A Study of Distribution and Sampling of Soybean Plants Naturally Infected with *Pseudomonas syringae* pv. *glycinea*

G. Poushinsky and P. K. Basu

Statistician and plant pathologist, respectively, at the Engineering and Statistical Research Institute and Ottawa Research Station, Agriculture Canada, Ottawa, Ontario, K1A 0C6. Contributions ESRI I-386 and ORS729, respectively.
Accepted for publication 30 August 1983.

ABSTRACT

Poushinsky, G., and Basu, P. K., 1984. A study of distribution and sampling of soybean plants naturally infected with *Pseudomonas syringae* pv. *glycinea*. Phytopathology 74:319-326.

Soybean (*Glycine max*) plants were grown in a field plot (42.7 by 53.9 m) at Ottawa in 1978, and similar sized areas from eight farmers' fields were examined in 1982 for the presence of bacterial blight caused by *Pseudomonas syringae* pv. *glycinea*. Tagged plants were assessed several times during the growing season. The pattern of disease occurrence was investigated by using three existing indices of nonrandomness and a fourth method was proposed and applied. Diseased plants were distributed nonrandomly from early to midseason. Samples of various sizes following three types of sampling paths were taken to estimate disease incidence. Simple random samples were inadequate to assess disease percentage when the underlying disease distribution was nonrandom.

Bacterial blight caused in soybean (*Glycine max* (L.) Merr.) by *Pseudomonas syringae* pv. *glycinea* Young, Dye & Wilkie (11) is widespread in the temperate regions (2,18). The pathogen is seedborne (15), can remain in buds (19), and also can overwinter in infected leaf-debris (9,16,18). After the initial lesions are produced, the disease can spread rapidly with wind and rain storms (8), but its distribution pattern is not fully understood. The same field may include both "heavily" and "lightly" diseased areas (8). This obviously poses some problems in sampling to estimate the percentage of infected plants in a field, since many schemes assume a random distribution of the character of interest (5).

There are many methods available for assessing disease distribution in a field. The quadrat methods require excessive sample sizes and the results are affected by the relation of quadrat size to the underlying disease distribution (12,22–24). Attention has also been focussed on distance methods. Stauffer (26) examined the three widely used indices of nonrandomness of Pielou (23), Clark and Evans (6), Hopkins and Skellam (14) and concluded that Pielou's was the best. However, these depend on a complete list of diseased individuals being available and were not developed in the context of regularly spaced plants on the plane. An alternative is presented here which examines the distribution of distances between a plant and its nearest diseased neighbor obtained in a sample from the field.

The main objectives of this work were to study the distribution pattern of soybean plants naturally infected by the bacterial blight pathogen and to evaluate the shape of sampling paths (X, W, or random) and the sizes of samples used to estimate the percentage of diseased plants in a field.

MATERIALS AND METHODS

Soybeans (cultivar Maple Arrow) were planted in a 42.7 × 53.9-m field plot at Ottawa on 16 May 1978. This plot was divided by access roads into four quadrants each 13.3 × 24.4 m. In each quadrant, there were 26 rows, 71 cm apart; and the in-row distance between plants was 5–7 cm. Three border rows on each side of a quadrant and ten plants at each end of a row were excluded to minimize border effects. In each of the remaining 20 rows, 30 plants ~75 cm apart were tagged for recording the presence or absence of the disease (600 per quadrant). The selected plants were in a lattice pattern. The presence or absence of the disease was recorded five times during the growing season (27 June, 11 and 25 July, 10 and 20 August) for all 2,400 tagged plants. In 1982, eight farm fields planted to the same cultivar (cultivar Maple Arrow) and with the same row spacing were selected from three counties in eastern Ontario (Ottawa-Carleton, Dundas, and Stormont) and two areas similar in size and shape to the earlier described quadrants were laid out randomly in each field at the beginning of the season. The 600 plants per quadrant (chosen in the same manner as in 1978) were assessed for the presence of disease symptoms twice during the growing season. For each quadrant, the true incidence of disease was taken to be the proportion of the 600 tagged plants which were diseased. All analyses presented here were performed using the data from the selected plants. This lattice of plants was judged to adequately represent the true field conditions.

The diagnosis of the disease was confirmed by isolating the pathogen (*Pseudomonas syringae* pv. *glycinea*) and testing its pathogenicity on the same cultivar by standard methods (7).

The disease incidence data from the tagged plants were analyzed to discover if there was any pattern in the distribution of infected plants by examining graphic plots of the data; by calculating the indices of nonrandomness of Pielou, Hopkins and Skellam, and Clark and Evans; and by using a new method which assesses the distribution of diseased plants in terms of distances to the nearest diseased plant in a sample. The indices of nonrandomness use two of the following three items:

(a) An estimate of the density of diseased plants;
(b) A sample of diseased-plant to nearest-diseased-plant distances; and
(c) A sample of random-point to nearest-diseased-plant distances.

Pielou's index (henceforth referred to as P) is calculated by using items (a) and (c):

$$P = \pi\rho\bar{\omega}$$

in which ρ = density of diseased plants (per unit area) and $\bar{\omega}$ = average for the sample of the squared distance from a random point to the nearest diseased plant.

