Modeling of Superimposed Spatial Patterns of Bacterial Brown Spot of Snap Bean

B. D. Hudelson, M. K. Clayton, K. P. Smith, and C. D. Upper

Department of Plant Pathology (all authors), Department of Statistics (second author), and Agricultural Research Service, U. S. Department of Agriculture (fourth author), University of Wisconsin, Madison 53706.

Cooperative research of the College of Agricultural and Life Sciences, University of Wisconsin, Madison, and the Agricultural Research Service, U. S. Department of Agriculture.

Research supported in part by USDA Competitive Grant 88-37151-3898.

We thank K. Hoffman and J. Armstrong for valuable technical assistance and S. Vicen for preparation of the figures.

We also thank the snap bean growers and food processors of Wisconsin for their cooperation in providing commercial snap bean fields for this work.

Mention of a trademark, proprietary product, or vendor does not constitute a guarantee or warranty of product by the U. S. Department of Agriculture and does not imply its approval to the exclusion of other products or vendors that may also be suitable.

Accepted for publication 15 December 1992.

ABSTRACT

Hudelson, B. D., Clayton, M. K., Smith, K. P., and Upper, C. D. 1993. Modeling of superimposed spatial patterns of bacterial brown spot of snap bean. Phytopathology 83:430-438.

Snap bean plants within seven-row segments that ranged from 30 to 108 m were sampled using a 2,4-systematic sampling plan: Two adjacent plants were sampled and evaluated for bacterial brown spot, four adjacent plants were skipped, two adjacent plants were sampled, etc. This sampling plan was developed using a priori knowledge of disease patterns in 5-m row segments. Every leaflet on every sampled plant was assessed for bacterial brown spot, and disease on these plants was quantified as the proportion of diseased leaflets per plant. Arcsine square root-transformed disease-incidence values were analyzed for spatial patterns using autoregressive integrated moving average (ARIMA) modeling. Three of seven data sets were described by a generalized ARIMA(1 0 1) model. These data sets exhibited the expected patterns of disease, based on previous descriptions of disease patterns in 5-m row segments. Of the remaining four data sets, one was modeled by an ARIMA(1 0 2) model, two by an ARIMA(1 0 3) model, and one by an ARIMA(1 0 4) model. These data sets exhibited the expected patterns, as well as additional patterns that recurred at intervals of six to 11, 12 to 17, or 18 to 23 plants, respectively. Significant autocorrelations at distances of six to 23 plants were also found in residuals from 24 of 55 5-m row segments evaluated for bacterial brown spot and modeled using ARIMA models. These patterns appear to occur frequently in commercial snap bean fields.

Additional keywords: adaptive sampling, *Pseudomonas syringae*.

In the past decade there has been substantial interest in spatial patterns of plant disease, pathogens, and development of disease in time and space (5,6,8). From the body of knowledge gained from this research, these general conclusions can be drawn: 1) Disease rarely is arranged randomly in space. 2) Current methodology in plant pathology literature is sufficient to determine whether disease is arranged randomly or nonrandomly in space but is less well suited for quantifying or modeling spatial patterns. There have been efforts to adapt existing statistical techniques to quantify and model spatial patterns (7,13,17,18). These efforts have clearly demonstrated the importance of such studies, but also that the utility of existing methods for quantifying spatial patterns is limited. 3) Applications of methodology for determining the presence or absence of multiple spatial patterns at different scales also are limited. We are aware of only a limited number of reports that have suggested there can be more than one pattern at two different scales in a single field (11,14,15,21).

In recognition of these concerns, we began developing a strategy that would be capable of detecting and modeling spatial patterns. The strategy that we developed in our studies of spatial patterns of bacterial brown spot is called adaptive sampling. This strategy is designed to detect and describe multiple spatial patterns that may occur at several different scales within a single field. Adaptive sampling involves the iterative application of two basic steps:

Step 1: Sampling and description of spatial patterns within an area of defined size in a field.

Step 2: Use of spatial-pattern information obtained in step 1 to develop new sampling plans for sampling larger areas in the field.

In our work with bacterial brown spot, the first step of the first iteration of the process was to search for disease patterns within short-row segments (approximately 5 m) (11). Nonrandom brown spot variability in these row segments occurred on at least two scales. The most prominent disease pattern could be described as a positive correlation between disease incidence on adjacent plants that corresponded to a slowly undulating change in disease levels within row segments. In addition to this dominant pattern, a component of negative correlation gave a jagged appearance to graphs of disease vs. distance. The disease patterns detected in short snap bean-row segments were described using ARIMA modeling (3,10,11)—a technique that is prevalent in the statistical time series literature. For 35 of the 38 samples discussed in our previous paper (11) and for 17 of 17 5-m samples collected subsequently, we found that bacterial brown spot spatial patterns could be described using a generalized ARIMA(1 0 1) model (11):

$$Y_t = \phi_{Y,1} Y_{t-1} + \epsilon_t - \theta_{Y,1}\epsilon_{t-1} + \delta$$

Y_t and Y_{t-1} are the proportions of diseased leaflets associated with the plants at positions t and $t-1$ in the row and transformed using an arcsine square root transformation. The terms ϵ_t and ϵ_{t-1} are random errors (with a variance of σ_ϵ^2) associated with the plants. The terms $\phi_{Y,1}$ and $\theta_{Y,1}$ are constants, similar to regression coefficients, that quantify the relationship between the Ys and the ϵs. Finally, δ is a constant related to the mean level of disease in the row segment. In the generalized ARIMA(1 0

1) model, $\phi_{Y,1} Y_{t-1}$ describes the slowly undulating change in disease levels within a row; $\theta_{Y,1}\epsilon_{t-1}$ describes disease patterns that give graphs of disease vs. distance their jagged appearance. In all of our samples in which brown spot patterns could be described by a generalized ARIMA(1 0 1) model, $0 \leq \phi_{Y,1} \leq 1$ and $0 \leq \theta_{Y,1} \leq \phi_{Y,1}$. Considering the frequency with which this model occurred, we feel that a common underlying mechanism might be involved in the development of these brown spot patterns.

The second step of adaptive sampling involved considering how brown spot patterns described using a generalized ARIMA(1 0 1) model could be used to develop sampling plans needed to detect patterns at larger scales. The work of Iachan (12) indicated that systematic samples should be more efficient than random samples for estimating means from positively autocorrelated processes (such as the ones we had observed). We considered the possibility that systematic samples might also be useful for detecting and modeling spatial patterns.

The general class of systematic samples we investigated were *A,B*-systematic samples, in which *A* and *B* are positive integers and *B* is an integer multiple of *A*. The choice of *A* and *B* in our work was somewhat arbitrary. In most instances, however, constraints on the time and resources available to collect a sample, as well as the size of the population to be sampled, dictate the limits of the magnitude of *A* and *B*. Theoretically, a priori knowledge of patterns in a population to be sampled could be used to tailor design an optimal *A,B*-systematic sample or group of *A,B*-systematic samples to detect other patterns. We are unaware of any published work in this area.

