

Spatial Distribution of Cabbage Black Rot and the Estimation of Diseased Plant Populations

James Strandberg

Assistant Plant Pathologist, University of Florida, Agricultural Research and Education Center, Sanford 32771.

Florida Agricultural Experiment Station Journal Series Paper No. 4671.

Accepted for publication 15 February 1973.

ABSTRACT

Spatial and temporal distributions of cabbage plants infected by the bacterium *Xanthomonas campestris*, and the distribution of lesions among cabbage plants were examined. The patterns of distribution of lesions per plant and of infected cabbage plants per unit area were highly aggregated. Patterns of distributions for both lesions and infected plants were best described by the negative binomial distribution and common k values of 0.15 and 0.50 were obtained for lesions per plant and

infected plants per unit area, respectively.

A modified sequential sampling method, based on the spatial distribution of lesions and infected plants, was used to obtain average infection rates for the black rot disease with a large reduction in time and effort. An average infection rate of 0.57 was observed. Evidence to verify the spread of black rot by farm machinery, was also obtained.

Phytopathology 63:998-1003.

Additional key words: epidemiology, sequential sampling, *Xanthomonas campestris*.

The study of plant diseases in relation to agro-ecosystems is receiving new and well-deserved attention. Recent publications by van der Plank have done much to encourage the quantitative study of plant disease epidemics (7, 8, 9), while a simulated epidemic designed for computer use has pointed out additional uses and requirements for quantitative studies (10). Additional knowledge of the quantitative aspects of pathogen increase and dispersal in relation to the ecosystem will be required if we are to better understand or predict disease epidemics. However, quantitative studies of epidemics require extensive field data that is both accurate and timely, usually obtained with a minimum expenditure of time and cost. New concepts in ecology and pest management, especially in the field of arthropod research, can provide ideas and methods of great value in advancing the study of plant disease epidemics, and expediting the acquisition of field data, but many of these have not been exploited.

One such concept is spatial distribution or dispersion. It is well known that many organisms are seldom distributed within their environment on a random or regularly spaced basis (1, 6, 12, 13) and that distributional patterns of organisms per unit area are often aggregated. Such aggregated patterns are better described by mathematical frequency distributions other than the normal or Poisson (2, 6, 12, 13). The negative binomial distribution has been particularly applicable in describing the spatial distributions of several types of organisms (2, 6). Patterns of distribution can often contribute to a better understanding of the organism, particularly its reproduction, behavior and dispersal (2, 6, 7, 13). Such patterns also affect approaches to estimating population means, and much practical use has been made of this fact in formulating sampling procedures for estimating plant and arthropod populations (1, 2, 3, 5, 6, 12, 13). Sequential sampling is a method based upon observed spatial distributions which allows the observer to stop sampling when enough

units have been examined to obtain a population estimate of the desired precision, often at a great savings of time and effort (5, 6, 12). Sequential sampling has been used extensively in ecological and pest management programs (6, 12) and is appropriate for use in the study of plant diseases. Observed spatial distributions including those other than the negative binomial are applicable to sequential sampling as well as a wide variety of sampling and statistical methods (1, 5, 6, 12).

Bacterial diseases have received little attention in regard to quantitative epidemiological work. Black rot of cabbage caused by *Xanthomonas campestris* (Pam.) Dows. is well suited to such studies since the bacterium is endemic to most cabbage-growing areas and is spread by infected seed, water, and farm equipment (11). Cabbage plants are compact, short lived and furnish discrete, regularly spaced sampling units. Thus, this pathogen-host system offers good opportunities for the study of spatial distribution.

In this study, spatial and temporal distributions of black rot-infected cabbage plants were examined and information on the dispersion of *Xanthomonas campestris* was obtained. Improved methods of estimating population means were developed by adapting sequential sampling techniques based on the observed distribution patterns of lesions and diseased plants. Advantages of considering both spatial as well as temporal distributions of pathogens or diseased plants are demonstrated.

