A Loss Model for Crops

L. V. Madden, S. P. Pennypacker, C. E. Antle, and C. H. Kingsolver

Former graduate research assistant and associate professor, Department of Plant Pathology; professor, Department of Statistics; and adjunct professor, Department of Plant Pathology, The Pennsylvania State University, University Park 16802.

Current address of the senior author: Department of Plant Pathology, Ohio Agricultural Research and Development Center, Wooster 44691.

Contribution 1191, Department of Plant Pathology, The Pennsylvania Agricultural Experiment Station. Authorized for publication 13 May 1980 as Journal Series Paper 5976.

We thank T. M. Abdel-Hak for supplying a portion of the data.

Accepted for publication 20 November 1980.

ABSTRACT

Madden, L. V., Pennypacker, S. P., Antle, C. E., and Kingsolver, C. H. 1981. A loss model for crops. Phytopathology 71:685-689.

Contemporary models for describing and predicting yield reduction caused by plant diseases were selected, usually with ordinary least-squares regression, to best fit the data and not necessarily to represent biological reality. A generalized nonlinear model with inherent flexibility was developed to characterize the relationship between crop loss and plant disease. In addition to providing loss predictions at various levels of disease, the model incorporates: a threshold disease level below which no loss occurs, a maximum level of loss that may occur prior to the maximum amount of disease, and a large family of curve shapes to depict disease-loss relationships. The model was fit to simulated and actual loss data sets. The model numerically approximated the critical-point loss models reported in the literature and also described loss for two disease-host systems over a 2-yr period. The availability of a single model has the advantage that losses caused by different diseases on different crops can be compared directly by an analysis of the model parameters.

Additional key words: epidemiology, crop loss, modeling.

The primary concern in applying plant protection measures is to reduce crop loss to an acceptable level—acceptable under biological, economic, and physical constraints. Loss is defined as the measurable reduction in quantity and/or quality of yield. To reduce losses, we must know how much loss occurs. The knowledge of loss is needed also to: evaluate control strategies, make optimal pest management decisions, make yield predictions, and evaluate the need for research on particular pests and crops (5,8,13). The relationship between loss and disease has not been intensively studied for many crops or many plant diseases (8,13).

Contemporary models for describing disease-loss relationships use one or more epidemic variables to predict loss. Common variables include: disease severity at one or more points in time, time or growth stage at which a certain disease level is reached, area under the disease progress curve (ADPC), and slope of a linearized disease progress curve (7,8,13). These models were selected, usually with ordinary least-squares regression, to best fit the data and not necessarily to represent biological reality. They were developed for specific cases without inherent flexibility.

In the following study, a generalized nonlinear model was developed to characterize the relationship between crop loss and plant disease. The model was evaluated to determine if it could numerically approximate the other loss models reported in the literature and also if it could represent actual yield-loss data.

DEVELOPMENT OF THE MODEL

In 1961, Tammes suggested a theoretical relationship between yield and stress (16). A stress in this case is defined as any potentially injurious biotic or abiotic environmental factor (11). Plant disease is a stress. Since loss is inversely related to yield, the theoretical relationship between loss and stress can be depicted by Fig. 1. Loss can be represented in many ways; for comparison purposes, loss is expressed as a proportion. The term $g(X)$ represents a function of stress factors, specifically plant disease. Properties of this theoretical relationship include: a disease threshold (ie, disease level below which no loss occurs due to compensation of the crop or plant), a maximum level of loss where an increase in the disease level does not cause an incremental increase in loss, and a middle region of the curve within which there is a positive correlation between loss and disease although a linear relationship is not anticipated. A nontransformed linear relationship in this middle region could occur but there is no biological reason why this should be so. The disease threshold could be zero and there may or may not be a maximum level of loss prior to maximum disease level. For the most part, these theoretical considerations have not been used by investigators who attempted to model the yield-loss process.

