Underestimation of Disease Progress Rates with the Logistic, Monomolecular, and Gompertz Models When Maximum Disease Intensity is Less Than 100 Percent

Deborah A. Neher and C. Lee Campbell

Department of Plant Pathology, Box 7616, North Carolina State University, Raleigh 27695-7616.
We thank I. Chet, R. Cohen, S. Freeman, C. M. Kenerley, D. O. Koch, L. V. Madden, A. F. Nash, B. D. Nelson, P. B. Shoemaker, and V. L. Smith, who provided empirical data for this study.
Accepted for publication 26 May 1992.

Classic growth curve models used in the analysis of plant disease epidemics include several implicit assumptions. If these assumptions are ignored, erroneous conclusions about disease progress may follow. One of the five assumptions listed by Campbell and Madden (4, p. 204) is that there is a constant host area to be infected (or diseased). A corollary to this assumption is that all host tissue can be infected and, thus, a disease intensity of 100% can be achieved (4,24). For simplicity, many applications of disease progress models assume an asymptote and a maximum level of disease intensity (K_{max}) of 100% or 1.0, but this assumption is not valid for many diseases (1,11,19). For example, many leaf spots and rusts have a maximum disease severity of 25–40% (11,14). The true asymptote or carrying capacity of the host for disease is a function of the amount of susceptible host tissue available and the extent to which infections from existing diseased tissue can expand into healthy tissue. Both of these factors can change with time (24).

Growth curve models can be altered in various ways to represent plant disease epidemics more realistically. A parameter representing actual K_{max} on host tissue can be used with growth curve models, such as the monomolecular, Gompertz, and logistic models. A K_{max} parameter may be as valuable as the rate parameter, r_*, in characterizing epidemics (4,15,19). Analytis (1,2) first suggested that the value of K_{max} could affect the calculation of the r_* of growth functions applied to plant disease epidemics. Park and Lim (19) illustrated mathematically how rates of disease increase are underestimated when calculated with traditional models of disease progress under the assumption that $K_{max} = 1.0$ when actual $K_{max} < 1.0$. They presented the problem of making the assumption that $K_{max} = 1.0$, but they did not calculate the magnitude of underestimation for a wide range of values or a variety of growth models. Our objective was to quantify the effect of assuming that $K_{max} = 1.0$, when it was actually lower, on the calculation of the rate of disease increase (r_*) for a range of K_{max} and rate values for the monomolecular, Gompertz, and logistic models (Fig. 1). The effect is examined from both theoretical and empirical perspectives, and guidelines are provided for two practical approaches to solve the problem using regression analysis.

Procedure. Theoretical predictions of r_* at varying K_{max} were determined from the linear forms of the monomolecular, Gompertz, and logistic models by implementing a Turbo BASIC program for use on a personal computer. The following linear expressions were solved for y (levels of disease) at two epidemic durations (i.e., 30 or 45 days).

Monomolecular: $$\ln[K/(K-y)] = \ln[K/(K-y_0)] + r_M t \quad (1)$$

Gompertz: $$-\ln[-\ln(y/K)] = -\ln[-\ln(y_0/K)] + r_G t \quad (2)$$

Logistic: $$\ln[y/(K-y)] = \ln(y_0/(K-y_0)] + r_L t \quad (3)$$

A range of initial disease ($y_0 = 0.01$ to 0.0001) and standardized rates (r_*) of disease progress were used. Standardized rates of disease increase were weighted mean absolute rates (21) ranging from $\rho = 0.00166$ to $\rho = 0.05$ in increments of 0.005 transformed to give equivalent rates for the monomolecular (r_M), Gompertz (r_G), and logistic (r_L) models. Equivalent rates were calculated as $r_* = \rho\,(2m + 2)/K$, with m as the parameter of shape for the disease progress curve (0, 1, and 2 for the monomolecular, Gompertz, and logistic models, respectively [21]). Values for K ($=K_{max}$) ranged from 0.25–1.0 in increments of 0.05.