MATERIALS AND METHODS.—Cabbage seeds var. 'Little Rock' were hot water treated (50 C for 35 min) to help eliminate black rot bacteria, planted in seedbeds and 40 days later plants were transplanted into the field. Normal methods of plant bed sanitation were observed (11). All sites were located on Leon fine sand at Sanford, Florida, and were fumigated with a dichloropropane-dichloropropene mixture to control nematodes. Plants were grown in rows 0.76 m apart and were spaced 0.3 m in-the-row. Plot size varied as did that of quadrats, but the

smallest quadrats used contained at least 10 plants. Quadrats used to obtain any one set of count data were always of equal size and contained the same number of plants. Numbers of black rot lesions per plant and numbers of infected plants per quadrat were determined by direct counts at weekly intervals. In studies of dispersion, the exact locations by row and plant of infected and noninfected plants within the field were plotted and analyzed by the doublet analysis method described by van der Plank (7). For distributional data, numbers of lesions per plant and numbers of infected plants per quadrat were tabulated and arranged into frequency classes including zeros (Table 3), and the data were analyzed for fitness to several frequency distributions with the FORTRAN IV program for fitting count data to discrete frequency distributions developed by Gates & Ethridge (4). For data fitting the negative binomial distribution, this program supplied the exponent k which is an index of aggregation (13). A value of k common to all sets of such data (common k or k_c) was estimated by the regression method of Bliss & Owens (3).

The sequential sampling method described by Kuno (5) was adapted for use with the common k derived from the analysis of distributional data for several separate counts of lesions and infected plants. Sampling curves (stop lines) were calculated and drawn for estimating populations of black rot lesions and diseased plants with a level of precision of 10% of the mean. An infinite population was assumed since all fields sampled contained several times the number of plants examined. Using procedures described by van der Plank (8) to calculate $\log_{10} \frac{x}{1-x}$ values for percentages of diseased plants, data from three fields were analyzed to obtain an estimate of the average infection rate. Transformed data derived from direct counts of all plants in 178 quadrats containing ten plants each in a single cabbage field were compared with data taken at random from that field on the same dates using sequential sampling methods (5, 12).

RESULTS.—Spread by equipment.—In many fields investigated, initial infection sites occurred in a strikingly similar manner. Infected plant density was usually higher in analogous corners of several

different fields, always at the end near the access road. It was also observed that farm equipment always entered the fields at these sites. This was considered circumstantial evidence for spread by equipment. To examine this method of spread quantitatively, five plots of uniformly spaced plants were grid-sampled and the exact location of infected and noninfected plants were analyzed to determine whether diseased plants were more numerous within the row (the direction of equipment travel) or within linear associations of plants across rows (perpendicular to equipment travel). The results demonstrated that a greater percentage of plants within rows were infected than across rows (Table 1). However, considering the average value for five plots, these differences were not highly significant (Table 1). Differences were greater at lower levels of infection and decreased with increasing disease levels. It is likely that other methods of disease spread may obscure the effect of equipment spread at higher disease levels.

Doublet analysis, which provides a measure of disease spread from infection sources within fields is also a test for aggregation (7). In doublet analysis the observed number of adjacent pairs of infected plants is compared with a value expected if infections occurred at random, or from outside sources. Numbers of adjacent pairs are greater than this value if infections have spread from foci within the field. Doublet analysis was applied to the data for the five plots described above (Table 1). Modest differences were obtained between observed and expected values across rows, but much larger differences were observed in the case of plants within rows. This data provides additional evidence for spread by equipment, and furthermore, demonstrates the aggregated pattern of distribution of the infected plants. In the above conclusions, differences in distances between plants within rows and between plants across rows were recognized as complicating, but not prohibitory factors.

Patterns of distribution.—Forty seven sets of black rot count data were analyzed for goodness of fit to the Poisson, negative binomial, Thomas double Poisson, Neyman Type A, Poisson with zeros, Poisson

TABLE 1. Analysis of distribution of black rot infected plants within rows (in the direction of equipment movement) and across rows (perpendicular to equipment movement) in five plots of regularly spaced cabbage plants

| Plot | Percent of plants infected | Percent of total plants infected | | Expected | Doublet analysis | |
|------|----------------------------|----------------------------------|-------------|------------|------------------|-------------|
| | | Within rows | Across rows | | Observed | Observed |
| | | | | | Within rows | Across rows |
| A | 8.7 | 11.7 | 5.9 | 5.3 | 33 | 22 |
| B | 16.5 | 16.5 | 6.5 | 16.5 | 43 | 42 |
| C | 12.6 | 14.0 | 13.3 | 12.6 | 74 | 44 |
| D | 14.5 | 32.0 | 31.9 | 14.5 | 166 | 127 |
| E | 21.9 | 44.0 | 46.0 | 21.9 | 280 | 225 |
| | Mean ^a | 23.6 | 20.7 | Total 70.8 | 596 | 460 |

^aMeans are significantly different at the $P = .05$ level as determined by Students' t -test applied to angular transformations of the percentage values.

