

## Predicting the Rate of Development of Black Sigatoka (*Mycosphaerella fijiensis* var. *difformis*) Disease in Southern Taiwan

T. Y. Chuang and M. J. Jeger

Associate professor, Department of Plant Pathology and Entomology, National Taiwan University, Taipei, Taiwan, and associate professor, Department of Plant Pathology and Microbiology, Texas Agricultural Experiment Station, College Station, Texas 77843. Present address of second author: Tropical Development and Research Institute, London WC1X 8LU, United Kingdom. Research supported in part by a grant from the National Science Council of Taiwan (ROC). Accepted for publication 7 May 1987 (submitted for electronic processing).

### ABSTRACT

Chuang, T. Y., and Jeger, M. J. 1987. Predicting the rate of development of black Sigatoka (*Mycosphaerella fijiensis* var. *difformis*) disease in southern Taiwan. *Phytopathology* 77:1542-1547.

A multiple regression equation relating disease incidence, accumulated rainfall, and days of RH  $\geq$  90% to the rate of disease increase of banana leaf spot caused by *Mycosphaerella fijiensis* var. *difformis* was developed from 5 yr of field data in three locations of southern Taiwan. The model is  $y = 0.124 - 0.00489 Z_1 + 0.0000462 A_1 + 0.0000515 I_{16}$ , in which  $y$  is the relative rate of disease increase at any given time;  $Z_1$  is disease incidence represented as percentage of leaf infected at the time of the previous assessment;  $A_1$  is the square of  $Z_1$ ; and  $I_{16}$  is a composite variate, the product of  $P_4$  and  $R_1$ ,

*Additional key word:* epidemiology.

where  $P_4$  is the accumulated precipitation in a 4-wk period and  $R_1$  is the number of days with RH  $\geq$  90% in a 1-wk period before the calculated relative rate. The equation accounted for 76% of the total variation in relative rate and performed best at two of the three locations and in years where there was some fluctuation in rate. Model selection and validation were based on the use of different regression analysis techniques including Mallows's  $C_p$  statistics, the variance inflation factor, and methods of data splitting.

Banana leaf spot diseases of economic importance include Sigatoka, black leaf streak, and black Sigatoka, caused by three very closely related pathogens: *Mycosphaerella musicola* Leach ex Mulder, *M. fijiensis* Morelet, and *M. fijiensis* var. *difformis*, respectively. The disease most predominant in Taiwan is black Sigatoka, caused by the most virulent of the three pathogens (14). Usually the disease first appears in early May and becomes severe throughout the summer season under conditions of high temperature and humidity in southern Taiwan. Field trials have shown that banana yields increase 4–7 kg per bunch, about 18.4–31.7% of the average yield, in sprayed compared with unsprayed plots (5). Commercial banana plantations receive about 8–10 fungicide applications from July to December each year to control the disease.

The epidemiology of black Sigatoka has not been studied in detail with respect to climatic factors. Based on field observations, Stover (22) suggested that black Sigatoka may respond to moisture and temperature in a similar way to Sigatoka, which occurs severely under conditions of high moisture and temperature. Younger leaves are more susceptible to infection than the older leaves (12,21). Several attempts have been made to develop a predictive system for the Sigatoka disease that relates change in disease incidence to weather (3,8,9), so that fungicide spray schedules may be timed appropriately. None of these disease forecasting systems, however, consider the amount of inoculum available if weather becomes favorable for infection. In addition, the relationship between rate of disease increase and climatic and other factors has not been quantified.

The rate of disease increase is a useful parameter to describe and quantify disease development (23). The rate may be used to predict future disease progress and determine whether fungicide spray schedules will achieve disease control effectively and economically. The rate is likely to depend on certain historical patterns of weather, host, and disease development (1,17,18). There is also a certain inertia in the response of the rate to the immediate climatic conditions. This is both a strength and a weakness in the use of the rate as a predictive parameter; but if used judiciously with damage

and economic thresholds (24), it provides a sound basis for tactical disease management decisions. In this paper, we use multiple regression analysis with specified constraints to select a best model for representing the relationship between the rate of disease increase and independent variables encompassing climatic factors, host growth, and disease incidence and severity. The relationship between incidence and severity of black Sigatoka was evaluated in a companion paper (6). Unlike most modeling studies, we also consider those conditions that lead to a decrease in disease (i.e., a rate that is negative). As pointed out by Kranz (10), the decline phase of epidemics should be of equal interest to epidemiologists, especially with tropical crops and diseases.

The methodology used follows closely that outlined recently by Coakley et al (7) and emphasizes procedures for developing, selecting, and validating a regression model.

### MATERIALS AND METHODS

**Experimental.** Data were collected every 2–4 wk from commercial banana plantations in southern Taiwan for 5 yr from 1980 through 1984 as described by Chuang and Jeger (6). Three locations, WEI, KAN, and RING, and the period from July to December in each year were selected for the analysis because of severe disease in these locations and the standard practices of fungicide treatments at the locations during these months. There were no unsprayed areas in these commercial plantations that could be used for observations. Fifty-one observations were made at each location over 5 yr. The geographical characteristics of these three locations, the cultural practices used, and disease assessment protocols were described previously (6). In addition to disease incidence (proportion of diseased leaves) and severity (calculated according to Stover's international scale [20]), records of leaf production and numbers of healthy leaves were made.