Disease levels (y) at the end of a specified epidemic duration were calculated for each combination of values for y_0, r_*, and K_{max}. Then, the linear expressions for each model (eq. 1–3) were solved for r_*, assuming $K = 1.0$. The r_* used to solve for y in linear equations 1–3 with varying actual K_{max} values (0.25–1.0) was compared with r_* calculated with an assumed $K_{max} = 1.0$ to quantify the percentage of underestimation (eq. 4).

$$[(r_{*(\text{actual } K\text{max})} - r_{*(K\text{max}\,=\,1.0)})/r_{*(\text{actual } K\text{max})}] \times 100 \quad (4)$$

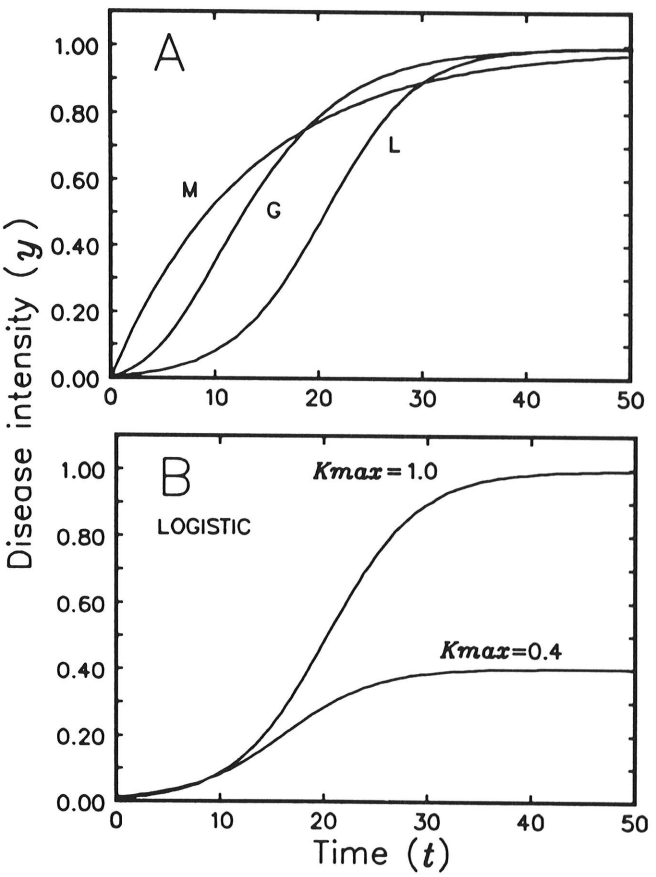

Fig. 1. A, Curves of data from the monomolecular (M), Gompertz (G), and logistic (L) models calculated with a standardized mean rate (ρ = 0.0375) (21), $y_0 = 0.01$, and $K_{max} = 1.0$ for an epidemic duration of 45 days. **B,** Curves of data from the logistic model calculated with K_{max} = 1.0 and $K_{max} = 0.4$ using equivalent y_0 and ρ values for an epidemic duration of 45 days.

© 1992 The American Phytopathological Society

TABLE 1. Sources of empirical disease progress data from published studies that were used to estimate the magnitude of underestimation of rate of disease increase when maximum disease (K_{max}) was assumed to equal 1.0

Figure 3 denotation	Model[a]	Disease/pathogen	Host	Reference
camp	L	Black shank	Tobacco	3
chet	M	*Trichoderma harzianum* with *Rhizoctonia solani* or *Sclerotium rolfsii*	Bean	7
cohen	M	Fusarium wilt	Melon	5
free	L	Fusarium wilt	Watermelon	8
ken	G	Phytophthora root rot	Fraser fir	12
koch	G	Phymatotrichum root rot	Cotton	13
madden	L	Tobacco etch virus and tobacco vein mottling virus	Tobacco	16
nash	M	Early blight	Tomato	17
nelson	G	Sclerotinia wilt	Sunflower	18
shoe	L	Early blight	Tomato	PBS[b]
smith	M	Southern blight	Carrot	26

[a]L = logistic, M = monomolecular, G = Gompertz. The model was chosen for the purposes of the current study and was not necessarily the one chosen by the original author.
[b]Data received from P. B Shoemaker, some of which are published (25).