TABLE 2. Summary of counts of bacterial lesions-per-plant and infected-plants-per-unit area and their fitness to the Poisson, negative binomial and other distributions

| Sampling method | No. of counts | Distributions fitted | | | Range of <i>k</i> values | |
|-------------------------------|---------------|----------------------|---------|-------|--------------------------|----------|
| | | Negative binomial | Poisson | Other | | |
| Lesions-per-plant | 21 | 12 | 1 | 1 | 7 | 0.05-1.5 |
| Infected-plants-per-unit area | 27 | 19 | 7 | 1 | 0 | 0.1-11 |

binomial, and logarithmic with zeros, distributions. Most counts of lesions per plant, or infected plants per quadrat, were fitted by the negative binomial distribution at the .01 level as determined by the chi-square test (Table 2). Some counts were best fitted by the Poisson (which describes a population distributed at random) or by other distributions such as the Neyman Type A or logarithmic with zeros, which describe aggregated populations. Seven sets of the lesion per plant data fitted none of the above distributions, but for these counts the variance was much greater than the mean, indicating a significant departure from random distributions. Some counts were fitted by more than one distribution at the *P*.05 or .01 levels, but this is a common occurrence in analyses of this type of data (13). Results (Table 2)

TABLE 3. Expected and observed frequencies of black rot-infected plants per quadrat for the negative binomial and Poisson distributions as applied to a single representative set of count data

| Class ^a | Observed frequency | Poisson expected | Negative binomial expected |
|--------------------|--------------------|------------------|----------------------------|
| 0 | 39 | 21.76 | 42.56 |
| 1 | 56 | 54.79 | 59.38 |
| 2 | 70 | 68.99 | 54.66 |
| 3 | 45 | 57.93 | 41.68 |
| 4 | 23 | 36.47 | 28.48 |
| 5 | 9 | 18.37 | 18.11 |
| 6 | 10 | 7.71 | 10.94 |
| 7 | 7 | 2.77 | 6.36 |
| 8 | 3 | 0.87 | 3.59 |
| 9 | 4 | 0.24 | 1.98 |
| 10 | 2 | 0.06 | 1.07 |
| 11 | 1 | 0.01 | 0.57 |
| 12 | 1 | 0.00 | 0.62 |
| 13 | 0 | 0.00 | ... |

| | | |
|-------------------------------------|-------------------|---------------------------------|
| Calculated chi-square value | 113.73 | 14.35 |
| Probability of exceeding this value | 0.00 ^b | 0.1103 ^c |
| General statistics | Mean 2.518 | Variance 4.853 <i>k</i> = 3.128 |

^aNumber of infected plants per quadrat.

^bSignificantly different than the expected distribution at the *P* = .01 or .05 level.

^cNot differing significantly from the expected distribution at the *P* = .01 level.

demonstrated that the spatial distribution of black rot lesions or infected plants is not random, but highly aggregated in nature and that it may best be described by the negative binomial distribution. Two explanations are possible for counts fitting the Poisson distribution. It is known that large populations sometimes result in random distributions (13) and that insufficiently small quadrat size may yield similar results.

For counts fitting the negative binomial, the mean, variance, and exponent *k* were provided by the computer print-outs. Results for one such count are shown in Table 3. Values of *k* common to all such counts for the lesion-per-plant or infected-plant-per-quadrat data were 0.15 and 0.50, respectively. The fitting of a common *k* to these data was justified by the criteria described by Bliss & Owen (3).

Sequential sampling procedure.—To adapt a sequential sampling procedure to estimate diseased plant populations, sampling curves were drawn for both lesions-per-plant and infected-plant-per-quadrat data following the methods of Kuno (5) and using the above common *k* values. The resulting sampling lines which appear as hyperbolic curves are shown in Fig. 1 and 2.

