A GRAPHICAL TEST TO DETECT INTERFERENCE IN SELECTING NEST SITES

BY Richard F. Green

Technical Report No. 44

Department of Statistics
University of California
Riverside, California 92521

December, 1978

A GRAPHICAL TEST TO DETECT INTERFERENCE IN SELECTING NEST SITES

RICHARD F. GREEN

Department of Statistics
University of California
Riverside, California 92521

Abstract. Birds nesting in boxes may interfere with other members of the same species in the sense that nearby boxes are unlikely to be occupied by members of the same species. Interspecific interference may also occur. In this paper simple methods are described to test for intraspecific and interspecific interference in selecting nest sites. These tests are based on looking at neighboring pairs of nest boxes to see whether they are occupied by the same species, different species or if one or both boxes is empty. A similar procedure is suggested to test for nest site interference in digger wasps. The tests are applied to some data on Great and Blue Tits nesting in Marley Wood, near Oxford, England and on Great Golden Digger Wasps digging their nest burrows in a planter on the campus of the University of Michigan, Dearborn. The wasp data was not sufficient to show interference but intraspecific interference was shown in both Great and Blue Tits. Interspecific interference was not shown for the tits.

Key words: England, graph theory, intra- and interspecific interference, Michigan, nest sites, *Parus caeruleus*, *Parus major*, *Sphex ichneumoneus*, statistical tests.

INTRODUCTION

Individual plants and animals may interfere with each other in the sense that the presence of one individual repels others or interferes with their growth. Plants may retard the growth of neighbors in several ways including shading and chemical interference. Animals often exclude others from their vicinity by territorial behavior. Interference may result in individuals spacing themselves out. Interference may be within species, between species or both.

A number of procedures exist to test whether points are distributed at random in space or whether several types of points are distributed at random with respect to each other. Many of these procedures of importance in ecology are described by Pielou (1969, Chapters 7-16), who is mainly concerned with plants, which have fixed locations. Many of the procedures test whether the individuals are distributed completely at random, that is, Poissonwise. For animals seeking sites for nests or burrows only certain locations may be possible and these possible locations may not be randomly distributed. In this paper a method is described to test whether nest sites are chosen at random from among a number of fixed locations.

The method is illustrated with two examples.

(1) In a study of two species of birds nesting in boxes the locations of the boxes are known and it is known which species nest in which boxes and which boxes are empty. There are two questions here. (a) Are boxes chosen at random within species? (Is there intraspecific interference?). (b) Are boxes chosen at random between species? (Is there interspecific interference?).

(2) Female digger wasps dig their burrows, provision them and then fill them in. The locations of the burrows and the time during which they are active may be observed. The question is whether burrows that are

active at the same time are less likely to be close together than those that are not active at the same time.

The method described in this paper to test for nest site interference is similar to the method used by Knox (1964) in his epidemiological study of childhood leukemia.

METHOD

Leukemia Cases and Digger Wasp Burrows

In his study of childhood leukemia Knox (1964) used data on 96 cases of childhood leukemia recorded in northern England from 1951 to 1960. Knox considered the number of pairs of cases that occurred close together in space (1 km.) and in time (60 days). Knox's method has been given a general formulation in terms of graph theory by Barton and David (1966).

In graph theory a graph is a collection of points and lines which connect some of the points. In the case of Knox's leukemia data the 96 cases used to test for epidemicity could be considered as points with lines connecting all pairs of cases that occurred within 1 km. distance of each other. This would give a graph which we could call the space graph. Again, the 96 cases could be considered as points with lines connecting all pairs of cases whose dates of onset were within 60 days of each other. This would give a graph we could call the time graph. The statistic Knox suggests to test for epidemicity is the number of pairs of cases close together both in space and time. Barton and David (1966) interpret Knox's statistic in terms of the intersection of the time graph and the space graph. (The intersection of two graphs consists of the points and the lines that the two graphs have in common.)

The observed value of Knox's statistic can be compared with the expected value calculated under the assumption that there is no space-time interaction. Using the notation of David and Barton (1966):

X = the observed number of cases that are close in both space and time (Knox's statistic)

N_{1S} = the number of pairs of cases close in space (the number of lines in the space graph)

N_{1T} = the number of pairs of cases close in time (the number of lines in the time graph), and

n = the number of cases (the number of points in each graph),

we have the expected value of X given by

(1) $E(X) = 2N_{1S}N_{1T}/n^{(2)}$, where $n^{(2)} = n(n-1)$ is the second factorial power of n. The r*th* factorial power of n is $n^{(r)} = n(n-1)(n-2)\ldots(n-r+1)$.