**Weather data collection.** A meteorological station located at the Chisan sugarcane factory was within a distance of 5 km from each of the three sites where disease assessments were made. Data collected daily for 1980–1984 were available and complete. These data included maximum and minimum temperature, relative humidity, and precipitation recorded at 0900 every day.

**Selection of independent variates.** Although the effects of weather on black Sigatoka have not been studied, the effects of

rainfall, relative humidity, and minimum temperature on Sigatoka disease have been reviewed (12,21). Maximum temperature is not important for infection by the black Sigatoka pathogen. Therefore, data of minimum temperature, days of RH  $\geq$  90%, and total amount and frequency of precipitation were selected as meteorological variables for the study. In addition, as the susceptibility of the leaf to the pathogen is closely related to leaf age, the change in leaf number during a given period, and the number of healthy leaves were also considered as host variables that may affect disease progress. Disease variables of incidence and severity were included as independent variables representing the source of inoculum available for infection. The first small lesion of black Sigatoka appears about 10–14 days after infection given favorable conditions, but symptom development may be delayed by as much as 30 days during a dry, cool season (12,21). Therefore, each weather, host, and disease variable was constructed over various intervals in a 4-wk period before the reference time when predictions for the dependent variables were made. Independent variates used in the study are shown in Table 1, and included three classes of variables corresponding to environment, host, and disease. Each independent variable was constructed to represent various intervals before the date of assessment of the dependent variable. These a priori constraints were set to avoid simultaneity in time between independent and dependent variables and to provide time for decisions on modified disease control options to be made.

**Selection of dependent variates.** Dependent variables were chosen to represent changes in disease incidence, either as discrete increments or as calculated rates of disease development. Disease increments were expressed in absolute and relative terms (Table 2).

**Model development, selection, and validation.** Multiple regression analysis was used to examine the influence of environment, host, and disease (as independent variables) on subsequent changes, either increments or rates of disease (as dependent variables). Model development followed several distinct stages with constraints to reduce the number of models examined and to provide an adequate means of selecting from and validating the smaller number that remained. In the early stages of model development, the SAS STEPWISE procedure (16) was used to screen for important independent variables.

To help select the regression models, the proportion of variance accounted for, Mallow's  $C_p$ , and the variance inflation factor (VIF) were used as criteria. The proportion of variance accounted for (total mean square-residual mean square)/(total mean square) takes into account the number of independent variables fitted, which is critical when comparing multiple regression equations. Mallow's  $C_p$  is used as a criterion for the goodness of fit of regression equations with different numbers of independent variables. A model is less subject to bias when the  $C_p$  value is closest

TABLE 1. Summary of independent variables used for analyzing the relationship between the rate of disease increase and environmental, host, and disease variables for banana leaf spot in southern Taiwan

Variable	Symbol
Environmental variable <sup>a</sup>	
Mean minimum temperature	$T(1-7)$
Accumulated precipitation	$P(1-7)$
Days of precipitation	$D(1-7)$
Days of RH $\geq$ 90%	$R(1-7)$
Host variable <sup>a</sup>	
Number of leaf change	$L(1-7)$
Number of healthy leaves	$H(1-7)$
Disease variable <sup>b</sup>	
Disease incidence	$Z(1-7)$
Disease severity	$W(1-7)$

<sup>a</sup> Numbers (1–7) represent data from the 1st, 2nd, 3rd, and 4th week and the 2-, 3-, and 4-week intervals, respectively, before the date of disease recorded.

<sup>b</sup> Numbers (1–7) represent previous assessments at time  $t-1$ ,  $t-2$ ,  $t-3$ , and  $t-4$ , and the mean of assessments at times  $(t-1)+(t-2)$ ,  $(t-1)+(t-2)+(t-3)$ , and  $(t-1)+(t-2)+(t-3)+(t-4)$ , respectively.

to the number of parameters in the model (4,11). The VIF is used to measure the effect of multicollinearity between independent variables on the variance of estimated coefficients and model stability. If a VIF is larger than 5, it may indicate that the associated coefficient is poorly estimated (13,19).

A form of data splitting was then used to validate the selected model. One observation was removed from each year at each location with random numbers; this resulted in three observations removed each year from the pooled data. The remaining data ( $n-3$ ) were then used to develop a new regression equation (based on the same variables), which was used to predict the rates of disease increase that had been removed from the data set. The procedures were repeated for each of the 5 yr, and the 15 predicted values obtained were compared with the corresponding observed values. Additionally, the proposed model was used to predict values for each separate location over 5 yr and for each year over the three locations.

TABLE 2. Summary of dependent variables used for analyzing the relationship between the rate of disease increase and meteorological, host, and disease variables for banana leaf spot in southern Taiwan

Absolute increments	Relative increments	Rates of disease increase
$Y_t - Y_{t-1}$ <sup>a</sup>	$(Y_t - Y_{t-1})/Y_{t-1}$	$\ln(Y_t - Y_{t-1})/\Delta t$
$(Y_t - Y_{t-1})/\Delta t$	$(Y_t - Y_{t-1})/\Delta t \cdot Y_{t-1}$	$\ln[Y_t(1 - Y_{t-1})/Y_{t-1}(1 - Y_t)]/\Delta t$

<sup>a</sup>  $Y$  = Disease incidence;  $\Delta t$  = number of days between time interval  $t$  and  $t-1$ .