Hopkins and Skellam's index (HS) is calculated by using items (b) and (c):

$$HS = \bar{\omega}/\bar{\omega}_1$$

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.

in which $\bar{\omega}_1$ = average squared distance from a random diseased plant to the nearest diseased plant and Clark and Evans index (CE) is calculated by using items (a) and (b):

$$CE = 2-2\sqrt{\rho \bar{r}}$$

in which $\bar{r}$ = average distance from a random diseased plant to the nearest diseased plant.

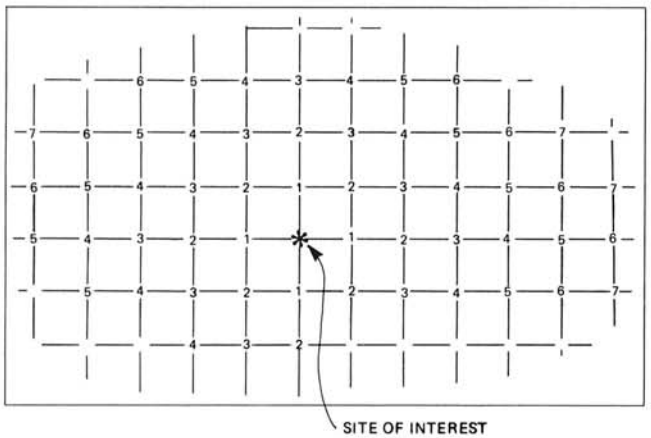

Fig. 1. Distances ("steps") of neighbors in the field from a site of interest.

These are used to test the hypothesis that diseased plants are distributed randomly. If the index value falls below a critical point (depending on the method and the sample size) the distribution is judged to be regular (R); if it falls above another critical point it is judged to be clustered (C); otherwise, it is judged to be random (r). The critical values are tabulated (26). We computed the indices using samples of 200 distances.

These three methods suffer from the drawback that they were developed for the situation where plants (diseased or not) can lie anywhere in the field. In our situation the plants were located in a lattice pattern imposed by the rows of the field; thus, the underlying assumption of a continuous (rather than discrete) distance measure is not met.

Recent literature (3) has looked at fitting an "auto-model" to the data, but the assumption is usually made that the bulk of usable information is contained in the four nearest plants (diseased or not). Diggle et al (10) and Besag and Gleaves (4) proposed "T^2" sampling, but this requires a sophisticated sampling procedure. Runs (21), doublets (27), and more complicated procedures of this type (24,25) have been used but these require a full map of disease incidence in a field. These seem unnecessarily restrictive and time consuming.

In the present situation of a lattice pattern of tagged plants another approach was used to assess the randomness of the underlying disease distribution. If a proportion t ($0 \leqslant t \leqslant 1$) of the plants in a field is diseased and these plants are distributed independently and randomly throughout the field, then for any randomly selected plant the probability that it is diseased is t. In the

