Comparison of Statistical Methods for Studying Spatial Patterns of Soilborne Plant Pathogens in the Field

P. C. Nicot, D. I. Rouse, and B. S. Yandell

First two authors, Department of Plant Pathology; third author, Departments of Horticulture and Statistics, University of Wisconsin, Madison 53706.

The authors wish to thank C. D. Upper for critically reviewing the manuscript. The assistance of Steve Vicen in the preparation of the figures is gratefully acknowledged.

Accepted for publication 24 September 1984.

Spatial patterns of diseased plants and of soilborne plant pathogens in fields or research plots have recently gained well-deserved attention. Knowledge from studies of such patterns can be applied to improve sampling or survey methods and may also provide a better understanding of the biology and ecology of the pathogen and of the disease (1–3,8–11,13,14,16,18,21,22,25–27). Past studies of spatial patterns of soilborne inoculum have relied mostly on methods based exclusively on the examination of the mean and variance or of the frequency distribution of observed inoculum levels. Since these methods do not take into account the actual location of the sampling sites in the study area, the following questions arise: Do these methods actually provide information on the "spatial pattern" of the pathogen in the field and, if so, at what scale? Can additional information be obtained with methods that take into account the location of samples?

We will illustrate in this letter to the editor the point that frequency distribution analysis of propagule counts in soil samples provides only limited information on spatial patterns. There are several alternative methods of statistical analysis that utilize information on actual location of samples with respect to each other. An example of one of these methods will be described to illustrate their advantage.

Methods most commonly used for studying the spatial patterns of soilborne plant pathogens. Soil samples are usually taken at random or at regularly spaced sites over the study area (2,8,10,11,14,21–23,27). These samples usually consist of a number of individual soil cores bulked together. The number of propagules per volume or mass unit of soil in each sample is then estimated, and the mean m and variance V are estimated. A frequency table is compiled to show the number of samples that contained 0, 1, 2, ... propagules per unit of soil, and the observed frequency distribution is compared with theoretical frequency distributions such as Poisson, negative binomial, Neyman type A, etc. Depending on which of these distributions fit the observed data, conclusions are drawn about the pattern of the pathogen in the study area (2,8,11,14,21,22,27). Other criteria commonly used in plant pathology to characterize patterns of soilborne inoculum include Fisher's variance-to-mean ratio V/m, David and Moore's index of clumping $IC = V/m - 1$, or Lloyd's indices of mean crowding $m^* = m + (V/m - 1)$ and patchiness m^*/m (8,10,11,14,21,23,27).

Theoretical basis for these methods. If the inoculum is randomly dispersed throughout the soil of the study area, the frequency distribution of propagule counts in the samples should be Poisson (4,17). Under this hypothesis, $V/m = 1$, $IC = 0$, and $m^*/m = 1$, because the mean and the variance are equal for a Poisson distribution. A good fit of a Poisson model to observed data is, therefore, an indication that the inoculum is likely to be randomly scattered throughout the soil, but this may not necessarily be true, as will be illustrated. A poor fit of a Poisson model and a good fit of "contagious distributions" such as negative binomial, Neyman type A, etc, suggest that at least one of the assumptions underlying the Poisson process is violated; for example, the propagules may occur in clumps. Whereas a random pattern is well defined by a Poisson process, an aggregated pattern can be aggregated in many different ways that result from very different biological phenomena. Two particular types of aggregated patterns are well documented (4,17). The "true contagion process" (ie, a generalized Poisson process) describes a pattern in which the pathogen occurs in small clusters, the clusters themselves being randomly scattered throughout the soil (4,17). Such a pattern could result, for example, from the decomposition of fragments of host tissue randomly dispersed in the soil and carrying groups of propagules of a pathogen. The "apparent contagion process" (ie, compound Poisson process) describes a pattern in which the propagules are randomly scattered within each sample but the average number of propagules within a sample is also a variable that takes different values from sample to sample (4,17). An example of such a situation would be one in which the pathogen is randomly dispersed in a given soil type and the size of the population is a function of the type of soil. The apparent contagion would then correspond to large-scale variations of soil type in a field. In the first case, the aggregation occurs at a scale smaller than that of the soil sample, whereas in the latter case it occurs at a larger scale.