To calculate the variance of X the idea of the degree of a point is needed. The degree of a point in a graph is the number of lines to that point. If we denote the degree of the i*th* point in a graph by L_i then we define

(2) $$N_2 = \sum_{i=1}^{n} L_i(L_i - 1)/2.$$

Notice that $N_1 = \sum_{i=1}^{n} L_i/2$ since N_1 is the total number of lines in a graph and each line connects two points.

All that need be observed in order to calculate the mean and variance of Knox's statistic is the degree of each point in the space graph and in the time graph. That is, we must see how many neighbors each point has in space and in time.

For the variance we have

$$\text{(3)}\qquad \mathrm{Var}(X) = 2N_{1S}N_{1T}/n^{(2)} + 4N_{2S}N_{2T}/n^{(3)} - (2N_{1S}N_{1T}/n^{(2)})^2$$
$$+ 4(N_{1S}^{\ 2} - N_{1S} - 2N_{2S})(N_{1T}^{\ 2} - N_{1T} - 2N_{2T})/n^{(4)}.$$

Formulas (1) and (3) give the mean and variance of X under the null hypothesis that the space graph and the time graph are independent. This is equivalent to assigning the observed locations and times to the points at random. It is not necessary to assume that the locations or times themselves are random. In fact, in Knox's leukemia study the cases are more frequent in cities than in the country and are more frequent in summer than in winter.

Exactly the same method as described for leukemia can be used to test whether digger wasps interfere with each other in the choice of nest sites. For the wasps the time graph will have lines connecting each pair of burrows that are active at the same time. For both leukemia cases and wasp burrows the null hypothesis is that the space and time graphs are independent. The only difference is that the alternative hypotheses predict more leukemia cases close together in space and time than under the null hypothesis (epidemicity) and fewer wasp burrows close together in space and time than under the null hypothesis (interference). To test whether the observations are sufficiently different from the expected we can calculate

$$\text{(4)}\qquad z = (X - E(X))/(\mathrm{var}(X))^{\frac{1}{2}}$$

which will be approximately normal with mean 0 and variance 1 under the null hypothesis if the number of burrows is large enough.

Bird nests: Intraspecific Interference

The method suggested in this paper to test for interference between birds in the selection of nest boxes is quite similar to that used in the case of wasp burrows. If we only consider one species of bird, for example, the Great Tit, then the nest boxes are the points and we will have a space graph with lines connecting boxes within some specified distance of each other (say 50 m.). However, in place of a time graph we will have a graph with lines between each pair of nest boxes with Great Tits nesting in them. The test statistic will again come from taking the intersection of two graphs. In this case X will be the number of "neighbor" nest box pairs with Great Tits using each box.

Under the null hypothesis that the Great Tits choose their nest boxes at random the mean and variance of the number of Great-Great neighbor box pairs, X, may be obtained from (1) and (3) where the time graph is replaced by a "species graph" with lines connecting boxes occupied by Great Tits. The calculations will be a bit simpler using the expressions

$$(5) \qquad E(X) = N_{1S}g^{(2)}/n^{(2)}, \text{ and}$$

$$(6) \qquad Var(X) = N_{1S}g^{(2)}/n^{(2)} + 2N_{2S}g^{(3)}/n^{(3)} + (N_{1S}{}^2 - N_{1S} - 2N_{2S})g^{(4)}/n^{(4)} - (N_{1S}g^{(2)}/n^{(2)})^2$$

where g is the number of nest boxes occupied by Great Tits. Again, if E(X) is reasonably large the value of z calculated from (4) will be approximately standard normal under the null hypothesis and may be used to test for intraspecific interference in choosing nest sites.

The same procedure may be followed for Blue Tits as well, merely substituting the number of Blue-Blue neighbor box pairs, Y, for X in (4) and the number of nest boxes occupied by Blue Tits, b for g in (5) and (6).