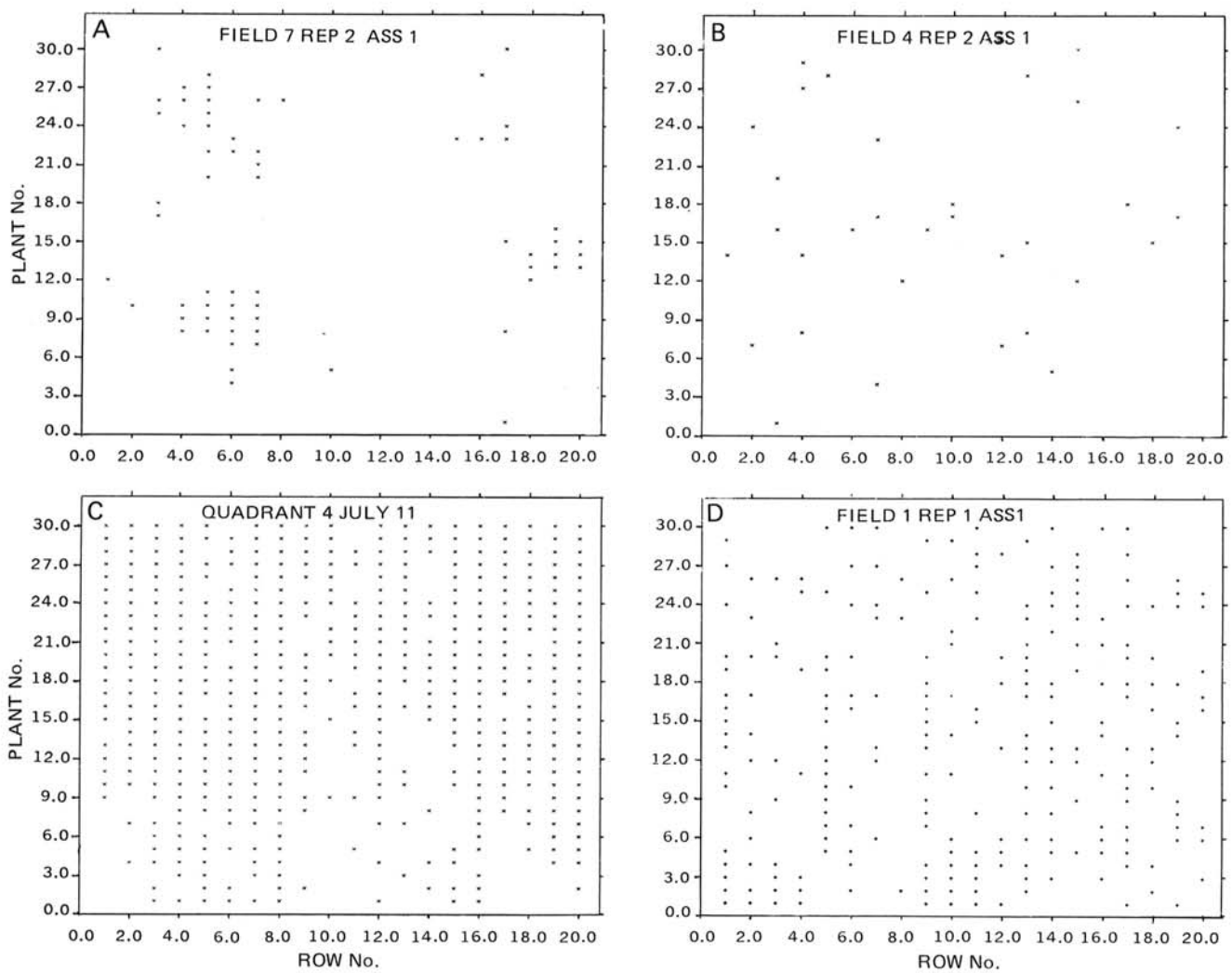

Fig. 2. Patterns of infected plants (×) in several quadrants: **A,** 1982, Field 7, Rep 2, Assessment 1 (11% diseased); **B,** 1982, Field 4, Rep 2, Assessment 1 (5% diseased); **C,** 1978, Quadrant 4, July 11 (78% diseased); **D,** 1982, Field 1, Rep 1, Assessment 1 (39% diseased).

field we can identify plants by row number and the position within a row (ie, plant *ij* is the *j*th plant in the *i*th row) and we define the "distance" between plants to be:

$$\text{distance from plant } ij \text{ to plant } st = \text{abs}(i - s) + \text{abs}(j - t)$$

in which "abs" is the absolute value. That is, the distance between locations is the number of "steps" one would take to go from one location to another where these steps must be along a row or at right angles to a row. Fig. 1 shows diagrammatically how to obtain these distances. Then, under the assumption that the disease is distributed independently and randomly, the probability (*Pr*) that the distance from any plant to its nearest diseased neighbor is *i* "steps" is:

Pr (nearest diseased neighbor is *i* "steps" away) =

TABLE 1. Pielou's (P), Clark and Evans' (CE), and Hopkins and Skellam's (HS) indices of nonrandomness, sample size 200, for 1978 soybean field sites affected by bacterial blight; P/CE/HS[a]

Date	Quadrant 1	Quadrant 2	Quadrant 3	Quadrant 4
27 June	1.84/0.35/0.87 C*/R*/r 66	1.19/0.19/0.38 C/R*/R* 82	1.86/0.38/0.80 C*/R*/R 63	0.85/0.18/0.32 R/R*/R* 83
11 July	0.76/0.29/0.29 R*/R*/R* 80	0.80/0.13/0.23 R*/R*/R* 88	0.79/0.15/0.22 R*/R*/R* 86	1.07/0.22/0.41 r/R*/R* 78
25 July	0.78/0.19/0.28 R*/R*/R* 82	0.95/0.22/0.34 r/R*/R* 79	0.88/0.21/0.32 r/R*/R* 80	0.72/0.03/0.19 R*/R*/R* 97
10 August	0.60/0.01/0.18 R*/R*/R* 98	0.63/0.02/0.19 R*/R*/R* 98	0.65/0.01/0.19 R*/R*/R* 99	0.60/0.01/0.21 R*/R*/R* 99
22 August	0.68/0.00/0.18 R*/R*/R* 100	0.67/0.00/0.19 R*/R*/R* 100	0.58/0.00/0.17 R*/R*/R* 100	0.56/0.00/0.19 R*/R*/R* 100

[a] Following each date, the upper row of figures gives the numerical value of the indices, the middle row gives the classification (r = random, R = regular [5%], C = clustered [5%], R* = regular [1%], and C* = clustered [%]), and the number in the lower row is true percent diseased.

