Field Runner: A Disease Incidence, Severity, and Spatial Pattern Assessment System

B. R. DELP, Former Graduate Research Fellow, L. J. STOWELL, Former Postgraduate Researcher, and J. J. MAROIS, Assistant Professor, Department of Plant Pathology, University of California, Davis 95616

ABSTRACT

Delp, B. R., Stowell, L. J., and Marois, J. J. 1986. Field runner: A disease incidence, severity, and spatial pattern assessment system. Plant Disease 70:954-957.

A computer software system titled Field Runner was developed to facilitate the assessment of plant disease incidence, severity, and spatial pattern in fields. Field Runner used the stratified random sampling design to provide an unbiased sample and a lower percent error of the disease incidence estimates than previously used sampling designs (diagonal, "X," or "W"). The computer directed the operator to each sample site, stored the data, and provided an immediate analysis. Analyses included 1) the estimated mean and variance of disease incidence, 2) the variance-to-mean ratio, 3) an estimate of the *k* parameter of the negative binomial distribution, 4) an estimate of Lloyd's indices of mean crowding and mean patchiness, and 5) the *Z*-score from an ordinary runs analysis. Fields could be assessed for severity of one disease or for incidence of one to several diseases simultaneously. Lettuce and alfalfa fields were sampled to test the performance of Field Runner under actual conditions. The incidences and aggregation indices for anthracnose and drop of lettuce and severity of alfalfa plant damage caused by alkali soil are reported. Lettuce anthracnose and alfalfa plant damage were aggregated in foci; lettuce drop occurred randomly in the field.

Knowledge of plant disease incidence (number of infected plants or plant units expressed as a percentage of the total number assessed [12]), severity, and spatial pattern is becoming increasingly important as the economics of agriculture require more critical decisions at all levels. Government, public, and private institutions use this information to evaluate their long-term research goals and resource allocations (4,12). Growers and agricultural advisors use it to make pest management decisions. It is an initial factor for an epidemiologist to study disease development. Disease incidence, severity, and spatial pattern depend on data obtained from field samples. The accuracy of these data, as well as the time and effort required to obtain them, is affected by the sampling technique used. Therefore, a desirable sampling technique would provide the most accurate data for a minimum associated cost.

Present address of first author: BASF Corporation, 100 Cherry Hill Road, Parsippany, NJ 07054; of second author: Mycogen Corp., 5451 Oberlin Dr., San Diego, CA 92121.

Use of trade or company names does not imply endorsement by the University of California of the products named or criticism of similar ones not mentioned.

Accepted for publication 21 May 1986 (submitted for electronic processing).

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.

MATERIALS AND METHODS

Field Runner, the computer software system described in this paper, simplifies the task of sampling fields. The system is designed to aid in the rapid and accurate assessment of disease incidence, severity, and spatial pattern in the field. Field Runner uses the stratified random sampling design (SRSD) with single-stage cluster sampling (5), in which the field is divided into equal-sized sectors and a randomly located sample is collected within each sector. Stratified random sampling provides accurate estimates of disease incidence (8), but it was too cumbersome in the past to be practical. The system, however, incorporates a field-portable microcomputer, and as a result, this sampling design can now be used more efficiently. The system prompts the operator for necessary information about the field and disease(s) to be sampled. It then directs the operator to each sample site, stores the data, and provides an immediate analysis. As a result, agricultural advisors and research scientists can rapidly collect important information about plant disease incidence, severity, and spatial pattern with a system that is efficient and achieves levels of accuracy not practically possible with conventional sampling systems.

Sampling strategy. Previously used sampling techniques included "W," "X," and diagonal designs across a portion or all of the field (2,13) (Fig. 1A–D). Usually, plants were evaluated at random intervals along these paths. Lin et al (13) and Basu et al (2) compared the relative accuracies of these sampling designs and determined that partial-field samples provided the least accurate estimates of disease. There was little difference in the accuracies of the designs across the whole field if the disease was randomly distributed; however, "X" and "W" designs were more accurate than the diagonal design if disease was aggregated. Maximum dispersal of sample sites along the sampling design was the most important factor to obtain accurate estimates of disease incidence if disease was aggregated (13).