The proposed loss model can take the following form:

$$L_i = 1 - \exp(-g(X_i)^s) + u_i \quad (1)$$

where L_i is the level of loss for the i-th experimental unit (field or plant) and is expressed as a proportion. The term $g(X)$ is the

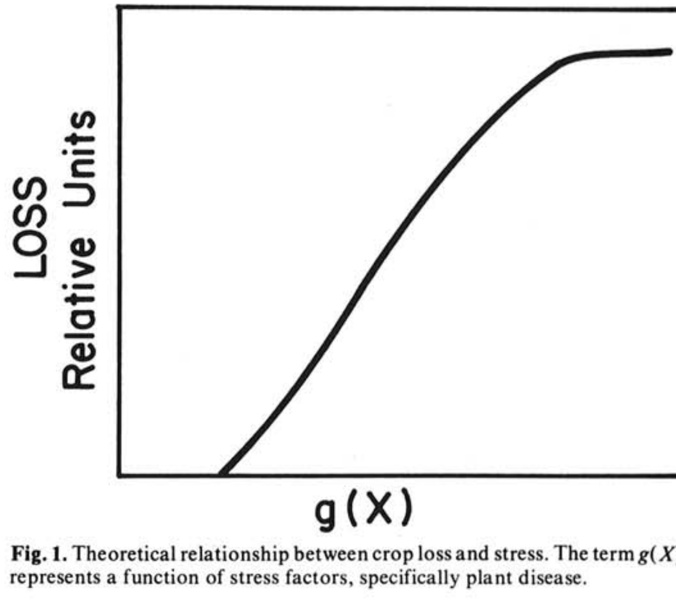

Fig. 1. Theoretical relationship between crop loss and stress. The term $g(X)$ represents a function of stress factors, specifically plant disease.

0031-949X/81/07068505/$3.00/0

function of disease; s is a shape parameter giving the form of the curve and must be greater than zero; and u is the unexplained variability (error term). The function of disease could take many forms; the simplest yet useful form can be written as:

$$g(X) = (X-d)/b$$

In this equation, X is the proportion of disease or some other univariate expression of disease stress (eg, area under the disease progress curve); d is the disease threshold; b is a slope parameter.

The loss equation can thus be written as:

$$L_i = 1 - \exp(-((X_i-d)/b)^s) + u_i \tag{2}$$

The predicted level of loss is given by equation 2 without the error term and subscripts:

$$L = 1 - \exp(-((X-d)/b)^s) \tag{3}$$

This equation, with different interpretations of the variables and parameters, has been used extensively as a cumulative distribution in life-testing and reliability studies (2,19). Example forms of the loss equation are given in Fig. 2 for several values of the shape parameter s with d equal to 0.08 and b equal to 0.40. This model, as well as the theoretical considerations of Tammes (16), does not take into account the possibility of "negative loss" at low levels of stress.

Our proposed model was compared with the basic models suggested by previous workers. Large and Doling (10) proposed a square-root equation to predict loss in barley and oats caused by powdery mildew. Their prediction equation can be written as:

$$L = 0.25(X^{1/2}) \tag{4}$$

where X is proportion of disease at growth stage 10.5. Teng et al (17) recently investigated the appropriateness of a square-root model for describing barley loss due to leaf rust severity at several growth stages of the crop. Romig and Calpouzos (15) proposed a logarithmic equation to predict loss in wheat due to stem rust severity at growth stage 11.2. Their prediction equation can be written as:

$$L = -0.25 + (27\ ln\ (100X)/100.0) \tag{5}$$

Several workers have proposed a straight-line relationship between loss and disease (3,4,9,14,17). The equation can take the following form:

$$L = cX \tag{6}$$

where X is the proportion of disease at some point in time. The coefficient c (eg, 1.0) varies with different crops.

The utility of the proposed model was evaluated through the analysis of both field data and generated data sets. Data were generated for equations 4, 5, and 6 from X equal 0.05 to 0.95 in steps of 0.025. The three parameters of the proposed general loss model were estimated from each of these generated data sets.

A FORTRAN program was written to estimate the parameters; calculate parameter variances and covariances, coefficient of determination (R^2), and error mean square; and print predicted and actual loss data (12). The program uses an iterative least-squares procedure to estimate the parameters.