TABLE 2. Pielou's (P), Clark and Evans' (CE), and Hopkins and Skellam's (HS) indices of nonrandomness, sample size 200, for 1982 soybean field sites affected by bacterial blight; P/CE/HS[a]

Field	Replicate 1, Assessment 1	Replicate 1, Assessment 2	Replicate 2, Assessment 1	Replicate 2, Assessment 2
1	0.85/0.69/0.60 R/R*/R* 39	0.60/0.02/0.20 R*/R*/R* 98	0.73/0.33/0.33 R*/R*/R* 69	0.57/0.00/0.17 R*/R*/R* 100
2	0.88/0.44/0.49 r/R*/R* 58	0.80/0.24/0.28 R*/R*/R* 77	1.56/0.73/1.28 C*/R*/C 31	1.58/0.68/0.90 C*/R*/r 39
3	1.12/0.82/0.80 r/R*/R 29	0.93/0.54/0.43 r/R*/R* 51	2.85/0.89/2.18 C*/R*/C* 23	0.88/0.33/0.36 r/R*/R* 68
4	1.46/0.89/1.22 C*/R*/C 23	0.60/0.68/0.69 C/R*/R* 45	1.20/0.88/0.94 C*/R*/r 5	1.04/0.87/0.89 r/R*/r 20
5	1.00/0.34/0.41 r/R*/R* 64	0.66/0.15/0.28 R*/R*/R* 85	1.07/0.40/0.54 r/R*/R* 61	0.79/0.21/0.26 R*/R*/R* 80
6	1.69/0.94/1.36 C*/r/C* 14	1.10/0.80/1.23 r/R*/C 22	1.79/0.95/1.75 C*/r/C* 16	1.26/0.74/0.93 C*/R*/r 19
7	1.86/0.94/1.56 C*/r/C* 19	1.89/0.93/1.76 C*/r/C* 20	2.40/1.12/2.41 C*/C*/C* 11	1.30/0.95/1.06 C*/r/r 13
8	1.82/0.97/1.33 C*/r/C* 7	1.46/1.04/2.32 C*/r/C* 11	2.04/1.18/2.26 C*/C*/C* 7	1.83/0.97/1.74 C*/r/C* 16

[a] Following each date, the upper row of figures gives the numerical value of the indices, the middle row gives the classification (r = random, R = regular [5%], C = clustered [5%], R* = regular [1%], and C* = clustered [%]), and the number in the lower row is true percent diseased.

$$(1-t)^{4\sum_{k=0}^{i-1} k} \quad (1-(1-t)^{4i}) \qquad (1)$$

Given a sample of such distances the likelihood (1) of the observed data compared to the theoretical distribution can be used to assess whether the data came from a random distribution of diseased plants. The χ^2 obtained using the likelihood is a measure of whether more than just an assumption of randomness of diseased plants is required to explain the data. A large χ^2 indicates a departure from random distribution. This method was applied to all disease data except the final two readings in 1978 which were all close to 100% diseased. Note that the decision made is whether the underlying distribution is random or nonrandom—not the regular-random-clustered decision given by the three indices described earlier. To assess the true underlying distribution, the distance to the nearest diseased neighbor was calculated for all plants. To minimize edge effects, only plants 5–25 in rows 5–15 (231 plants) were considered. The distances to nearest diseased neighbor were also calculated for all samples drawn (as described below) to investigate whether less than the full field gave adequate information about the underlying disease distribution.

To assess three sampling schemes (random sampling, X-shaped path, W-shaped path [20]) for estimating percent infection, samples of different sizes and shapes (random: 10, 20, 30, and 40 plants; X-shape: 20, 28, and 40 plants; W-shape: 20, 28, and 40 plants) were drawn from all 600 tagged plants for each assessment at each date and estimated disease incidence was expressed as a percentage. The differences of the sample values from the "true" values based on all 600 tagged plants were analyzed to investigate differences between schemes and to obtain estimates of precision. These differences will be referred to as DFT values (DFT = 'true' incidence [%] − sample incidence [%]).

TABLE 3. Distribution of number of "steps" to the nearest diseased neighbor plant for selected disease levels and sample sizes

Diseased (%)	Probability or sample size	Proportion or number in sample expected to have a nearest diseased neighbor this many "steps" away							
		1	2	3	4	5	6	7	8
5	Prob	0.185	0.274	0.248	0.163	0.082	0.033	0.010	0.003
	40	7.4	11.0	9.9	6.5	3.3	1.3		
	200	37.1	54.8	49.7	32.7	16.5	6.5	2.1	0.5
10	Prob	0.344	0.374	0.203	0.065	0.013			
	40	13.8	14.9	8.1	2.6	0.5			
	200	68.8	74.7	40.5	13.0	2.6			
15	Prob	0.478	0.380	0.122	0.019				
	40	19.1	15.2	4.9	0.7				
	200	95.6	76.0	24.4	3.7				
20	Prob	0.590	0.341	0.064	0.005				
	40	23.6	13.6	2.6	0.2				
	200	118.1	68.2	12.8	0.9				
25	Prob	0.684	0.285	0.031					
	40	27.3	11.4	1.2					
	200	136.7	56.9	6.1					
30	Prob	0.760	0.226	0.014					
	40	30.4	9.1	0.5					
	200	152.0	45.3	2.7					
35	Prob	0.821	0.173	0.006					
	40	32.9	6.9	0.2					
	200	164.3	34.6	1.1					
40	Prob	0.870	0.127						
	40	34.8	5.1						
	200	174.1	25.5						
50	Prob	0.938	0.062						
	40	37.5	2.5						
	200	187.5	12.5						
60	Prob	0.974	0.026						
	40	39.0	1.0						
	200	194.9	5.1						
70	Prob	0.992	0.008						
	40	39.7	0.3						
	200	198.4	1.6						
80	Prob	0.998							
	40	39.9							
	200	199.7							
90	Prob	1.00							
	40	40							
	200	200							

RESULTS AND DISCUSSION

Disease distribution. Fig. 2 shows the distribution pattern of infected plants for four selected experimental areas. The values of P, CE, and HS for the field plots (Table 1) and grower's fields (Table 2) indicated that when disease incidence was relatively low the distribution pattern was predominantly clustered (C) and with an increase of incidence it became random (r) or regular (R) depending on the index used. Fig. 3 graphs these index values against the actual disease levels. It was conjectured that in the present context (observed plants in a rectangular lattice pattern) the values of the indices are influenced more by the proportion of plants diseased than by their pattern of distribution. Simulations performed by one of us (G. Poushinsky, *unpublished*) confirmed this.

