

An Algorithm for Fitting Seinhorst Curves to the Relationship between Plant Growth and Preplant Nematode Densities

H. Ferris, W. D. Turner, and L. W. Duncan¹

Quantitative description of the relationship between plant growth and preplant nematode densities is an essential prerequisite for nematode pest management decisions in annual crops (2,3). A common technique to describe this relationship is to use linear regressions of plant growth and log-transformed nematode population densities (1,7). The linearization produced by the transformation empirically accounts for the decreased damage per nematode, due to competitive interactions and damage overlap, as nematode densities increase. Seinhorst (9) derived an explanatory model for the relationship based on Nicholson's competition curve.

The model is based on sound biological theory and is of the form $y = m + (1-m)$

$z^{(P-T)}$. The minimum yield (m) (on a 0-1 relative scale) represents the residual plant growth at extremely high nematode population densities. It assumes that there may be a point below which yield cannot be further reduced by nematodes. Of course, for any crop or environmental situation, m may be zero. P is the nematode population density per prescribed unit of soil and/or roots. The tolerance limit (T) is the nematode population density below which yield reduction cannot be measured. Again, for any crop variety or environmental situation, T may be zero. The slope-determining parameter (z) of the damage function is defined as the proportion of the plant undamaged in the presence of parasitism by one nematode. Since z has a value slightly smaller than 1, raising it to the exponent $(P-T)$ results in less effect per nematode as each new nematode is added to the system.

The linear approach has been charac-

Received for publication 29 September 1980.

¹Associate Nematologist, Department of Nematology, and Graduate Research Assistants, Departments of Statistics and Nematology, University of California, Riverside, CA 92521.

teristically used, since linear regression techniques are readily available. Seinhorst curves are often fitted by the use of transparent overlays of standard curves with varying values of m and z . We have developed a computerized algorithm to derive the values of m , z , and T , at which the curve best fits the data points, as measured by minimization of the residual sum of squares and maximization of the correlation coefficient. A potential problem with this type of derivation is that the algorithm assumes that some minimum yield value is represented among the data set. In practice, population densities may not have been high enough to approach the minimum yield of the crop under that nematode and environmental stress. The algorithm allows the user to decide whether the minimum yield should be calculated based on the available data, or to specify some value of the minimum yield and avoid overestimation of predicted yield at high nematode densities. If there is no obvious leveling trend for yield values at high nematode densities among the available data, it may be wise to set $m = 0$ and calculate z and T values accordingly.

Using the model—

$$y_i = m + (1-m)z^{(P_i-T)} + \epsilon_i \dots \dots \dots \text{if } P_i > T$$

$$y_i = 1 + \epsilon_i \dots \dots \dots \text{if } P_i \leq T$$

where y_i is the relative yield of the i th plant or plot, determined by dividing the actual yield by the average yield of all plots with $P_i \leq T$, the objective is to minimize the residual sum of squares with respect to the parameters T , m , and z . From the residual sum of squares obtained with this model—

$$RSS = \sum \epsilon_i^2 = \sum [y_i - m - (1-m)z^{(P_i-T)}]^2 + \sum (y_i - 1)^2 \tag{i}$$

the least squares estimator of m is calculated by determining partial derivatives. Since the second term in equation (i) does not depend on m or z —

$$\frac{\partial}{\partial m} \sum (y_i - 1)^2 = \frac{\partial}{\partial z} \sum (y_i - 1)^2 = 0$$

then the partial derivative of the first term of equation (i) can be set equal to zero—

$$\frac{\partial RSS}{\partial m} = \sum \{2[y_i - m - (1-m)z^{(P_i-T)}] \cdot [-1 + z^{(P_i-T)}]\} = 0 \tag{ii}$$

and solved for m —

$$\hat{m} = \frac{\sum y_i - \sum z^{(P_i-T)} - \sum y_i z^{(P_i-T)} + \sum z^{2(P_i-T)}}{n - 2 \sum z^{(P_i-T)} + \sum z^{2(P_i-T)}} \tag{iii}$$

where n is the number of points with P_i greater than T . Summation is only over these n points in recognition that their population densities are high enough to contribute to yield reduction. Having obtained an estimate of m by minimizing residual sum of squares with respect to m , we now minimize the residual sum of squares with respect to z —

$$\frac{\partial RSS}{\partial z} = \sum \{2[y_i - m - (1-m)z^{(P_i-T)}] \cdot [m-1] \cdot [P_i-T] z^{(P_i-T)-1}\} \tag{iv}$$

again, summation is only over the points with P_i greater than T . Unlike equation (iii), equation (iv) cannot be solved analytically. A numerical solution is to establish an interval of uncertainty within which z is known to lie, and to successively reduce the range by evaluating equation (iv) at the midpoint of each new interval. Ultimately the z value is found at which the residual sum of squares is minimized for a given T .

Calculus cannot be used to find the least squares estimator of the tolerance limit (T) because the residual sum of squares is not a continuous function of T . Hence an iterative procedure is used to evaluate the residual sum of squares over an increasing range of T values and to calculate the m and z values at each of these levels. This allows selection of the value of T that minimizes the residual sum of squares and provides the associated z and m values.