Logistically, an *A,B*-systematic sample involves sampling a group of *A* adjacent plants and evaluating these plants for a characteristic of interest (in our case bacterial brown spot disease incidence), skipping the next *B* adjacent plants, evaluating the next *A* adjacent plants, etc. The measured values for each group of *A* adjacent plants (or an appropriate transformation thereof) are averaged to yield a single data point for each group of *A* plants sampled. Data from samples of this type are amenable to analyses such as runs tests, autocorrelation analysis, and ARIMA modeling. More importantly, however, the type of ARIMA model that will describe disease patterns in an *A,B*-systematic sample can be predicted from statistical theory, based on knowledge of underlying patterns. For example, as noted above, bacterial brown spot patterns observed in short-row segments (5 m) were described by a generalized ARIMA(1 0 1) model. One can show that an *A,B*-systematic sample collected from the same row should also be well described by a generalized ARIMA(1 0 1) model if no additional brown spot patterns are present in the systematic sample (20,22, B. D. Hudelson, *unpublished data*). Deviations from this predicted model would be an indication of the presence of additional, previously undescribed brown spot patterns. Similar predictions are possible for more complex ARIMA processes (20,22, B. D. Hudelson, *unpublished data*).

The work we describe represents the application of a 2,4-systematic sampling plan (i.e., an *A,B*-systematic sampling plan with $A = 2$ and $B = 4$) to elucidate bacterial brown spot patterns. This work is an example of step 1 of the second iteration of adaptive sampling for bacterial brown spot of snap beans.

TABLE 1. Characteristics of fields from which 2,4-systemic samples and total-census samples were collected

Field number	County	Nearest village[a]	Field area (ha)	Cultivar[b]
11	Columbia	Arlington	4	BBL 109
14	Waushara	Almond	64	BBL 94
15	Columbia	Arlington	8	BBL 109
18	Green Lake	Markesan	14	Peak
19	Green Lake	Fairwater	31.2	Peak
20	Green Lake	Fairwater	60	Peak
22	Columbia	Rio	5.2	Peak

[a] In Wisconsin.
[b] BBL = Bush Blue Lake.

MATERIALS AND METHODS

Commercial snap bean fields. 2,4-systematic samples were collected, over a 3-yr period (1987–1989), from seven commercial snap bean fields, each planted with one of three snap bean cultivars and located in three diverse areas of Wisconsin. More detailed information on the characteristics (location, cultivar, area, etc.) of these fields is provided in Table 1.

Sampling and disease assessment. Within each snap bean field, a 2,4-systematic sample was collected from within a single-row segment. A 2,4-systematic sample involved sampling two adjacent plants and evaluating those plants for bacterial brown spot, skipping four adjacent plants, sampling two plants, etc. The leaflets of each selected plant were evaluated for the presence or absence of bacterial brown spot as described previously (11). In fields with normal snap bean populations, the row segments sampled were 30- to 45-m long. In one field with a light stand, the row segment was 108-m long. Samples collected during 1987, 1988, and on 24 July 1989 (samples 11.24, 14.24, and 15.24 in Table 2) were assessed by two workers. For these samples, the individual who assessed a given plant was recorded. One worker assessed all other samples.

In addition to 2,4-systematic samples, bacterial brown spot was evaluated on all plants in one or more short-row segments (approximately 5 m) from within the same row segment from which the 2,4-systematic sample was collected (fields 15, 18–20, and 22) or from an adjacent row (fields 11 and 14). Total-census

TABLE 2. Summary of characteristics of 2,4-systematic and total-census samples

2,4-Systematic sample[a]	Sample			Total-census sample	Sample		
	Date collected	Length (m)	Number of plants		Date collected	Length (m)	Number of plants
11.24	7/21/87	45	678	11.1	7/22/87	5.0	81
				11.2	7/22/87	5.0	82
14.24	8/24/88	36	492	14.1	8/24/88	5.0	100
15.24	7/24/89	30	684	15.1	7/24/89	4.2	116
				15.2	7/24/89	3.6	96
				15.3	7/24/89	4.2	100
				15.4	7/24/89	ND[b]	67
18.24	8/09/89	108	630	18.1	8/09/89	6.8	111
19.24	8/30/89	41	594	19.1	8/30/89	ND	100
20.24	9/21/89	45	582	20.1	9/21/89	7.0	92
22.24	7/17/89	30	858	22.1	7/17/89	2.5	70

[a] Numbers to the left of the decimal point refer to the field from which 2,4-systematic and total-census samples were collected. The number 24 to the right of the decimal point denotes a 2,4-systematic sample. Single digits to the right of the decimal point distinguish total-census samples collected from the same field. Data sets 15.3, 15.4, 18.1, and 20.1 were collected from the same row segments from which the corresponding 2,4-systematic samples were collected.
[b] Not determined.

samples were collected to verify that the underlying pattern of disease could be described by a generalized ARIMA(1 0 1) model. Total-census samples were assessed by one worker, as described previously (11). Detailed characteristics of the 2,4-systematic samples and total-census samples are provided in Table 2.

Data analysis. Total-census data from short-row segments were analyzed as described previously (11). Disease associated with each plant was expressed as the proportion of diseased leaflets per plant; proportions were transformed using an arcsine square root transformation to stabilize variance (19). The resulting transformed disease-incidence values were used in subsequent runs, autocorrelation, and ARIMA analyses.

The treatment of data from 2,4-systematic samples was more complex. Trained assessors had a high probability of correctly identifying bacterial brown spot. Thus, the assessed number of diseased leaflets varied little between assessors. However, deciding whether an emerging (typically undiseased) leaflet had attained sufficient size to be included in the tally did vary between assessors, which led to variability in the total number of leaflets per plant evaluated by different assessors. Therefore, for 2,4-systematic samples assessed by more than one worker, an attempt was made to correct the data for effects resulting from assessor differences.

In 1987 and 1988 (2,4-systematic samples 11.24 and 14.24), data were first transformed using an arcsine square root transformation. Such a transformation is recommended for proportion data to stabilize variance (19). An assumption of homogeneity of variance is inherent in both the regression technique we used to correct data for assessor differences (as described below) and for ARIMA modeling we used to detect spatial patterns.

Resulting transformed disease-incidence values of each pair of adjacent plants were averaged. An assessor's average transformed disease-incidence values were standardized by subtracting the mean average transformed disease incidence for that assessor. Finally, the largest mean average transformed disease incidence obtained among the assessors was added to all standardized average transformed disease-incidence values. This yielded nonnegative values. These values (Z_t) were used in runs, autocorrelation, and ARIMA analyses.

In 1989, only a single 2,4-systematic sample was evaluated by more than one assessor (sample 15.24). To correct this sample for assessor differences, a group of 14 bean plants was evaluated for brown spot by both assessors. In a regression analysis, the total number of leaflets evaluated by one assessor was used as a predictor for the total leaflets evaluated by the second assessor.