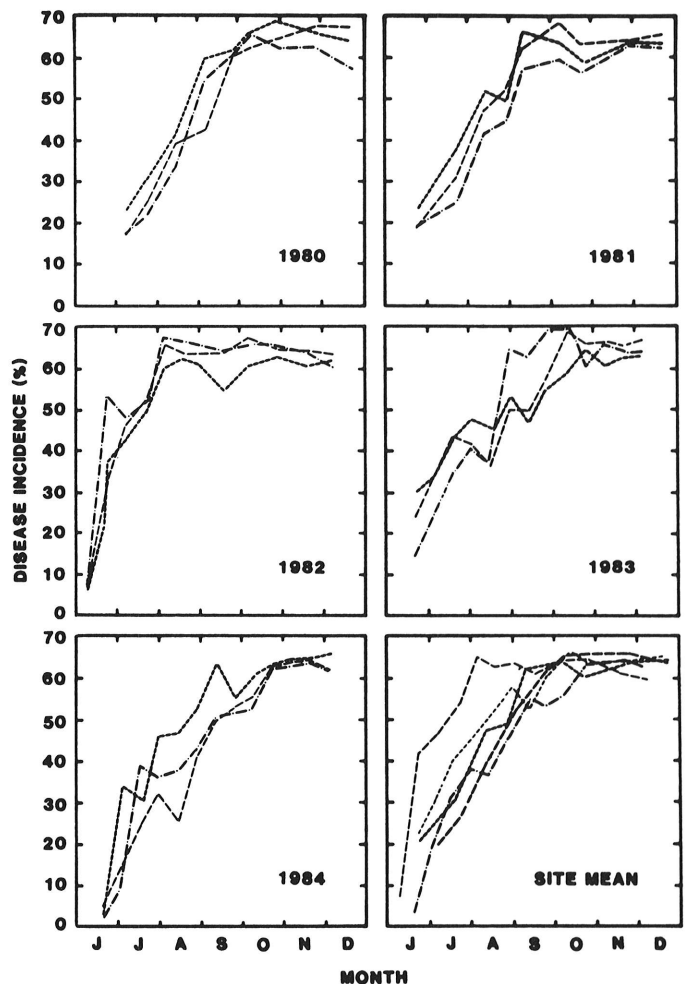


Fig. 1. Disease incidence of black Sigatoka in southern Taiwan at three locations WEI (---), KAN (-.-) and RING (· · ·) in 1980–1984. Site means are shown for 1980 (—), 1981 (----), 1982 (---), 1983 (----), and 1984 (—).

## RESULTS

**Model development.** Disease incidence at each site is plotted in Figure 1 for each year of the study. Models to predict the rates of disease increase were developed in several stages. In the first stage, each dependent variable was regressed on the 28 environmental variables and the seven host variables giving the change in leaf number over respective time intervals. The SAS STEPWISE procedure was used throughout these preliminary stages. The  $R^2$  values of the best equations were used as a guide to indicate the level of precision that could be attained for any given dependent variable, rather than as a basis for comparing equations. More appropriate statistics for that purpose were used at later stages of the analysis. The  $R^2$  values ranged from 0.11 to 0.61, 0.29 to 0.51, and 0.36 to 0.48, respectively, at locations WEI, KAN, RING, and all locations pooled. There was no consistency among sites in the independent variables that were selected for the best