Cabbage plants in one field were sampled by both a complete inspection of all plants within 178 quadrats containing 10 plants each and by the modified sequential method. In the sequential sample, plants were examined at random covering all parts of the field. Cumulative numbers of infected plants (T_n of Fig. 2) were plotted against number of plants examined (*n* of Fig. 2). When the plotted cumulative count line (see theoretical count lines for means of 0.1 and 0.5, which appear as diagonal lines in Fig. 2) intersected the stop line, that point estimated the population mean with the predetermined level of precision; in this case, 10% of the mean. The mean was obtained from T_n/n at this point on the curve and further sampling was not required. It should be noted that the stop lines of Fig. 1 and 2 are described by the following formula for the special case of data fitting the negative binomial distribution (5):

$$T_n = \frac{1}{Do^2 - \frac{1}{nk_c}} \quad (1)$$

TABLE 4. Comparison of direct counts and sequential sampling for obtaining transformed data from percentage of black rot-infected plants (data for curves C and D, Fig. 3)

| Percent infected plants by | | Log ₁₀ $\frac{x}{1-x}$ values by | | Number of plants examined | |
|----------------------------|------------|---|------------|---------------------------|------------|
| Direct | Sequential | Direct | Sequential | Direct | Sequential |
| 5.0 | 5.0 | 2.72 | 2.72 | 1,780 | 2,000 |
| 17.9 | 16.4 | 1.34 | 1.30 | 1,780 | 820 |
| 27.3 | 31.2 | 1.58 | 1.66 | 1,780 | 510 |
| 66.4 | 68.0 | 0.29 | 0.32 | 1,780 | 360 |
| 70.5 | 71.5 | 0.38 | 0.40 | 1,780 | 340 |
| 72.2 | 73.5 | 0.41 | 0.44 | 1,780 | 330 |
| 82.5 | 82.0 | 0.67 | 0.65 | 1,780 | 320 |
| Total | | | | 12,460 | 4,680 |

where T_n =cumulative number of lesions or infected plants; D_0 =predetermined level of precision (0.1), n =number of plants examined; i.e., sample size and k_c = common k (calculated above). The stop line was also calculated for estimating number of lesions per plant (Fig. 1). A comparison of Fig. 1 and 2 demonstrates the effect of the value of k_c on the sampling plan. It is obvious that the smaller value of k_c , obtained for lesions per plant, and which indicates a more highly aggregated population, requires that much larger samples be taken to estimate the true mean with a given precision.

Estimation of average infection rate.—In most fields sampled, counts were made on a weekly or

biweekly basis. Such data were suitable for transformation by the methods of van der Plank (8) to obtain estimates of average infection rates. Results from three fields obtained at different times of the year are shown in Fig. 3. There was good agreement between all three average infection rates, as indicated by a similarity between slopes of lines, A, B, and C (Fig. 3). An estimated average infection rate of 0.57 was derived by averaging the slopes of the three curves.

Comparison of sequential and standard sampling methods.—The data transformed by $\log_{10} \frac{x}{1-x}$ to draw curve C and D (Fig. 3) were utilized to make a comparison between the use of sequential sampling methods and complete counts of infected plants to estimate the mean population of diseased plants. Complete counts made at weekly intervals in large areas of the field provided accurate measurements of the percentage of black rot-infected plants. Similar percentage values were obtained by applying the sequential sampling method and the sampling curve of Fig. 2. Values obtained by sequential sampling

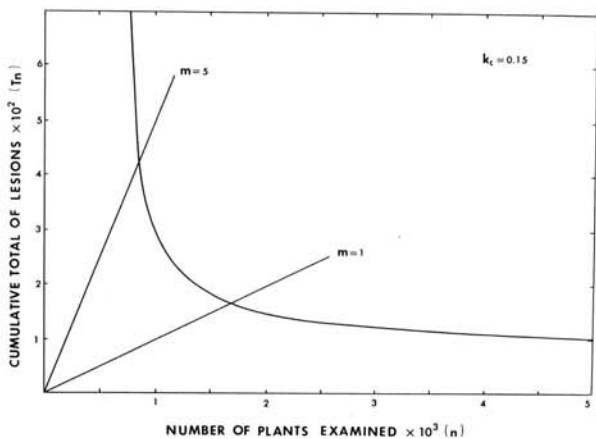


Fig. 1. Sampling curve (stop line) for sequential sampling to estimate the average number of black rot lesions per plant with theoretical plots (shown as diagonal lines) for population means (m) of 1 and 5 lesions per plant. The cumulative number of lesions observed (T_n) are successively plotted against the number of plants examined (n). If a line is drawn connecting these points, the point on the sampling curve intersected by that line estimates the population mean with a predetermined level of precision, in this case 10% of the mean. The mean (m) = T_n/n at the point of intersection with the sampling curve and further sampling is unnecessary. The curves are constructed by calculating T_n for several values of n using the observed common k and Equation 1.

