Temporal and Spatial Analysis of Maize Dwarf Mosaic Epidemics

L. V. Madden, Raymond Louie, and J. K. Knoke

First author, associate professor, Department of Plant Pathology, The Ohio State University (OSU), Ohio Agricultural Research and Development Center (OARDC); second and third authors, research plant pathologist and research entomologist, respectively, Agricultural Research Service (ARS), U. S. Department of Agriculture (USDA), Wooster, OH 44691.

Cooperative investigations of OSU and USDA/ARS. Salaries and research support provided by State and Federal funds (especially USDA Crop Loss Grant 83-CRSR-2-2235) appropriated to the OARDC, OSU. Journal Article 28-86.

We thank S. S. Mendiola, R. J. Anderson, and J. J. Abt for technical assistance.

This paper reports the results of research only. Mention of a commercial or proprietary product does not constitute endorsement by the USDA.

Accepted for publication 1 July 1986 (submitted for electronic processing).

ABSTRACT


Epidemics of maize dwarf mosaic (MDM) were monitored in experimental field plots of susceptible maize at three locations in Ohio (Wooster, Marietta, and Portsmouth). Based on regression analysis, initial disease incidence ($y_0$) was less than 0.003 in all plots. The weighted mean rate of disease increase ($\rho$) ranged from 0.001 to 0.062 per day. The lowest rates, and the lowest final disease incidences ($y_f$) were in the northern Ohio plots; these three plots were best described by the monomolecular model and the flexible Weibull model with a shape parameter ($c$) equal to 1. At the other locations, disease progression was best described by the logistic and Gompertz models, and by the Weibull model with $c > 1$. Disease incidence reached a level of $y_f \geq 0.995$ in nine of 12 plots at Marietta and Portsmouth. MDM incidence also was accurately described by a nonlinear function of the cumulative number of aphids trapped in all plots. Spatial patterns of diseased plants changed in most plots during the season. Ordinary runs declined in 10 of 13 plots; a random pattern usually was detected during the early part of the epidemic and a clustered pattern later. Patchiness also declined in nine of 13 plots. Most plots exhibited random and clustered patterns of MDM, depending on the time. Six of the plots had an overall clustered pattern as determined by Taylor's power relation; this did not always agree with the individual patchiness calculations. Additionally, analysis based on ordinary runs and patchiness (or the variance-to-mean ratio) often led to different conclusions about randomness or clustering. There was a significant ($P \leq 0.05$) negative correlation between the final level of patchiness in the plots and both $\rho$ and $c$. Initial level of aggregation was not correlated with the temporal characteristics of the MDM epidemics.

Additional key words: comparative epidemiology, disease progress curves, dispersion, quantitative epidemiology, spatial distribution, vectors, Zea mays.

Quantitative aspects of the epidemiology of diseases caused by viruses remain largely unreported in the phytopathological literature. Although several researchers have collected detailed data on virus disease progression (36-39), few have taken the extra step to model the epidemics (1) and the results obtained have seldom been analyzed in sufficient detail to obtain the maximum possible information on the various factors influencing the sequence and rates of spread observed (39). One must quantify disease progression temporally and spatially to compare treatments, viruses and crops, predict future disease levels, understand the dynamics of virus disease epidemics, and determine the mechanisms of disease spread (19).

Plant disease epidemics can be quantified by modeling the level of disease ($y$) as a function of time ($t$) (17-19). After a regression analysis, estimated model parameters are used to describe the initial level of disease and the rate of disease increase. Some models have additional parameters that permit a more thorough description of an epidemic.

The spatial pattern of virus-diseased plants has not been well studied. This information is critical for sampling as well as for understanding disease progression (21,28). One method for assessing patterns is to record the position of infected plants within
rows and analyze the data with doublet or ordinary runs analysis (4, 8, 21, 40). A different approach to assessing aggregation is based on the analysis of the number of virus-diseased plants per sample unit. The determination of randomness or clustering depends on the mean (m) and variance (V) for the sampling units (5, 6, 19, 30).