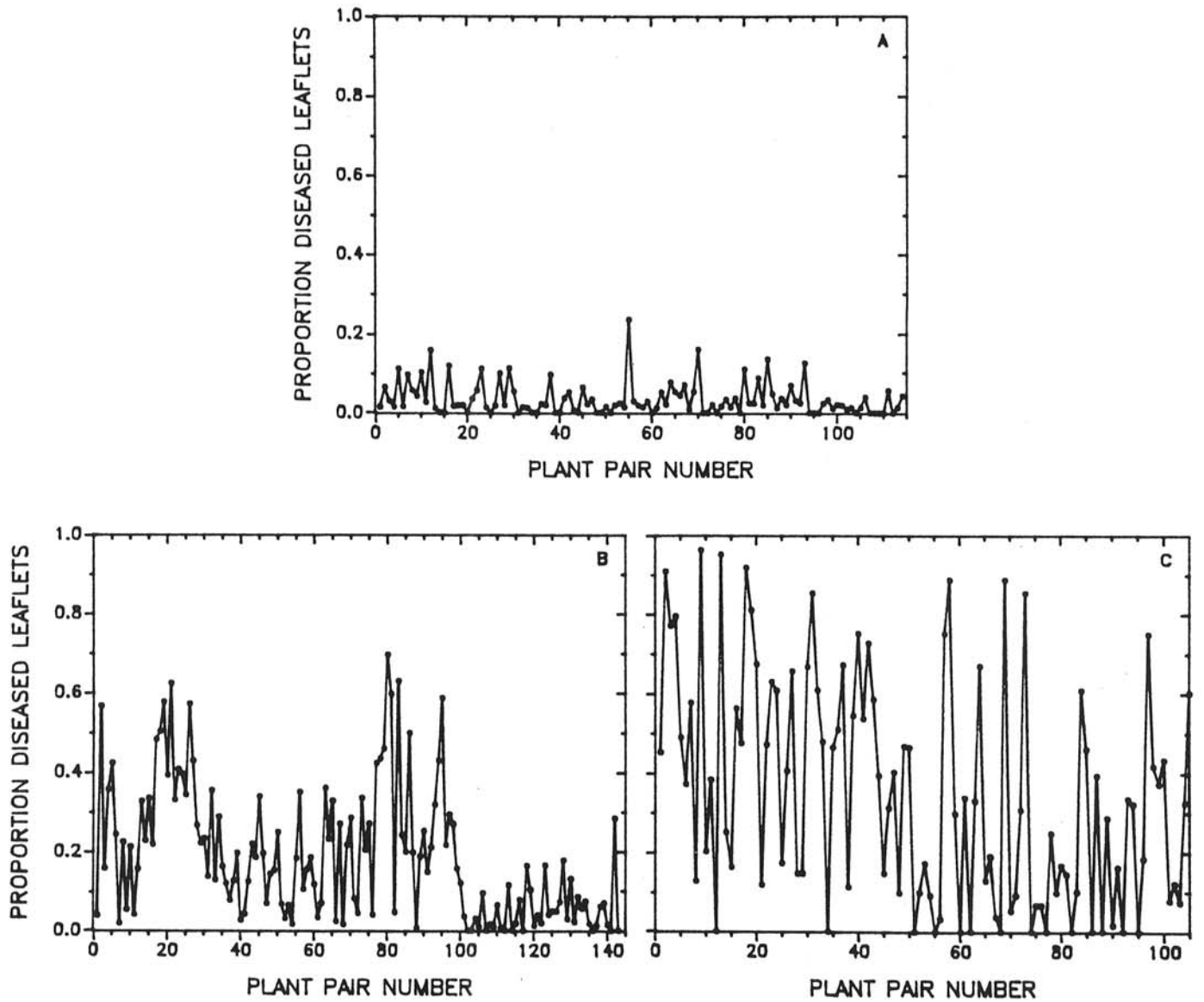

Fig. 1. 2,4-systematic samples exhibiting the spatial patterns of brown spot expected from patterns detected in total-census data sets but exhibiting no additional patterns. **A,** Data set 15.24 collected 24 July 1989 from a 8-ha snap bean field (cv. BBL 109) near Arlington, WI. Data modeled using an ARIMA(0 0 0) model. **B,** Data set 22.24 collected 17 July 1989 from a 5.2-ha snap bean field (cv. Peak) near Rio, WI. Data modeled using an ARIMA(1 0 1) model. **C,** Data set 18.24 collected 9 August 1989 from a 14-ha snap bean field (cv. Peak) near Markesan, WI. Data modeled using an ARIMA(0 1 1) model.

The resulting prediction equation was used to correct the total number of leaflets associated with those plants in the 2,4-systematic sample evaluated by the first assessor. The resulting corrected total-leaflet values were used to calculate corrected proportions of diseased leaflets per plant, which were in turn transformed using an arcsine square root transformation. The transformed disease values associated with adjacent plants were averaged to yield corrected average transformed disease values (Z_t) used for subsequent runs, autocorrelation, and ARIMA analyses.

Runs, autocorrelation, and initial ARIMA analyses for 2,4-systematic samples were conducted, as detailed previously (11), using MINITAB, release 5.1.1 (The Pennsylvania State University, University Park). The SCA Statistical System, release III.3.2 (Scientific Computing Associates, DeKalb, IL), was used to confirm the ARIMA models computed using MINITAB. For graphic illustration (Figs. 1 and 2), the inverse of the original transformation (i.e., the squared sine transformation) was applied to the corrected average transformed disease values to yield values that ranged between zero and one.

RESULTS

Total-census samples. Results from the analyses of total-census samples collected in conjunction with 2,4-systematic samples are presented in Table 3. Data sets 11.1 and 11.2 have been described previously (11). The 11 samples had mean disease values ranging from 2.7 (sample 15.2) to 60.9% (sample 20.1). In runs analyses, six of the samples had fewer runs than expected (four with Z values of -1.96 or less [$P \leq 0.05$]). Two samples had more runs than expected but not significantly so ($P = 0.63$ and $P = 0.84$, respectively). For the remaining three samples, a runs analysis could not be performed because the median disease value for these samples was zero.

Sample autocorrelation functions (3,10), $r_Y(s)$, in which s is the lag value, also varied from sample to sample. Five of the 11 samples had sample autocorrelation functions that had no values of $r_Y(s)$ ($1 \leq s \leq 20$) that lay outside a 95% confidence interval (CI) calculated under the assumption of a random disease pattern. These samples, along with a sample that had a single value of $r_Y(s)$ that lay outside a 95% CI, were adequately described