TABLE 2. Equations[a] used in nonlinear regression for analysis of empirical disease progress data

Model	$K_{max} = 1.0$	Actual K_{max}	B^b
Monomolecular	$y = 1 - B\exp(-r_M t)$	$y = K[1-B\exp(-r_M t)]$	$(K - y_0)/K$
Gompertz	$y = \exp[-B\exp(-r_G t)]$	$y = K[\exp(-B\exp(-r_G t)]$	$-\ln(y_0/K)$
Logistic	$y = 1/[1 + \exp(B - r_L t)]$	$y = K/[1 + \exp(B - r_L t)]$	$\ln[y_0/(K - y_0)]$

[a]K ($=K_{max}$) = maximum level of disease (y) or asymptote of disease progress curve; y = disease at time of observation; y_0 = level of disease at first observation; r_M, r_G, and r_L = rate of disease increase for the specific model; t = epidemic duration.
[b]Constant of integration, i.e., with no implied biological importance.

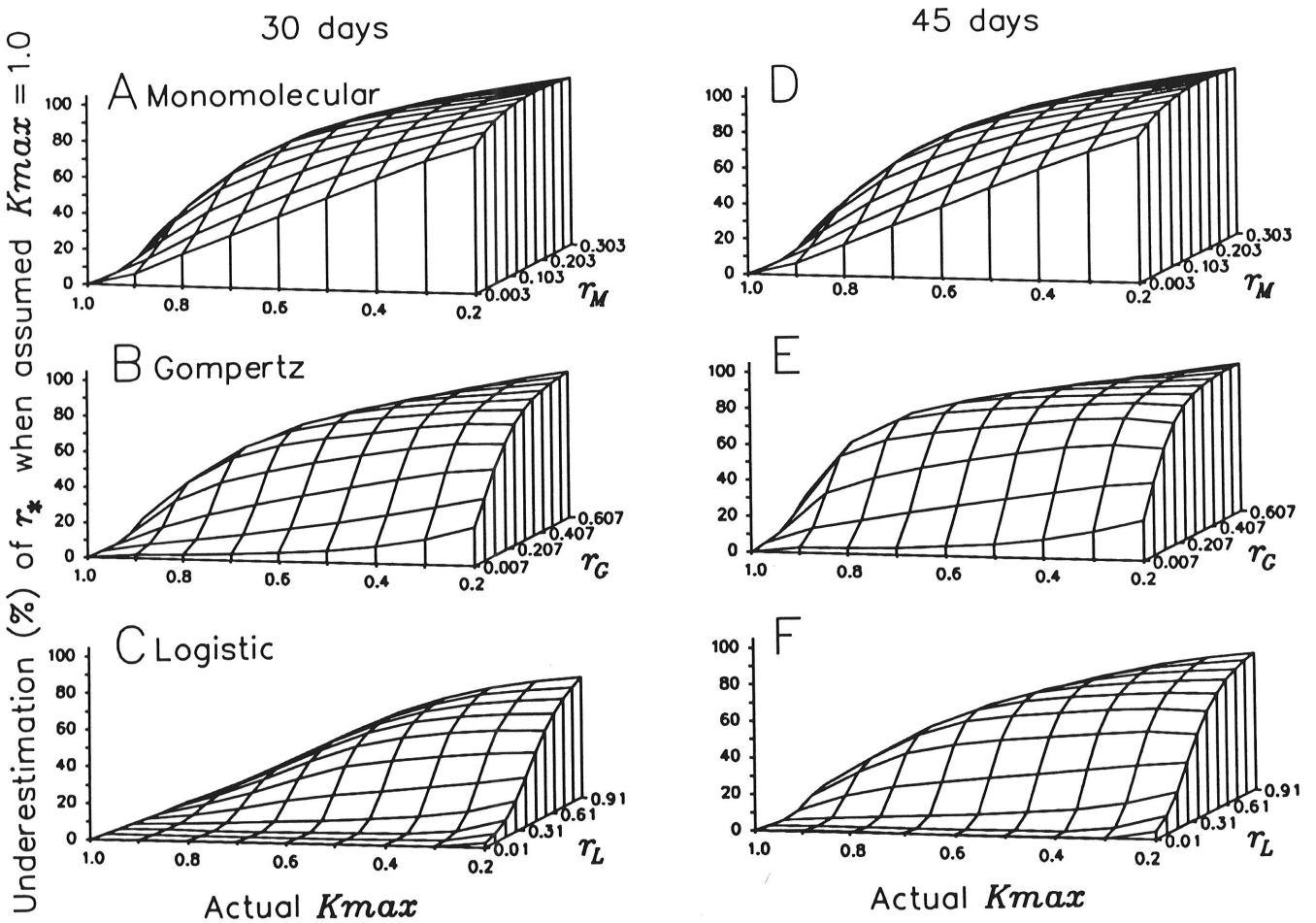

Fig. 2. Percent underestimation of r_* ($=[\rho\ (2m + 2)]/K_{max}$ [21]) when r_* values with assumed $K_{max} = 1.0$ were compared with r_* values with actual K_{max} for monomolecular (**A** and **D**), Gompertz (**B** and **E**), and logistic (**C** and **F**) models. Epidemics of 30-day (**A–C**) and 45-day (**D–F**) duration with $y_0 = 0.0001$ are illustrated. The rotation of x and y axes must be considered to accurately read percentage values on the y axis.

The effect of a longer epidemic duration on the underestimation of r_* was assessed by comparing the underestimation of r_* for epidemics of 30- and 45-day durations. The results were illustrated using a smoothed spline estimation in PROC G3GRID and PROC G3D of SAS ver. 6.04 for microcomputers (23).

To validate theoretical predictions, original disease progress data were requested from authors for published studies of soilborne and foliar epidemics of fungal and viral diseases (Table 1). These studies were selected mostly from the 1986–1990 volumes of *Plant Disease* and *Phytopathology*. The principal criteria for selection of data sets were 1) a satisfactory number of data points in time to permit fitting of nonlinear models, i.e., ≥9 (15); 2) a satisfactory spacing of data to permit estimation of asymptotes and rate parameters (15); and 3) the attainment of a definite asymptote. Only 10 data sets were found in the literature to satisfy the established criteria; in addition, one unpublished data set was found to satisfy the criteria. Of