For selected values of disease incidence, Table 3 gives the theoretical probabilities obtained by using equation 1, that the closest diseased plant is i "steps" away and the number expected in samples of size 40 and 200. Tables 4 and 5 give the distribution of nearest diseased plants for the 231 plants in each experimental area after excluding a border to minimize edge effects, and the χ^2 value obtained by using the likelihood. These were used to classify the underlying distribution of diseased plants as random or nonrandom. The 1978 quadrants gave some indication that at early stages the underlying disease distribution was nonrandom. This result was borne out by the likelihood values for the 1982 data. In general, the distribution of diseased plants in the fields exhibited a nonrandom pattern unless the proportion of diseased plants was very high. It is worth mentioning that in Minnesota (17) bacterial blight lesions developed at first on a few isolated seedlings, and then with the onset of cool, damp weather, the disease appeared suddenly on most plants, suggesting an early clustered and later random pattern. The probable causes of such early nonrandomness may be traced to initial groups of infected seeds in a lot and/or nonrandom distribution of the primary inoculum in the soil.

Sampling. It is of interest to see how well samples (rather than all the data as in Tables 4 and 5) perform in assessing the underlying distribution. The distances of the nearest diseased plant from random samples of size 200 and from samples of size 40 following W- and X-shaped paths for all experimental areas were tabulated and the likelihood method (based on the distribution of steps) outlined above was applied to these sets of data. Tables 6 and 7 summarize these results. Table 8 shows the numbers and types of misclassifications of the various methods compared to the true classification of disease distribution (Tables 4 and 5). It is apparent that the likelihood method is superior to use of the three indices, even when smaller sample sizes were used to compute the likelihood. In addition, the likelihood method does not require a complete list of all diseased sites as do analyses based on the three indices.

Table 9 summarizes the results of drawing random, X- and W-shaped samples to estimate the percentage of plants that were diseased. Only random samples of size 10 and 20 were statistically different from zero. The overall interpretation is that precision increases with increasing sample size and that X or W paths perform better than random samples. The fact that some distributions were nonrandom explains the latter (20). Table 10 gives a summary of the sampling results when the experimental areas were subdivided into random and nonrandom underlying distributions by using the classification implied in Tables 4 and 5.

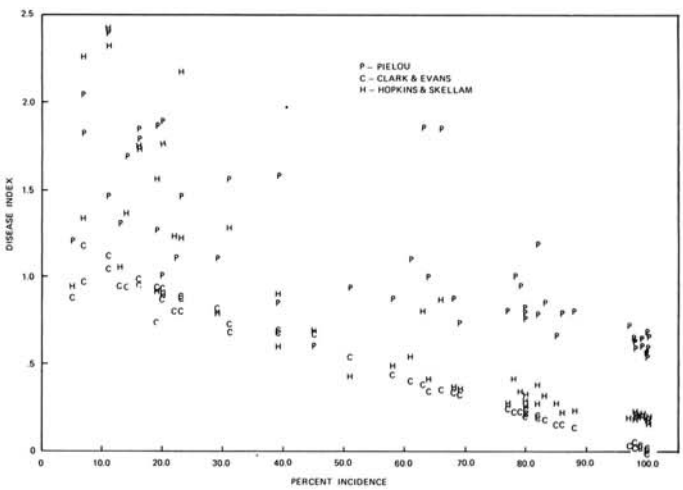

Fig. 3. Plotted values of three disease indices (P = Pielou's, C = Clark and Evans's and H = Hopkins and Skellam's) plotted against percent incidence in the field.

TABLE 4. Observed and expected distribution of number of steps to nearest diseased neighbour for 1978 experimental areas and likelihood value for the hypothesis of an underlying random disease distribution

Quadrant	Date	Sample size or expected no.	Steps 1	2	3	4	True percent diseased	χ^2(d.f.)[a]
1	27 June	231	207	20	4		66	104.88(3)**
		Exp	228	3				
	11 July	231	230	1			80	.82(2)
		Exp	231					
	25 July	231	231				82	.46(1)
		Exp	231					
2	27 June	231	231				82	.46(1)
		Exp	231					
	11 July	231	230	1			88	3.87(2)
		Exp	231					
	25 July	231	231				79	.92(1)
		Exp	231					
3	27 June	231	197	23	10	1	63	238.23(4)**
		Exp	227	4				
	11 July	231	231				86	0(1)
		Exp	231					
	25 July	231	230	1			80	.65(2)
		Exp	231					
4	27 June	231	230	1			83	1.49(2)
		Exp	231					
	11 July	231	223	8			78	27.08(2)**
		Exp	230	1				
	25 July	231	231				97	0(1)
		Exp	231					

[a] A large χ^2 indicates a departure from an underlying random disease distribution. * Prob $\leq$0.05 and ** Prob $\leq$.01.