Limitations of these methods. Depending on the distribution of numbers of propagules per cluster for the generalized Poisson process and on the distribution of the average propagule numbers in the samples for the compound Poisson process, well-defined distributions including negative binomial and Neyman type A, may fit the frequency distribution for the number of propagules per sample. However, both of these distributions can be obtained from either a generalized or a compound Poisson process. Therefore, the frequency distribution of inoculum levels cannot be used to determine which process accounts for the observed aggregation unless additional information is available (4,17).

Another limitation of methods that rely only on frequency distributions comes from the fact that the information on the location of each sampling site is ignored when the frequency tables are compiled. Consider hypothetical square fields divided into contiguous quadrats of equal size and infested with varying levels of a given soilborne plant pathogen. Suppose that soil samples are taken from each quadrat and that a reliable method is available to quantify the pathogen in field soil. Assuming that records are kept of the precise location of each sample, suppose that the inoculum is found dispersed over three fields as described in Fig. 1. For these three fields the frequency of inoculum levels is identical and a Poisson model fits well the observed data (Fig. 2). From this result, it might be concluded that no departure from randomness can be detected in any of those three fields, based on the analysis of frequency distributions. However, it can easily be seen that the pattern of dispersion of the pathogen, *at a scale larger than the quadrat* is not random for all three fields of Fig. 1. Suppose now that the hypothetical fields sampled are similar to those described in Fig. 3. The frequency distribution of inoculum levels in all three fields is identical; the observed data give a poor fit for the Poisson model and a good fit for a negative binomial model (Fig. 4). This suggests some aggregation in the pattern of the pathogen in all three fields at some unknown scale. However, the examination of frequency distributions alone does not allow us to distinguish the very different spatial arrangements of inoculum levels represented in Fig. 3.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. § 1734 solely to indicate this fact.