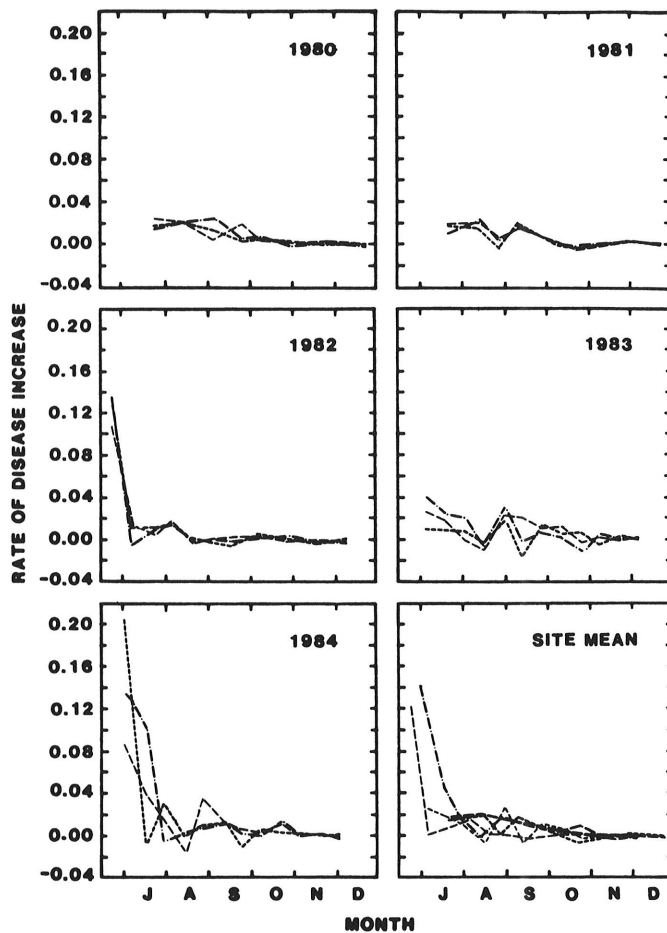


Fig. 2. The relative rate of disease increase of black Sigatoka in southern Taiwan, calculated as  $\ln(Y_t/Y_{t-1})/\Delta t$ , where  $Y$  is disease incidence,  $\Delta t$  gives the time between assessments  $t$  and  $t-1$ . Line (---) represents WEI; (—), KAN; and (---), RING in 1980–1984. Site means are shown for 1980 (---), 1981 (----), 1982 (---), 1983 (-----), 1984 (—).

equations. There were more independent variables selected for pooled data (8–11) than for any given site (1–7). In the second stage of the analysis the number of dependent variables was reduced; disease incidence, severity, and number of healthy leaves were included as independent variables. Again with the STEPWISE procedure, there was an improvement in the  $R^2$  values that could be attained for any given dependent variable, reaching 0.75 but with 13–15 independent variables required to reach this level of precision. The variables that regularly were selected in the STEPWISE procedure were  $Z_1$ ,  $H_1$ ,  $L_3$ ,  $T_2$ ,  $P_4$ , and  $R_1$ , where the symbols and subscript are those given in Table 1. In no analysis did disease severity, as an independent variable, rather than disease incidence, improve the  $R^2$  value of an equation. It is noteworthy that these independent variables represented each class of measurement: environment, host, and disease (Table 1). These variables, their squared terms, and all possible pairwise products were used as independent variables and analyzed further, again using STEPWISE. Dependent variables were restricted to the incremental changes and rates of disease increase shown in Table 3.

The model for the relative rate of disease development,  $\ln(Y_t/Y_{t-1})/\Delta t$ , had the highest  $R^2$  and contained only four independent variables (Table 3); only models for this dependent variable were developed further. Plots of the relative rate of disease increase at each location for each year are shown in Figure 2. The rates of increase were initially much greater in 1982 and 1984, and to a lesser extent in 1983, than in 1980 and 1981. The frequency distributions of the relative rate and selected independent variables are shown in Figure 3. The distribution of the rate was slightly skewed, although there were very few high rates. The distribution of precipitation was skewed, although there were a large number of zero classes. The distribution of days  $RH \geq 90\%$  had a larger number of zero entries; thus, the composite variate of their product ( $I_{16}$ ) is essentially a  $\pm$  type of influence. The composite variate  $I_{10}$  is the product of the mean number of healthy leaves in the preceding 4-wk period and the mean minimum temperature 2 wk before the reference time; this variate was approximately normally distributed. The distribution of incidence was interesting in that, except for one incidence class, squared incidence followed a uniform distribution.

The intercorrelations among the variables are shown in Table 4. All independent variables, except  $I_{20}$ , were significantly correlated with the dependent variable. The variable  $I_{20}$  was removed and the remaining variables were used to obtain all combinations of one-, two-, three-, and four-variable regression models, using more stringent criteria to determine the best model.

**Model selection.** Selection of the best model was based on criteria shown in Table 5. There were no one- or two-variable models with  $C_p < p$ . There was only one three-variable model (J), which accounted for a high proportion of variance and  $C_p < 3$ ; the four-variable model accounted for 0.76 of the variance and  $C_p = 3.0$ . However, when the VIF was used as a criterion to evaluate multicollinearity between variables, based on the variance of estimated coefficients, it showed high collinearity for those models including the variable  $A_1$  ( $VIF > 5$ ). This was not surprising as  $A_1$  is the square of  $Z_1$ , and although mathematically they are independent, there was high correlation between them over the range observed (Table 4.) There was no improvement in the four-variable model (K) over the three-variable model, and the model had slightly higher standard errors in the coefficients estimated, indicating that  $I_{10}$  was not contributing significantly to the model.

TABLE 3. Independent variables, consisting of single variables, squared terms, and pairwise products associated with each dependent variable selected by using the SAS STEPWISE procedure (at  $F = 0.15$ ) for pooled data of banana leaf spot at three locations in southern Taiwan

Dependent variable	Independent variable <sup>a</sup>	$R^2$
$Y_t - Y_{t-1}$	$Z_1, I_{15}, I_8, P_4, I_{19}, A_7, I_{13}, A_3, I_{11}, I_{16}$	0.69
$(Y_t - Y_{t-1})/\Delta t$	$Z_1, I_{16}, I_{15}, I_8, A_6, P_4, A_7, I_{11}, I_{10}$	0.69
$\ln(Y_t/Y_{t-1})/\Delta t$	$Z_1, A_1, I_{16}, I_{20}$	0.77
$\ln[Y_t(1 - Y_{t-1})/Y_{t-1}(1 - Y_t)]/\Delta t$	$Z_1, A_1, I_{16}, I_{20}, I_{10}$	0.71

<sup>a</sup>  $A_1 = Z_1^2$ ;  $A_3 = L_3^2$ ;  $A_6 = T_2^2$ ;  $A_7 = D_2^2$ ;  $I_8 = H_7 \times P_4$ ;  $I_{10} = H_7 \times T_2$ ;  $I_{13} = L_3 \times R_1$ ;  $I_{15} = L_3 \times D_2$ ;  $I_{16} = P_4 \times R_1$ ;  $I_{19} = R_1 \times T_2$ ;  $I_{20} = R_1 \times D_2$ , where all individual variables are described in Table 1.