the 11 data sets, two represented pathosystems with airborne fungal inoculum, eight represented pathosystems with soilborne fungal inoculum, and one represented a viral pathosystem; no data sets involving bacterial pathogens were found that satisfied the criteria for selection. No criticism of original methods of analysis or conclusions drawn by the authors is intended.

The linear forms of the logistic, Gompertz, and monomolecular models were examined for goodness-of-fit to each data set. Appropriateness of model selection for each data set was appraised by plotting standardized residuals versus predicted values and examining unadjusted coefficients of determination (r^2) from linear regression analyses (4). All linear regression analyses were performed using the General Linear Models procedure (PROC GLM) of SAS ver. 6.04 (22). Model selection was based on statistical fit; therefore, no inference of the biological nature of the pathosystem was intended by the model selected (4,10,20).

After the most appropriate model was selected, r_* was estimated for models with assumed $K_{max} = 1.0$ and compared with r_* calculated from the same general model with actual K_{max} values using nonlinear regression. Data for each replicate of each treatment were analyzed as a separate curve. Disease progress data were analyzed using the integrated, nonlinear form of the appropriate growth model (Table 2). Nonlinear regression permitted simultaneous estimations of r_* and K_{max} parameters. Initial disease (y_0) also could have been estimated, but we felt that three parameters were too many, given the number of observations in the data sets. Thus, y_0 was incorporated into a constant of integration (B) (Table 2).

All data were proportional (not percentage) and not transformed. Transformations were not appropriate for fitting nonlinear curves. Parameters and associated statistics were estimated using a least squares, nonlinear regression procedure (PROC NLIN) with Marquardt's compromise method (6) in SAS ver. 6.04 (22). The lower and upper bounds were specified as >0.0 and <0.3 for r_* and >0.0 and ≤1.0 for K_{max}, respectively. If one of the parameters failed to converge at a value between its limits, at least two different initial estimates were attempted.