In this study, epidemics caused by maize dwarf mosaic virus (MDMV) on maize (Zea mays L.) were quantified temporally and spatially. MDMV, a potyvirus, is the most widespread virus of maize in the United States (9). The virus is transmitted nonpersistently by at least 25 aphid species (14-16). Most of the species responsible for virus transmission are transients and do not colonize maize. The virus is spread into maize in the spring and summer from johnsongrass (Sorghum halepense L.) and, probably, other grass species (16). Little has been published on the temporal and spatial aspects of MDMV epidemics (15, 16).

Objectives of this study were to: describe MDMV disease progression over time; describe the spatial patterns of MDMV-diseased plants over time and compare techniques to assess spatial patterns; attempt to relate temporal aspects of MDMV epidemics to spatial patterns; and relate MDMV progression to the number of aphids landing in the experimental plots.

**MATERIALS AND METHODS**

**Field plots.** Three areas of Ohio were chosen to monitor MDMV epidemics. These areas were near Wooster (40° 47'=N, 81° 55'=W), Marietta (39° 25'=N, 81° 26'=W), and Portsmouth (38° 44'=N, 82° 58'=W). The known overwintering source of MDMV, johnsongrass, was common at the latter two locations but rarely found at the former. Plots measured 30.5 × 30.5 m with 3.5-m fallow borders. Seeds were planted in a 50.8 × 50.8-cm grid resulting in 3,600 plants per plot with 60 plants on a side. Except in two plots, each plot was divided into five rows and tested for disease incidence. All plants were sampled in disease assessments (about 62-75 days after planted).

Aphids were collected in yellow pan traps by techniques described previously (20). At least two traps were placed within plots at each location. Trapped aphids were removed daily but sometimes 2-3 days elapsed between removals. Number of aphids per week at each location was then calculated for later analyses.

No disease developed in plots I-4 at Wooster (W-1 to W-4, planted on day 136 of the year), and plants were killed or removed. New plots were then established at the sites of W-1 to W-3 on day 198 (W-5 to W-7), and W-5 and W-7 were planted adjacent to 16×16- and 8×8-plant maize plots, respectively, that were mechanically inoculated.

**Data analysis.** To obtain precise descriptions of MDMV disease progression, the following epidemic models were fitted to the data on aggregates incidence of each plot: 1) exponential, 2) monomolecular, 3) logistic, and 4) Gompertz (17, 35). All models were linearized using the method shown in Madden (17) and then fit to the data with ordinary least squares regression. The parameters were estimated by computer and then transformed into disease increase (r) and initial level of disease (y0). One cannot compare directly the r parameters from different models. For comparisons, one can calculate the weighted mean rate as p = r/n, in which n = 2, 4, and 6 for the monomolecular, Gompertz, and logistic models, respectively (35).

After fitting each model to the data, their appropriateness was determined using standardized residual plots, coefficients of determination (R²), and additional statistics as described elsewhere (17, 18, 23). It is not possible to compare directly R² from two models with different transformations of y; the predicted transformed values of disease must be detransformed to obtain predicted y's and then a new R² is calculated. This was done to compare models (10).

After the most appropriate model was chosen, the autocorrelation of the residuals was calculated, as described by Madden (18). A Student's t test then was used to determine if the autocorrelation coefficient was different from 0 (18).

The flexible Weibull model also was fitted to the disease progression data by a maximum likelihood technique (25). The scale (b) and shape (c) parameters were estimated after the location parameter a was assigned a value equal to the total assessment time with any infected plants, minus 1 day; this has previously been proven to be acceptable for estimating b and c when the first level of y is low (3).

The pattern of diseased plants on each assessment date was determined using ordinary runs (8) and Lloyd's patchiness index (27). Ordinary runs were calculated by starting arbitrarily at the current row, the next, and so on. To determine if there was significant clustering of diseased plants, the standardized statistic (Zc) was calculated. A clustered pattern is indicated if Zc ≤-1.64 at P = 0.05 or Zc ≤-2.33 at P = 0.01.