The linear equation that described the three-variable model was:

$$y = 0.124 - 0.00489 Z_1 + 0.0000462 A_1 + 0.0000515 I_{16} \quad (1)$$

in which  $y$  = relative rate of disease increase represented as  $\ln(Y_t/Y_{t-1})/\Delta t$  at any given time;  $Z_1$  = disease incidence represented as percentage of leaves infected at the time of the previous assessment;  $A_1 = Z_1^2$ ; and  $I_{16}$  = a composite

meteorological variate of  $P_4$  and  $R_1$ , where  $P_4$  is the accumulated precipitation 4 wk before the estimated rate and  $R_1$  is the days of RH  $\geq 90\%$  1 wk before the estimated rate. This equation is proposed as the best one describing the relationship of the relative rate of disease increase with disease, environmental, and host variables at some previous time. About 76% of total variance in the relative rate of disease increase was accounted for by this linear model.

TABLE 4. Correlation matrix of independent<sup>a</sup> and dependent variables used in regression analysis of rate of disease increase of banana leaf spot in Southern Taiwan<sup>b</sup>

	$Z_1$	$A_1$	$I_{10}$	$I_{16}$	$I_{20}$	$\ln(Y_t/Y_{t-1})/\Delta t$
$Z_1$	1.000	0.977 (0.001) <sup>a</sup>	-0.736 (0.001)	-0.193 (0.017)	-0.196 (0.015)	-0.710 (0.001)
$A_1$		1.000	-0.717 (0.001)	-0.198 (0.014)	-0.225 (0.005)	-0.592 (0.001)
$I_{10}$			1.000	0.021 (0.789)	0.046 (0.577)	0.540 (0.001)
$I_{16}$				1.000	0.750 (0.001)	0.279 (0.001)
$I_{20}$					1.000	0.136 (0.095)
$\ln(Y_t/Y_{t-1})/\Delta t$						1.000

<sup>a</sup>  $A_1 = Z_1^2$ ;  $I_{10} = H_7 \times T_2$ ;  $I_{16} = P_4 \times R_1$ ;  $I_{20} = R_1 \times D_2$ ; where all individual variables are defined in Table 1.

<sup>b</sup> Values in parentheses give the significance levels.