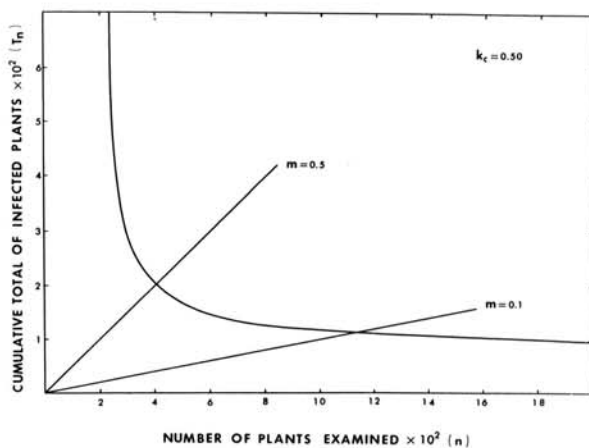


Fig. 2. Sampling curve (stop line) for sequential sampling to determine the percentage of black rot-infected plants in a population with theoretical plots for means of 0.1 and 0.5 (10 and 50% of the plants infected).

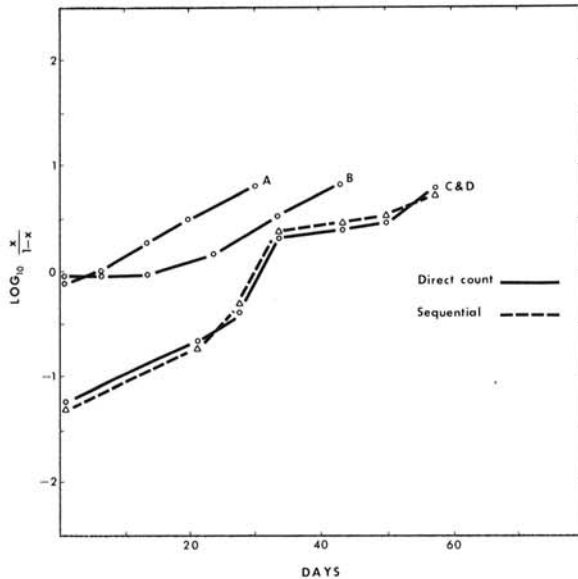


Fig. 3. Infection rates for black rot of cabbage observed on variety 'Little Rock' during the periods April 13 - May 10 (curve A), March 8 - June 11 (curve B) and June 28 - August 25 (curve C and D). An average infection rate of 0.57 was estimated. Curves C and D compare values obtained by a direct count of all plants in a given area and by sequential sampling which resulted in a large reduction in the number of plants sampled.

were very close to those obtained from complete field counts (Fig. 3 and Table 4), and approximately two-thirds fewer plants were examined to obtain them, thus demonstrating the utility of the sequential method. In addition, a known level of precision was assigned to the data.

DISCUSSION.—It seems proper to assume that spatial patterns of distribution of plant pathogens, like other natural populations of organisms, are seldom truly random or regular in nature. Results of this study verify that in the case of black rot of cabbage, such patterns are not random, but are best described by the negative binomial distribution. It is recognized that, in this study, the distribution of diseased plant material and not the pathogen itself has been described. This does not seem to be a serious obstacle since most epidemiological studies are based on similar, if not identical, assumptions that lesions, disease damage, or diseased plants all provide an estimate of localized densities of pathogen populations. The fact that pathogens are highly aggregated in diseased plants or lesions demonstrates that their spatial distribution is highly aggregated and not random. The parameter k , which is a measure of aggregation, has been called a "dispersion parameter" in the case of insects (1, 6, 13). Small values of k (ca. 0.1-1) indicate extreme aggregation; larger values (ca. 1-8) are common for most arthropods, while very large values which may approach infinity define a random distribution (1, 6, 13). The k value obtained for number of black rot lesions per plant is small

($k_c=0.15$) and indicates extreme aggregation. This is expected since bacteria increase and spread within infected plants, and the probability is high that new lesions will be formed on previously infected plants. *X. campestris* is spread by rain, water or farm equipment all of which involve relatively small distances. The k value for infected-plants-per-unit area ($k_c=0.50$), while somewhat greater than that for lesions-per-plant is also indicative of a high degree of aggregation. Considering the methods of spread within fields, this is fully explainable by the fact that new infections will occur very close to previously infected plants.