TABLE 5. Observed and expected distribution of number of "steps" to the nearest bacterial blight-diseased neighbor soybean plant in 1982 experimental areas and the likelihood value for the hypothesis of an underlying random disease distribution

Field	Rep	Assessment	Sample size or expected no.	Steps 1	2	3	4	5	6	7	8	9	True percent diseased	χ^2 (d.f.)[a]
1	1	1	231	199	32								39	1.27(2)
			Exp	199	31	1								
		2	231	231									98	0(1)
			Exp	231										
	2	1	231	226	5								69	3.08(2)
			Exp	229	2									
		2	231	231									100	0(1)
			Exp	231										
2	1	1	231	216	15								58	6.01(2)*
			Exp	224	7									
		2	231	229	2								77	1.98(2)
			Exp	230	1									
	2	1	231	170	46	15							31	27.31(3)**
			Exp	179	50	3								
		2	231	176	49	6							39	28.65(3)**
			Exp	199	31	1								
3	1	1	231	164	60	7							29	2.41(3)
			Exp	172	55	4								
		2	231	212	19								51	1.96(2)
			Exp	218	13									
3	2	1	231	109	67	39	16						23	137.71(4)**
			Exp	150	71	10								
		2	231	226	5								68	2.02(2)
			Exp	229	2									
4	1	1	231	126	81	22	2						23	20.02(4)**
			Exp	150	71	10								
		2	231	188	35	8							45	54.56(3)**
			Exp	210	21									
	2	1	231	46	76	61	29	14	5				5	8.76(6)
			Exp	43	63	57	38	19	8	2	1			
		2	231	112	97	22							20	14.53(3)**
			Exp	136	79	15	1							
5	1	1	231	221	10								64	7.34(2)*
			Exp	227	4									
		2	231	230	1								85	2.69(2)
			Exp	231										
	2	1	231	210	19	2							61	42.87(3)**
			Exp	226	5									
		2	231	231									80	.93(1)
			Exp	231										
6	1	1	231	103	86	32	10						14	3.28(4)
			Exp	105	88	32	6	1						
		2	231	139	70	19	2	1					22	13.51(5)*
			Exp	145	74	11	1							
	2	1	231	97	68	35	17	10	4				16	132.01(6)**
			Exp	116	87	25	3							
		2	231	113	88	22	8						19	18.19(4)**
			Exp	132	81	17	1							
7	1	1	231	74	73	44	26	11	3				19	303.38(6)**
			Exp	132	81	17	1							
		2	231	81	80	50	18	2					20	150.22(5)**
			Exp	136	79	15	1							
	2	1	231	57	36	38	44	31	15	8	2		11	390.17(8)**
			Exp	86	88	43	12	2						
		2	231	81	65	46	28	11					13	84.13(5)**
			Exp	86	88	43	15	2						
8	1	1	231	51	72	61	25	12	4	2	2	2	7	29.60(9)**
			Exp	58	76	56	28	10	3					
		2	231	68	75	55	26	6	1				11	20.87(6)**
			Exp	86	88	43	12	2						
	2	1	231	72	51	55	34	14	4	1			7	14.85(7)*
			Exp	58	76	56	28	10	3					
		2	231	122	68	36	5						16	10.43(4)*
			Exp	116	87	25	3							

[a] A large χ^2 indicates a departure from an underlying random disease distribution. Asterisks * and ** signify statistically significant values $P \leq 0.05$ and $P \leq 0.01$, respectively.

All sampling methods (except a random sample of size 10) gave adequate results when the underlying distribution was random. When it was nonrandom the random samples were wrong by a statistically significant amount, but the X- and W-shaped paths have a mean DFT that was not statistically different from zero except for sample size 40. This latter result was unexpected, and no pattern was evident to explain it.

The present results indicate that bacterial blight of soybean is distributed in a nonrandom pattern in fields during early to midseason. The assessment of pattern in the present context using distances to nearest-diseased-neighbor was more reliable than the methods based on the indices of Pielou, Hopkins and Skellam, or Clark and Evans. The likelihood method is of general applicability in assessing random distribution of plant diseases.

The information obtained from a W- or X-shaped sample is adequate to assess disease incidence. This substantiates the results of Lin et al (20) and Hau et al (13) that in the presence of nonrandomness of disease distribution, disease incidence is best estimated by designed sampling paths rather than by simple random sampling.