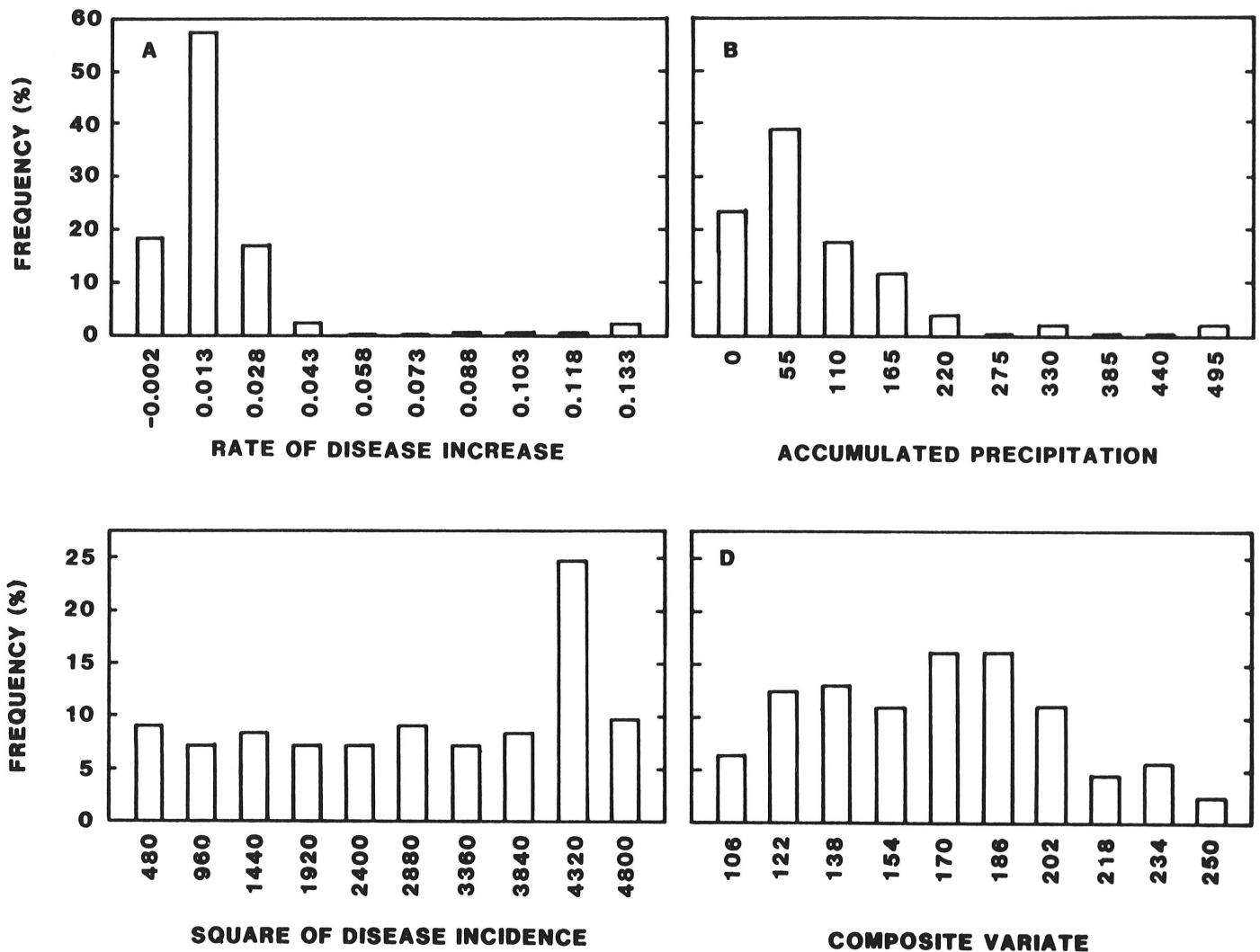


Fig. 3. Frequency distribution of dependent and independent variables selected for regression analysis: A, Relative rate of disease increase of black Sigatoka; B, Accumulated precipitation at the 4th wk before disease assessment; C, Square of disease incidence at the time of the previous assessment; D, Composite meteorological variate  $H_7 T_2$ , where  $H_7$  is the average number of healthy leaves in the 4 wk before the date of disease assessment, and  $T_2$  is the mean of daily minimum temperatures in the 2nd wk before the date of disease assessment.

**Model validation.** To evaluate further the applicability of this linear model, the pooled data of WEI, KAN, and RING were separated by locations or years. The predicted values for each site or year, using the model, were regressed on the actual values observed. Results are shown in Table 6. A slope of unity would indicate a good relationship between predicted and observed values. The  $R^2$  values for the linear regression of predicted against observed values were 0.83 and 0.81 in WEI and KAN, respectively, but only 0.19 for RING. The data of RING were removed from the pooled data, and a new multiple regression equation was generated using the same variables as equation 1; this was then used to predict the actual values, but the accuracy was not improved. When we compared the restricted model with equation 1, the coefficients were not significantly different, although the standard errors of the estimates were always higher for the two-site model.

The accuracy of prediction with the proposed model also varied according to years (Table 6). Higher  $R^2$  values were obtained in 1982, 1983, and 1984, than in 1980 and 1981. The epidemics in

1982, 1983, and 1984 were characterized by a lower minimum disease incidence and by a higher maximum rate of disease increase; the epidemics in 1980 and 1981 were characterized by higher minimum disease incidence and by a lower maximum rate of disease increase (Table 7). In the later years (1982–1984), there was also a much wider range (max–min) in the rate of disease increase. These results indicate that the proposed model may be more useful when the disease incidence is low and relative rates correspondingly higher during the early stage of an epidemic.

Finally, three observations were removed randomly from each year, one for each location, and the remaining observations were then used to develop a new regression equation to predict the removed values. With three locations and 5 yr, this gave 15 pairs of predicted and observed values to be compared. There was no significant difference ( $P = 0.05$ , based on Student's  $t$  test) between the predicted and observed values. The observed values were each within one standard error of the predicted values in 14 of 15 observations.

## DISCUSSION

A model for describing the relationship between the relative rate of disease increase and preceding independent variables of disease incidence and weather (accumulated precipitation, and days of RH  $\geq 90\%$ ) was developed in this study. The regression model provided good predictions for locations WEI and KAN and good predictions for 1982–1984 especially. The model accounted for about 76% of the variation in the observed rate of disease increase. Although this is a relatively high value for this kind of study, it indicates that some important factors may still not be accounted for, probably more sensitive assessments of available inoculum, amount of susceptible host tissue, and periods of leaf wetness. Stover (21) suggested that the amount of leaf spotting (disease severity) may be correlated with available inoculum, but we know of no evidence to support this. Disease incidence was more important as an independent variable than severity but may inadequately represent the inoculum available for infection. Spore production, germination, and host penetration by the pathogen only occur when the leaf surface has a film of free water or is at saturated or near-saturated conditions (12,21). No information on leaf wetness at the three sites was available at the time of the study, however, and this is a major defect in the development and selection of the model.

Our previous results (6) indicated that microenvironmental factors may have subtle effects at these locations and play an important role in black Sigatoka epidemiology. Weather data were not available at each location, and the lack of detailed information for each site may account for the poor predictions for RING.