FIELD VERIFICATION

The proposed model was fit to the loss data of two disease-host systems over a 2-yr period. These disease-host systems were chosen as examples for evaluating the model; no attempt was made to evaluate cultivars or control techniques. Epidemics of Cercospora leaf spot of peanut (*Arachis hypogeae* L. 'Alpha'), caused by a complex of *Cercospora arachidicola* Hori and *Cercospora personata* (Berk. and Curt.) Ell. and Ev. (= *Cercosporidium personatum*), were generated in 6 × 7-m field plots in the Sharkia province of Egypt during 1977 and 1978 (1). Standard cultural practices were used throughout the experiment. Treatments consisted of spraying the plots either zero, one, two, or three times with the protectant fungicide thiophanate-methyl (0.7 kg a.i./ha). The experiment was replicated six times in a randomized complete block design. Disease severity was estimated using the keys and methods of James (6). Entire plots were harvested for determining yield and loss with the exception of one row along each edge of the plots and also 50 cm from the end of each row.

Potato (*Solanum tuberosum* L. 'Star') early blight epidemics, caused by *Alternaria solani* (Ell. and G. Martin) Sor., were generated in 6×7-m field plots in the Tawila-Dakahlia province of Egypt during the 1977–1978 and 1978–1979 growing seasons (1). Standard cultural practices were used throughout the study. The experiment in each growing season consisted of seven treatments that were applied from zero to six times. The protectant fungicide captafol was sprayed on the plants at a rate of 1.7 kg a.i./ha. The experiments were replicated six times in a randomized complete block design. During the 1978–1979 season, loss and disease values were averaged across blocks resulting in only seven data points. Disease severity was estimated using the technique of Townsend and Heuberger (18). Entire plots were harvested except for one border row at the edge of each plot and 50 cm from the end of each row.

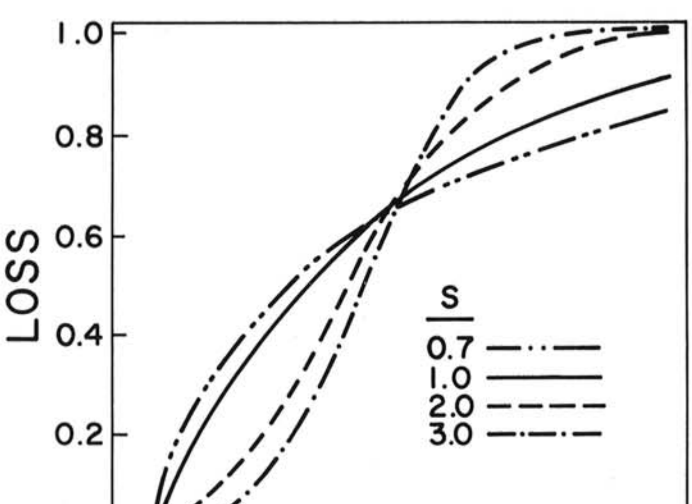

Fig. 2. Example forms of the loss model (equation 3) for four values of the shape parameter s. The disease threshold (d) equals 0.08 and the slope parameter (b) equals 0.40.

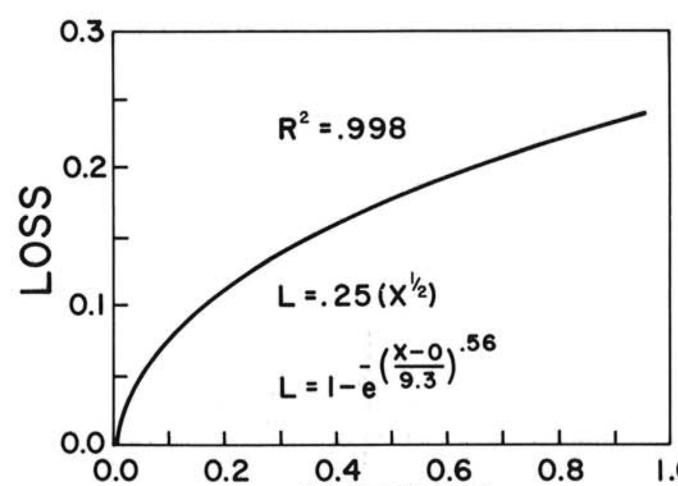

Fig. 3. Loss values generated from the square-root model (equation 4) and predicted losses of the proposed general model (equation 3) based on the parameters estimated from the generated data. The generated and predicted values fall on the same line.