Results. From our analyses with theoretical data, the underestimation of r_* increased as actual K_{max} decreased most for the monomolecular model and less for the Gompertz and logistic models, respectively (Fig. 2). An increase in degree of underestimation of r_* with greater standardized mean rates of disease progress also occurred. The monomolecular model was least sensitive and the Gompertz and logistic models were most sensitive to changes in standardized rates (Fig. 2). When epidemic duration was 45 days, rather than 30 days, effects of decreased K_{max} and increased standardized mean rates of disease progress resulted in greater underestimation of r_*, especially for the Gompertz and logistic models (Fig. 2).

The underestimation of r_M was affected more by decreasing K_{max} than by increasing r_* for disease progress in the monomolecular model; the opposite was true for the value of r_L with the logistic model. Changes in K_{max} and r_* affected the underestimation of r_G similarly in the Gompertz model. The degree of underestimation of rate of disease progress depends on the specific model, the actual rate of disease increase, and the closeness of the actual value of K_{max} to 1.0. The problems were accentuated with lower actual K_{max} values, higher rates of disease progress, and longer epidemics.

Theoretical predictions were confirmed by examination of the empirical data. Underestimation of rates of disease progress in curves calculated from empirical data with $K_{max} = 1.0$ compared with actual K_{max} decreased linearly with greater estimated K_{max} for all growth models (Fig. 3). The epidemics analyzed represented a range of hosts, pathogens, and disease types, e.g., foliar, systemic, wilt, and root diseases (Table 1). These consistent results confirm that Analytis (1,2) and Park and Lim (19) were correct when they stated that it often is inappropriate to assume that $K_{max} = 1.0$, especially for disease progress curves that have low asymptotic values.

Recommendations. Estimation of K_{max} can be perplexing, especially for diseases that do not have a predictable or repeatable maximum intensity. Presently, it is not possible to classify pathosystems with regard to potential K_{max} based solely on biological criteria. For example, not all diseases caused by soilborne pathogens would be expected to have $K_{max} < 1.0$. The