Bird Nests: Interspecific Interference

Interspecific interference may be tested for by using the number of neighbor nest box pairs with Great Tits occupying one box and Blue Tits occupying the other. Call this number Z.

If there is no intraspecific interference then the null hypothesis of no interspecific interference may be tested by using (4) to compare Z with its mean and variance where

(7) $$E(Z) = 2N_{1S}bg/n^{(2)}, \text{ and}$$

(8) $$\mathrm{Var}(Z) = 2N_{1S}bg/n^{(2)} + 2N_{2S}bg(b + g - 2)/n^{(3)}$$
$$+4(N_{1S}^2 - N_{1S} - 2N_{2S})b^{(2)}g^{(2)}/n^{(4)} - (2N_{1S}bg/n^{(2)})^2.$$

If intraspecific interference exists it must be taken into account when testing for interspecific interference. If there is intraspecific interference but not interspecific interference and if a given box is occupied by Great Tits, say, then a nearby box would be more likely to be occupied by Blue Tits than by chance simply because Great Tits tend to be excluded. Thus if there is intraspecific but not interspecific interference then Great-Blue neighbor box pairs would be more likely than if all boxes are occupied completely at random.

In order to calculate the mean and variance of Z under the null hypothesis of no interspecific interference but conditioned on the observed number of Great-Great and of Blue-Blue neighbor box pairs, X and Y, respectively, it is necessary to calculate the covariances: Cov(X,Y), Cov(X,Z), and Cov(Y,Z). These are given by

(9) $$\mathrm{Cov}(X,Y) = (N_{1S}^2 - N_{1S} - 2N_{2S})b^{(2)}g^{(2)}/n^{(4)} - N_{1S}^2 b^{(2)}g^{(2)}/(n^{(2)})^2, \text{ and}$$

(10) $$\mathrm{Cov}(X,Z) = 2N_{2S}bg^{(2)}/n^{(3)} + 2(N_{1S}^2 - N_{1S} - 2N_{2S})bg^{(3)}/n^{(4)} - 2N_{1S}^2bg^2(g-1)/(n^{(2)})^2.$$

The expression for Cov(Y,Z) will be the same as that for Cov(X,Z) given in (10) but with b and g interchanged.

To test the null hypothesis of no interspecific interference conditional on the values of X and Y observed we compare the observed value of Z with its conditional mean $E(Z|X,Y)$ and variance $\mathrm{Var}(Z|X,Y)$. If we denote the means of the variables X, Y, Z by μ_x, μ_y, μ_z, their variances by a_{xx}, a_{yy}, a_{zz}, and their covariances by a_{xy}, a_{xz}, a_{yz} then we have

(11) $$E(Z|X,Y) = \mu_z + [a(X - \mu_x) + b(Y - \mu_y)]/c, \text{ and}$$

(12) $$\mathrm{Var}(Z|X,Y) = a_{zz} + (a(a_{xz}) + b(a_{yz}))/c$$

where

$$a = (a_{xz}a_{yy} - a_{yz}a_{xy})$$

$$b = (a_{xx}a_{yz} - a_{xy}a_{xz}), \text{ and}$$

$$c = (a_{xx}a_{yy} - a_{xy}^2).$$

RESULTS

The methods described in the preceding section are illustrated using two sets of data.

Marley Wood Tit Data for 1977

Since 1947 a population study of Great Tits, *Parus major*, and Blue Tits, *Parus caeruleus*, has been carried on by members of the Edward Grey Institute of Field Ornithology in Marley Wood, a part of Wytham Wood, near Oxford, England. The birds nest is boxes which have been provided for them. Each year it is observed which boxes are occupied by which birds and nesting success is recorded. Much of this work has been described by Lack (1966).

The locations of the nest boxes in Marley Wood have recently been accurately mapped by Ed Minot (pers. comm., 1978) who found that earlier maps were in error. In 1977 there were 214 nest boxes in Marley Wood. Thirty pairs of Great Tits nested there and 86 pairs of Blue Tits (I have not counted boxes used by the five pairs of Blue Tits that renested in the area and I have only counted once the box that was used twice). The calculations have been done four times, successively defining nest boxes within 20, 30, 40 and 50 meters of each other as neighbors. The results are shown in Table 1.