RESULTS

There was almost perfect agreement between the proposed loss model and the square-root loss model (Fig. 3). The data points generated from equation 4 and the predicted values from equation 3 fall on the same line. The estimated shape parameter equaled the low value of 0.56. There also was very good agreement between the logarithmic model (equation 5) and the proposed general loss model as depicted in Fig. 4. Only slight deviations between the generated and predicted values were detected. The value of s in this case equaled 0.84. The agreement between the straight-line model (equation 6) and the proposed general model is shown in Fig. 5. The proposed general model could not correspond to a straight line for the full range of disease severity values. The fit was good, however, for losses less than 0.6, which is the range of interest in a practical setting. The shape parameter (1.5) corresponding to the straight-line equation had the highest value of the three models taken from the literature. In the above comparisons it was assumed that the models in the literature are correct; an analysis of the raw data would be required to determine if these models are correct.

Table 1 contains the estimated parameters of the loss model, their respective standard deviations, coefficients of determination, degrees of freedom, and error mean square for the peanut and potato loss data. Peanut loss in relation to final disease severity of Cercospora leaf spot was accurately described by the model (Table 1). The peanut loss data together with the predicted losses for the 1977 season are plotted in Fig. 6; Fig. 7 contains the 1978 data. Estimated shape parameters had the low values of 0.58 and 0.40 for 1977 and 1978, respectively. The point estimates of the disease threshold (d) were close to zero (0.02 and 0.03, respectively) for both seasons.

The potato loss data also were described well with the model (Table 1, Figs. 8 and 9). Unlike the peanut system, the shape of the potato loss curves differed for the two growing seasons; s equaled 0.51 and 1.69 for 1977–1978 and 1978–1979, respectively. The early blight disease threshold for the 2 yr equaled 0.17 and 0.12. The

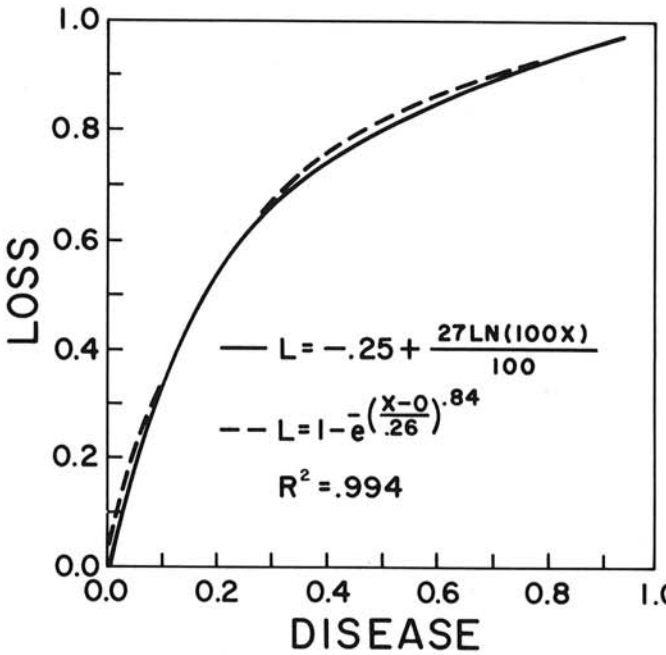

Fig. 4. Loss values generated from the logarithmic model (equation 5) and predicted losses of the proposed loss model (equation 3) based on the parameters estimated from the generated data.