In the SRSD, fields are divided into sectors of equal size (stratification) and plants are evaluated at a random location within each sector (Fig. 1E) (5). Thus, sample sites are distributed thoroughout the field without bias to any section of the field. Preliminary results from application of the SRSD to plant diseases have been reported (7). This sampling technique was tested and compared with the "W" and diagonal designs for accuracy of estimating the mean disease incidence for the entire field (8). The SRSD was more accurate than the "W" or diagonal designs if sample intensity was greater than 0.2%. In addition to increased accuracy of the estimate of the mean, stratified random sampling also provided

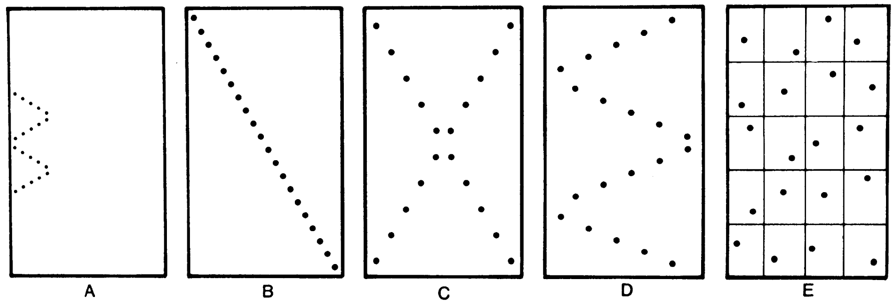

Fig. 1. Designs used to sample fields for disease: **(A)** partial field "W," **(B)** diagonal, **(C)** "X," **(D)** "W," and **(E)** stratified random. Points represent sampling sites.

data for analysis of among-sector variance to estimate the degree of disease aggregation.

Single-stage cluster sampling is an extension of the SRSD, in which the sample unit consists of a group or cluster of smaller units (areas, individual plants, or leaves) (5). Sample units are located randomly within each sector as described. Each sample unit is composed of a "cluster" of adjacent plants, which will be referred to as a transect because of the linear arrangement of the plants (Fig. 2). The sampling system implemented both SRSD and single-stage cluster sampling to provide accurate estimates of disease incidence, severity, and spatial pattern in the field.

Sampling system. Field Runner is a computer software system developed by the authors to make the SRSD more feasible. The software was written in MBASIC and developed for the Epson HX-20 (Epson America, Inc., Torrance, CA). This microcomputer was chosen because it is light enough to carry through the field and it has a video screen to display prompts for the operator, a printer to provide an immediate copy of the analysis, a tape drive to store data for later use, and an internal clock for elapsed time calculations. A manual was written to explain operation, sampling strategy, and analyses. The program source code was included in the manual to aid programmers with modifications.

Field layout. Field Runner requires field dimensions and plant spacing to divide the field into uniform sectors. This information is easily obtained by calibrating the operator's pace (as prompted by the system) and walking the width and length of the field. The direction of water flow in the furrows is required for orientaton with respect to subsequent directions provided by the system. The operator also specifies the transect size (number of plants sampled per sector) and the sample intensity (the proportion of the entire population that the operator wishes to evaluate). The number of sectors in the field is determined by the system as a function of the total number of plants in the field (calculated from field dimensions and plant spacing), the transect size, and the sample intensity.

Data collection. Field Runner directs the operator to each randomly located sample site with a series of instructions, e.g., FACE 90 DEGREES, WALK 29 PACES, FACE 180 DEGREES, WALK 7 BEDS. Once at a sample site, the operator is prompted to enter data for each plant in the transect and is then directed to the next site. One such path is illustrated in Figure 2. Plant evaluations are entered as codes that represent the condition of the plant. These codes are defined by the operator at the beginning of each sampling session. Plants can be evaluated for presence or absence (incidence) of disease or for disease severity. Multiple diseases can be rated simultaneously for incidence; evaluations for severity are limited to one disease at a time (Table 1).