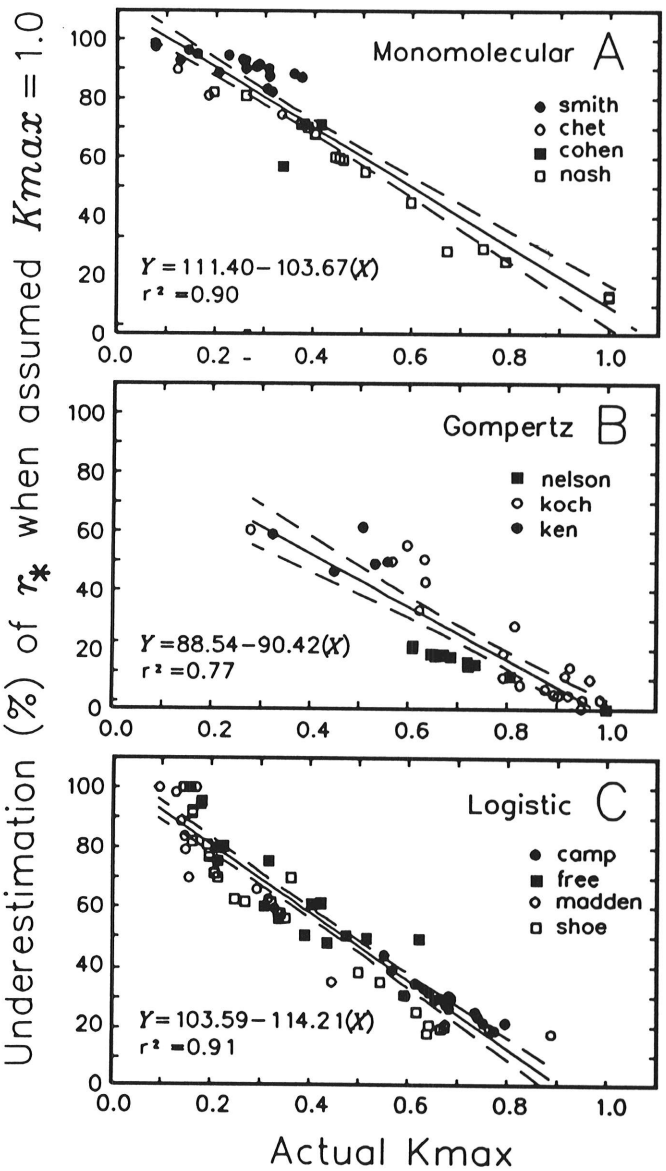

Fig. 3. Linear regression between percent underestimation of r_*, the rate of disease increase if it is assumed that $K_{max} = 1.0$, and actual K_{max} values for published data (see Table 1 for references associated with symbols) that fit the **A,** monomolecular, **B,** Gompertz, and **C,** logistic, models. The 95% confidence intervals around the regression lines are illustrated.

ideal approach to solve this problem would be to use nonlinear regression analysis to describe the data sets obtained. K_{max} can then be estimated as one of the unknown parameters. However, this procedure has the limitation of requiring at least three observations of disease through time for each parameter of the model estimated, i.e., the equations solved for this study required a minimum of nine observations through time. Without this minimum number of observations, convergence of the model for the empirical data may become impossible. Knowledge of the partial derivatives for each parameter of each model is generally needed to run nonlinear regression.

In cases with too few observations through time to permit nonlinear regression analyses, the approach of linear regression is available. However, linear regression requires estimation of K_{max} from empirical data before analysis (9), which requires some professional judgment. If the maximum level of disease were constant for a given pathosystem, measured disease severity simply could be divided by the maximum disease severity and y analyzed as the proportion of host tissue diseased relative to the maximum. This procedure would in effect provide a scaled or proportional K_{max} value of 1.0. Otherwise, K_{max} may represent the maximum level of disease 1) across treatments or among epidemics, 2) by treatment or location, or 3) for individual plots. The advantage of using method 2 or 3 is that actual K_{max} may vary by treatment or experimental location, which may result in selection of K_{max} that is too high for some treatments. If the estimate of K_{max} is too high, the problem of underestimation of the rate parameter will be repeated.

The next step for either analytical approach is to compare rate parameters calculated from disease progress curves with different K_{max} values and growth models. Weighted mean absolute rates (ρ) (21) of disease increase can be used to compare rate parameters among models with contrasting K_{max} values and shape parameters (m). These rate values can be calculated as $\rho = r_*K/(2m + 2)$, where m is the parameter for the shape of disease progress curve (0, 1, and 2 for the monomolecular, Gompertz, and logistic models, respectively), r_* is the rate of disease increase for the specific model, and K ($=K_{max}$) is the maximum level of disease (4,21). Gray et al (9) illustrated the use of ρ for comparing epidemics with different shape parameters. Values of ρ should be compared by analysis of variance procedures in agreement with the experimental design of the experiment. When K values are equivalent, the K parameter can be dropped from the equation. If K values are different but the same model holds (i.e., fixed shape parameter), then r_*K could be calculated as an overall (mean) measure of the absolute rate (4).

Our study reemphasizes and demonstrates clearly the importance of including actual K_{max} in asymptotic models of disease progress as originally proposed by Analytis (1,2) and demonstrated mathematically by Park and Lim (19). The effect of using incorrect K_{max} values occurs over a wide range of values for y_0 and r_* and for the three models used widely in plant pathology to describe disease progress curves. Practical approaches to including K_{max} in growth models are available and should be used whenever maximum or asymptotic values of disease intensity are less than 100% or 1.0.