TABLE 1. Estimated parameter values and their respective standard deviations, coefficient of determination (R^2), degrees of freedom (df), and error mean square (MSE) for the loss model fit to four data sets

Crop/ Disease/Year	Parameters[a] d	b	s	R^2	df	MSE ($\times 10^{-3}$)
Peanut/Cercospora leafspot						
1977	0.018	0.244	0.578	0.842	21	7.37
	(8.9×10^{-5})	(0.039)	(0.079)			
1978	0.033	1.954	0.400	0.934	21	1.95
	(7.2×10^{-4})	(0.364)	(0.034)			
Potato/early blight						
1977/1978	0.173	0.570	0.513	0.829	39	6.08
	(5.0×10^{-3})	(0.066)	(0.061)			
1978/1979	0.125	0.217	1.692	0.978	4	0.43
	(3.4×10^{-2})	(0.021)	(0.760)			

[a] d = disease threshold, b = a slope parameter, and s = the shape parameter. Standard deviations appear below in parentheses.

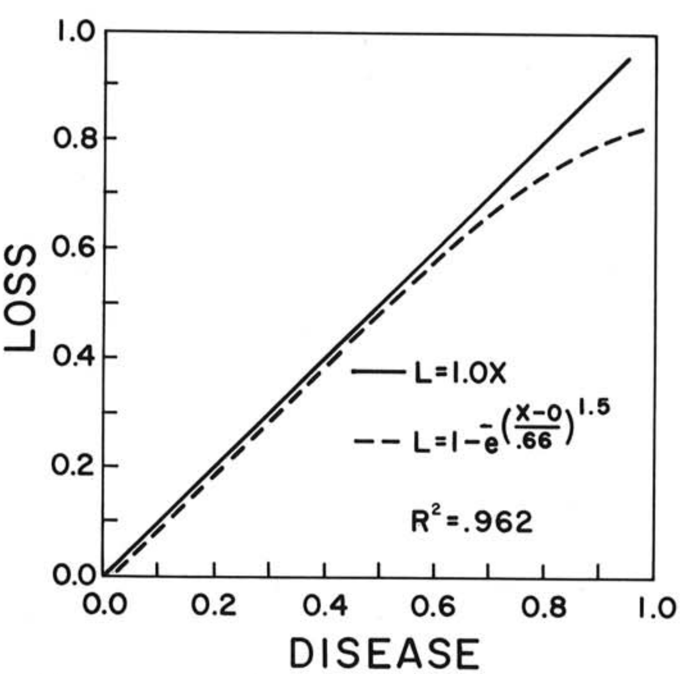

Fig. 5. Loss values generated from the straight-line model (equation 6) and predicted losses of the proposed loss model (equation 3) based on the parameters estimated from the generated data.

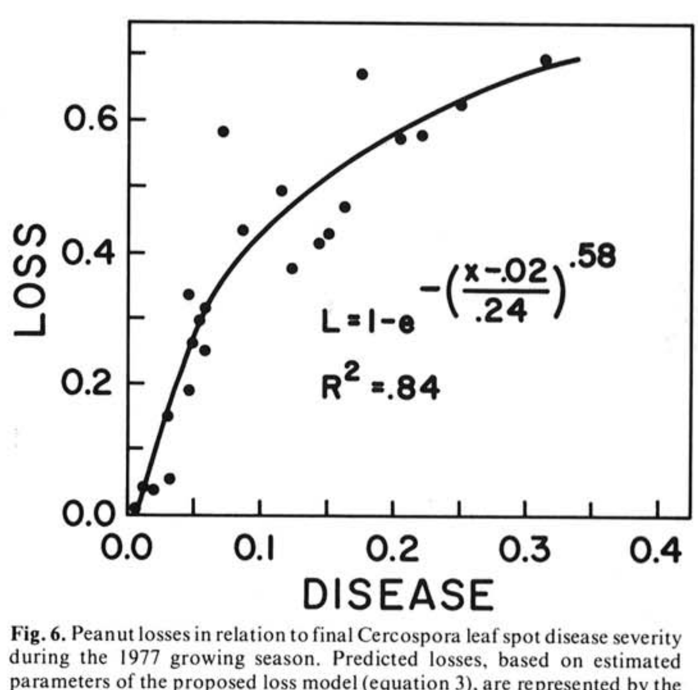

Fig. 6. Peanut losses in relation to final Cercospora leaf spot disease severity during the 1977 growing season. Predicted losses, based on estimated parameters of the proposed loss model (equation 3), are represented by the solid line. Disease severity was estimated 96 days after planting. Loss and disease are expressed on a per unit basis.

potato crop was, therefore, capable of "withstanding" appreciable levels of disease before loss occurred or before it was possible to detect that loss.

