JOURNAL OF NEMATOLOGY

VOLUME 17 JULY 1985 NUMBER 3

Journal of Nematology 17(3):245–251. 1985.

Overestimation of Yield Loss of Tobacco Caused by the Aggregated Spatial Pattern of *Meloidogyne incognita*[1]

J. P. Noe and K. R. Barker[2]

Abstract: Overestimation of yield loss caused by *Meloidogyne incognita* on tobacco was calculated as a function of the statistical frequency distribution of sample counts. Sampling frequency distributions were described by a negative binomial model, with parameter k, and the resulting probability generating function was used to calculate discrete damage probabilities. Negative binomial damage predictions were compared to mean-density estimates of damage. Predictions based on mean density alone overestimate yield loss by values ranging from 300% at a k of 0.1 to less than 10% at a k of 1.0. Damage overestimation was described as an exponential function of k and mean density. Preplant sampling data for *M. incognita* were used to derive a linear model for the estimation of k from mean density, allowing the calculation of yield-loss overestimation based on one parameter, the field mean density. Overestimation of damage ranged from 288% at a density of 50 juveniles/500 cm^3 soil, to 5% at a density of 1,000 juveniles/500 cm^3 soil.

Key words: crop-loss estimation, negative binomial, nematode damage functions.

Management decisions aimed at minimizing losses caused by plant-parasitic nematodes are based on estimates of nematode population density and resulting crop-loss predictions. The validity of economic crop-management decisions is limited by the accuracy and precision of the information on which they are based. Accuracy, or bias, is a measure of how close the expected value of an estimator is to the true population parameter, and precision is a measure of the repeatability, or agreement in estimates obtained from independent samples (7). Precision of nematode density estimates derived from field samples is affected by spatial patterns of plant-parasitic nematodes (2,15), and techniques are available to estimate sampling precision as a function of distributional parameters (8,13–15,21). The precision of crop-loss estimates also is influenced by spatial patterns, as well as other sampling factors. Techniques for estimation of the precision of nematode damage functions have been presented by Ferris (10).

Precision of estimates may be calculated to provide an indication of the confidence one may place in a sample parameter, and precision may be increased by improved experimental and sampling techniques. The accuracy of estimates also may be increased by improved techniques. Additionally, however, inaccuracy (i.e., predictable bias) may be compensated for by knowledge of the source and amount of bias. Seinhorst (20) has demonstrated that aggregated spatial patterns of plant-parasitic nematodes caused consistent overestimation of yield losses where only mean densities were used in crop-loss models. In that study, yield losses attributable to nematode densities in more than 100 locations in a field were summed to derive an estimate of total yield loss in the field. This total then was compared to an estimate obtained by using the arithmetic mean density of the individual samples in the same yield loss model. In a different approach, Perry (17) simulated the effects of an aggregated spatial pattern on crop root damage. It was shown that serious errors could occur in estimates of yield loss if only the field mean density was used. Perry suggested that samples be taken from different locations with-

---

Received for publication 23 March 1984.

[1] Paper No. 9229 of the Journal Series of the North Carolina Agricultural Research Service, Raleigh, North Carolina. This study was supported by U.S. Agency for International Development Contract AID-ta-C-1234 to J. N. Sasser.

[2] Research Associate and Professor, Department of Plant Pathology, North Carolina State University, Raleigh, NC 27695-7616.

The JOURNAL OF NEMATOLOGY for April (17:81–244) was issued 22 April 1985.

in a field, and that the nematode density at each location be input to a root-damage model.

This paper presents an approach to estimating inaccuracy in crop-loss predictions resulting from aggregated nematode spatial patterns with a mathematical analysis of frequency distributions and damage functions. As an example, the response of tobacco (*Nicotiana tabacum* L.) to *Meloidogyne incognita* (Kofoid & White) Chitwood will be analyzed.

Rationale for the mathematical approach: Least-squares parameter estimation has been used to quantify the relationship between nematode density and yield loss (5). This method can be used to fit a number of different curves, with each function specifying a different crop–nematode relationship. Population density–yield loss relationships then can be applied to nematode assay results from a particular field, in order to project damage caused by plant-parasitic nematodes (4,9). The nematode density reported in a nematode assay, however, represents the mean number of nematodes in a composite sample of individual cores taken from a number of locations within a field (3). If the quantitative relationships were derived in greenhouse, microplot or field plot studies with artificially uniform patterns of nematode inoculum, then for the yield-loss model to apply to field assay results one of two assumptions must be met:

1. The nematodes must be distributed uniformly throughout the field, so that each host plant will encounter the mean density of parasites.
2. If the nematodes do not have a uniform spatial distribution, then each nematode must cause the same amount of yield loss regardless of how many other nematodes are nearby.