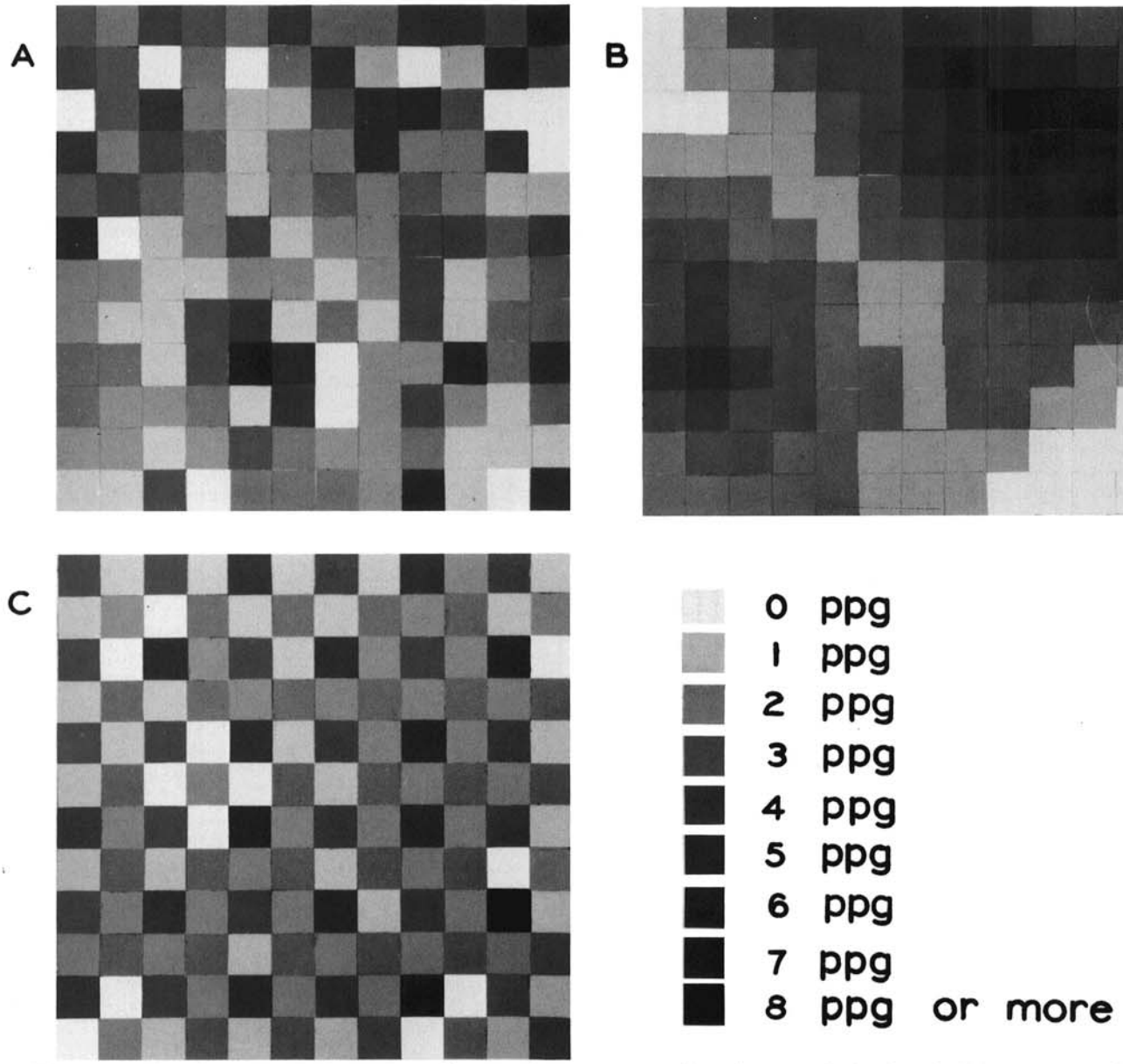

Fig. 1. Three hypothetical fields corresponding to three different spatial arrangements of one frequency distribution of soil inoculum levels of a plant pathogen. The frequency distribution is described by a Poisson with mean $m = 2.5$ propagules per gram of soil (ppg). The spatial arrangements appear: **A,** random; **B,** aggregated; and **C,** regular.

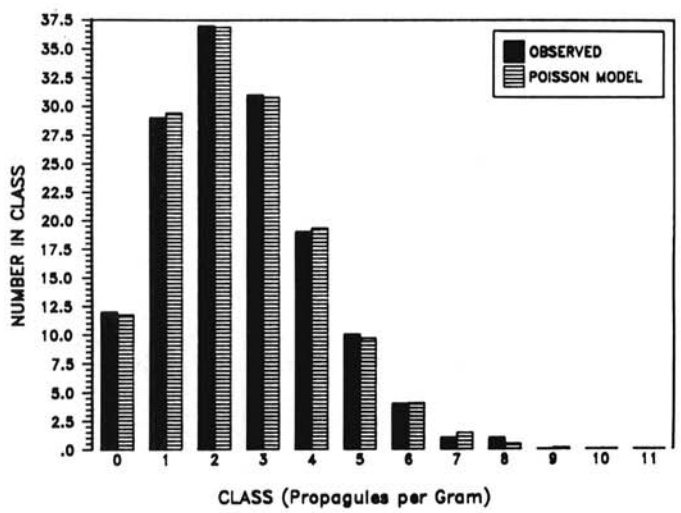

Fig. 2. Fit of a Poisson model with mean $m = 2.5$ on the observed frequency distribution of inoculum levels from the fields of Fig. 1. $P(\chi^2 > 0.031) > 0.999$.

