



Pm 360

Fonds Documentaire IRD

Cote: B\* 22174 Ex: Antiquaire

## Ecology and Epidemiology

**Analysis of Temporal Disease Progress of African Cassava Mosaic Virus**

D. Fargette, M. Jeger, C. Fauquet, and L. D. C. Fishpool

First and third authors, Laboratoire de Phytovirologie, ORSTOM, Adiopodoumé BP V 51, Abidjan, Ivory Coast, Africa; second and fourth authors, NRI, Chatham, Kent ME4 4TB, United Kingdom.

Present address of the first author to whom correspondence should be sent: LPRC, CIRAD, BP 5035, Montpellier cedex, 1 France. We are grateful to B. Monteny and J. P. Raffailac for helpful discussions, to J. McNicol for advice on the statistics, and to J. M.

Thresh for constructive criticisms and detailed reviews of the manuscript. This work was supported in part by the Commission of the European Communities grants TSD-102 and TS2A-0137-C (CD).

Accepted for publication 7 September 1993.

**ABSTRACT**

Fargette, D., Jeger, M., Fauquet, C., and Fishpool, L. D. C. 1994. Analysis of temporal disease progress of African cassava mosaic virus. *Phytopathology* 84:91-98.

Forty-nine disease progress curves of African cassava mosaic virus recorded in Ivory Coast (West Africa) of monthly plantings between 1981 and 1986 were analyzed. The Gompertz model was the most appropriate to describe the epidemics, and analysis of the parameters of the fitted models indicated that the maximum rate of disease increase was reached an average of 2 mo after planting and that the rate of disease progress has a seasonal component. There was a large increase in disease incidence

from November to June and a relatively small increase between July and October. About 70% of the variation was related to changes in whitefly numbers and to fluctuations in temperature and radiation. Other possible causes were changes in whitefly activity, virus concentration in plant reservoirs, and plant susceptibility to infection. By contrast, in this tropical, humid climate with a short dry season, the impact of the rain-induced parameters was limited. Whatever the overall disease incidence, a reduction in the rate of spread with age occurred.

*Additional keywords:* epidemiology, geminivirus, Africa.

African cassava mosaic virus (ACMV) causes a widespread disease of cassava in Africa and is responsible for serious yield losses. The pathogen is a geminivirus that is transmitted by the whitefly

*Bemisia tabaci* (Homoptera: Aleyrodidae) and is disseminated through infected cuttings (9). At Adiopodoumé, in the southern part of Ivory Coast (West Africa), there is a high rate of infection by whiteflies of healthy cassava fields throughout the year. The spatial pattern of spread has been reported (6,7). We analyzed

the pattern of change of disease incidence with time. Special emphasis was placed on the primary spread (caused by inoculum introduced from outside the plot), because earlier studies showed that primary spread was overwhelming (7) compared to the secondary spread (caused by inoculum from inside the plot).

Disease incidence was monitored in a series of 49 sequential plantings over 6 yr, and the results were analyzed. First, inflexible models of the family of growth curves typified by the generalized logistic model were fitted to disease progress curves in order to summarize, compare, and classify the epidemics and to bring out features that were not obvious from the data alone (11,13,15). Then, changes of disease incidence with age and season were characterized, and evidence was obtained on the biotic and abiotic factors involved, in particular on the respective roles of cassava age, whitefly populations, and climatic variables.

## MATERIALS AND METHODS

**Experimental trials.** Between May 1981 and May 1986, 49 plantings of healthy cassava were made at the ORSTOM experimental station at Adiopodoumé (20 km west of Abidjan in the lowland forest zone of Ivory Coast). The infection of each plot was monitored. Organic fertilizer was applied before planting, and those plots planted from December to February inclusive (the dry season) usually had to be irrigated. In 1981 and 1982, each planting consisted of 10 plots of 100 plants each, with plants spaced at 1 × 1 m. Successive plantings were made next to each other along a south-north orientation, each separated by a 10-m gap. From April 1983, plantings were made monthly and comprised seven plots of 100 plants each, oriented along a southwest-northeast axis (Fig. 1). Each planting was isolated from diseased cassava fields by at least several tens of meters.