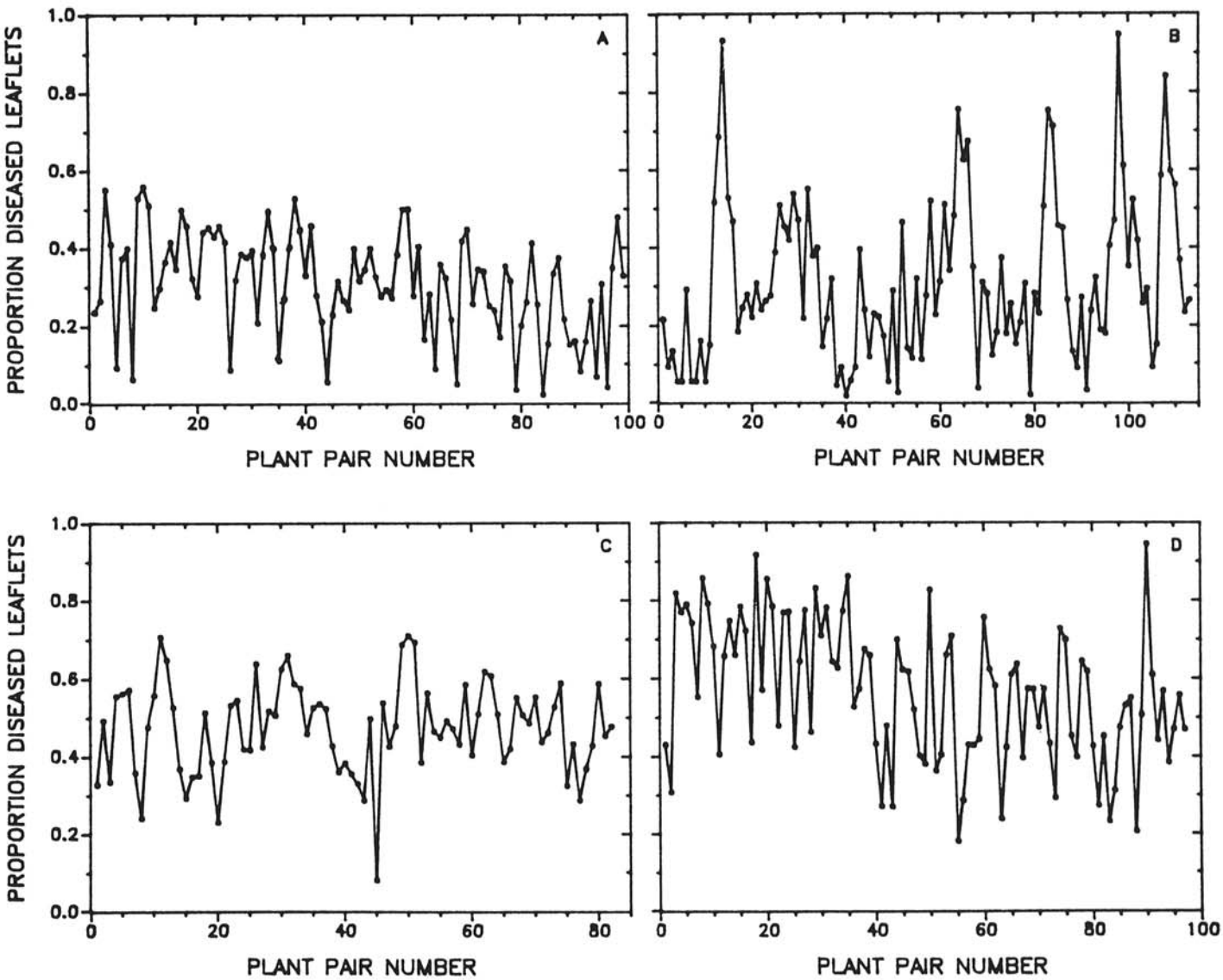

Fig. 2. 2,4-systematic samples exhibiting the spatial patterns of brown spot expected from patterns detected in total-census data sets as well as exhibiting additional patterns. **A,** Data set 19.24 collected 30 August 1989 from a 31.2-ha snap bean field (cv. Peak) near Fairwater, WI. Data modeled using an ARIMA(0 1 2) model. **B,** Data set 11.24 collected 21 July 1987 from a 4-ha snap bean field (cv. BBL 109) near Arlington, WI. Data modeled using an ARIMA(1 0 4) model. **C,** Data set 14.24 collected 24 August 1987 from a 64-ha snap bean field (cv. BBL 94) near Almond, WI. Data modeled using an ARIMA(1 0 3) model. **D,** Data set 20.24 collected 21 September 1989 from a 60-ha snap bean field (cv. Peak) near Fairwater, WI. Data modeled using an ARIMA(1 0 3) model.

TABLE 3. Results from ARIMA modeling, autocorrelation analyses, and runs analyses of bacterial brown spot incidence from total-census data sets collected in conjunction with 2,4-systematic samples

Data set[a]	Mean disease (%)	ARIMA model	$\phi_{Y,1}$ estimate[b]	$\theta_{Y,1}$ estimate[b]	σ_ϵ^2 estimate[b]	Significant ACF lags[c]	Runs-test Z-value
11.1	25.4	0 0 0	0 (ND)	0 (ND)	0.056	NONE	−2.48
11.2	34.2	1 0 1	0.941 (0.056)	0.648 (0.118)	0.043	1–7,9,19	−4.22
14.1	58.3	0 0 0	0 (ND)	0 (ND)	0.021	NONE	−2.41
15.1	3.0	0 0 0	0 (ND)	0 (ND)	0.021	NONE	CND[d]
15.2	2.7	1 0 0	0.220 (0.101)	0 (ND)	0.017	1,8	CND
15.3	3.8	0 0 0	0 (ND)	0 (ND)	0.024	14	CND
15.4	6.1	0 0 0	0 (ND)	0 (ND)	0.029	NONE	−0.12
18.1	29.6	1 0 0	0.388 (0.088)	0 (ND)	0.049	1,2	−3.27
19.1	37.3	0 1 1	1 (ND)	0.875 (0.048)	0.032	1,2,4–7,11	+0.20
20.1	60.9	0 1 1	1 (ND)	0.899 (0.044)	0.048	1,3,4,9,13	−1.47
22.1	33.8	0 0 0	0 (ND)	0 (ND)	0.050	NONE	+0.48

[a] Numbers to the left of the decimal point refer to the field or area within a field from which the data sets were collected. Numbers to the right of the decimal place distinguish between data sets collected from the same field or area within the field.
[b] Parameter estimates for the generalized ARIMA(1 0 1) model (11). A $\phi_{Y,1}$ value of 1 implies that differenced data were modeled. Numbers in parentheses are standard errors of the parameter estimates. (ND) = Not determinable.
[c] Autocorrelation function (ACF) lags (s), among the first 20 lags, for which autocorrelation-function values, $r_Y(s)$, lay outside a 95% confidence interval.
[d] Could not determine. Median disease was zero.

TABLE 4. Results from autocorrelation and runs analyses of 2,4-systematic samples

Data set[a]	Mean disease (%)	Significant ACF lags[b]	Runs-test Z-value
11.24	31.1	1–3	−2.28
14.24	47.1	1,10	−1.11
15.24	3.5	NONE	−0.86
18.24	35.2	14	−1.97
19.24	30.7	1,12	−2.44
20.24	56.3	1,3,4,8,11,15	−2.05
22.24	18.6	1–15	−8.42

[a] Numbers to the left of the decimal point refer to the field or area within a field from which the data sets were collected. Numbers to the right of the decimal point indicate that the data sets are 2,4-systematic samples.
[b] Autocorrelation function (ACF) lags (s), among the first 20 lags, for which autocorrelation-function values, $r_Y(s)$, lay outside a 95% confidence interval.

by an ARIMA(0 0 0) model. In contrast, five samples exhibited two or more values of $r_Y(s)$ that were outside a 95% CI. These samples were adequately described by an ARIMA(1 0 0) model (two samples), by an ARIMA(1 0 1) model (one sample), or by an ARIMA(0 1 1) model (two samples). The parameter estimates for these models are given in Table 3.