(Insert Table 1 here)

Testing the data shows significantly too few Blue-Blue neighbor pairs for all four distances considered. There are also too few Great-Great neighbor pairs for the greater distances 40 and 50 meters and possibly for 30 meters as well. Since only 30 boxes were occupied by Great Tits the expected number of Great-Great neighbor pairs within 20 m. is only about 1 and the fact that no such pairs were observed is not statistically significant.

These significance tests indicate that intraspecific interference in the selection of nest boxes exists for both Great Tits and Blue Tits. Tests of interspecific nest site interference do not show significance for any neighbor distance. Although further analyses may yield more difinitive results, the 1977 data suggest that if interspecific interference does occur its effect is small compared to that of intraspecific interference.

Michigan Digger Wasp Data for 1974

The Great Golden Digger Wasp, Sphex icheumoneus provides another example of possible interference in the choice of nest site. The female wasp digs a burrow which she provisions with prey. She lays her eggs on the prey and fills in the burrow. In some cases burrows are abandoned before the process is completed.

Brockmann (1976) has studied digger wasps for several years. In one study she observed wasps digging their burrows in a planter on the campus of the University of Michigan, Dearborn. She observed a total of 79 burrows being dug over the six week study period. At any one time some of the burrows were active, some were abandoned, and some had been filled in completely. The location of each burrow and the time during which it was active were recorded. The question here is whether the position of active nests influenced where a wasp dug a new nest.

Brockmann's 1974 Michigan wasp data are tested for nest site interference using formula (4) where the mean and variance are calculated using formulas (1) and (3). The calculations have been done four times, successively defining burrows within 4, 6, 8 and 10 cm. of each other as neighbors in space. Burrows active at the same time are considered to be neighbors in time. The results are shown in Table 2.

(Insert Table 2 here)

The observed numbers of burrows close together in space and simultaneous in time are not significantly different from those expected by chance. Since the total number of burrows is small the differences would have to be striking in order to be statistically significant.

Several burrows were excavated by the same individual wasps. It might be interesting to see whether each wasp tended to clump or to space out her own burrows. This could be tested by using the same general method used here with neighbor burrows defined in the same way but with the other graph connecting burrows dug by the same individual rather than burrows active at the same time. I have not looked at the data in this way.

DISCUSSION

The Results

The 1974 wasp burrow data suggest there may be nest site interference but the differences observed were not statistically significant. It would take several times as much data as analyzed here to detect interference if it is no stronger than these data suggest.

The 1977 test data do show significant intraspecific interference for both Great Tits and Blue Tits but interspecific has not been shown.

More useful than statistical tests, however, is the direct comparison of the observed with the expected number of intraspecific neighbor pairs. There are neither Great-Great nor Blue-Blue neighbor pairs for the shortest distance considered (20 m.). But while there are still no Great-Great neighbor pairs for 30m. there are 12 Blue-Blue pairs for this distance (with about 27 expected under the null hypothesis). By the time we consider neighbors within 50 m. there are 72 such Blue-Blue neighbor pairs (with about 90 expected) while there are only 3 Great-Great neighbor pairs (with about 11 expected). Thus for a distance of 50 meters there are about 80% of the number of Blue-Blue neighbor pairs expected under the null hypothesis while there are fewer than 30% of the expected number of Great-Great neighbor pairs. This difference suggests that the effect of nest box interference extends further for Great Tits than for Blue Tits. This is certainly reasonable because the Great Tits are substantially larger birds.

Krebs (1971) presented evidence that in years of low Great Tit density the interference effect extended as far as 50 meters. In years of high density the effect did not extend as far. By Krebs' criterion 1977 would be a low density year for Great Tits.

Knox's Method and Other Methods

The methods I have described and illustrated in this paper are based on that used by Knox (1964) in his study of childhood leukemia. Knox's method can be applied directly to test for nest site interference in digger wasps. For wasp burrows there will be a space graph (where neighbor burrows are joined) and a time graph (where burrows active at the same time are joined). For bird nests there will be a space graph (where neighbor nest boxes are joined) and a "species" graph (where boxes occupied by the same species are joined). In each case X will be the number of lines (pairs of nests or burrows) the two graphs have in common. For bird nests the variance of X under the hypothesis of random occupation of boxes will be easier to calculate using formula (6) than using formula (3). I have extended Knox's method to the case of interspecific interference when it is known that intraspecific interference may occur.