Accordingly, the spatial distribution of black rot lesions and infected plants describes a situation in which the infection of new plants is highly dependent upon their proximity to high bacterial densities; i.e., infected plants. Similar observations have already been reported for fungal diseases which are known to increase from foci (8) and van der Plank has discussed spatial distribution for some viral diseases and considered them to be aggregated in nature (7).

Of the several explanations proposed to account for aggregated distributions of organisms, method of reproduction probably best describes the reason for the observed patterns in black rot and may apply to plant diseases in general, although other explanations are possible (1, 6, 7, 13).

Sequential sampling is basically a procedure which considers the observed patterns of distribution of the object sampled. As applied in this study, it was necessary to sample only enough units to obtain an estimate of the mean with the desired precision. In making complete counts, an arbitrary number of plants were examined and compared with the number required for sequential sampling. It should be pointed out that methods are available which use other criteria to estimate required sample size (6) and that sequential sampling more commonly reduces expenditures of time and effort by about one-third.

Sampling methods, especially statistical ones, are greatly affected by spatial distribution. This paper does not explore this important area, but does help to point out that the true distribution can be of help in designing improved sampling and analysis procedures and that caution should be observed in assuming randomness when sampling pathogens or diseased plants. Spatial distribution can also provide or verify useful information on the etiology of diseases or their method of spread through a field or area. There are many observations pertaining to plant diseases which are known and accepted as valid, but as yet have not been quantified. Suitable techniques applicable to these problems may already exist in related fields of research seldom explored by plant pathologists.

LITERATURE CITED

1. BLISS, C. I. 1958. The analysis of insect counts as negative binomial distributions. Proc. 10th Intl. Congr. Entomol. 2:1015-1032.
2. BLISS, C. I., & R. A. FISHER. 1953. Fitting the negative binomial distribution to biological data and note on

- the efficient fitting of the negative binomial. *Biometrics* 9:176-200.
3. BLISS, C. I., & A. R. G. OWEN. 1958. Negative binomial distribution with a common k . *Biometrika* 45:37-58.
 4. GATES, C. E., & F. G. ETHRIDGE. 1972. A generalized set of discrete frequency distributions with FORTRAN program. *Math. Geol.* 4:1-24.
 5. KUNO, E. 1969. A new method of sequential sampling to obtain the population estimates with a fixed level of precision. *Res. Popul. Ecol.* 11:127-136.
 6. SOUTHWOOD, T. R. E. 1966. *Ecological methods*. Chapman-Hall, London. 391 pp.
 7. VAN DER PLANK, J. E. 1960. Analysis of epidemics. Vol. 3. p. 229-287. *In* J. G. Horsfall & A. E. Dimond [ed.]. *Plant pathology, an advanced treatise*. Academic Press, N.Y.
 8. VAN DER PLANK, J. E. 1963. *Plant diseases: epidemics and control*. Academic Press, N.Y. 349 pp.
 9. VAN DER PLANK, J. E. 1972. Basic principles of ecosystem analysis. p. 109-118. *In* *Pest control: strategies for the future*. National Academy of Sciences, Washington, D.C. 376 pp.
 10. WAGGONER, P. E., & J. G. HORSFALL. 1969. Epidem-A simulator of plant disease written for a computer. *Conn. Storrs Agric. Expt. Stn. Bull. No. 698*. 80 pp.
 11. WALKER, J. C. 1953. *Diseases of vegetable crops*. McGraw-Hill, N.Y. 525 pp.
 12. WATERS, W. E. 1955. Sequential sampling in forest insect surveys. *For. Sci.* 1:69-79.
 13. WATERS, W. E. 1959. A quantitative measure of aggregation in insects. *J. Econ. Entomol.* 52:1180-1184.