LITERATURE CITED

1. Analytis, S. 1973. Zur Methodik der Analyse von Epidemien dargestellt am Apfelschorf (*Venturia inaequalis* (Cooke) Aderh.). Acta Phytomed. 1:1-75.
2. Analytis, S. 1979. Die Transformation von Befallswerten in der quantitativen Phytopathologie. Phytopathol. Z. 96: 156-171.
3. Campbell, C. L., Jacobi, W. R., Powell, N. T., and Main, C. E. 1984. Analysis of disease progression and the randomness of occurrence of infected plants during tobacco black shank epidemics. Phytopathology 74:230-235.
4. Campbell, C. L., and Madden, L. V. 1990. Introduction to Plant Disease Epidemiology. John Wiley & Sons, New York.
5. Cohen, R., Riov, J., Lisker, N., and Katan, J. 1986. Involvement of ethylene in herbicide-induced resistance to *Fusarium oxysporum* f. sp. *melonis*. Phytopathology 76:1281-1285.
6. Draper, N. R., and Smith, H. 1981. Applied Regression Analysis. 2nd ed. John Wiley & Sons, New York.
7. Elad, Y., Chet, I., and Katan, J. 1980. *Trichoderma harzianum:* A biocontrol agent effective against *Sclerotium rolfsii* and *Rhizoctonia solani*. Phytopathology 70:119-121.
8. Freeman, S., and Katan, J. 1988. Weakening effect on propagules of *Fusarium* by sublethal heating. Phytopathology 78:1656-1661.
9. Gray, S. M., Moyer, J. W., Kennedy, G. G., and Campbell, C. L. 1986. Virus-suppression and aphid resistance effects on spatial and temporal spread of watermelon mosaic virus 2. Phytopathology 76:1254-1259.
10. Hau, B., and Kranz, J. 1977. Ein vergleich verschiedener Transformationen von Befallskurven. Phytopathol. Z. 88:53-68.
11. Jeger, M. J. 1982. Using growth curve relative rates to model disease progress of apple powdery mildew. Prot. Ecol. 4:49-58.
12. Kenerley, C. M., and Bruck, R. I. 1987. Distribution and disease progress of Phytophthora root rot of Fraser fir seedlings. Phytopathology 77:520-526.
13. Koch, D. O., Jeger, M. J., Gerik, T. J., and Kenerley, C. M. 1987. Effects of plant density on progress of Phymatotrichum root rot in cotton. Phytopathology 77:1657-1662.
14. Kranz, J. 1977. A study on the maximum severity in plant diseases. Travaux dédiés à G. Viennot-Bourgin. pp. 169-173.
15. Madden, L. V., and Campbell, C. L. 1990. Nonlinear disease progress curves. Pages 181-229 in: Epidemics of Plant Diseases: Mathematical Analysis and Modeling. J. Kranz, ed. Springer-Verlag, Berlin.
16. Madden, L. V., Pirone, T. P., and Raccah, B. 1987. Temporal analysis of two viruses increasing in the same tobacco fields. Phytopathology 77:974-980.
17. Nash, A. F., and Gardner, R. G. 1988. Tomato early blight resistance in a breeding line derived from *Lycopersicon hirsutum* PI 126445. Plant Dis. 72:206-209.
18. Nelson, B. D., Hertsgaard, D. M., and Holley, R. C. 1989. Disease progress of Sclerotinia wilt of sunflower at varying plant populations, inoculum densities, and environments. Phytopathology 79:1358-1363.
19. Park, E. W., and Lim, S. M. 1985. Empirical estimation of the asymptotes of disease progress curves and the use of the Richards generalized rate parameters for describing disease progress. Phytopathology 75:786-791.
20. Pfender, W. F. 1982. Monocyclic and polycyclic root diseases: Distinguishing between the nature of the disease cycle and the shape of the disease progress curve. Phytopathology 72:31-32.
21. Richards, F. J. 1959. A flexible growth function for empirical use. J. Exp. Bot. 10:290-300.
22. SAS Institute Inc. 1988. SAS/STAT User's Guide, release 6.03 ed. SAS Institute Inc., Cary, NC.
23. SAS Institute Inc. 1990. SAS/GRAPH Software, reference version 6. 1st ed., Vol. 2. SAS Institute Inc., Cary, NC.
24. Seem, R. C. 1988. Measurement and analysis of the effects of crop development on epidemics. Pages 51-68 in: Experimental Techniques in Plant Disease Epidemiology. J. Kranz and J. Rotem, eds. Springer-Verlag, Berlin.
25. Shoemaker, P. B. 1991. Alternatives to EBDC fungicides and resistance to tomato early blight, 1990. Fungic. Nematicide Tests 46:192.
26. Smith, V. L., Campbell, C. L., Jenkins, S. F., and Benson, D. M. 1988. Effects of host density and number of disease foci on epidemics of southern blight of processing carrot. Phytopathology 78:595-600.