2,4-systematic samples. Within samples 11.24, 14.24, and 16.24 (evaluated by two workers), the mean disease incidences assessed by individual workers differed by 8.4, 13.1, and 8.0%, respectively. Results of runs, autocorrelation, and ARIMA analyses of the 2,4-systematic samples (after correction for assessor effects) are summarized in Tables 4 and 5. Based on all three types of analyses, we conclude that disease was not randomly arranged within the row segments. For all seven sets, the observed number of runs was smaller than the expected number. In five of seven samples, runs-test Z-values were −1.96 or less ($P \leq 0.05$). Similarly, six of the seven samples exhibited at least one value of $r_Z(s)$ that lay outside a 95% CI calculated assuming a random disease pattern. In five of these six samples, the lag 1 value of $r_Z(s)$, $r_Z(1)$, lay outside this CI. In addition, six of seven samples could be modeled with nonrandom ARIMA models. The single sample for which a random model, an ARIMA(0 0 0) model, provided an adequate fit (sample 15.24) was the sample that had no values of $r_Z(s)$ that lay outside the 95% CI described above. In addition, this sample also did not show a statistically significant ($\alpha = 0.05$) number of runs (Tables 4 and 5).

Among those samples that exhibited some type of nonrandom pattern of disease, two (samples 18.24 and 22.24) were best fit by a generalized ARIMA(1 0 1) model (11). A third sample (19.24) was most adequately modeled by an ARIMA(1 0 2) model:

$$Z_t = \phi_{Z,1} Z_{t-1} + \epsilon_t - \theta_{Z,1}\epsilon_{t-1} - \theta_{Z,2}\epsilon_{t-2} + \delta_Z$$

As noted earlier, Zs represent average transformed disease values (corrected for assessor if appropriate) associated with each sampled plant pair in the row segment. Similarly, the ϵs are random errors (with a variance of σ_ϵ^2) associated with each sampled plant pair. The ϕs, θs, and δ_Z are constants that quantify the relationship between the Zs and ϵs.

Two additional samples (14.24 and 20.24) were adequately modeled by an ARIMA(1 0 3) model:

$$Z_t = \phi_{Z,1} Z_{t-1} + \epsilon_t - \theta_{Z,1}\epsilon_{t-1} - \theta_{Z,2}\epsilon_{t-2} - \theta_{Z,3}\epsilon_{t-3} + \delta_Z$$

For both of these samples, $\theta_{Z,2}$ was not significantly different from zero, and thus, the $-\theta_{Z,2}\epsilon_{t-2}$ term could be dropped from the ARIMA(1 0 3) model, leading to a more specialized model:

$$Z_t = \phi_{Z,1} Z_{t-1} + \epsilon_t - \theta_{Z,1}\epsilon_{t-1} - \theta_{Z,3}\epsilon_{t-3} + \delta_Z$$

Similarly, data from the final systematic sample (11.24) could

be modeled by a specialized version of an ARIMA(1 0 4) model:

$$Z_t = \phi_{Z,1} Z_{t-1} + \epsilon_t - \theta_{Z,1}\epsilon_{t-1} - \theta_{Z,4}\epsilon_{t-4} + \delta_Z$$

Parameter estimates for the ARIMA models are provided in Table 5.

DISCUSSION

In previous work (11), we described patterns of bacterial brown spot that occurred within short snap bean-row segments (approximately 5 m). Brown spot patterns in these row segments were adequately described by a generalized ARIMA(1 0 1) model. A total census of disease in short-row segments, taken in conjunction with the 2,4-systematic samples collected in this study, yielded data that also were described adequately by a generalized ARIMA(1 0 1) model. We used this information to predict the type of ARIMA model that would describe a 2,4-systematic sample collected from the same field.

The type of ARIMA model that will describe an *A,B*-systematic sample can be predicted based on knowledge of the ARIMA model that describes the underlying process from which the sample is collected (20,22, B. D. Hudelson, *unpublished data*). If the underlying process from which an *A,B*-sytematic sample is collected is an ARIMA(1 0 q) process, then the *A,B*-systematic sample will be described by an ARIMA(1 0 q') model in which q' is the greatest integer in $1 + (q + A - 2)/(A + B)$ (20,22, B. D. Hudelson, *unpublished data*). In the context of our bacterial brown spot work, these theoretical results implied that if a total census sample from a snap bean field was described by a generalized ARIMA(1 0 1) model, then in the absence of additional patterns, a 2,4-systematic sample from the same field should also follow a generalized ARIMA(1 0 1) model.

For three of the seven fields (15,18,22), the patterns of disease were described by a generalized ARIMA(1 0 1) model. The patterns observed in 2,4-systematic samples 15.24, 18.24, and 22.24 were consistent with the patterns predicted from a complete census of disease within short-row segments. However, four of the seven 2,4-systematic samples were not adequately described by the predicted, generalized ARIMA(1 0 1) model. These samples were better fit by an ARIMA(1 0 2), ARIMA(1 0 3), or ARIMA(1 0 4) model. Disease in these samples exhibited not only the patterns of disease that were expected, based on disease patterns previously observed in short-row segments (Table 3; [11]), but also exhibited patterns characterized by correlations in disease associated with plant pairs two, three, or four sampled plant pairs apart, respectively.

Given knowledge of the ARIMA model that describes an *A,B*-systematic sample, one can attempt to determine the ARIMA model that will describe the underlying process from which that sample was collected. For the 2,4-systematic sample described in this work, $A = 2$ and $B = 4$. The value of q' is 2, 3, or 4 for the ARIMA(1 0 2), ARIMA(1 0 3), or ARIMA(1 0 4) processes, respectively. Thus, for sample 19.24, described by an ARIMA(1 0 2) model, we seek q, such that 2 is the greatest integer in $1 + q/6$ (i.e., $6 \leq q \leq 11$). Thus, 2,4-systematic sample 19.24 appears to exhibit a nonrandom relationship in brown spot incidence at a distance of approximately six to 11 plants. This plant spacing corresponds to a physical distance of approximately 0.41–0.75 m. Similarly, those samples described by an ARIMA(1 0 3) model (14.24 and 20.24) exhibit a nonrandom relationship in brown spot incidence at a distance of 12–17 plants. For sample 14.24, this plant spacing is approximately 0.89–1.26 m; the plant spacing in sample 20.24 is approximately 0.97–1.37 m. Finally, sample 11.24, which was described by an ARIMA(1 0 4) model, exhibits a relationship in disease incidence at a distance of 18–23 plants or 1.18–1.51 m. We should note that there may be many ARIMA(p 0 q) processes ($p > 1$, $q > 1$) that could potentially lead to the observed ARIMA(1 0 q') processes (20). However, for bacterial brown spot, a $p > 1$ has not been detected in the 55 total-census samples collected to date. As a result, models with $p > 1$ have not been considered further.