Methods similar to Knox's have been used by statistical geographers for some time. This work is discussed in the book by Cliff and Ord (1973). Cliff and Ord give conditions for the asymptotic normality of (4).

Several alternatives to Knox's method have been proposed.

1. Krebs (1971) investigated nest site interference by considering nearest neighbor distances. Analyzing Great Tit data from Wytham Wood, Krebs measured the distance from each nest box occupied by Great Tits to the nearest box also occupied by Great Tits. Years were broken into high and low density years and the observed yearly distributions of nearest neighbor distances were compared with simulated distributions obtained by choosing boxes at random to be occupied. Simulations were done by using one number of boxes for "low" density years and a different number of boxes for "high" density years. The "observed" nearest neighbor distances (obtained by combining nearest neighbor distances for low density years in one group and distances for high density years in another group) were compared with the "theoretical" nearest neighbor

distances by using a Chi-squared test. There are several reasons why this procedure is unsatisfactory. Using a simulation to obtain the theoretical distribution and combining several sets of observations both tend to produce high values for a Chi-squared statistic and nearest neighbor distances are not independent of each other since if two nearby nests are both occupied by Great Tits then each box is likely to have the other as its nearest neighbor and the distance is counted twice. This violates the assumption in the Chi-squared test that observations are independent. Krebs' procedure is more likely to show a significant difference when none actually exists than the stated significance level suggests.

Simulations may, however, be used to produce an exact test for nest site interference. Such procedures are known as Monte Carlo tests (see Cliff and Ord (1973, p. 50) or Hope (1968)). For example, to test for nest site interference among Great Tits using the 1977 Marley Wood data one could pick a test statistic, say the number of Great-Great neighbor pairs within 50 meters. The observed value is: $X = 3$. Then a simulation is done, choosing 30 boxes at random from among the 214 and imagining these are occupied by Great Tits. The statistic is calculated for the simulation and the process is repeated until there are, say 99 simulated populations each with 30 boxes occupied by Great Tits. The statistic, X, is calculated for each of the 99 simulations. If the observed value $X = 3$ is one of the five smallest values when combined with those from the 99 simulations then the null hypothesis that the nest boxes were chosen at random is rejected.

Monte Carlo tests have the advantage of not requiring knowledge of the distribution of the test statistic, X. If the simulation can be done (in this case by choosing 30 boxes at random) the test can be performed.

2. Morris (1975) tested whether the pattern of development of retinal cells was random or nonrandom. She constructed maps of labelled and unlabelled principal cones in the retinae of embryo chicks. The cones were labelled or not according to whether or not they were developing when they were exposed to a radioactive marker. The maps were tested for "contagion" by grid-analysis (Greg-Smith, 1964) or by graphical methods where neighboring pairs of cones were looked at to see whether they were of the same type (both labelled or both unlabelled).

Morris concluded that the graphical methods were generally more successful than grid-analysis in detecting "contagion." One of the difficulties with grid-analysis is that the quadrat boundaries are arbitrary and contagion is likely to act across boundaries as well as within them. This is especially important when local effects are of interest and the grids must be fine.

3. The method of Knox has been generalized by Mantel (1967). Instead of simply counting the number of cases that occur close together in time and in space it is possible to weight the pairs of cases according to how close they are in space and by how close they are in time. If the proper weights are used such a procedure may be more powerful in detecting contagion or interference than Knox's method. Even with Knox's method the distance and time used to define neighbors are arbitrary. The advantage of Knox's method is its conceptual simplicity.

SUMMARY AND CONCLUSION

In this paper a simple method is described to test for interference in the choice of nest sites. Similar methods have been used for some time by epidemiologists and statistical geographers. The method is illustrated with two sets of data. Intraspecific interference in the choice of nest sites has been shown for both Great Tits and Blue Tits but interspecific interference has not been shown. If it exists it is almost certainly weaker than intraspecific interference.

Interference was not shown in the choice of burrow sites in the Great Golden Digger Wasp.