Given the limitations on the development of the model, however, the results of the study showed that the coefficients of the regression model were quite stable. When the data of RING were removed (51 out of 153 observations) and a restricted regression equation was developed or three observations were removed randomly in each year at each location (15 out of 153 observations) and new regression equations were developed, the coefficients were not significantly different to those of equation 1.

Based on the criteria used for model selection and validation, the proposed model proved satisfactory except for the correlation between disease incidence and its squared term. Removing the squared term, however, reduced the predictive power of the model. Butt and Royle (2) pointed out that if a proposed model is used for predicting the development of an epidemic, the principal requirements are to account for a high proportion of variance and the reliability of the equation in future. Based on these pragmatic considerations, we believe the proposed model is adequate for predicting the rate of disease increase of banana leaf spot.

This proposed model has at least two advantages: First, prediction can be made for each site by using only the disease incidence of previous assessment and a composite variate of accumulated rainfall and days of RH  $\geq 90\%$ . This procedure is simple, and all variables can be calculated or measured easily using simple apparatus. Second, because disease incidence is measured

TABLE 5. The proportion of variance accounted for, Mallow's  $C_p$ , and the variance inflation factor (VIF) for models predicting the relative rate of disease increase of banana leaf spot in southern Taiwan from different combinations of selected independent variables

Model variable <sup>a</sup>	Proportion of variance	$C_p$	VIF
A $Z_1$	0.50	168.4	1.0
B $A_1$	0.35	267.7	1.0
C $I_{10}$	0.29	305.9	1.0
D $I_{16}$	0.07	154.9	1.0
E $Z_1, A_1$	0.74	18.5	22.5
F $Z_1, I_{10}$	0.50	170.0	2.2
G $Z_1, I_{16}$	0.52	156.9	1.0
H $Z_1, I_{10}, I_{16}$	0.52	157.1	2.3
I $Z_1, A_1, I_{10}$	0.74	22.4	22.9
J $Z_1, A_1, I_{16}$	0.76	2.5	22.5
K $Z_1, A_1, I_{10}, I_{16}$	0.76	3.0	22.9

<sup>a</sup>  $A_1 = Z_1^2$ ;  $I_{10} = H_7 \times T_2$ ;  $I_{16} = P_4 \times R_1$ ;  $I_{20} = R_1 \times D_2$ ; where all individual variables are defined in Table 1.

TABLE 6. Linear regression of predicted rate of disease increase on actual rate of disease increase, using regression equations generated from pooled data for the WEI, KAN, and RING sites or for the WEI and KAN sites only for banana leaf spot in Southern Taiwan<sup>a</sup>

Location	WEI KAN RING		WEI KAN	
	$R^2$	Slope	$R^2$	Slope
WEI	0.83	0.62 (0.040)	0.83	0.70 (0.045)
KAN	0.82	0.80 (0.055)	0.71	0.73 (0.067)
RING	0.19	0.11 (0.048)	No predicted data	
Year				
1980	0.50	1.42 (0.303)	0.41	1.43 (0.456)
1981	0.39	1.09 (0.290)	0.35	1.10 (0.404)
1982	0.93	0.83 (0.134)	0.94	0.89 (0.049)
1983	0.61	0.97 (0.134)	0.63	1.09 (0.177)
1984	0.82	0.71 (0.056)	0.91	0.74 (0.051)

<sup>a</sup> Standard errors on slope coefficients are shown in parentheses.

TABLE 7. Comparison of disease incidence and rate of disease incidence for banana leaf spot from 1980 to 1984 for the pooled data of three locations in southern Taiwan

Year	Disease incidence			Rate of disease increase		
	Mean	Maximum	Minimum	Mean	Maximum	Minimum
1980	0.508	0.688	0.176	0.0078	0.0229	-0.0026
1981	0.509	0.682	0.189	0.0063	0.0235	-0.0049
1982	0.532	0.673	0.063	0.0138	0.1344	-0.0073
1983	0.515	0.698	0.146	0.0063	0.0401	-0.0167
1984	0.488	0.656	0.028	0.0198	0.2063	-0.0148

2–4 wk before the date of the predicted relative rate, decisions can be made early enough to permit use of a fungicide. Values for the composite environmental variate  $I_{16}$  may then be used to modify the predicted rate, if necessary.

Plaut and Berger (15) studied infection rates in three diseases and concluded that low initial disease was compensated for by faster rates of disease increase. Our results confirm that general observation. The proposed model showed that under favorable conditions, disease increased at a much faster rate at low initial disease incidences than at high initial disease incidences. Our model also enables a prediction of when disease will cease to increase and possibly decrease. Setting the equation to zero and solving for  $Z_1$ , in cases where the environment is also unfavorable and hence  $I_{16}$  is zero, shows that the rate is negative at incidence values of about 54%. Any nonzero value for the environmental variate  $I_{16}$  will, of course, increase the rate. The predictions made with the proposed model gave more accuracy at low disease incidences and higher relative rates at the onset of an epidemic. This further supports use of the proposed model to assist earlier decisions concerning use of fungicide sprays when these are most needed. Rate prediction can be used in conjunction with the incubation period of this disease, about 10–14 days under favorable conditions (22). Further research, therefore, on the timing of application and postinfection effectiveness of fungicides is needed to improve further our ability to manage this disease.