Data analysis. Data can be analyzed immediately by Field Runner, or the stored data can be transferred to a microcomputer for more detailed analysis by a related program, Data Runner. Field Runner and Data Runner were designed to answer two important questions: 1) What is the incidence or severity of the disease in the field? 2) Is the spatial pattern of the disease uniform, random, or aggregated within the field and within the transects? The system provides five analyses of disease in a field: 1) esimates

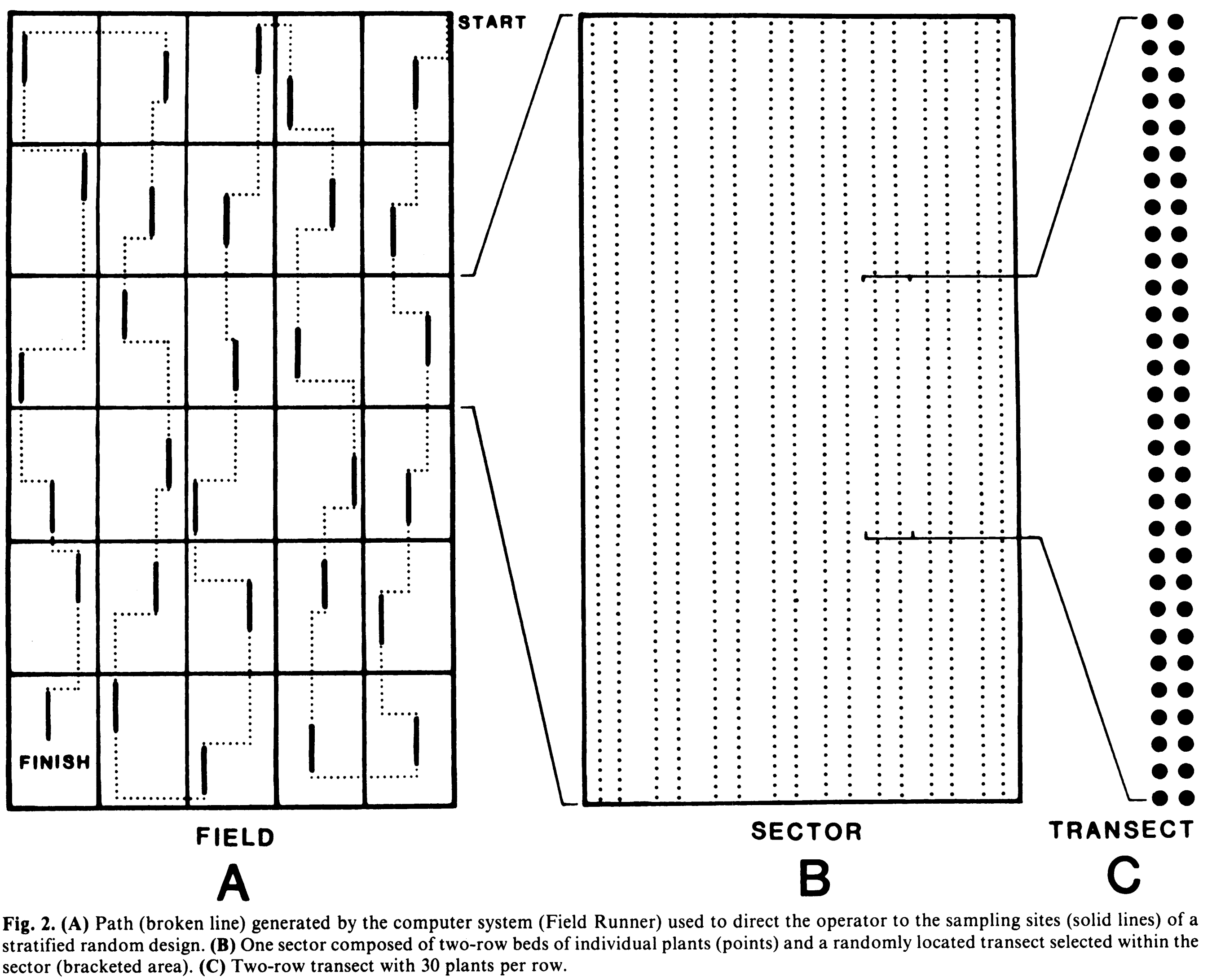

Fig. 2. **(A)** Path (broken line) generated by the computer system (Field Runner) used to direct the operator to the sampling sites (solid lines) of a stratified random design. **(B)** One sector composed of two-row beds of individual plants (points) and a randomly located transect selected within the sector (bracketed area). **(C)** Two-row transect with 30 plants per row.