Information provided by methods taking into account the location of each sample. Besides those discussed earlier, many statistical methods used for the analysis of spatial patterns, including nearest neighbor, Greig-Smith's method, spatial autocorrelation analysis, and two-dimensional spectral analysis take into account the location of each sample (4–7,12,17–20,24). Some of these methods, which were devised to analyze the patterns of discrete objects, assume that distances can be measured between individual objects (eg, nearest neighbor methods) or that the coordinates of every object are known over the study area (4–6,17,18); for practical reasons, therefore, they could not be applied to a study of patterns of soilborne inoculum. Most other methods discussed in the references cited above can be applied to both continuous data and counts of discrete objects, including soilborne inoculum, in samples or quadrats. Studies of patterns, based on the computation of coefficients of autocorrelation, have recently been reported in plant pathology (16,22) and are common in ecology (24). Analytical methods based on a coefficient of spatial autocorrelation, the "*I* statistic," were first introduced by Moran in 1950 (15) and further investigated by Cliff and Ord (4) and others (12,24). The *I* statistic is easily computed, and is very similar to the commonly used coefficient of correlation between two random

variables (4). With the I statistic, the inoculum level X at each sampling location i is compared to the values of X at locations neighboring i, instead of being compared to a second random variable Y as it would with a conventional coefficient of correlation. Practically, I is positive if X tends to be high in some groups of neighboring quadrats and low in other groups of neighboring quadrats (aggregated spatial pattern; eg, Figs. 1B and 3B). I is negative if high values of X tend to be located near low values of X and vice versa (regular pattern; eg, Figs. 1C and 3C). And finally, I is approximately equal to zero if no trend is present in

TABLE 1. Characterization of the spatial patterns of a soilborne fungus in six hypothetical fields, using three different methods

Fields[a]	Best-fitting frequency distribution	Indices			Moran's I[c]	
		V/m[b]	Lloyd's index of mean crowding	Lloyd's index of patchiness	$I \pm$ SE	P value [$I = E(I)$]
1-A	Poisson	1.01	2.52	1.00	0.017 ± 0.061	0.697
1-B	Poisson	1.01	2.52	1.00	0.874 ± 0.061	0.000
1-C	Poisson	1.01	2.52	1.00	-0.582 ± 0.061	0.000
3-A	Negative binomial	2.37	3.91	1.54	0.009 ± 0.060	0.787
3-B	Negative binomial	2.37	3.91	1.54	0.915 ± 0.060	0.000
3-C	Negative binomial	2.37	3.91	1.54	-0.526 ± 0.060	0.000

[a] Fields are as described in Figs. 1 and 3.
[b] Variance (V)-to-mean (m) ratio.
[c] Spatial autocorrelation coefficient: each observation is compared with its four immediate neighbors. The expected value of I under the hypothesis of randomness is $E(I) = -0.007$ for all six fields. The I statistic was assumed to be approximately normally distributed for the computation of the p values (see justification in Cliff and Ord [4]).

