University Park. Pennsylvania State University Press.

11. Gordon, A. D. (1999). *Classification*. Chapman & Hall, 2nd edition.