An attempt can also be made to interpret the regression model in terms of epidemic dynamics. The dependent variable is the mean relative rate of disease increase, defined as  $(1/Z)dZ/dt$ , during any time interval  $\Delta t$ . The independent variables are a quadratic function of disease incidence and an additional term representing an environmental influence of 'wet' conditions.

The regression model can then be written as a differential equation of the form

$$dZ/dt = rZ[1 - f(Z)] + g(Z, E) \quad (2)$$

where  $r$  is given by the intercept of the regression model,  $f(\cdot)$  is a quadratic function of  $Z$ , and  $g(\cdot)$  is a function of both  $Z$  and environment  $E$ . This equation is similar to the logistic equation, but with a more complex correction term  $1 - f(Z)$  and an additional term  $g(Z, E)$  representing variation caused by environment and its interaction with disease.

#### LITERATURE CITED

1. Analytis, S. 1973. Methodik der Analyse von Epidemien dargestellt am Apfeschorf *Venturia inaequalis* (Cook) Aderh. Acta Phytomedica 1:1-76.
2. Butt, D. J., and Royle, D. J. 1974. Multiple regression analysis in the epidemiology of plant disease. Pages 78-114 in: Epidemics of Plant Disease: Mathematical Analysis and Modeling. J. Kranz, ed. Springer-Verlag, New York.
3. Calpouzos, L., Colberg, C., Riollano, A., Ramos, C., and Theis, T. 1964. The control of Sigatoka disease of bananas by oil spray schedules based on rainfall data. J. Agric. Univ. P. R. 48:32-38.
4. Chatterjee, S., and Price, B. 1977. Regression Analysis by Example. John Wiley & Sons, New York.
5. Chuang, T. Y. 1981. Chemical control of banana leaf spot caused by *Mycosphaerella fijiensis* var. *difformis*. Plant Prot. Bull. (Taiwan) 23:87-94.
6. Chuang, T. Y., and Jeger, M. J. 1987. Relationship between incidence and severity of banana leaf spot in Taiwan. Phytopathology 77:1537-1541.
7. Coakley, S. M., McDaniel, L. R., and Shaner, G. 1985. Model for predicting severity of Septoria tritici blotch on winter wheat. Phytopathology 75:1245-1251.
8. Guyot, H., and Cuille, J. 1958. Essai de prevision des attaques de Cercospora en Guadalupe. Fruits 13:85-94.
9. Klein, H. H. 1960. Control of Cercospora leaf spot of bananas with application of oil sprays based on the disease cycle. Phytopathology 50:488-490.
10. Kranz, J. 1978. Comparative anatomy of epidemics. Pages 33-62 in: Plant Disease: An Advanced Treatise Vol. 2. How Disease Develops in Populations. J. G. Horsfall and E.B. Cowling, eds. Academic Press, New York.
11. Mallow, C. L. 1975. Some comments on  $C_p$ . Technometrics 15:661-675.
12. Meredith, D. S. 1970. Banana Leaf Spot Disease (Sigatoka) Caused by *Mycosphaerella musicola* Leach. No. 11. Commonw. Mycol. Inst., Kew, Surrey, England. 147 pp.
13. Montgomery, D. C., and Peck, E. A. 1982. Introduction to Linear Regression Analysis. John Wiley & Sons, New York. 504 pp.
14. Mulder, J. L., and Stover, R.H. 1976. *Mycosphaerella* species causing banana leaf spot. Trans. Br. Mycol. Soc. 67:77-82.
15. Plaut, J. L., and Berger, R. D. 1981. Infection rates in three pathosystem epidemics initiated with reduced disease severities. Phytopathology 71:917-921.
16. SAS Institute Inc. 1985. SAS User's Guide: Statistics. Version 5 Edition. Cary, NC. 956 pp.
17. Schroeder, H., and Ullrich, J. 1965. Untersuchungen zur Biometeorologie und Epidemiologie von *Phytophthora infestans* (Mont.) de By. auf mathematisch-statistischer Grundlage. Phytopathol. Z. 54:87-103.
18. Schroeder, H., and Ullrich, J. 1967. Eine Mathematisch-Statistische Lösung des Problems der Prognose von Epidemien mit Hilfe Meteorologischer Parameter, dargestellt an Beispiel der kartoffelkrautfaule (*Phytophthora infestans*). Agric. Meteorol. 4:119-135.
19. Snee, R. D. 1977. Validation of regression models: Methods and examples. Technometrics 19:415-428.
20. Stover, R. H. 1971. A proposed international scale for estimating intensity of banana leaf spot. Trop. Agric. (Trinidad) 48:185-196.
21. Stover, R. H. 1972. Banana, plantain and abaca diseases. Commonw. Mycol. Inst. Kew, Surrey, England. 316 pp.
22. Stover, R. H. 1980. Sigatoka leaf spot of bananas and plantains. Plant Dis. 64:750-756.
23. Vanderplank, J. E. 1963. Plant Diseases: Epidemics and Control. Academic Press, New York. 349 pp.
24. Zadoks, J. C., and Schein, R. D. 1979. Epidemiology and Plant Disease Management. Oxford University Press, Oxford. 427 pp.