These assumptions are, however, seldom fulfilled under field conditions. Plant-parasitic nematodes usually have an aggregated spatial pattern (2,12,15,16). Most of the nematodes occur in clusters of relatively high population densities, whereas large areas of the field have low population densities. Additionally, where nematodes occur in clusters the potential yield loss from an individual nematode is decreased. There are a limited number of feeding sites on a root, and plant-parasitic nematodes are restricted in their ability to move out of clusters and attack other hosts. There is a point beyond which additional nematodes cause no further damage (minimum yield, or death of host), and the damage per additional parasite becomes zero. Decreasing damage per parasite is inherent as density increases in log-linear or exponential yield-loss models (4,19). The slopes of yield-loss functions eventually decrease, and the curves become asymptotic with some value (minimum yield).

Concepts of decreasing damage per parasite can be quantified readily in existing crop-loss models. In order to compensate for this effect on a field basis, the proportion of a field likely to contain a given nematode density must be calculated. One approach to this calculation has been suggested (17,20), where data from discrete sampling locations were input to crop-loss functions. This is essentially a process of using field-specific sampling frequency distributions to estimate probability densities. However, frequency distributions derived from nematode field-sampling data often can be described mathematically by the negative binomial probability distribution (12,15,16).

The negative binomial distribution is positively skewed, which means that values with a density less than the mean are more frequent than values with a density greater than the mean. The negative binomial is described by the mean and a parameter k. The parameter k may be used as an index of dispersion to indicate the degree of aggregation in a field population (1,22); smaller k-values are associated with a more aggregated population and a more positively skewed frequency distribution. Negative binomial k-values are a function of mean density for most biological populations (22).

The spatial patterns of plant-parasitic nematodes should be considered and incorporated into yield-loss models in order to accurately predict potential crop loss. The bias induced by spatial patterns may be estimated through a combined analysis of crop-loss functions and frequency distributions.

MATERIALS AND METHODS

A coefficient of yield-loss overestimation was calculated as the ratio of two separately derived estimates of damage. The numerator of the ratio was a damage estimate based on arithmetic mean density alone. The denominator was an estimate based on frequency probabilities for population density categories calculated from a negative binomial model. Discrete population density probabilities were computed iteratively using a FORTRAN program. The zero probability class P(0) was calculated as:

$$P(0) = \left[1 + \frac{\text{mean}}{k}\right]^{-k}$$

Subsequent probability densities were calculated from the formula:

$$P(I) = P(I - 1)\left[\frac{k + I + 1}{I}\right] \times \left[\frac{\text{mean}}{\text{mean} + k}\right]$$

where P(I) represented the probability, under a negative binomial model, of a density of "I" nematodes occurring for a given combination of mean and k. This probability was multiplied by the yield loss caused by a density of "I" nematodes D(I) using a percent yield-loss damage function for *M. incognita* on tobacco (5):

$$D(I) = 0.1 \log_{10}(I)$$

The resulting yield-loss components were summed over all densities until the discrete probability of occurrence was less than 1.0×10^{-37}. The remaining probability in the tail of the negative binomial distribution was multiplied by the maximum yield loss expected from the damage function. This term was added to the negative binomial estimate of damage so that the probability densities summed to one.

The yield loss expected from a given population mean density assuming a uniform spatial pattern for *M. incognita* was computed from the damage function as:

$$D(\text{mean}) = 0.1 \log_{10}(\text{mean})$$

This estimate can be derived from a process similar to the negative binomial algorithm. Since a uniform spatial pattern is assumed, all locations in a field would have a nematode density equal to the mean. The probability of any density other than the mean occurring is zero, whereas the probability of the mean occurring is one. The summation takes the form:

$$\begin{aligned} P(n)D(n) &= P(0)D(0) + P(1)D(1) + \ldots \\ &\quad + P(\text{mean})D(\text{mean}) + \ldots \\ &\quad + P(n)D(n) \\ &= 0 + 0 + \ldots + (1)D(\text{mean}) \\ &\quad + \ldots + 0 \\ &= D(\text{mean}) \end{aligned}$$

since all probability terms other than the mean are zero.