TABLE 6. Likelihood (χ^2 value) for the hypothesis of an underlying random disease (bacterial blight of soybeans) distribution in 1978 experimental areas tested by using three sampling schemes: X/W/random[a]

Quad	27 June	11 July	25 July
1	0.3(2)/38.2(3)**/348.5(4)** 66%	0.1(1)/3.8(2)/13.4(2)** 80%	0.1(1)/4.3(2)/1.4(2) 82%
2	32.3(3)**/4.6(2)/5.7(2) 82%	0.1(1)/0.1(1)/4.1(2) 88%	0.1(1)/0.1(1)/0.7(2) 79%
3	16.4(3)**/40.1(3)**/130.6(4)** 63%	0(1)/0(1)/0(1) 86%	0.1(1)/0.1(1)/12.4(2)** 80%
4	0.1(1)/0.1(1)/1.7(2) 83%	0.2(1)/14.7(2)**/47.1(3)** 78%	0(1)/0(1)/0(1) 97%

[a]The upper numbers give the value of χ^2 (with degrees of freedom in brackets) for X- and W-shaped sampling paths with 40 sample sites, and a random sample of size 200; the number centered beneath each set of values is the true percent disease incidence. ** $P \leq 0.01$.

TABLE 7. Likelihood (χ^2 value) for the hypothesis of an underlying random disease (bacterial blight of soybeans) distribution in 1982 experimental areas tested by using three sampling schemes: X/W/random[a]

	Replicate 1		Replicate 2	
Field	Assessment 1	Assessment 2	Assessment 1	Assessment 2
1	0.3(2)/1.2(2)/1.6(2) 39%	0(1)/0(1)/0(1) 98%	0.8(2)/0.8(2)/15.1(2)** 69%	0(1)/0(1)/0(1) 100%
2	1.7(2)/3.8(2)/10.3(2)**[b] 58%	0.2(1)/2.7(2)/1.2(1) 77%	11.0(3)*/4.7(3)/14.2(3)** 31%	19.4(3)**/3.4(3)/39.6(3)** 39%
3	1.6(2)/13.1(3)**/4.2(3) 29%	2.4(2)/8.6(3)*/0.9(2) 51%	33.0(4)**/60.2(6)**/131.9(4)** 23%	3.1(2)/6.7(2)*/4.7(2) 68%
4	6.5(3)/5.5(3)/36.9(4)** 23%	23.4(3)**/1.5(2)/8.1(2)* 45%	3.1(6)/21.2(8)**/13.6(6)* 5%	1.9(3)/3.4(3)/25.5(3)** 20%
5	0.2(2)/17.3(3)**/0.8(2) 64%	0.1(1)/0.1(1)/0.4(1) 85%	5.6(2)/32.0(3)**/0.3(2) 61%	0.2(1)/10.1(2)**/0.8(1) 80%
6	15.8(5)**/14.3(5)*/3.3(4) 14%	3.6(3)/7.9(4)/9.8(4)* 22%	64.2(6)**/10.2(5)/95.2(6)** 16%	8.6(4)/10.3(4)**/14.7(4)** 19%
7	13.4(4)**/10.2(4)*/146.2(5)** 19%	15.8(5)**/3.3(3)/111.6(4)** 20%	57.8(7)**/34.3(7)**/195.4(7)** 11%	2.8(5)/17.1(5)**/86.3(5)** 13%
8	41.0(8)**/45.2(8)**/16.1(6)* 7%	9.0(5)/20.8(6)**/6.8(4) 11%	6.9(6)/63.9(9)**/10.7(6) 7%	2.7(4)/79.1(7)**/10.5(4)* 16%

[a]The upper numbers give the value of χ^2 (with degrees of freedom in parentheses) for X- and W-shaped paths with 40 sample sites, and a random sample of size 200; the number centered beneath each set of values is the true percent disease incidence. *$P \leq 0.05$, and **$P \leq 0.01$.

TABLE 8. Comparison of the number and type of misclassifications made by using six methods for determining the distribution of bacterial blight in soybean fields

Method	Random classified as nonrandom	Nonrandom classified as random	Misclassified No.	Misclassified %
Pielou	16	6	22	50
Clark and Evans	20	7	27	61
Hopkins and Skellam	20	5	25	57
Likelihood (X, 40 sites)	2	13	15	34
Likelihood (W, 40 sites)	6	9	15	34
Likelihood (random sample, 200 sites)	3	4	7	16

TABLE 9. Summary of the results of applying various sampling schemes to all soybean fields assessed for bacterial blight

Sampling scheme[a]	Number of samples of this type	Mean percent true disease − sample estimate (S.E.)
Random (10)	44	−5.26 (1.403)
Random (20)	44	−2.30 (1.192)
Random (30)	44	−0.94 (1.213)
Random (40)	44	1.22 (1.435)
X or W (20)	88	0.14 (0.941)
X or W (28)	88	−0.03 (0.722)
X or W (40)	88	−1.05 (0.753)

[a]Sampling schemes (random or X and W sampling paths are followed by number of sample sites in parentheses.

TABLE 10. Summary of the results of applying various sampling schemes to determine the percent incidence of bacterial blight in soybean fields with diseased plant distributions assessed as random or nonrandom

Disease distribution	Sampling scheme[a]	Number of samples of this type	Mean percent true disease and (sample estimate)[b]
Random	Random (10)	21	−4.74 (1.809)
	Random (20)	21	0.26 (1.964)
	Random (30)	21	0.35 (1.015)
	Random (40)	21	−0.81 (0.937)
	X or W (20)	42	1.81 (1.330)
	X or W (28)	42	1.15 (0.970)
	X or W (40)	42	0.92 (0.997)
Nonrandom	Random (10)	23	−5.73 (2.151)
	Random (20)	23	−4.65 (1.259)
	Random (30)	23	−3.56 (1.328)
	Random (40)	23	−4.97 (1.124)
	X or W (20)	46	−1.38 (1.302)
	X or W (28)	46	−1.10 (1.090)
	X or W (40)	46	−2.85 (1.057)

[a] Sampling schemes (random or X and W sampling paths) are followed by number of sample sites in parentheses.
[b] Mean percent of true disease with sample estimate (S.E.) in parentheses.