The cassava variety CB, which was used throughout the trials, is considered to be moderately susceptible to ACMV (C. Fauquet and D. Fargette, unpublished) and, when infected, exhibits clear

mosaic symptoms. Healthy cuttings were obtained from the Toumodi experimental station, 200 km north of Abidjan in the Guinea savannah zone. All trials were inspected shortly after planting, and any infected cuttings (i.e., those showing mosaic symptoms on the first leaves) were removed and replaced by noninfected plants of the same age taken from a reserve plot. This was done so that all later infection would result exclusively from whitefly transmission.

**Surveys.** The virus was introduced by viruliferous whiteflies that invaded the field. Disease incidence was monitored fortnightly in 1981 and 1982 and weekly thereafter. After disease incidence was recorded, diseased cassava plants were removed to eliminate most of the secondary spread. Because symptom expression in the CB variety is closely associated with virus content (8), this regular removal of diseased plants efficiently inhibited secondary spread. The resulting disease progress curves were studied in several steps.

**Analysis of the disease progress curves.** The following models were tested to fit the observed disease progress curves. All analyses were made with the statistical software Genstat V (18).

Ordinary exponential:  $y = a + c \exp(-bt)$

Ordinary logistic:  $y = a + c / \{1 + \exp[-b(t - m)]\}$

Gompertz:  $y = a + c \exp\{-\exp[-b(t - m)]\}$

Generalized logistic:  $y = a + c / \{1 + t \times \exp[-b(t - m)]\}^{1/n}$

For the ordinary exponential model (also known as the negative exponential, monomolecular, or Mitscherlich model), the parameter  $a$  is an asymptote,  $c$  (negative for the curves investigated) determines the range of possible values of  $y$  for positive values of  $t$  (time), and  $b$  is a rate parameter that describes the rate of change of  $y$ . No constraints were imposed on  $a$ .

Each of the other growth curve models has a point of inflection (or maximum rate of change when  $t > 0$ ):  $a$  is the lower asymptote,  $m$  is the point of inflection,  $b$  is a rate parameter, and  $a + c$  is the upper asymptote. A sigmoid curve results. In the generalized logistic (also known as the Von Bertalanffy-Richards model),  $n$  is a power parameter; each of the other models can be obtained from the generalized logistic model by the insertion of particular values of  $n$ . Data for the steep central regions of curves and for both flat extremes of the curves are necessary for regression with the growth curve models (18). In all analyses, the lower asymptote  $a$  was set to 0. Initially, no constraints were imposed on the upper asymptote  $c$ . For each regression, the following parameters were obtained: the coefficient of determination ( $R^2$ ), the coefficient of determination calculated on back-transformed fitted values ( $R^{*2}$ ), the mean square error, the values of the parameters and their standard error estimates, and the autocorrelation of the residuals (15).

**Variation of disease progress with age and season.** Special attention is paid first to the variation of disease incidence with age (at a specific period) and second to the fluctuation of disease incidence with season (at a given age). To assess the relationship between disease incidence and age, the disease increment between November and December in plots of different ages (2–6 mo old) was assessed. This timing was chosen because in November through December healthy plants remained in plots of different ages in proportions high enough to permit meaningful comparisons. By contrast, during other periods (e.g., March through April) comparisons were not possible because most cassava of all plots were infected, whatever their age. To assess the relationship between disease incidence and season, disease increment 2 mo after planting ( $Y$ ) for each planting date was determined. In all analyses,  $Y$  is expressed after angular transformation of the proportion  $p$  of disease plants with  $Y = \arcsine \sqrt{p}$ .