Bacterial brown spot patterns for plants that are six to 23 plants apart can be detected not only in 2,4-systematic samples, but also in the 55 total-census samples collected from short-row segments (approximately 5 m) between 1985 and 1989 ([11]; Table 6). In total, 52 of these 55 total-census samples were modeled by a generalized ARIMA(1 0 1) model. Sample autocorrelation functions (ACFs) and partial autocorrelations (PACFs) (3) of the residuals obtained after modeling these total-census samples were evaluated for evidence of disease patterns in addition to those described by the ARIMA(1 0 1) model. For the three total-census samples that could not be modeled using ARIMA models, sample ACFs and PACFs of the arcsine square root transformed disease-incidence values were also reevaluated for the presence of disease patterns. For all 55 total-census samples, a sample ACF or PACF value that lay outside an appropriate 95% CI (4) was considered possible evidence of an additional disease pattern. A summary of the results of this analysis is given in Table 6. Almost half of the samples (24 of 55) had one or more residual ACF or PACF values that were outside an appropriate 95% CI. This frequency is similar to the frequency with which the six- to 23-plant disease patterns were observed in the 2,4-systematic samples (four of seven). Additional nonrandom

TABLE 5. Results from ARIMA modeling of 2,4-systematic samples

Data set[a]	ARIMA model	$\phi_{Z,1}$ estimate[b]	$\theta_{Z,1}$ estimate[b]	$\theta_{Z,2}$ estimate[b]	$\theta_{Z,3}$ estimate[b]	$\theta_{Z,4}$ estimate[b]	σ_ϵ^2 estimate[b]
11.24	1 0 4	0.757 (0.124)	0.293 (0.166)	0 (ND)	0 (ND)	0.261 (0.095)	0.039
14.24	1 0 3	0.817 (0.083)	0.687 (0.111)	0 (ND)	0.316 (0.096)	0 (ND)	0.012
15.24	0 0 0	0 (ND)	0 (ND)	0 (ND)	0 (ND)	0 (ND)	0.013
18.24	0 1 1	1 (ND)	0.924 (0.032)	0 (ND)	0 (ND)	0 (ND)	0.123
19.24	0 1 2	1 (ND)	0.772 (0.096)	0.166 (0.105)	0 (ND)	0 (ND)	0.024
20.24	1 0 3	0.596 (0.234)	0.475 (0.246)	0 (ND)	−0.183 (0.098)	0 (ND)	0.033
22.24	1 0 1	0.937 (0.044)	0.658 (0.090)	0 (ND)	0 (ND)	0 (ND)	0.036

[a] Numbers to the left of the decimal point refer to the field from which the data sets were collected. Numbers to the right of the decimal point indicate the data sets are 2,4-systematic samples.

[b] Parameter estimates from ARIMA modeling. A $\phi_{Z,1}$ value of 1 implies that differenced data were modeled. Numbers in parentheses are standard errors of the parameter estimates. (ND) = Not determinable.

variability in the total-census samples occurred most frequently (15 of 24 cases) at a distance of 12 to 17 plants and approximately one-half as frequently at distances of six to 11 or 18 to 23 plants (seven of 24 and four of 24 samples, respectively). These frequencies are consistent with frequencies observed in the 2,4-systematic samples (two of four, one of four, and one of four samples, respectively). Thus, evidence for the frequent occurrence of the additional patterns of brown spot observed in the 2,4-systematic samples appears to be found in the total-census samples. We believe that these are general patterns, and they might be expected to occur on a regular basis in other snap bean fields and in other growing seasons.

The underlying mechanisms generating the disease patterns we observed in total-census and 2,4-systematic samples are unclear. Patterns described by a generalized ARIMA(1 0 1) model could arise in many ways. They simply may reflect a similar pattern in *Pseudomonas syringae* population sizes associated with the plants. Alternately, using the theory of ARIMA models, one can show that a generalized ARIMA(1 0 1) pattern can also arise through the combination of several processes: One of which can be described by an ARIMA(1 0 0) model; the rest of which are random processes (3). The random processes could include factors such as the random initial distribution of *P. syringae* or random variability in the genotype of plants within the row segment. Possible candidates for the ARIMA(1 0 0) process, represented by a relatively smooth, undulating change in the process along the row, might be microenvironmental factors that affect disease development or soil nutritional factors that affect plant susceptibility. Many other potential explanations for the observed patterns described by the generalized ARIMA(1 0 1) model are possible.

The origins of the six- to 23-plant patterns are also open to speculation. Overall, within total-census samples, the average distance at which the six- to 23-plant patterns occur is 76.2 cm (SD = 28.9 cm). This mean value is interesting because a common row spacing for several agricultural row crops in Wisconsin is 76.2 cm. If at least some growers alternate the direction of row orientation from year to year, then the patterns that center on 76.2 cm might represent some sort of carry-over effect related to past cropping history. This carry-over might include *P. syringae* pv. *syringae* associated with plant debris or could represent the effects of furrow-incorporated fertilizers or pesticides. Because of the relatively large standard error associated with the 76.2-cm mean, this explanation is not adequate to explain all of the six- to 23-plant patterns detected in total-census or 2,4-systematic samples. Other cultural practices carried out across rows (or even at an oblique angle) might be the source of these patterns. Potential causes of the six- to 23-plant patterns (including possible cultural origins) are currently under investigation.

In addition to its ability to detect larger scale patterns than those detected in total-census samples, 2,4-systematic samples may also provide an added ability to detect the undulating and jagged patterns originally detected in short-row segments by total-census sampling and described by a generalized ARIMA(1 0 1) model (11). Consider, for example, total-census samples 11.1, 14.1, and 22.1 that appeared to exhibit a random brown spot pattern and were adequately described by an ARIMA(0 0 0) model (Table 3)—the special case of a generalized ARIMA(1 0 1) model in which $\phi_{Y,1} = \theta_{Y,1}$ (11). Corresponding 2,4-systematic samples 11.24, 14.24, and 22.24, however, exhibited both the undulating and jagged patterns and were fit with ARIMA(1 0 1) models (Table 5) in which $\phi_{Z,1} > \theta_{Z,1} > 0$. Results from the 2,4-systematic samples are consistent with the hypothesis that the undulating and jagged patterns may have been present in the total-census samples but could not be detected easily.

Difficulty in detecting the undulating and jagged patterns can arise when the magnitudes of these two patterns are approximately equal. In the context of the generalized ARIMA(1 0 1), this means that $\phi_{Y,1} \approx \theta_{Y,1}$. In such a situation, a total-census data set of small size collected from such a process may appear random and be modeled as an ARIMA(0 0 0) process.