LITERATURE CITED

1. Barnett, V. 1975. Comparative Statistical Inference. John Wiley & Sons, London. 287 pp.
2. Basu, P. K. 1979. Occurrence of soybean foliage diseases in eastern Ontario, 1979. Can. Plant Dis. Surv. 60:23-24.
3. Besag, J. E. 1974. Spatial interaction and the analysis of lattice systems. J. R. Statist. Soc. B 36:192-225.
4. Besag, J. E., and Gleaves, J. T. 1973. On the detection of spatial pattern in plant communities. Bull. Int. Statist. Inst. 45(1):153-158.
5. Chiarappa, L. (editor). 1970. Crop loss assessment methods. FAO Manual on the evaluation and prevention of losses by pests, diseases and weeds. AGP:CP/22. Food Agric. Organ. U.N., Rome, Italy.
6. Clark, P. J., and Evans, F. C. 1954. Distance to nearest neighbor as a measure of spatial relationship in populations. Ecology 35:445-453.
7. Cross, J. E., Kennedy, B. W., Lambert, J. W., and Cooper, R. L. 1966. Pathogenic races of the bacterial blight pathogen of soybeans, *Pseudomonas glycinea*. Plant Dis. Rep. 50:557-560.
8. Daft, G. C., and Leben, C. 1972. Bacterial blight of soybeans: Epidemiology of blight outbreaks. Phytopathology 62:57-62.
9. Daft, G. C., and Leben, C. 1972. Bacterial blight of soybeans: Seedling infection during and after emergence. Phytopathology 62:1167-1170.
10. Diggle, P. J., Besag, J. E., and Gleaves, J. T. 1976. Statistical analysis of spatial point patterns by means of distance methods. Biometrics 32:659-667.
11. Dye, D. W., Bradbury, J. F., Goto, M., Hayward, A. C., Lelliott, R. A., and Schroth, M. N. 1980. International standards for naming pathovars of phytopathogenic bacteria and a list of pathovar names and pathotype strains. Rev. Plant Pathol. 59:153-168.
12. Greig-Smith, P. 1957. Quantitative Plant Ecology. Butterworth's Scientific Publications, London. 198 pp.
13. 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.
14. Hopkins, B. 1954. A new method for determining the type of distribution of plant individuals. (with an appendix by J. G. Skellam) Ann. Bot. Lond. N.S. 18:213-227.
15. Kendrick, J. B., and Gardner, M. W. 1921. Seed transmission of soybean's bacterial blight. Phytopathology 11:340-342.
16. Kennedy, B. W. 1969. Detection and distribution of *Pseudomonas glycinea* in soybean. Phytopathology 59:1618-1619.
17. Kennedy, B. W., and Ercolani, G. L. 1979. Soybean primary leaves as a site for epiphytic multiplication of *Pseudomonas glycinea*. Phytopathology 68:1196-1201.
18. Kennedy, B. W., and Tabachina, H. 1973. Bacterial diseases. Pages 491-504 in: Soybeans: Improvement, Production, and Uses. B. E. Caldwell, ed. Am. Soc. Agron., Madison, WI.
19. Leben, C., Rushch, V., and Schmitthenner, A. F. 1968. The colonization of soybean buds by *Pseudomonas glycinea* and other bacteria. Phytopathology 58:1677-1681.
20. Lin, C. S., Poushinsky, G. P., and Mauer, M. 1979. An examination of five sampling methods under random and clustered disease distributions using simulations. Can. J. Plant Sci. 59:121-130.
21. Madden, L. V., Louie, R., Abt, J. J., and Knoke, J. K. 1982. Evaluation of tests of randomness of infected plants. Phytopathology 72:195-198.
22. Pielou, E. C. 1957. The effect of quadrat size on the estimation of the parameters of Neyman's and Thomas's distributions. J. Ecol. 45:31.
23. Pielou, E. C. 1959. The use of point-to-point distances in the study of the pattern of plant populations. J. Ecol. 47:607-713.
24. Pielou, E. C. 1962. Runs of one species with respect to another in transects through plant populations. Biometrics 18:579-583.
25. Roach, S. A. 1968. The Theory of Random Clumping. Methuen, London. 94 pp.
26. Stauffer, H. B. 1977. Application of the indices of nonrandomness of Pielou, Hopkins and Skellam, and Clark and Evans. Rep. BC-X-166. Dep. Fish. & Environ., Can. For. Serv., Pac. For. Res. Cen., Victoria, B.C.
27. Vanderplank, J. E. 1946. A method for estimating the number of random groups of adjacent diseased plants in a homogeneous field. Trans. R. Soc. S. Africa 3:269-278.