DISCUSSION

Incorporation of the theoretical considerations of Tammes (16) for yield-stress phenomena led to the development of a loss model with the following properties: loss predictions at various levels of disease severity, a disease threshold, a maximum level of loss, and large family of loss curve shapes. Five points can be made concerning the development of an ideal model of this type. First, the model should be *general,* ie, correspond to several crops and diseases by simple parameter changes. This generality was demonstrated by the ability of the loss model to represent the previously reported models in the literature and also accurately describe the four crop loss data sets. All of these situations were statistically described by this one model with only three parameters.

Second, an ideal loss model should be *flexible*; ie, account for a variable disease threshold, maximum loss, and many shapes. The general loss model accounts for all of these factors. Disease threshold and shape are described explicitly by the parameters d and s, while the maximum loss is depicted by a combination of the parameters. The disease threshold can range from zero to the minimum X; thus the calculated threshold, d, can indicate a disease level less than that observed in the field.

Third, an ideal loss model should be *expandable* to other stress factors; eg, insects and weeds. Although the analysis in this paper used only a single stress factor (X), the general loss model can be expanded by altering $g(X)$. The original form of $g(X)$ can be written as:

$$g(X) = (X-d)/b = B_0 + B_1X$$

in which $B_0 = -d/b$ and $B_1 = 1/b$. The first expression is more convenient when considering only one stress factor, while the second expression is more useful for increasing the number of factors. The function $g(X)$ can be expanded to:

$$g(X) = B_0 + B_1X_1 + B_2X_2 + B_3X_3 + \cdots$$

in which the Xs are stress variables (factors) and the Bs are parameters. Since $g(X)$ appears as an exponent (equation 1), interactions of the stress variables are incorporated in the model.

Fourth, it should be reasonably easy to estimate the parameters of the ideal loss model. Standard statistical computer packages are not available that will estimate readily the parameters of equation 3. For this reason, a FORTRAN program has been written that operates with FORTRAN G and H, and WATFIV compilers. The program takes less than 5 sec to run on most data sets. The FORTRAN algorithm should be adaptable to many types of minicomputers.

Fifth, the ideal loss model must be *precise*; ie, the model must do a "good" job of summarizing raw data. The coefficient of determination (R^2) is one measurement of a model's precision; R^2 represents the proportion of total variability "explained" by a model. The values of R^2 for the four data sets were all greater than 0.80, which we consider very high for field data of this type.

The loss model as written in equation 2 is not the only model that can describe the crop-loss data of Figs. 2–9. Other equations could be written. Any model, however, should be evaluated on the basis of its generality, flexibility, precision, expandability, and ease with which the parameters may be estimated.

The proposed loss model in its simple form (equation 3) and the other three models reported in the literature (equations 4–6) have been called critical-point models. These models use only one

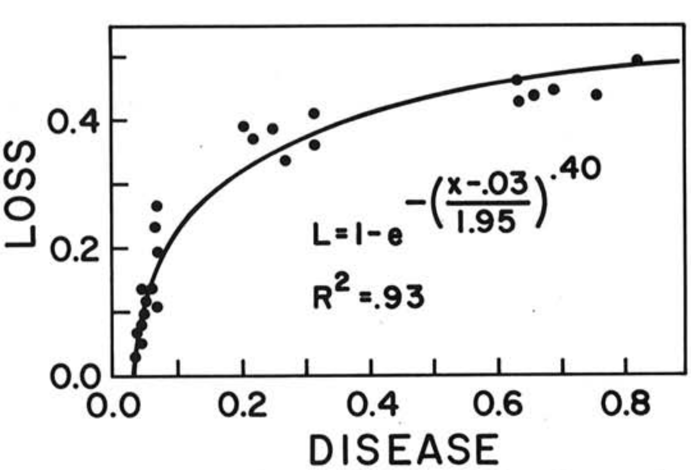

Fig. 7. Peanut losses in relation to final Cercospora leaf spot disease severity during the 1978 growing season. Predicted losses, based on estimated parameters of the proposed loss model (equation 3), are represented by the solid line. Disease severity was estimated 108 days after planting. Loss and disease are expressed on a per unit basis.