ACKNOWLEDGMENTS

Most of the work on this paper was done while I was on sabbatical from the University of California, Riverside, at the Department of Zoology, Oxford. I would like to thank Dr. C. M. Perrins, director of the Edward Grey Institute of Field Ornithology, for permission to use the Wytham tit nesting, Mr. E. O. Minot for the use of his data on nest locations, and Dr. H. J. Brockmann for the use of her data on the digger wasps. I would like to especially thank Prof. F. N. David, through whom I first heard of the problem of interspecific nest site interference. She suggested the method of solution and has given me a great deal of help and encouragement. An earlier draft of this paper was read and criticized by H. J. Brockmann, E. O. Minot, M. P. Sloan, and F. C. Vasek.

LITERATURE CITED

Barton, D. E. and F. N. David. 1966. The random intersection of two graphs. In: F. N. David (ed.), Research Papers in Statistics, pp. 455-469. John Wiley & Sons, New York.

Brockmann, H. J. 1976. The control of nesting behavior in the Great Golden Digger Wasp, *Sphex ichneumoneus* L. (Hymenoptera, Sphecidae). Ph.D. dissertation, Univ. of Wisconsin, Madison.

Cliff, A. D. and J. K. Ord. 1973. Spatial Autocorrelation. Pion Limited, London.

David, F. N. and D. E. Barton. 1966. Two space-time interaction tests for epidemicity. Brit. J. prev. soc. Med. 20: 44-48.

Grieg-Smith, P. 1964. Quantitative Plant Ecology. Second edition. Butterworths, London.

Hope, A. C. A. 1968. A simplified Monte Carlo significance test procedure. J. Roy. Stat. Soc., Ser. B. 30:582-598.

Knox, G. 1964. Epidemiology of childhood leukemia in Northumberland and Durham. Brit. J. prev. soc. Med. 18: 17-24.

Krebs, J. R. 1971. Territory and breeding density in the great tit, *Parus major* L. Ecology. 52: 2-22.

Lack, D. 1966. Populations Studies of Birds. Clarenden Press, Oxford.

Mantel, N. 1967. The detection of disease clustering and a generalized regression approach. Cancer Research. 27: 209-220.

Morris. V. B. 1975. Non-randomness in the sequential formation of principal cones in small areas of the developing chick retina. J. Comp. Neur. 164: 95-104.

Pielou, E. C. 1969. An Introduction to Mathematical Ecology. John Wiley & Sons, New York.

Table 1

1977 Marley Wood tit data and analysis

n = 214 boxes

g = 30 Great Tit nesting pairs

b = 86 Blue Tit nesting pairs

N_{1S} = the number of neighbor pairs of boxes (unordered)

N_{2S} = half the sum of the degree x (degree - 1) of the points

Distance (meters)	20	30	40	50
L_{1S}	49	170	341	561
L_{2S}	22	255	1059	2900
X	0	0	1	3
Y	0	12	38	72
Z	8	20	41	64
E(X)	.94	3.24	6.51	10.71
E(Y)	7.86	27.26	54.69	89.97
E(Z)	5.55	19.24	38.60	63.51
Var(X)	.92	3.11	6.20	10.15
Var(Y)	6.59	20.78	43.73	71.96
Var(Z)	4.91	16.50	33.02	54.13
Cov(X,Y)	-.15	-.44	-.91	-1.54
Cov(X,Z)	-.10	-.50	-.95	- 1.44
Cov(Y,Z)	-.89	-3.18	-6.28	-10.14
z(X)	-.98	-1.84	-2.21	-2.42
z(Y)	-3.06	-3.27	-2.53	-2.12
$z(Z \mid X,Y)$	.55	-.51	-.17	-.45

Table 2

1974 Dearborn digger wasp data and analysis

$n = 79$ burrows

N_{1S} = number of neighbor pairs (within chosen distance)

N_{1T} = number of simultaneous pairs of burrows

$$N_{2S} = \sum_{i=1}^{n} l_i^{(2)}/2$$

$$N_{2T} = \sum_{i=1}^{n} k_i^{(2)}/2, \text{ where}$$

l_i = local degree of the ith burrow in space

k_i = local degree of the ith burrow in time

X = number of simultaneous neighbor pairs.

Distance (cm.)	4	6	8	10
N_{1S}	8	17	26	42
N_{1T}	800	800	800	800
N_{2S}	3	11	26	74
N_{2T}	18513	18513	18513	18513
X	3	3	4	8
EX	2.08	4.41	6.75	10.91
Var(X)	1.57	3.35	5.18	8.66
z	.74	-.77	-1.21	-.99