**Whitefly surveys.** From April 1983, the whitefly populations were monitored weekly in each planting. Adults were counted on the undersides of the five uppermost leaves of 10 plants along a diagonal of each of the seven plots of each planting. These youngest leaves are important epidemiologically because they harbor most of the adult whiteflies (8) and are the most susceptible to infection (23). Sampling started 2–3 wk after planting when a sufficient number of leaves had developed to support whiteflies; the average weekly number of adult whiteflies for the first 2 mo

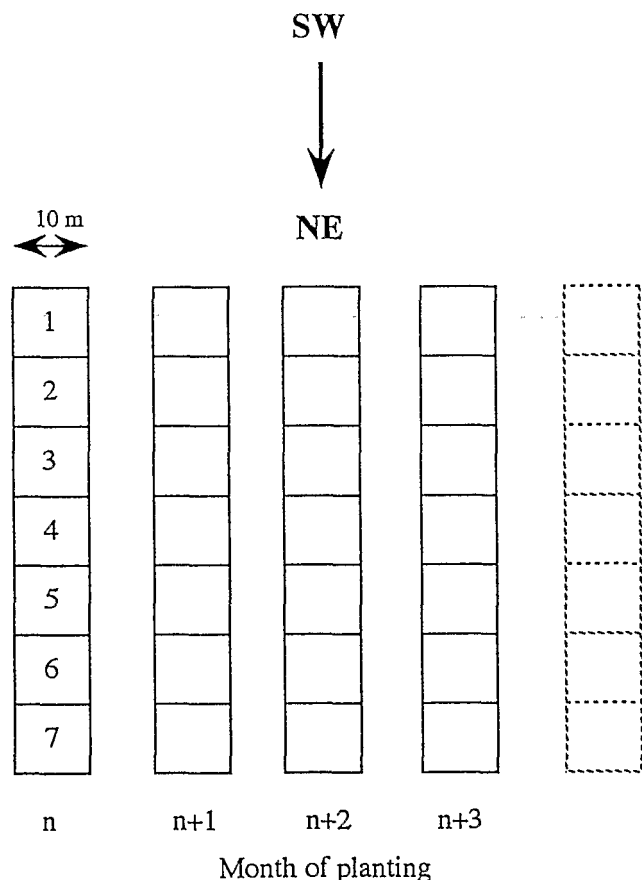
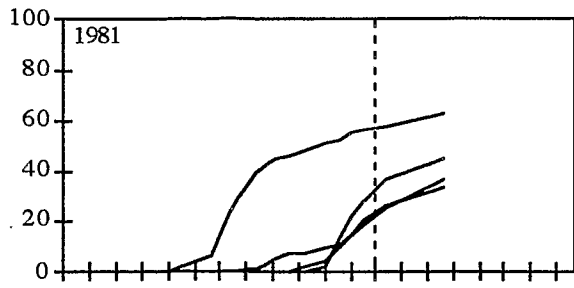


Fig. 1. Sketch showing the arrangement of the experimental plantings made between 1983 and 1986.



of growth was calculated.

**Climatic data.** The climate in Ivory Coast is characterized by a long rainy season from April to July followed by a short and relatively dry period in August and September. There is then a further short rainy season in October through November followed by a long (and comparatively) dry period from December to March (4). However, the patterns vary between years, and the seasons are not always well-defined. Climatic data were obtained from the ORSTOM meteorological station located adjacent to the experimental fields (16). The following variables were re-

partial correlation coefficients between two variables suggest a direct relationship, whereas partial correlation coefficient values lower than the multiple correlation coefficient suggest that the relationships between the two variables are due to the effect of a third variable correlated with the other two (22).

## RESULTS

**Disease progress curves.** Figure 2 illustrates the 49 disease progress curves obtained. By the end of each trial (within 3–12 mo of growth, depending on the planting date), the incidence of disease was usually high, indicating high whitefly activity. The curves were generally sigmoid, although there was much variation in shape.

**Selection of the model and analysis of its parameters.** The Gompertz model was chosen as the best at describing the disease progress curves. For most curves, a better fit, based on higher coefficients of determination and a more random structure within the residuals, was observed with the Gompertz model than with the ordinary exponential and the ordinary logistic models. Moreover, the generalized logistic model converged to the Gompertz for most disease progress curves. The regression parameters of 37 of 49 disease progress curves were chosen for further analysis; these fittings were characterized by high correlation coefficients, realistic values of the parameters, and low standard errors of the estimates, especially when autocorrelation of the residuals was high. The upper values of acceptable standard errors for  $m$  and  $c$  were fixed to 3.5 to limit imprecision in the parameter estimates.