of the mean and variance of disease incidence, 2) the variance-to-mean ratio (5), 3) the k parameter of the negative binomial distribution (1), 4) Lloyd's indices of mean crowding and mean patchiness (14), and 5) an ordinary runs analysis (10,15). Analyses 2–4 indicate degree of within-field aggregation (among-transect variance), whereas analysis 5 relates to aggregation within transects.

The variance-to-mean ratio, where the variance and mean are estimated from the number of diseased plants in each transect, is an indicator of disease spatial pattern. A variance-to-mean ratio of 1 indicates that disease is randomly distributed. When disease is aggregated, the variance-to-mean ratio is greater than 1.

The negative binomial distribution has been used to describe data from a number of biological systems (3,17). The k parameter of this distribution describes the relative aggregation within the population. The negative binomial distribution converges to the Poisson (random) distribution if k is large and the mean is moderate (16). The value of k is not only dependent on the degree of aggregation but also on the population mean; therefore, k alone will not indicate the level of aggregation of the population. If the means of two populations are the same, however, the population with the smaller k is more aggregated than the population with the larger k. Relative aggregation is not as easily determined if the population means are different. The k parameter of the negative binomial distribution is calculated using the method of moments as described by Anscombe (1).

Lloyd's index of mean crowding is another measure of aggregation:

$$\text{Mean crowding} = \bar{x} + (s^2/\bar{x}) - 1,$$

where x is the estimated mean disease incidence and s^2 is the variance of the observation. Lloyd (14) states, "Any two populations are considered equivalent, in respect of the biological effects of crowding, if they have the same mean crowding." Mean crowding is a relative statistic that is dependent on the mean. When the distribution is random, the mean is equal to Lloyd's index of mean crowding. When diseased plants are more aggregated, mean crowding is greater than the mean.

Lloyd's index of mean patchiness:

$$\text{Mean patchiness} = [\bar{x} + (s^2/\bar{x}) - 1]/\bar{x}$$

is a measure of aggregation closely related to the k parameter of the negative binomial and is equal to mean crowding divided by the mean. Mean patchiness is a measure of "how many times as 'crowded' an individual is, on the average, as it would have to be if the sample population had a random distribution" (14). For example, a patchiness value of 3 implies that the diseased plants are three times more crowded than they would be if they were randomly distributed. Unlike mean crowding values, mean patchiness values from fields with different disease incidences can be compared. Therefore, only values for mean patchiness are reported in the subsequent portions of this paper.

The sampling technique used with Field Runner provides for a sequential series of codes to represent the condition of each plant in a transect. Ordinary runs analysis (10,15) can be performed on data in this form if 20 or more plants per transect are evaluated (15). The analysis produces a Z-score, which is used to accept or reject the hypothesis of randomness. If the Z-score is less than −1.64, there is a 95% chance that the spatial pattern of diseased plants in the transect is not random. The alternative hypothesis, that the disease is aggregated within the transect, is accepted. In some instances, this may be interpreted as an indication of plant-to-plant spread of the pathogen (15).

Table 1. Examples of disease rating codes used with computer system (Field Runner)

Rating code[a]	Incidence[b]	Severity[c] (%)
0	Healthy	0
1	Anthracnose	0–3
2	Drop	4–10
3	Both	11–25
4	···	26–50
5	···	51–75
6	···	76–90
7	···	91–97
8	···	98–100

[a] Codes used to enter data.
[b] Code set for simultaneous rating of the incidences of two diseases, anthracnose and drop of lettuce.
[c] Code set for rating the severity of one disease.

Data Runner. In addition to the analyses performed by Field Runner, Data Runner provides the several additional options:

1. Compositing codes is a useful procedure if a field has been rated for more than one disease. All codes relating to each disease are combined into a single code for subsequent analysis. For example, codes 1 and 3 from Table 1 (incidence) are combined if an analysis of anthracnose is desired.