The coefficient of damage overestimation (COD) was calculated as the ratio of the estimate of damage caused by a given mean density D(mean), where only the arithmetic mean density was input to the damage function, to the estimate from the summation of individual negative binomial probability terms:

$$\text{COD} = \frac{D(\text{mean})}{\sum P(I)D(I)} \times 100$$

A coefficient of 100 indicated no overestimation, whereas a coefficient of 200 indicated a 200 percent overestimation of potential yield loss. A ratio statistic was selected to indicate the degree of error, since it remains the same regardless of the actual yield loss levels. The actual levels of yield loss were a function of the nematode–host system selected as an example, and they would vary widely for different nematode–host combinations.

Data on mean density and k-values for *M. incognita* were collected from 10 tobacco fields. Soil samples were collected in a systematic grid sampling plan (ten 2.5-cm-d cores per quadrat, 64–100 quadrats per field) prior to planting in the spring. *M. incognita* juveniles were extracted from the soil by elutriation and centrifugation (6). A negative binomial distribution was fitted to frequency counts of *M. incognita* with a FORTRAN program (11), deriving maximum likelihood estimates of the parameter k.

Preplant mean densities and k-values from the 10 fields were used to establish a range of values which would be of practical significance in projecting yield losses caused by *M. incognita* on tobacco. A COD was calculated for combinations of mean density and k-values in the prescribed ranges. The calculated coefficient of damage overestimation over varying mean and k-values formed a response surface. An exponential relationship was suggested as:

$$\text{COD} = 100 + aY^{bk} + cZ^{\text{mean}}$$

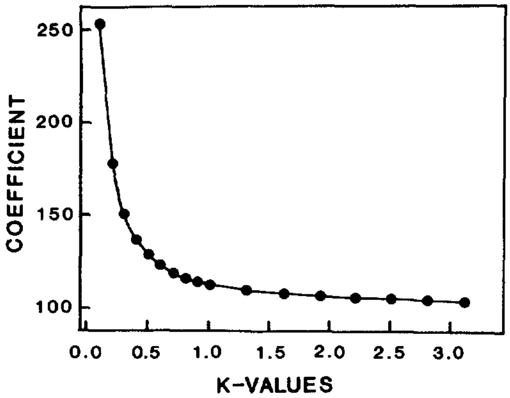

FIG. 1. Coefficient of overestimation of percent damage caused by *Meloidogyne incognita* on *Nicotiana tabacum* versus negative binomial k-values. Coefficient was calculated as the ratio of a damage estimate from population mean density to an estimate from the summation of individual negative binomial damage components multiplied by 100. Coefficient of 100 indicates no overestimation.

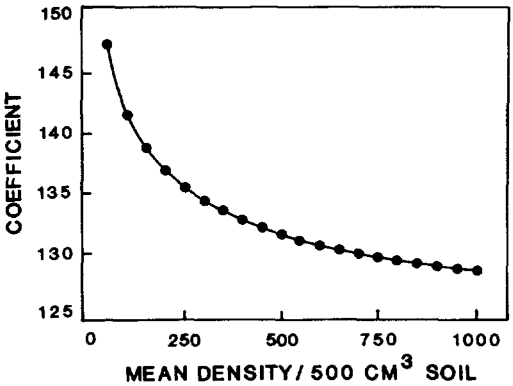

FIG. 2. Coefficient of overestimation of percent damage caused by *Meloidogyne incognita* on *Nicotiana tabacum* versus nematode population mean density. Coefficient was calculated as the ratio of a damage estimate from population mean density to an estimate from the summation of individual negative binomial damage components multiplied by 100. Coefficient of 100 indicates no overestimation.

where a, b, c, Y, and Z were parameters estimated by fitting the response surface with the NLIN procedure of the Statistical Analysis System (SAS) (18). Field data also were used to quantify the relationship of k to mean density for the *M. incognita*–tobacco system. This relationship was used to calculate the COD as a function of mean density alone, using predicted k-values. A mathematical formula has been derived from the negative binomial distribution (8), relating k to the mean and variance of a population. However, to use the equation an estimate of population variance is necessary, and this would require that multiple samples be assayed from a single field. The models in this study were derived to illustrate a method of computing the effect of spatial dispersion on yield-loss estimates with a single estimate of field mean density as input.

RESULTS

Preplant mean densities of *M. incognita* ranged from 23 to 1,890 juveniles/500 cm^3 soil, and k-values ranged from 0.09 to 2.6 in the 10 tobacco fields. Negative binomial distributions were fitted to frequency counts from all fields, with a chi-square test for goodness-of-fit not significant at the 0.05 probability level.