The point of inflection  $m$  is the time when maximum spread occurs. Figure 3 illustrates the distribution of the  $m$  values for the 37 disease progress curves described adequately by the Gompertz model. The  $m$  values ranged generally between 40 and 80 days, with a mode of 51 days, a median of 59, and a mean of 64, which indicates that the maximum rate of increase was generally reached about 2 mo after planting.

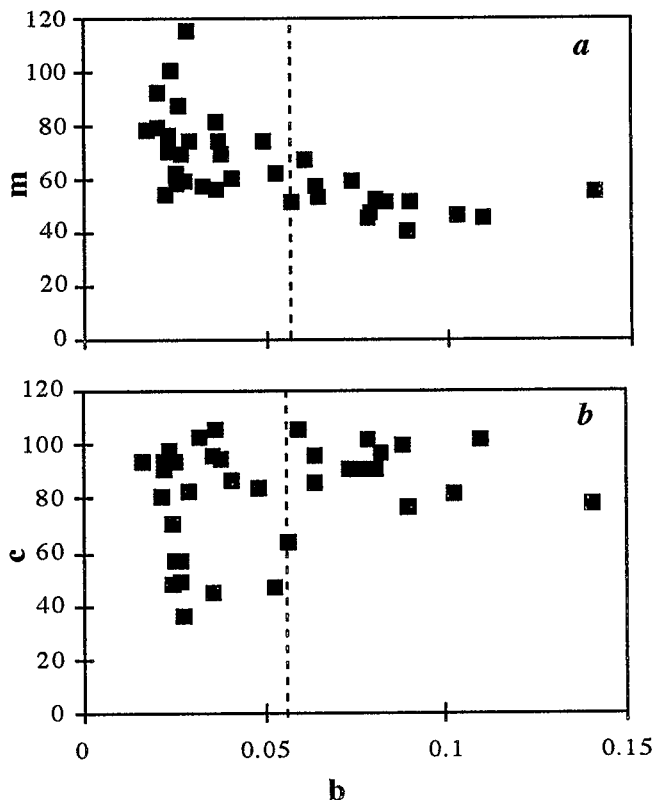


Fig. 4. Distribution of the slope parameter  $b$  plotted a, against  $m$  (the point of inflection) and b, against  $c$  (negative for the curves investigated) for curves fitted with the Gompertz model. The dotted line indicates  $b = 0.06$ .

The  $b$  values ranged between 0.02 and 0.14, which indicates a wide variation in the rates of increase of disease progress. The estimated  $b$  values were analyzed in relation to estimated  $m$  and  $c$  (3). Figure 4 illustrates the relationships between  $b$  and  $m$  and between  $b$  and  $c$ . There was no relationship between  $m$  and  $c$ . The epidemics could be separated into two groups on the basis of their parameter values. High values of  $b$  ( $b \geq 0.06$ ) were consistently associated with the inflection point  $m$  being reached within 60 days and with asymptotes  $c$  within the range of 80–100%. Most of these epidemics had high coefficients of determination with low autocorrelation of the residuals. The 13 disease progress curves in which  $b$  was above 0.06 were all for plots planted between October and April. Their asymptotes were approached within 4 mo. By contrast, no such relationships were apparent between the parameters from epidemics with lower values of  $b$  ( $b < 0.06$ ), where  $m$  was 50–120 days and  $c$  was 30–100%. These epidemics occurred mostly between May and October, and regression often revealed systematic deviation of the residuals.

**Disease increments in cassava plots of different ages.** There was a consistent decrease in the spread of disease as the plants aged, and Figure 5 illustrates the close relationship ( $R = 0.88$ ) between cassava age and monthly disease increment (after logarithmic transformation). However, the degree of the decrease varied between years, and the ratios of spread at 2 mo to spread at 6 mo ranged from 3 in 1985 to 10 in 1981.