The ability of 2,4-systematic samples to detect ARIMA(1 0 1)

TABLE 6. Analysis of total-census data sets for brown spot patterns that recur every six to 23 plants

Data set[a]	Length (m)	Number of plants	Significant ACF/PACF lag[b]	Lag distance (m)
1.1	5.00	78	13(A)	0.833
1.2	5.00	74	NONE	...
1.3	5.00	70	15(A/P)	1.071
1.4	5.00	73	NONE	...
2.1	5.00	102	10(A/P)	0.490
2.2	5.00	102	NONE	...
2.3	5.00	94	NONE	...
2.4	5.00	98	NONE	...
3.1	5.00	72	NONE	...
3.2	5.00	96	NONE	...
3.3	5.00	91	NONE	...
3.4	5.00	85	19(P)	1.118
4.1	5.00	97	15(A/P)	0.773
			17(A/P)	0.876
4.2	5.00	83	NONE	...
4.3	5.00	107	NONE	...
4.4	5.00	91	9(A/P)	0.494
5.1	5.00	65	NONE	...
5.2	5.00	82	NONE	...
5.3	5.00	76	17(A/P)	1.118
5.4	5.00	74	17(P)	1.149
5.5	5.00	75	NONE	...
6.1	5.00	101	8(A/P)	0.396
6.2	5.00	83	NONE	...
6.3	5.00	74	NONE	...
6.4	5.00	67	15(A/P)	1.119
7.1	5.00	96	16(A/P)	0.833
7.2	5.00	95	13(A/P)	0.684
7.3	5.00	60	NONE	...
7.4	5.00	107	14(P)	0.654
8.1	2.81	59	19(A/P)	0.905
8.2	1.89	45	NONE	...
8.3	2.71	88	6(P)	0.185
9.1	1.85	51	10(A/P)	0.363
10.1	12.29	264	11(P)	0.512
			13(A/P)	0.605
11.1	5.00	81	18(P)	1.111
11.2	5.00	82	NONE	...
12.1	5.00	60	NONE	...
13.1	5.00	60	13(A/P)	1.083
14.1	5.00	100	NONE	...
14.2	5.00	77	NONE	...
15.1	4.25	116	NONE	...
15.2	3.58	96	15(A/P)	0.559
15.3	4.21	100	14(A)	0.589
15.4	ND[c]	67	NONE	...
16.1	ND	100	16(P)	0.616[d]
16.2	ND	100	15(P)	0.578[d]
16.3	ND	100	NONE	...
17.1	6.52	98	9(P)	0.599
			19(A/P)	1.264
17.2	ND	102	NONE	...
18.1	7.00	111	NONE	...
19.1	ND	100	NONE	...
20.1	ND	100	NONE	...
20.2	ND	101	NONE	...
21.1	7.11	92	NONE	...
22.1	2.43	70	NONE	...

[a] Numbers to the left of the decimal point refer to the field or area within a field from which the data sets were collected. Numbers to the right of the decimal place distinguish between data sets collected from the same field or area within the field.

[b] Lags (s) for which residual autocorrelation-function (ACF) or partial autocorrelation-function (PACF) values lay outside a 95% confidence interval (CI), after fitting the data with an appropriate generalized ARIMA(1 0 1) model. (A) indicates that only the ACF value lays outside the CI. (P) indicates that only the PACF value lays outside the CI. (A/P) indicates that both the ACF and PACF values lay outside the CI.

[c] Not determined.

[d] Distances were based on the average plant densities for data sets 15.1, 15.2, and 15.3.

processes in which $\phi_{Y,1}$ and $\theta_{Y,1}$ are close but not equal may arise from several sources. The 2,4-systematic samples collected in this study contained approximately two times the number of plants found in the typical total-census sample collected in our previous work. This larger sample size should yield more precise parameter estimates and add power to detect ARIMA processes with small differences in $\phi_{Y,1}$ and $\theta_{Y,1}$.

Added sensitivity in 2,4-systematic samples may also be a result of the averaging of individual disease-incidence values that takes place before the sample is modeled. The averaging process reduces variability and potentially increases the ability to detect spatial patterns when $\phi_{Y,1} \approx \theta_{Y,1}$. In addition, averaging has more subtle effects. As noted above, a 2,4-systematic sample collected from an underlying ARIMA(1 0 1) process ($\phi_{Y,1} \neq \theta_{Y,1}$) should also be described as an ARIMA(1 0 1) process ($\phi_{Z,1} \neq \theta_{Z,1}$). If a total-census sample exhibits an undulating and a jagged disease pattern, then the corresponding 2,4-systematic sample will also exhibit both patterns. However, the strengths of these two patterns will not be the same in the total-census and in the 2,4-systematic samples, and more importantly, the relative magnitudes of the two patterns will differ in the two samples (B. D. Hudelson, *unpublished data*). For example, if $0.90 \leq \phi_{Y,1} < 1$ (the undulating pattern in the total-census sample is strong) and $0 < \phi_{Y,1} - \theta_{Y,1} < 0.1$ (the jagged pattern in the total-census sample is only slightly weaker than the undulating pattern), then the difference in the magnitudes of the undulating and jagged patterns in the 2,4-systematic sample will increase. Thus, both patterns are more likely to be detected.

For an empirical example of this phenomenon, consider 2,4-systematic sample 22.24. The total-census data set collected in conjunction with this 2,4-systematic sample appeared to have a random disease pattern (Table 3). However, parameter estimates for the ARIMA(1 0 1) model fit to the 2,4-systematic sample were $\hat{\phi}_{Z,1} = 0.937$ and $\hat{\theta}_{Z,1} = 0.658$, indicative of a nonrandom disease pattern. The estimates for $\phi_{Z,1}$ and $\theta_{Z,1}$ can be used to estimate the corresponding values of $\phi_{Y,1}$ and $\theta_{Y,1}$ (B. D. Hudelson, *unpublished data*). These values are $\hat{\phi}_{Y,1} = 0.989$ and $\hat{\theta}_{Y,1} = 0.887$. The difference between $\hat{\phi}_{Y,1}$ and $\hat{\theta}_{Y,1}$ is only 0.102. The ARIMA modeling techniques we have used in our studies have detected a difference of this magnitude or smaller in only one other sample (11). Thus, we should not be surprised that disease from the total-census sample was modeled with an ARIMA(0 0 0) or a random model.

In this paper, we described one form of the second iteration of an adaptive sampling strategy for detection and modeling of unknown spatial patterns. The first step, detection and modeling of patterns within small areas using a total census, has been described previously (11). A generalized ARIMA(1 0 1) model described patterns of disease in more than 90% of the row segments in which these censuses were taken. Knowledge that disease in short-row segments could be described using this model provided the basis for developing the 2,4-systematic sampling method for use in the second iteration of adaptive sampling (B. D. Hudelson, *unpublished data*). The 2,4-systematic sample allowed efficient sampling of plants for detection of patterns at larger scales. In addition, disease patterns in 2,4-systematic samples were described readily using ARIMA models. Finally, the type of ARIMA model that should fit data from a 2,4-systematic sample could be predicted, based on the patterns observed at smaller scales. Thus, inconsistencies between observed and predicted models were indicative of additional, larger scale disease patterns. Application of 2,4-systematic sampling in our work with bacterial brown spot has confirmed the usefulness of this sampling plan for detecting spatial patterns, at least when patterns at the smaller scale conform to the generalized ARIMA(1 0 1) model.