A coefficient of yield-loss overestimation calculated for k-values ranging from 0.1 to 3.1, at a mean density of 200 (Fig. 1), decreased as k increased, from 250% overestimation at a k of 0.1 to less than 5% overestimation at a k of 3.1. The coefficient decreased very rapidly from a k of 0.1 and was less than 10% for k-values greater than 1.0. The accuracy of yield-loss estimates increased rapidly with increasing k-values.

The coefficient was next calculated for mean densities ranging from 50 to 1,000 juveniles/500 cm^3 soil, at a constant k-value of 0.4 (Fig. 2). Overestimation decreased as the mean increased, from a value of 47% at a mean of 50 to 28% at a mean of 1,000. Accuracy of yield-loss estimates increased with increasing mean densities.

Yield-loss overestimation was more a function of changes in k than changes in mean density (Fig. 3). The slopes of response curves along the mean density axis are not as steep as the slopes along the k axis. Responses to changes in mean density are larger at smaller k-values, indicating interaction between these two negative binomial parameters.

The proposed exponential function was adequate to describe the response of yield-loss overestimation to varying mean and k-values (Table 1). The parameterized formula fits the computed data well and allows

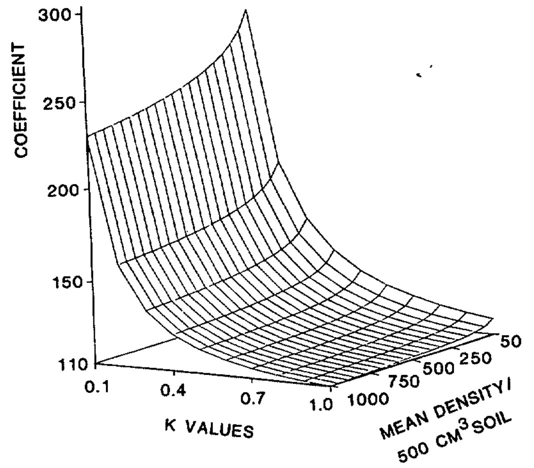

FIG. 3. Response surface of coefficient of overestimation of percent damage caused by *Meloidogyne incognita* on *Nicotiana tabacum* versus population mean density and k-values. Coefficient was calculated as the ratio of a damage estimate from population mean density to an estimate from the summation of individual negative binomial damage components multiplied by 100. Coefficient of 100 indicates no overestimation.

TABLE 1. Nonlinear least squares fit of model for response of yield-loss overestimation to mean population density and negative binomial k-value.

Parameter*	Estimate	Standard error
a	261	4.1
b	97	0.6
c	24	0.8
Y	0.93	0.001
Z	0.99888	0.00008

Least squares summary statistics:

Source	df	Sum of squares	Mean square
Regression	5	4,201,468	840,294
Residual	195	3,597	18
Total	200	4,205,065	

* Model parameters: COD = $100 + aY^{bk} + cZ^{m}$ where COD is the coefficient of damage overestimation, k is a parameter from the negative binomial distribution, m is mean population density/500 cm^3 soil.

the prediction of a coefficient of overestimation without going through the iterative algorithm for the *M. incognita*–tobacco system.

The relationship of k to mean density in the 10 selected tobacco fields was described by the linear model:

$$k = 0.0018(\text{mean})$$

with an *r*-squared of 0.82 and a standard error about the estimate of 0.0003. This simple model was adequate to describe the relationship of k to mean density within the range of values sampled and enabled the computation of yield-loss overestimation solely as a function of mean density.

The estimated yield loss for a mean density of 50 juveniles/500 cm^3 soil was 17% (Table 2). The estimate dropped to 6%, however, where yield loss was calculated with a negative binomial distribution, resulting in a difference of 11%. The estimated yield loss at a density of 400 was 26% where only the arithmetic mean was input, and 22% using negative binomial damage components. The resulting difference of only 4% demonstrated the decrease in overestimation as mean and k-values increased. Damage caused by *M. incognita* on tobacco would be overestimated by 288% at a mean density of 50 juveniles/500 cm^3 soil (Fig. 4), which corresponded to the 11% disparity in yield-loss estimates reported in Table 2 (17%/6%). The overestimation decreased rapidly as the mean increased and was less than 10% for nematode population densities greater than 400.