**Disease incidence 2 mo after planting.** Disease incidence 2 mo after planting varied considerably between planting dates (Fig. 6), ranging from 2% (August 1981) to 97% (March 1984). There was a consistent pattern within and between years. Greatest disease at 2 mo occurred in plantings made from November to June, while incidence was generally low in plantings between July and October. The autocorrelogram for the disease incidence series (Fig. 7) showed the highest positive correlation at a 12-mo time lag ( $R = 0.53$ ;  $P = 0.001$ ) and the highest negative correlations at a time lag of 6 mo ( $R = -0.44$ ;  $P = 0.01$ ).

**Disease incidence, temperature, and whitefly numbers.** The pattern of temperature fluctuations was similar to that of changes in disease incidence. More spread occurred between December and May when temperatures were high and less occurred between June and November when temperatures were low. There was a positive linear relationship (Fig. 8a) ( $R^2 = 39\%$ ) between mean maximum temperature of the first month and disease incidence 2 mo after planting.

The amount of ACMV incidence was also related to the size of the whitefly population. On average, cassava plantings infested with larger adult whitefly populations were more rapidly infected with ACMV than those harboring smaller populations, and there was a significant positive relationship ( $R^2 = 41\%$ ) between disease incidence and adult whitefly numbers (after logarithmic transformation) (Fig. 8b). Errors were normally distributed, were independent, and had a constant variance.

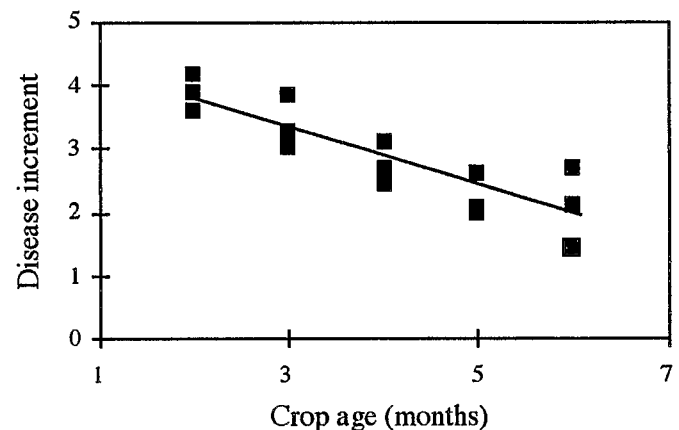
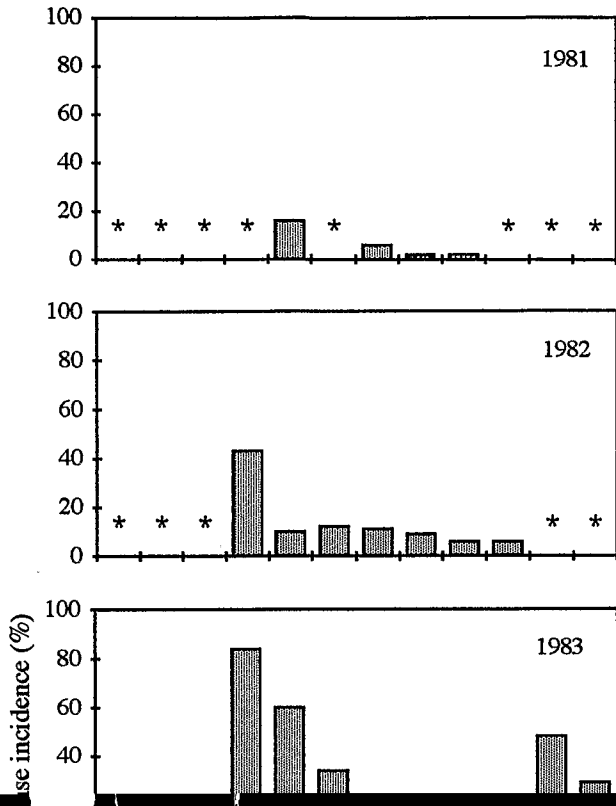


Fig. 5. Disease increment (after logarithmic transformation) in 2- to 6-mo-old cassava plots and linear regression fitting.