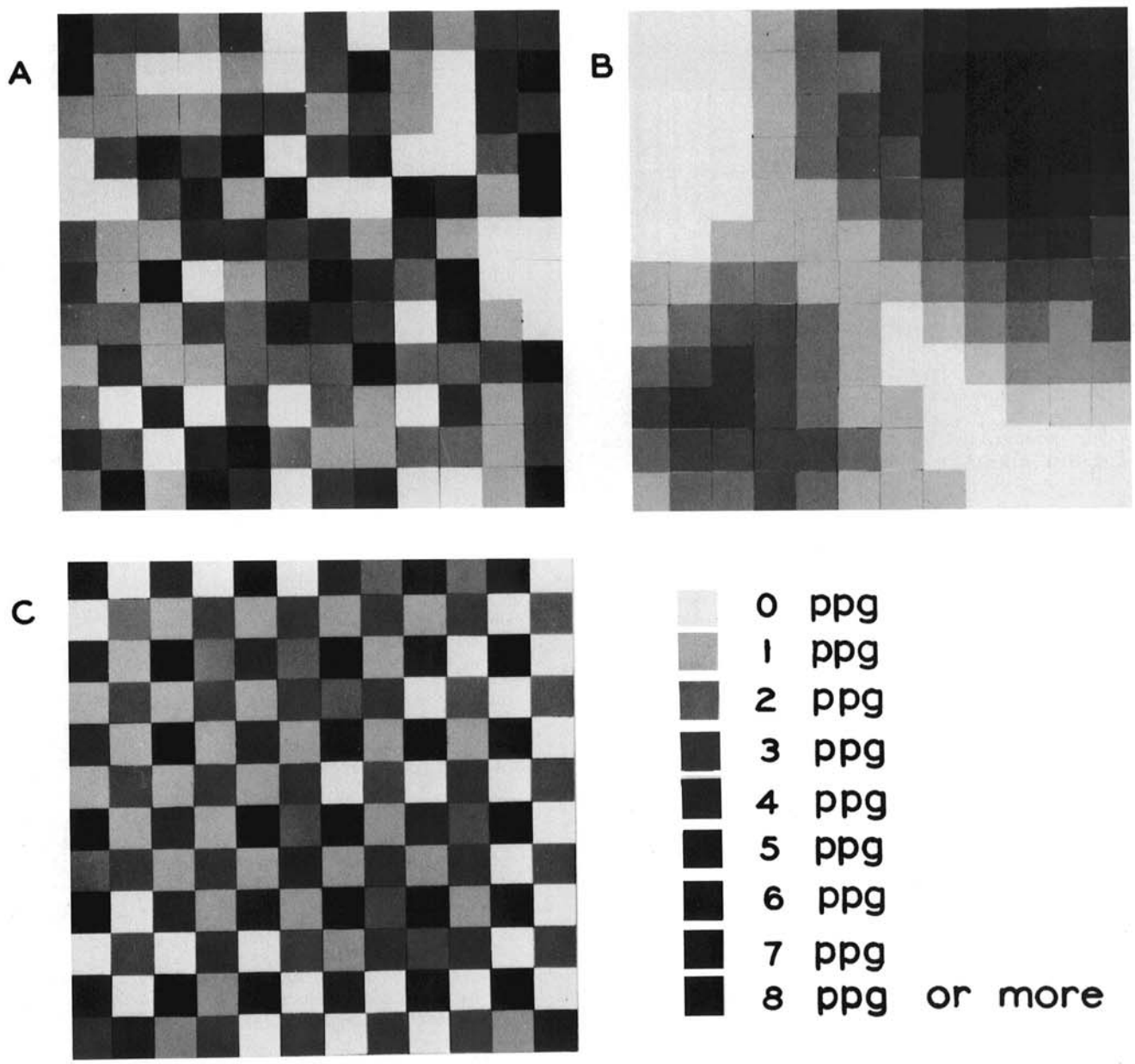

Fig. 3. Three hypothetical fields corresponding to three different spatial arrangements of one frequency distribution of soil inoculum levels of a plant pathogen. The frequency distribution is described by a negative binomial with mean $m = 2.5$ propagules per gram of soil (ppg) and parameter $k = 1.75$. The spatial arrangements appear: **A,** random; **B,** aggregated; and **C,** regular.

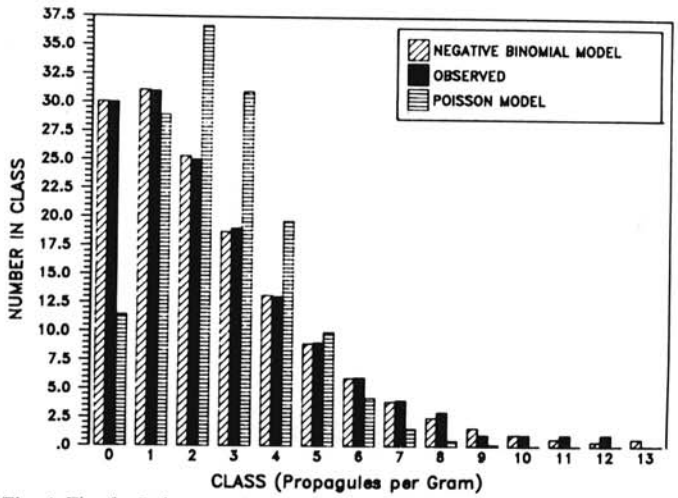

Fig. 4. Fit of a Poisson and of a negative binomial model on the observed frequency distribution of inoculum levels from the fields of Fig. 3. Poisson model with mean $m = 2.5$; $P(\chi^2 > 58.7) < 0.001$. Negative binomial model with mean $m = 2.5$ and parameter $k = 1.75$; $P(\chi^2 > 0.014) > 0.999$.

the spatial pattern.