2. Data codes can be converted to the percentages of severity that they represent (Table 1 [severity]). Percent loss is calculated from the frequency and proportion of each code.

3. Field data can be plotted to produce a "map" of the frequency of a particular code.

Field testing. Lettuce and alfalfa fields were sampled to test the performance of the Field Runner system. Lettuce is a row crop with discrete plants, whereas alfalfa is a forage crop with essentially continuous plant cover.

Four lettuce fields in the Salinas Valley of California were sampled for incidence of both anthracnose and drop caused by *Marssonina panattoniana* (Berl.) Magn. and *Sclerotinia minor* Jagger, respectively. The set of codes used is listed in Table 1 (incidence).

M. panattoniana apparently is a soilborne pathogen favored by cool, wet spring conditions (6). Spores of the fungus are borne in a slime mass and are spread locally (2–3 m) by splashing rain (B. R. Delp, *unpublished*). Based on this biology, anthracnose would be expected to occur in expanding foci throughout a field.

RESULTS

Incidence of anthracnose in sampled fields ranged from 3.7 to 78.8% (Table 2). The disease was aggregated within three of the fields as determined by all indices of within-field aggregation; the variance-to-mean ratios and Lloyd's mean patchiness indices were greater than 1 and the negative binomial k was small. The indices of disease spatial pattern, however, approached values indicating randomness for the Bluff field. This was attributed to the high percentage of disease, i.e., a field with 100% diseased plants has a uniform disease spatial pattern. Therefore, at 79% disease incidence, the indices of aggregation in the field approached uniformity. This resulted in an indication of near

Table 2. Lettuce anthracnose data collected and analyzed with computer system (Field Runner)

Field	Disease incidence (%)	Variance/mean[a,b]	Patchiness[b,c]	k[d]	Z-score[e]
Zabala	3.7	26.7	12.6	0.09	−3.40
Nixon	5.4	15.6	5.4	0.22	−1.99
Giberson	23.4	19.8	2.3	0.75	−2.64
Bluff	78.8	5.1	1.1	11.40	−2.86

[a] Ratio of within-field variance to mean number of infected plants per treatment.
[b] A value of 1 implies randomness and values >1 imply aggregation of disease within fields.
[c] Lloyd's index of patchiness.
[d] The k parameter of the negative bionomial distribution, which decreases as within-field aggregation increases.
[e] Z-score from ordinary runs analysis (values >−1.64 denote randomness and values <−1.64 denote aggregation of disease within transects).

randomness, which was intermediate between aggregation and uniformity. All Z-scores were less than −1.64, which indicated that there was within-row (within-transect) aggregation. These types of aggregation, foci (within-field aggregation), which were expanding because of plant-to-plant spread (within-row aggregation), were consistent with those to expected based on the biology of *M. panattoniana*.

S. minor is a soilborne sclerotial fungus that attacks plants if they are within a competence zone of several centimeters (11) and shows very little subsequent plant-to-plant spread. The incidence of drop ranged from 2.2 to 9.6% (Table 3). Indices of within-field aggregation indicated that disease occurred randomly in all fields. The Z-scores indicated that the disease occurred randomly within the rows also. This was consistent with previous reports on the occurence of lettuce drop (9).

Table 3. Lettuce drop data collected and analyzed with computer system (Field Runner)

Field	Disease incidence (%)	Variance/mean[a,b]	Patchiness[b,c]	k[d]	Z-Score[e]
Zabala	1.1	1.1	1.1	7.4	...[f]
Nixon	4.9	1.5	1.2	6.4	...
Giberson	6.7	1.2	1.1	20.0	−0.7
Bluff	9.6	3.1	1.4	2.8	0.6

[a] Ratio of within-field variance to mean number of infected plants per transect.
[b] A value of 1 implies randomness and values >1 imply aggregation of disease within fields.
[c] Lloyd's index of patchiness.
[d] The k parameter of the negative binomial distribution, which decreases as within field aggregation increases.
[e] Z-score from ordinary runs analysis (values >−1.64 denote randomness and values <−1.64 denote aggregation of disease within transects).
[f] Disease incidence was too low to calculate a reliable Z-score.