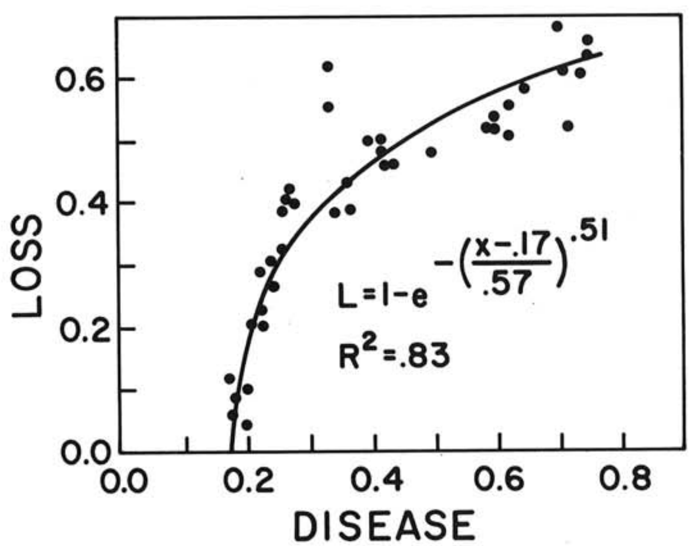

Fig. 8. Potato losses in relation to final early blight disease severity during the 1977–1978 growing season. Predicted losses, based on estimated parameters of the proposed loss model (equation 3), are represented by the solid line. Disease severity was estimated 100 days after planting. Loss and disease are expressed on a per unit basis.

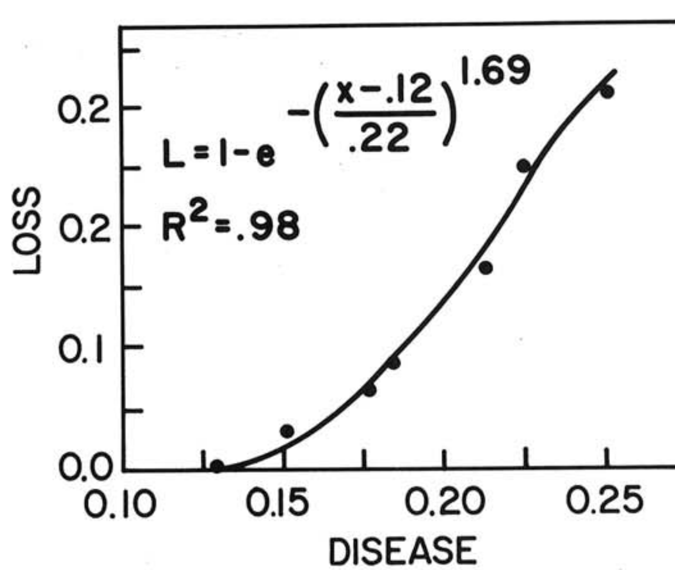

Fig. 9. Potato losses in relation to final early blight disease severity during the 1978–1979 growing season. Predicted losses, based on estimated parameters of the proposed loss model (equation 3), are represented by the solid line. Disease severity was estimated 100 days after planting. Loss and disease are expressed on a per unit basis.

measurement at one time of disease severity. The second major category of loss models is called multiple point. Multiple-point models use disease severity measurements at more than one point in time to predict loss. James and Teng (8) recently stated that the precision of the critical point can almost always be surpassed by a multiple-point model. A comparison of the general loss model with the multiple point will be published.