DISCUSSION

This study has presented a method for using information on the spatial patterns of nematodes to improve the accuracy of yield-loss estimation. Yield-loss overestimation caused by spatial aggregation was a significant factor in analysis of data for *M. incognita* on tobacco and could result in erroneous management recommendations. With a preplant mean density of 50 juveniles/500 cm^3 soil, the difference between the compensated (6%) (Table 2) and uncompensated (17%) estimates of yield loss could easily be across the economic threshold. Treatment could be recommended, based on a standard nematode assay, when in fact it would result in a net economic loss. Actual values for the yield-loss estimates depend on the dynamics of each nematode–host system, but the ratio should remain stable for a given probability distribution.

A negative binomial model was used to describe the frequency probabilities of nematode densities. Other discrete frequency distributions have been fitted to

TABLE 2. Two estimates of percent yield loss caused by *Meloidogyne incognita* on *Nicotiana tabacum.* One estimate was based on mean population density alone, and the other was adjusted for frequency distribution by a summation of damage components from negative binomial frequency probabilities (negative binomial parameter k predicted from mean density).

Mean number of *M. incognita* juveniles/500 cm^3 soil	Predicted k-value*	Estimate of percent yield loss	
		Mean density based	Frequency distribution adjusted
50	0.09	17	6
100	0.18	20	10
150	0.27	22	14
200	0.37	23	16
300	0.55	25	20
400	0.73	26	22
500	0.91	27	24
600	1.10	28	26
700	1.28	28	27
800	1.46	29	27
900	1.64	30	28
1,000	1.83	30	29

* k = 0.0018 (mean).

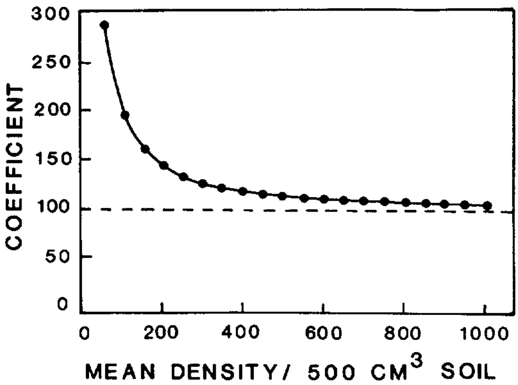

FIG. 4. Coefficient of overestimation of percent damage caused by *Meloidogyne incognita* on *Nicotiana tabacum* versus population mean density. Predicted k-values from linear least-squares model of relationship between preplant population densities and k-values of *M. incognita* in 10 fields. Coefficient was calculated as the ratio of a damage estimate from population mean density to an estimate from the summation of individual negative binomial damage components multiplied by 100. Coefficient of 100 indicates no overestimation.

frequency counts of plant-parasitic nematodes, including the Neyman type-A (15,16) and the Thomas double Poisson (16). If another distribution were fitted to the same data as was the negative binomial, and both fits were adequate, then the alternative distribution would be generating very similar probabilities. If the two distributions were generating essentially the same probabilities, then the summation of damage components would be the same, regardless of the distribution used in the algorithm. Thus, the effects of substituting another discrete distribution for the negative binomial in the algorithm would be minimal.

The degree of error in yield-loss estimation induced by the frequency distribution is largely a function of its skewness, or asymmetry. The increased percentage of counts below the mean and the relatively few very large counts cause overestimation. The skewness of the negative binomial is related to k. Thus, the coefficient responded strongly to changes in k over a critical region. As k increased beyond 3.0, the coefficient rapidly became asymptotic and the skewness was minimal.

A computational advantage of using the negative binomial derives from the fact that it is specified by mean density and only one other parameter, k. This parameter can be used as an index of dispersion, which provides a method to relate the frequency distribution back to spatial pattern. Also, for some biological populations, including *M. incognita* in tobacco fields, k is a quantifiable function of mean density. For advisory purposes, only mean density need be assessed, as is commonly the practice (3). Mean density can be used to predict a k-value, and then the mean and k can be used in an appropriate response model to predict a level of damage overestimation.

The degree of damage overestimation is an interaction between the shape of the probability distribution and the yield-loss model. The steepness of decline in yield and the population density at which the curve becomes asymptotic determine the effect of differing probability models on predicted damage. A relatively simple log-linear yield-loss model was selected to illustrate the techniques in this study. Other models could be substituted in the algorithm for damage estimation. Seinhorst's explanatory model (19) incorporates the concept of a damage tolerance level into the yield-loss relationship. Since no damage is predicted for population densities below the tolerance level, there can be no overestimation of damage at densities below this level. However, use of the Seinhorst model may result in an underestimation of damage. Where nematodes occur in clusters, host plants in certain parts of a field may sustain damage even though the overall mean density is below the tolerance level.