To illustrate the information provided by methods that take into account the location of samples, a value of the I statistic, its expected value, and its standard error were computed for each of the fields represented in Figs. 1 and 3 (Table 1). The inoculum level of each quadrat was compared with those of its four immediate neighbors. These calculations were performed on an Apple II plus microcomputer with a BASIC program, which is available from the authors. In contrast to the results obtained with the other methods presented in Table 1, a different value of the I statistic was obtained for each individual field of Figs. 1 and 3. Assuming that I is approximately normally distributed (see justification in [4]), fields 1-B and 3-B showed significant positive autocorrelation, fields 1-C and 3-C showed significant negative autocorrelation, and no significant autocorrelation was observed for fields 1-A and 3-A. This indicates that the patterns at a scale larger than the quadrat are aggregated, regular, and random for fields 1-B and 3-B, 1-C and 3-C, and 1-A and 3-A, respectively, which corresponds to the visual perceptions given by the graphic representation of the fields.

DISCUSSION

The frequency distribution of inoculum levels in soil samples is valuable information about a study area. However, the examples presented in Figs. 1 and 3 and Table 1 show that the examination of frequency distributions alone allows only limited conclusions about the spatial patterns of soilborne inoculum. Frequency distribution analysis does not adequately discriminate among random, aggregated, or regular dispersion of the pathogen over the field, at a scale larger than that of the quadrat or soil sample. On the other hand, methods that take into account the location of samples, such as those based on the analysis of spatial autocorrelation, may allow one to distinguish such patterns. They could also be particularly useful when no well-defined theoretical model adequately fits the observed frequency (23).

Regardless of the type of statistical method used to analyze the data, there still may be several difficulties in studying dispersion patterns of soilborne inoculum. For example, in a single soil core all the propagules of the pathogen could be located on the surface or they could be randomly scattered throughout the sample. Likewise, if a series of soil cores are bulked together in one sample, there is no way to determine how the propagules were dispersed among the cores. If the study area is divided into quadrats and samples are taken from each quadrat, one may ask how representative each sample is of its quadrat. How similar would the estimate of the inoculum level be if a second sample were taken from the same quadrat? Would the conclusions about the pattern be similar if a smaller or larger quadrat size were chosen?

Attempts to answer these questions in conjunction with the use of methods of analysis that take into account the location of sampling sites will provide a broader knowledge of the spatial patterns of soilborne plant pathogens.