**The regression model.** The equation of the stepwise regression model between disease incidence ( $Y$ ), whitefly numbers ( $w$ ), and the climatic variables was

$$Y = 0.18 \log(w + 1) + 0.10 T_x + 4.8 \cdot 10^{-4} R + 0.009 H - 3.39$$

Wind velocity, water vapor pressure, minimum temperature, and total radiation were not chosen during the stepwise regression process. Only the climatic variables for the first month of growth were selected; those for the second month of growth were not chosen. Four variables were selected;  $w$  and  $T_x$  were highly significant ( $P < 0.001$ ), and  $R$  and  $H$  were significant only at  $P =$

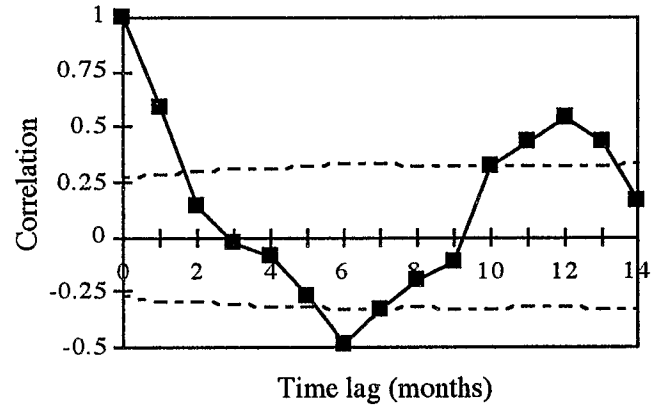


Fig. 7. Autocorrelogram of disease incidence (after angular transformation). Each point represents the correlation coefficient (y-axis) between

0.05. These four parameters accounted for 67% of the total variation. The calculated values of disease incidence corresponded closely to the observed values (Fig. 9); not only was the pattern of change reproduced over the whole survey period, but the quantitative values were also similar.

Table 1 lists the four steps of the regression. The increasing (adjusted) coefficients of determination associated with the introduction of each selected explanatory variable indicate that temperature, whitefly numbers, and rain-associated parameters are all necessary to explain the variation in disease incidence. Temperature and whitefly numbers contributed more to disease incidence than did rainfall and relative humidity, as indicated by the standard errors of the estimates of the regression coefficients of the whitefly numbers and of the temperature, which were comparatively small. Those for rain and relative humidity were large (Table 1).

**Cross-validation of the regression model.** A form of data splicing was used to validate the selected model. The roles of the whiteflies and temperature were thus validated; both these variables were retained, and the regression coefficients of the subset regression were within the confidence interval of the whole regression (Table 1). By contrast, the regression was not stable with the rain-associated parameters. Relative humidity failed to be selected three times and rainfall twice. The adjusted coefficient of determination of the subregression varied between 57 and 74% (compared to 67% for the whole model). Four times out of five, there were no significant differences between the predicted and the observed values ( $P = 0.05$ ), indicating that the regression model was valid and robust enough to withstand the removal (or addition) of 20% of the observations.

**Rain and radiation-associated variables and disease spread.** Attempts were made to dissociate the effects of the climatic variables from those of the whitefly populations by comparison of the multiple and partial correlation matrices. The results indicated that whitefly numbers had a direct effect on disease incidence. More interestingly, both multiple and partial correlation coefficients were similar for temperature, suggesting that its effect on the rate of spread is also direct and therefore not dependent on whitefly numbers. A direct effect of relative humidity and an indirect effect of rainfall on disease incidence was also observed.