It is not anticipated that this new loss model will be the most precise or the most realistic statistical model in all situations. It is anticipated, however, that the model will be useful as an analytical tool for research in many crop loss situations. This will enable researchers to compare crops, diseases, and treatments through a direct analysis of the three parameters and the predicted losses at various levels of disease.

LITERATURE CITED

1. Abdel-Hak, T. M., El-Refaei, M. I., Stino, M. N., El-Rafei, M. E., and El-Dauodi, Y. 1979. An appraisal of losses caused by important diseases in major crops in Egypt. Abstract No. 864 in: Abstracts of Papers, IX Int. Congr. Plant Prot., 5–11 August 1979, Washington, D.C.
2. Billman, B. R., Antle, C. E., and Bain, L. J. 1972. Statistical inference from censored Weibull samples. Technometrics 14:831-840.
3. Eyal, Z., and Ziv, O. 1974. The relationship between epidemics of Septoria leaf blotch and yield losses in spring wheat. Phytopathology 64:1385-1389.
4. Gregory, L. V., Ayers, J. E., and Nelson, R. R. 1978. Predicting yield losses in corn from southern corn leaf blight. Phytopathology 68:517-521.
5. Horsfall, J. G., and Cowling, E. B. 1978. Pathometry: The measurement of plant disease. Pages 119–136 in: J. G. Horsfall and E. B. Cowling, eds. Plant Disease, An Advanced Treatise. Vol. II. How Disease Develops in Populations. Academic Press, New York. 436 pp.
6. James, W. C. 1971. An illustrated series of assessment keys for plant diseases, their preparation and usage. Can. Plant Dis. Surv. 51:39-65.
7. James, W. C. 1974. Assessment of plant diseases and losses. Annu. Rev. Phytopathol. 12:27-48.
8. James, W. C., and Teng, P. S. 1979. The quantification of production constraints associated with plant diseases. Appl. Biol. 4:201-267.
9. Kingsolver, C. H., Schmitt, C. G., Peet, C. E., and Bromfield, K. R. 1959. Epidemiology of stem rust. II. Relation of quantity of inoculum and growth stage of wheat and rye at infection to yield reduction by stem rust. Plant Dis. Rep. 43:855-862.
10. Large, E. C., and Doling, D. A. 1962. The measurement of cereal mildew and its effects on yield. Plant Pathol. 11:47-57.
11. Levitt, J. 1972. Responses of Plants to Environmental Stresses. Academic Press, New York. 697 pp.
12. Madden, L. V. 1980. Modeling crop losses due to disease. Ph.D. thesis, The Pennsylvania State University, University Park. 62 pp.
13. Main, C. E. 1977. Crop destruction—The raison d'être of plant pathology. Pages 55–78 in: J. G. Horsfall and E. B. Cowling, eds. Plant Disease, An Advanced Treatise. Vol. I. How Disease is Managed. Academic Press, New York. 465 pp.
14. Mogk, M. 1973. Untersuchungen zur Epidemiologie von *Colletotrichum coffeanum* Noack sensu Hindorf in Kenia, eine Analyse der Wirt-Parasit-Umwelt-Beziehungen. Ph.D. thesis, University of Giessen, W. Germany.
15. Romig, R. W., and Calpouzos, L. 1970. The relationship between stem rust and loss in yield of spring wheat. Phytopathology 60:1801-1805.
16. Tammes, P. M. L. 1961. Studies of yield losses. II. Injury as a limiting factor of yield. Neth. J. Plant Pathol. 67:257-263.
17. Teng, P. S., Close, R. C., and Blackie, M. J. 1979. Comparison of models for estimating yield loss caused by leaf rust (*Puccinia hordei*) on Zephyr barley in New Zealand. Phytopathology 69:1239-1244.
18. Townsend, G. R., and Heuberger, I. W. 1943. Methods of estimating losses caused by diseases in fungicide experiments. Plant Dis. Rep. 27:340-343.
19. Weibull, W. 1951. A statistical distribution function of wide applicability. J. Appl. Mech. 18:293-297.