LITERATURE CITED

1. Basu, P. K., Lin, C. S., and Binns, M. R. 1977. A comparison of sampling methods for surveying alfalfa foliage diseases. Can. J. Plant Sci. 57:1091-1097.
2. Brewer, B. J., Campbell, C. L., and Beute, M. K. 1981. Inoculum distribution of *Sclerotium rolfsii* and incidence pattern of southern stem rot on peanut in North Carolina. (Abstr.) Phytopathology 71:863.
3. Campbell, C. L., and Pennypacker, S. P. 1980. Distribution of hypocotyl rot caused in snapbean by *Rhizoctonia solani*. Phytopathology 70:521-525.
4. Cliff, A. D., and Ord, J. K. 1981. Spatial Processes: Models and Applications. Pion Ltd., London. 266 pp.
5. Diggle, P. J. 1979. On parameter estimation and goodness-of-fit testing for spatial point patterns. Biometrics 35:87-101.
6. Diggle, P. J., and Cox, T. F. 1983. Some distance-based tests of independence for sparsely-sampled multivariate spatial point patterns. Int. Statist. Rev. 51:11-23.
7. Gilligan, C. A. 1982. Statistical analysis of the spatial pattern of *Botrytis fabae* on *Vicia faba*: A methodological study. Trans. Br. Mycol. Soc. 79:193-200.
8. Goodell, P. B., and Ferris, H. 1980. Plant parasitic nematode distributions in an alfalfa field. J. Nematol. 12:136-141.
9. Goodell, P. B., and Ferris, H. 1981. Sample optimization for five plant parasitic nematodes in an alfalfa field. J. Nematol. 13:304-313.
10. Griffin, G. J., Garren, K. H., and Taylor, J. D. 1981. Influence of crop rotation and minimum tillage on the population of *Aspergillus flavus* group in peanut field soil. Plant Dis. 65:898-900.
11. Hau, F. C., Campbell, C. L., and Beute, M. K. 1982. Inoculum distribution and sampling methods for *Cylindrocladium crotalariae* in a peanut field. Plant Dis. 66:568-571.
12. Jumars, P. A., Thistle, D., and Jones, M. L. 1977. Detecting two-dimensional spatial structure in biological data. Oecologia 28:109-123.
13. Lin, C. S., Poushinsky, G., and Mauer, M. 1979. An examination of five sampling methods under random and clustered disease distribution using simulation. Can. J. Plant Sci. 59:121-130.
14. Martin, S. B., Campbell, C. L., and Lucas, L. T. 1983. Horizontal distribution and characterization of *Rhizoctonia* spp. in tall fescue turf. Phytopathology 73:1064-1068.
15. Moran, P. A. P. 1950. Notes on continuous stochastic phenomena. Biometrika 37:17-23.
16. Nicot, P. C., Rouse, D. I., and Yandell, B. S. 1983. Spatial statistics as a tool to compare the distribution of *Verticillium dahliae* in the soil to that of diseased plants in a commercial potato field. (Abstr.) Phytopathology 73:803.
17. Pielou, E. C. 1977. Pages 113-165 in: Mathematical Ecology. J. Wiley & Sons, New York. 385 pp.
18. Poushinsky, G., and Basu, P. K. 1984. A study of distribution and sampling of soybean plants infected with *Pseudomonas syringae* pv. *glycinea*. Phytopathology 74:319-326.
19. Renshaw, E., and Ford, E. D. 1983. The interpretation of process from pattern using two-dimensional spectral analysis: Methods and problems of interpretation. Appl. Statist. 32:51-63.
20. Ripley, B. D. 1981. Spatial Statistics. J. Wiley & Sons, New York. 252 pp.
21. Roth, D. A., and Griffin, G. J. 1981. Cylindrocladium root rot of black walnut seedlings and inoculum pattern in nursery soil. Can. J. Plant Pathol. 3:1-5.
22. Shew, B. B., Beute, M. K., and Campbell, C. L. 1984. Spatial pattern of Southern stem rot caused by *Sclerotium rolfsii* in six North Carolina peanut fields. Phytopathology 74:730-735.
23. Smith, V. L., and Rowe, R. C. 1984. Characteristics and distribution of propagules of *Verticillium dahliae* in Ohio potato field soils and assessment of two assay methods. Phytopathology 74:553-556.
24. Sokal, R. R., and Oden, L. 1978. Spatial autocorrelation in biology. 1. Methodology. 2. Some biological implications and four applications of evolutionary and ecological interest. Biol. J. Linn. Soc. 10:199-249.
25. Strandberg, J. 1973. Spatial distribution of cabbage black rot and estimation of diseased plant populations. Phytopathology 63:998-1003.
26. Stanghellini, M. E., Stowell, L. J., Kronland, W. C., and von Bretzel, P. 1983. Distribution of *Pythium aphanidermatum* in rhizosphere soil and factors affecting expression of the absolute inoculum potential. Phytopathology 73:1463-1466.
27. Taylor, J. D., Griffin, G. J., and Garren, K. H. 1981. Inoculum pattern, inoculum density-disease incidence relationships, and population fluctuations of *Cylindrocladium crotalariae* microsclerotia in peanut field soil. Phytopathology 71:1297-1302.