The overall contribution of the rain-associated parameters may be somewhat diminished because the plots had to be irrigated during the main dry season (December through February) for growth to occur. However, if data for these months were excluded from the regression, the structure of the model was not changed (adjusted  $R^2 = 64\%$ ), because whitefly numbers and temperature were still highly significant ( $P < 0.001$ ), and rainfall was significant

## DISCUSSION

Combining the results of the analysis of the disease progress curves with the available biological information (28) and personal observations (D. Fargette and C. Fauquet, *unpublished*) suggests the following sequence of events in a field newly planted with healthy cassava. The first leaves are produced 2–3 wk after planting; these are colonized by viruliferous whiteflies that inoculate the virus (1–2 wk); visible symptoms appear after a further 3–5 wk (latency period). Therefore, it would take an average of about 8 wk for symptoms to be expressed at the maximum rate. This period is consistent with the modal, median, and mean  $m$  values of the Gompertz model of approximately 60 days. Since each of the three steps outlined above varies in duration, the maximum rate of symptom expression would be reached 6–10 wk after planting, a range that compares well with the variation of  $m$  values, which is found to be mostly between 40 and 70 days. However, values of  $m$  greater than 70 days suggest that variations in the duration of one or more of the steps may have been even larger than observed. Of the factors likely to affect these timings, periods with high rates of spread ( $b \geq 0.06$ ) were associated with a lower inflection point ( $m \leq 60$ ). If it is assumed that young emerging leaves are most susceptible to infection, then the time taken for symptoms to appear would be closely related to the maximum rate of disease spread. However, it is unlikely that this accounts for all the variation observed, and other sources of seasonal variation are likely to be of importance. For example, changes in host susceptibility to infection over the year are also associated with rates of leaf and branching production in cassava.

The rate of disease spread varied greatly between the epidemics. Analysis of the rate parameter  $b$  suggested that epidemics could be divided into two groups. For those that developed rapidly ( $b > 0.06$ ), maximum spread is always reached well within the first 2 mo of planting, and asymptotes of around 100% are reached within 4 mo. Such epidemics appear to be restricted to the time between October and April (inclusive). By contrast, epidemics that develop slowly ( $b < 0.06$ ) are associated with a wide range of  $m$  and  $c$  values. These epidemics occur mostly between May and October. This clearly indicates that the rate of spread has a seasonal component.

Thresh (26) distinguished two types of epidemics. In the first, many factors act independently to determine the course of the epidemic, so that predicting the amount and rate of spread is difficult. In the second, a few and/or interrelated factors influence the disease progress rate, so that understanding or predicting the course of such epidemics is feasible as long as the key or driving function(s) can be estimated. ACMV epidemics, at least at Adio-

rate of disease progress was higher in young cassava plots (2 mo old) than in older ones (6 mo old) by a factor of 3-10, depending on the year. This decreasing susceptibility likely reflects the reduced rate of growth of aging cassava (20) rather than the variation in whitefly numbers, which generally did not fall before 4 mo after planting (5).

A typical feature of vectorborne viruses is the dependence of the course of epidemics upon vector numbers. Such dependence has been found with many aphidborne viruses (26), as well as African cassava mosaic (14) and other whitefly-transmitted viruses (17). In this study, rates of disease progress were consistently associated with total adult whitefly numbers, although this gives only an indirect and a possibly biased assessment of the fraction of vector population actually involved in ACMV progress, which is due solely to those that are infective and active (26). The logarithmic relationship between disease incidence and whitefly number suggests that saturation occurs. This is to be expected because increases in whitefly numbers are not necessarily reflected by corresponding increases in the rates of spread when whiteflies are numerous, possibly because of the limited number of healthy plants left.

The analyses indicate an annual periodicity of the disease progress, despite a large year-to-year variation, with a strong seasonal fluctuation, which suggests that one or more climatic variables play a key role in the epidemiology. The close association between disease incidence and average monthly climatic variables does reflect the strong dependency on the overall macroclimate, but it is likely that microclimatic data, collected daily at the field level, would have given an even more detailed description of the comparison (2,19). Significant relationships were found with climatic values of the first month of growth but not of the second, despite the fact that disease incidence was assessed at the end of the second month. This is expected if the second month is mainly the time for plants infected in the first month to show symptoms.