Overall, adaptive sampling is relatively simple and is consistent with the constraints of small sample sizes. The number of plants assessed for disease needed to satisfy this sampling plan was designed to fit within the constraint that it should not exceed the capability of one trained individual working for one day (although some samples have been evaluated by more than one person). Analytical methodology is readily available in commercial software packages (e.g., MINITAB, release 5.1.1, or the SCA Statistical System, release III.3.2). Although use of ARIMA models requires some knowledge of statistics, it is neither conceptually difficult nor inaccessible (Cryer [10] gives a good introduction to the subject). Thus, with a modest amount of effort, the methods we used in the first two iterations of adaptive sampling for spatial patterns of bacterial brown spot should be readily available to plant pathologists attempting to understand patterns of other plant diseases. Although the focus in our work has been on a row crop, adaptive sampling also can be used to study spatial patterns in nonrow crops. For nonrow crops, however, added care must be taken to initially select an appropriate sampling unit. The advantages of adaptive sampling are that it is amenable to detection of multiple, superimposed patterns that occur at different scales. ARIMA modeling provides a model of disease patterns, rather than simply determining whether they are random, aggregated, or disaggregated. It also allows determination of the scale at which the patterns occur and avoids the detection limitations imposed by the use of quadrat sampling (16). Finally, ARIMA modeling provides a model that can be used to develop sampling plans for exploring the possible existence of patterns that may exist at still larger scales.

Adaptive sampling as we have developed it for the study of spatial patterns of bacterial brown spot is not without its problems. The 2,4-systematic samples used in this study imprecisely identify the distance at which larger scale patterns occur. For example, we are able to identify nonrandom variability in disease for plants that are 12 to 17 plants apart but cannot identify the distance within this range of values more precisely. This imprecision is inherent to *A,B*-systematic samples and will only be overcome by the development and use of more sophisticated sampling plans.

In addition, we have concentrated on identifying disease patterns in a single dimension. Disease development is a multidimensional process, and the use of multidimensional techniques would be preferable. We have opted for a unidimensional approach because the theory and application of unidimensional techniques (e.g., ARIMA modeling and associated sampling theory) are well developed and documented and are easily accessible. Our models of disease patterns, while not realistic mechanistically, do represent a relatively precise description of a unidimensional "slice" of the multidimensional brown spot patterns. This description provides a foundation that can be used to hypothesize about and test potential multidimensional mechanisms for pattern development. Any mechanistic hypothesis for bacterial brown spot must be consistent with the patterns we have observed.

The development of statistical methods for analyzing multidimensional spatial patterns continues to be an active area of research, and it must be acknowledged that all currently available techniques have their strengths and weaknesses. The most sophisticated multidimensional techniques used by plant pathologists have been kriging (13) and STARIMA (spatio-temporal autoregressive integrated moving average) modeling (17,18). Kriging is an extension of the use of semivariograms for describing spatial patterns. It is difficult to find a way to incorporate such methods into the type of adaptive plan discussed in this paper, although some work on sampling design has been discussed (9). STARIMA models represent a logical extension of multidimensional autocorrelation functions. Although STARIMA models have been well described theoretically (2,17), parameter estimates for such models cannot be obtained, except for models describing the simplest disease patterns (1,18). The work of Basu and Reinsel (1) provides hope that methods for estimating parameters for more complex models will be available in the near future. Such a development would make adaptive sampling more amenable to use in multiple dimensions.

LITERATURE CITED

1. Basu, S., and Reinsel, G. C. 1990. Analysis of first-order spatial bilateral ARMA models. University of Wisconsin, Madison, Dep. Stat. Tech. Rep. 862.

2. Bennett, R. J. 1979. Spatial Time Series. Pion Limited, London. 674 pp.
3. Box, G. E. P., and Jenkins, G. M. 1976. Time Series Analysis. Forecasting and Control. Rev. ed. Holden-Day, San Francisco. 575 pp.
4. Box, G. E. P., and Pierce, D. A. 1970. Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. J. Am. Stat. Assoc. 65:1509-1526.
5. Campbell, C. L., and Madden, L. V. 1990. Introduction to Plant Disease Epidemiology. John Wiley & Sons, New York. 532 pp.
6. Campbell, C. L., and Noe, J. P. 1985. The spatial analysis of soilborne pathogens and root diseases. Annu. Rev. Phytopathol. 23:129-148.
7. Chellemi, D. O., Rohrbach, K. G., Yost, R. S., and Sonoda, R. M. 1988. Analysis of the spatial pattern of plant pathogens and diseased plants using geostatistics. Phytopathology 78:221-226.
8. Clayton, M. K., and Hudelson, B. D. 1991. Spatial patterns in microbial activity. Pages 111-131 in: Microbial Ecology of Leaves. J. H. Andrews and S. S. Hirano, eds. Springer-Verlag, New York. 499 pp.
9. Cressie, N. A. C. 1991. Statistics for Spatial Data. John Wiley & Sons, New York. 900 pp.
10. Cryer, J. D. 1986. Time Series Analysis. Duxbury Pr., Boston, MA. 286 pp.
11. Hudelson, B. D., Clayton, M. K., Smith, K. P., Rouse, D. I., and Upper, C. D. 1989. Nonrandom patterns of bacterial brown spot in snap bean row segments. Phytopathology 79:674-681.
12. Iachan, R. 1983. Asymptotic theory of systematic sampling. Ann. Stat. 11:959-969.
13. Lecoustre, R., Fargette, D., Fauquet, C., and de Reffye, P. 1989. Analysis and mapping of the spatial spread of African cassava mosaic virus using geostatistics and the kriging technique. Phytopathology 79:913-920.
14. Mihail, J. D., and Alcorn, S. M. 1987. *Macrophomina phaseolina*: Spatial patterns in a cultivated soil and sampling strategies. Phytopathology 77:1126-1131.
15. Nicot, P. C. 1985. Epidemiology of Verticillium wilt of potato: Spatial patterns of soilborne *Verticillium dahliae* and incidence of stem colonization. Ph.D. thesis. University of Wisconsin, Madison, Dep. Plant Pathol. 134 pp.
16. Nicot, P. C., Rouse, D. I., and Yandell, B. S. 1984. Comparison of statistical methods for studying spatial patterns of soilborne plant pathogens in the field. Phytopathology 74:1399-1402.
17. Reynolds, K. M., and Madden, L. V. 1988. Analysis of epidemics using spatio-temporal autocorrelation. Phytopathology 78:240-246.
18. Reynolds, K. M., Madden, L. V., and Ellis, M. A. 1988. Spatio-temporal analysis of epidemic development of leather rot of strawberry. Phytopathology 78:246-252.
19. Snedecor, G. W., and Cochran, W. G. 1989. Statistical Methods. 7th ed. Iowa State University Press, Ames. 503 pp.
20. Stram, D. O., and Wei, W. W. S. 1986. Temporal aggregation in the ARIMA process. J. Time Ser. Anal. 7:279-292.
21. Thal, W. M., and Campbell, C. L. 1986. Spatial pattern analysis of disease severity data for alfalfa leaf spot caused primarily by *Leptosphaerulina briosiana*. Phytopathology 76:190-194.
22. Weiss, A. A. 1984. Systematic sampling and temporal aggregation in time series models. J. Economet. 26:271-281.