The general procedure (Fig. 1) is to start with a coarse range of T values which covers a wide span of the observed P values by setting a relatively large T increment. An initial value for T and a starting value for z are selected. A value of $z = 0.99$ is convenient, since z is generally close to 1. Using the estimated z and T values, an m value is derived from equation (iii). This is the best possible value of m for a given value of z and T . Using this m value, the partial derivative of the residual sum of squares with respect to z is calculated from equation (iv). Since equation (iv) cannot be solved analytically, an iterative procedure is used to

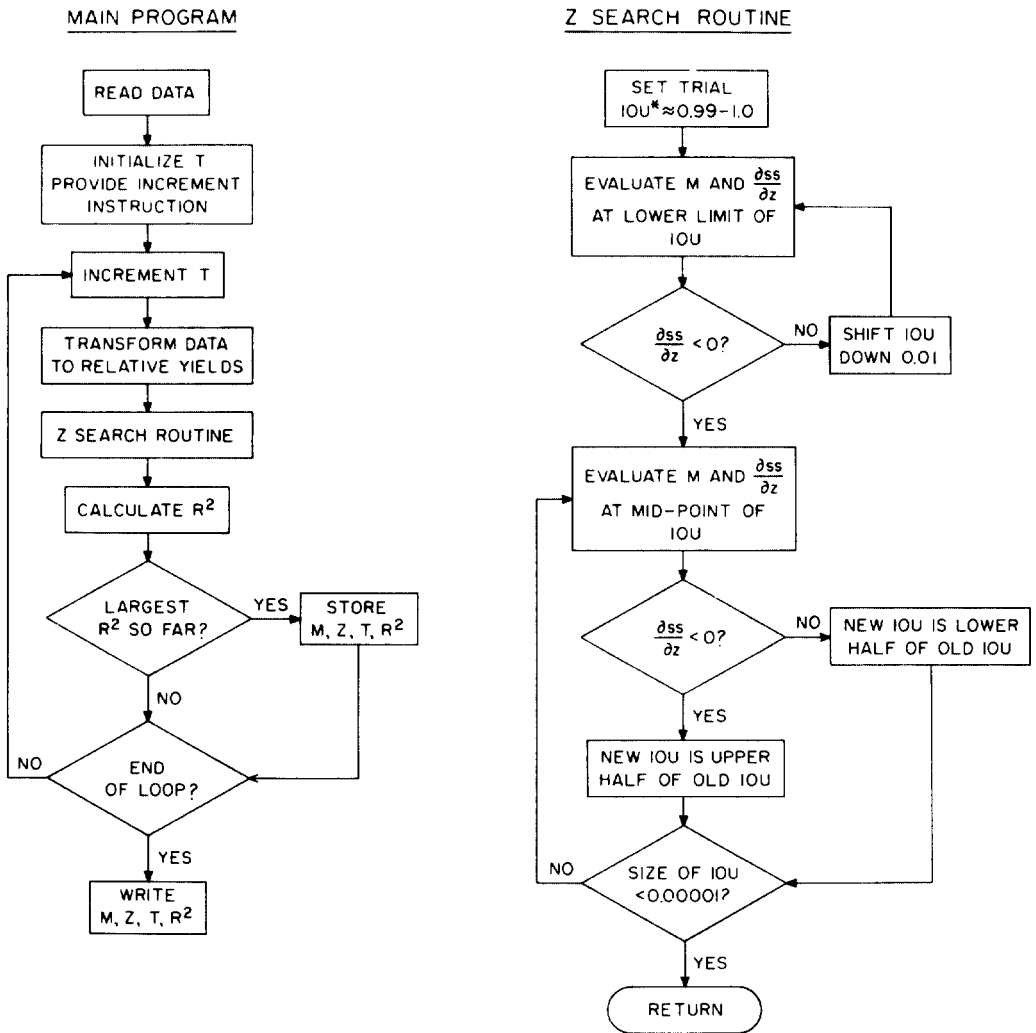


Fig. 1. Flow charts of main program and "z search" routine for determining T, z, and m values for the least squares fit of $y = m + (1-m)z^{(P-T)}$ when $P > T$ and $y = 1$ when $P \leq T$ for a data set.

converge the partial derivative of the residual sum of squares with respect to z on 0. If the sign of the calculated partial derivative is initially positive, the lower limit of the interval of uncertainty is decreased by a value of 0.01 and the upper limit becomes the previous lower limit—in the first case, 0.99. It may be necessary to limit the lowest allowable value of z, since by definition, z is close to 1.0. However, the units in which the population is measured should be considered in this limit. If the population is expressed per 200 g soil in determining the damage function, the z value will be lower

than if the population is expressed per 1,000 g soil since the perceived growth reduction is being attributed to fewer nematodes. Consistency in use of the model is necessary. By definition, z is the proportion of the plant undamaged by a single nematode. To maintain the biological integrity of the model, a lower limit of about z = 0.97 may be reasonable if nematodes are expressed per 1,000 g soil.