In temperate regions, seasonal variation of disease progress has often been observed and generally linked to variations in measures of temperature such as mean temperature, accumulated day-degrees, and number of frost days (26). In contrast, seasonal variations in tropical regions were found for maize streak and

content in cassava because the rate of ACMV multiplication is dependent on temperature (21). This would result in more rapid spread because cassava is the main virus reservoir (10).

Temperature and global radiation patterns in Ivory Coast are closely associated, and selection of temperature and not global radiation through the stepwise regression process reflects the colinearity of the two climatic variables but does not preclude an effect of global radiation on disease progress. In particular, cassava growth is dependent on global radiation and exhibits a strong seasonal component (19). Cassava is less susceptible to ACMV infection when its growth rate is reduced. This phenomenon, recorded for many virus-host combinations (25), was first described by Storey and Nichols (24), who showed that growing leaves were readily infected by viruliferous whiteflies, whereas mature leaves were not. Then changes in cassava growth rate, linked to fluctuations in global radiation, would further contribute to the seasonal variation in the rate of virus spread. The effect of rain-associated parameters on virus spread, although significant, remains more limited and more speculative. Heavy rains may affect virus spread by depressing whitefly populations (3). Prolonged droughts slow cassava growth, cause leaf abscission, and possibly result in a reduced susceptibility to infection. However, under conditions such as those in Adiopodoumé, where the relatively dry season is comparatively short and where irrigation is carried out, a limited effect of rain is to be expected.

#### LITERATURE CITED

1. Butt, D. J., and Royle, D. J. 1990. Multiple regression analysis in the epidemiology of plant diseases. Pages 163-180 in: *Epidemics of Plant Diseases*. J. Kranz, ed. Springer-Verlag, Berlin.
2. Campbell, C., and Madden, L. V. 1990. *Introduction to Plant Disease Epidemiology*. Wiley Interscience, New York.
3. Cock, M. J. W. 1986. *Bemisia tabaci*, A Literature Survey. Chameleon Press, London.
4. Eldin, M. 1971. Le milieu naturel en Côte d'Ivoire. Pages 75-108 in: *Mémoires No. 50*. Editions de l'ORSTOM, Paris.
5. Fargette, D. 1987. *Epidémiologie de la mosaïque africaine du manioc*. Ph.D. thesis. Université des Sciences et Techniques du Languedoc. Editions de l'ORSTOM, Paris.
6. Fargette, D., Fauquet, C., and Thouvenel, J.-C. 1985. Field studies

20. Raffailac, J. P., and Nedelec, G. 1986. Comportement de 10 clones de manioc dans les conditions édapho-climatiques du sud de la Côte d'Ivoire. Editions de l'ORSTOM, Paris.
21. Sequeira, J. C., and Harrison, B. D. 1982. Serological studies on cassava latent virus. *Ann. Appl. Biol.* 101:33-42.
22. Sokal, R. R., and Rohlf, F. J. 1981. *Biometry*. W. F. Freeman, New York.
23. Storey, H. H., and Nichols, R. F. W. 1936. A field experiment in the transmission of cassava mosaic. *East Afr. Agric. J.* 6:446-449.
24. Storey, H. H., and Nichols, R. F. W. 1938. Studies on the mosaic of cassava. *Ann. Appl. Biol.* 25:790-806.
25. Thresh, J. M. 1983. Progress curves of plant virus disease. *Adv. Appl. Biol.* 8:1-85.
26. Thresh, J. M. 1986. Plant virus disease forecasting. Pages 359-386 in: *Plant Virus Epidemics, Monitoring, Modelling and Predicting Outbreaks*. G. D. Mc Lean, R. G. Garret, and W. G. Ruesink, eds. Academic Press, New York.
27. Thresh, J. M. 1991. The ecology of tropical plant viruses. *Plant Pathol.* 40:324-339.
28. Waggoner, P. E. 1986. Progress curves of foliar diseases: Their interpretation and use. Pages 3-37 in: *Plant Disease Epidemiology*. K. J. Leonard and W. E. Fry, eds. Macmillan, New York.
29. Wilkinson, L. 1989. *SYSTAT: The System for Statistics*. SYSTAT, Evanston, IL.