The form of a yield-loss model is important to overestimation, but the relative level of damage is not. The coefficient is calculated as a ratio. Changes in scale for the damage function will have no effect on the calculated overestimation, but may affect the practical significance of the results. The difference between 1% yield loss and 2% yield loss is of little economic importance, even though it represents a 200% overestimation. In the example given in this study, the actual damage estimates and differences in the estimates were significant, but other nematode–host combinations would need to be analyzed individually to measure the importance of spatial pattern compensation to those systems. After selection of an appropriate damage function and probability distribution, the derived response model can be used to predict a level of yield-loss overestimation. This technique should provide more accurate damage estimates and allow improved economic decision making.

LITERATURE CITED

1. Anscombe, F. J. 1950. Sampling theory of the negative binomial and logarithmic series distribution. Biometrika 37:358–382.

2. Barker, K. R., and C. L. Campbell. 1981. Sampling nematode populations. Pp. 451–472 *in* B. M. Zuckerman and R. A. Rohde, eds. Plant parasitic nematodes, vol. 3. New York: Academic Press.

3. Barker, K. R., and C. J. Nusbaum. 1971. Diagnostic and advisory programs. Pp. 231–301 *in* B. M. Zuckerman, W. F. Mai, and R. A. Rohde, eds. Plant parasitic nematodes, vol. 1. New York: Academic Press.

4. Barker, K. R., and T. H. A. Olthoff. 1976. Relationships between nematode population densities and crop responses. Annual Review of Phytopathology 14:327–353.

5. Barker, K. R., F. A. Todd, W. W. Shane, and L. A. Nelson. 1981. Interrelationships of *Meloidogyne* species with flue-cured tobacco. Journal of Nematology 13:67–79.

6. Byrd, D. W., Jr., K. R. Barker, H. Ferris, C. J. Nusbaum, W. E. Griffin, R. H. Small, and C. A. Stone. 1976. Two semi-automatic elutriators for extracting nematodes and certain fungi from soil. Journal of Nematology 8:206–212.

7. Cochran, W. G. 1977. Sampling techniques. New York: J. Wiley & Sons.

8. Elliott, J. M. 1977. Some methods for the statistical analysis of samples of benthic invertebrates. Freshwater Biological Association Scientific Publication 25.

9. Ferris, H. 1981. Dynamic action thresholds for diseases induced by nematodes. Annual Review of Phytopathology 19:427–436.

10. Ferris, H. 1984. Nematode damage functions: The problems of experimental and sampling error. Journal of Nematology 16:1–9.

11. Gates, C. E., and F. G. Ethridge. 1970. A generalized set of discrete frequency distributions with FORTRAN program. Mathematical Geology 4:1–24.

12. Goodell, P. B., and H. Ferris. 1980. Plant-parasitic nematode distributions in an alfalfa field. Journal of Nematology 12:136–141.

13. Goodell, P. B., and H. Ferris. 1981. Sample optimization for five plant-parasitic nematodes in an alfalfa field. Journal of Nematology 13:304–313.

14. Karandinos, M. G. 1976. Optimum sample size and comments on some published formulae. Bulletin of the Entomology Society of America 22:417–421.

15. McSorley, R., and J. L. Parrado. 1982. Estimating relative error in nematode numbers from single soil samples composed of multiple cores. Journal of Nematology 14:522–529.

16. Noe, J. P., K. R. Barker, and D. P. Schmitt. 1981. Comparison of four soil sampling methods for estimating field populations of plant-parasitic nematodes. Journal of Nematology 13:452 (Abstr.).

17. Perry, J. N. 1983. Effects of spatial heterogeneity on Jones's model for cyst-nematode population dynamics and crop root damage. Journal of Applied Ecology 20:849–856.

18. SAS Institute Inc. 1982. SAS user's guide: Statistics 1982 edition. Cary, NC: SAS Institute Inc.

19. Seinhorst, J. W. 1961. Plant nematode interrelationships. Annual Review of Microbiology 15:177–196.

20. Seinhorst, J. W. 1973. The relation between nematode distribution in a field and loss in yield at different average nematode densities. Nematologica 19:421–427.

21. Southwood, T. R. E. 1978. Ecological methods. New York: J. Wiley & Sons.

22. Taylor, L. R., I. P. Woiwod, and J. N. Perry. 1979. The negative binomial as a dynamic ecological model for aggregation, and the density dependence of k. Journal of Animal Ecology 48:289–304.