The partial derivative is recalculated successively until the lower limit of the interval of uncertainty becomes negative. This indicates that the true value of z is now

within the interval of uncertainty. Once the true interval of uncertainty has been identified, its midpoint is calculated and used to evaluate the partial derivative. The sign of the partial derivative at this new value is determined, indicating whether the true z value is included within the top half or the bottom half of the interval of uncertainty. Thus a new interval of uncertainty is obtained with half the previous range.

The iterative procedure is repeated until the upper and lower limits of the interval of uncertainty are separated by less than .00001, or another appropriately small number. At this stage the r^2 value is calculated and tested against previously calculated r^2 values. The tolerance limit is then incremented by the preselected increment and the iterative procedure repeated over the initial coarse grid. Each time a new T value is selected by incrementation from the previous one it is necessary to recalculate the relative yields throughout the whole data set using the maximum yield as an average of the yield at those population densities less than T . This assumes that below T the yield is unaffected by the nematode population. The residual sum of square values are scanned for each T value to select the region of the population densities at which the residual sum of squares approaches a minimum. A new search is made over a range of T values in this region with finer increments between them. The procedure is continued until the best value of T is known to a desired level of accuracy.

The procedure described allows determination of the descriptive parameters of a

Seinhorst curve for any data set. Further, it allows subjective judgment in the determination of appropriate m and T values. However, it removes the absolute subjectivity involved in selection of the parameters by overlaying curves. The algorithm simplifies use of a model based on sound biological theory.

We have used the algorithm to calculate T , z , and m values for some published data and compared r^2 values with those for linear regressions on the log transformed nematode populations (Table 1). The data were used by Seinhorst (9,11) as examples (4,5,6,8,11). In the cases where there were relatively few data points (5,6,8), or the data were from an extremely careful greenhouse study (11), both Seinhorst and linear models produced high r^2 values. In one data set (6,9) the predicted z value was apparently low, indicating considerable damage per nematode. However, the nematode counts were expressed per 100 ml soil, thereby inflating the perceived damage per nematode.

It is possible that a minimum yield has not been reached in data sets used with this program. Calculating a value of m and fitting a least squares curve based on this may constitute a false biological assumption. Accordingly, the program allows the user to subjectively assign a value of m for cases where the true value may not occur within the range of the data set. Trial runs with the algorithm indicate that the magnitude of m strongly influences T and z , underscoring the assertion of Seinhorst (10) that plant growth should be tested against a wide range of nematode densities.

Table 1. Comparison of damage function parameters of Seinhorst model† and linear model‡ for several data sets.

Source of data set§	No. of observations	Seinhorst model				Linear model		
		T	m	z	r^2	T	slope(b)	r^2
11	43	85	.242	.99964	.934**	295	-.143	.925**
4,9	67	78	.734	.98007	.382**	9	-.057	.377**
5,9	4	20	.220	.99603	.999**	38	-.174	.997**
6,9	7	10	.347	.95523	.979**	10	-.188	.844**
8,9	6	10	.240	.99680	.994**	30	-.144	.980**

† $y = m + (1-m)z^{(P-T)}$ for $y > T$, $y = 1$ for $P \leq T$.

‡ $y = 1 + b \log (P-T)$ for $y > T$, $y = 1$ for $P \leq T$.

§Data sets from references in Literature Cited.

LITERATURE CITED

1. Barker, K. R., P. B. Shoemaker, and L. A. Nelson. 1976. Relationships of initial population densities of *Meloidogyne incognita* and *M. hapla* to yield of tomato. *J. Nematol.* 8:232-239.
2. Ferris, H. 1978. Nematode economic thresholds: derivation, requirements, and theoretical considerations. *J. Nematol.* 19:341-350.
3. Ferris, H. 1980. Nematology—status and prospects: practical implementation of quantitative approaches to Nematology. *J. Nematol.* 12:164-170.
4. Jones, F. G. W. 1956. Soil populations of beet eelworms (*Heterodera schachtii* (Schm.) in relation to cropping. II. Microplot and field plots results. *Ann. Appl. Biol.* 44:25-26.
5. Mountain, W. B., and Patrick, Z. A. 1959. The peach replant problem in Ontario. VII. The pathogenicity of *Pratylenchus penetrans* (Cobb 1917) Filip. & Stek. 1941. *Can. J. Bot.* 37:459-470.
6. Oostenbrink, M. 1964. Harmonious control of nematode infestations. *Nematologica* 10:49-56.
7. Oostenbrink, M. 1966. Major characteristics of the relation between nematodes and plants. *Meded. Landbouwhogeschool Wageningen*, 66(4):1-46.
8. Peters, B. G. 1961. *Heterodera rostochiensis* population density in relation to potato growth. *J. Helm. R. T. Leiper Suppl.*:141-150.
9. Seinhorst, J. W. 1965. The relation between nematode density and damage to plants. *Nematologica* 11:137-154.
10. Seinhorst, J. W. 1970. Dynamics of populations of plant parasitic nematodes. *Ann. Rev. Phytopath.* 8:131-156.
11. Seinhorst, J. W. 1972. The relationship between yield and square root of nematode density. *Nematologica* 18:585-590.