Alfalfa fields posed an interesting navigational problem for the system in that there were no rows or plants to count; however, paces were substituted for these. That is, the operator walked the indicated number of paces in one direction, turned to the right or left, and walked the indicated number of paces (not rows or beds) in that direction. A transect consisting of square areas of defined dimensions rather than individual plants was evaluated. This necessitated the use of a severity scale.

An alfalfa field was evaluated for plant damage caused by alkali soil. One-pace-square areas were rated for damage severity. Data were entered as codes listed in Table 1 (severity). Codes were converted to percent damage. The average plant damage in the field was 11.4%, the variance-to-mean ratio was 10.3, and the Z-score was −4.3, which indicated that the damaged areas were aggregated within the field and within the transect.

DISCUSSION

The Field Runner system provided rapid estimates of disease incidence, severity, and spatial pattern. Results obtained from sampling diseased fields were consistent with the types of aggregation one would expect from the pathogens' respective biological mechanisms. The system also was useful to assess damage from abiotic causes.

The system has a broad range of potential applications. It could aid in the development of integrated pest management programs, evaluation of weed populations in agricultural and natural ecosystems, location of soil samples, and preharvest yield evaluations. In general, it could be useful for most studies that require unbiased estimates from field samples. This flexibility makes the system a powerful information gathering tool for researchers, agricultural consultants, disease scouts, and many others.

LITERATURE CITED

1. Anscombe, F. J. 1950. Sampling theory of the negative binomial and logarithmic series distribution. Biometrika 37:358-382.
2. Basu, P. K., Lin, C. S., and Binns, M. R. 1977. A comparison of sampling methods for surveying alfalfa foliage diseases. Can. J. Plant Sci. 57:1091-1097.
3. Campbell, C. L., and Pennypacker, S. P. 1980. Distribution of hypocotyl rot caused in snap bean by *Rhizoctonia solani*. Phytopathology 70:521-525.
4. Chiarappa, L., Moore, F. J., and Strickland, A. H. 1971. Crop Loss Assessment Methods. Alden and Mowbray Ltd. Oxford, England.
5. Cochran, W. G. 1977. Sampling Techniques. 3rd ed. John Wiley & Sons, New York. 428 pp.
6. Couch, H. B., and Grogan, R. G. 1955. Etiology of lettuce anthracnose and host range of the pathogen. Phytopathology 45:375-380.
7. Delp, B. R., Stowell. L. J., and Grogan. R. G. 1984. Distribution of lettuce anthracnose in the field. (Abstr.) Phytopathology 74:871.
8. Delp, B. R., Stowell, L. J., and Marois, J. J. 1986. Evaluation of field sampling techniques for accuracy of disease incidence estimation. Phytopathology 76:In press.
9. Dillard, H. R., and Grogan, R. G. 1985. Relationship between sclerotial spatial pattern and density of *Sclerotinia minor* and the incidence of lettuce drop. Phytopathology 75:90-94.
10. Gibbons. J. D. 1971. Nonparametric Statistical Inference. McGraw-Hill, New York. 306 pp.
11. Grogan, R. G., Sall, M. A., and Punja, Z. K. 1980. Concepts for modeling root infection by soilborne fungi. Phytopathology 70:361-363.
12. James, W. C. 1974. Assessment of plant diseases and losses. Annu. Rev. Phytopathol. 12:27-48.
13. Lin, C. S., Poushinsky, G., and Mauer, M. 1979. An examination of five sampling methods under random and clustered disease distributions using simulation. Can. J. Plant Sci. 59:121-130.
14. Lloyd. M. 1967. Mean crowding. J. Anim. Ecol. 36:1-30.
15. Madden, L. V., Louie. R., Abt. J. J., and Knoke, J. K. 1982. Evaluations of tests for randomness of infected plants. Phytopathology 72:195-198.
16. Pollard, J. H. 1979. Numerical and Statistical Techniques. Cambridge University Press, New York. 349 pp.
17. Shew, B. B., Beute, M. K., and Campbell, C. L. 1984. Spatial pattern of southern stem rot caused by *Sclerotium rolfsii* in six North Carolina peanut fields. Phytopathology 74:730-735.