Lloyd's patchiness was determined by dividing the plots into 26-40 square contiguous quadrats that each contained 100 plants. Preliminary results indicated maximum variance with this unit size. The mean (m) and variance (V) of the number of diseased plants per quadrat was then determined. Lloyd's patchiness was calculated as m²/m in which m² represents mean crowding and is given by m²/(P²/V-1). Significant clustering was indicated by a chi-square test of the variance to mean ratio: χ² = (m² - 1)/V/m, in which m is the number of quadrats. A clustered pattern was determined when χ² > 49.80 (P = 0.05, 35 df). Aggregation was determined for each disease assessment time except when P > 0.95. It was felt that artificial results would result by using cases where y was closer to unity.

A further index of aggregation based on V' and m was calculated for each point by using Taylor's power relation (31). The relation can be written in the linear form as:

\[
\ln(V) = \ln(b_0) + b_1 \ln(m)
\]

in which b₀ and b₁ are parameters. The regression was performed for the variances and means calculated at the different assessment times for each plot separately. The slope parameter of equation 1 (b₁) was used as the aggregation index.

The relationship between an aggregation index and time in each plot was quantified with empirical regression models in which the dependent variable was either m²/m or the standardized ordinary runs (Zc) statistic. Independent variables consisted of t, t², or b₁².

The association among disease progression characteristics and aggregation at the beginning and end of the diseases was determined. Because we did not assume any functional relationships (e.g., linear, curvilinear), Spearman's rank correlations, nonparametric analogs of Pearson's product moment correlations, were calculated for pairs of variables (8). Empirical regression equations were developed to describe some of the significant relationships. Appropriateness of the regressions was determined by standard procedures (23).

The relationship between y and the cumulative number of aphids (A) was described by a generalization of a model proposed by Jeger (12):

\[
y = y_0 \exp(-f(A)/y_0)
\]

in which y₀ is a parameter representing the maximum level of disease and f(A) is a linear function of A. Two forms were attempted: f(A) = k₀ + k₁A, and if this did not describe the relation, then f(A) = k₀ + k₁A + k₂A² was used. The former expression for f(A) is equivalent to Jeger's model for disease in relation to cumulative number of sires. Equation 2 was linearized to:

\[
\ln(1/(y_0 - y)) = f(A)/y_0
\]

for the analysis in this study.
RESULTS

Disease incidence. MDM incidence increased in a continuous fashion in all plots (Figs. 1—3). At Wooster, there was little disease increase and only one of the three plots had a final observed level of disease \( y \) greater than 0.10 (Fig. 1, Table 1). Except for three plots, disease progressed to \( y > 0.59 \) at Marietta and Portsmouth (Figs. 2 and 3). M-1 had the earliest planting (120 = 30 April) and only reached \( y = 0.38 \). P-5 and P-6 were planted with a mixture of three genotypes and their final disease level was less than 0.60; plants of all three genotypes became diseased (data not shown). The remaining plots reached levels of 0.75 by 50 days after planting.

No single disease progress model was sufficient to describe all of the epidemics (Table 1). All three plots at Wooster were best described by the monomolecular model. At the other two locations, the logistic and Gompertz were the most appropriate. Because \( y \) reached a level of about 1 in several plots, and there was no apparent leveling off at some maximum < 1 in the other plots, it was considered appropriate to use 1 as the maximum in the transformations (17). At P-4 and P-7, 100% disease was reached by the third assessment time; therefore, it was not possible to determine the most appropriate model and the data are not presented.

As indicated by the regression analysis, the estimated \( y_0 \) was extremely low in all plots (Table 1), suggesting that the epidemics were not started by a large influx of viruliferous aphids. The lowest values of \( p \) were for the Wooster plots, as expected from the low levels of \( y \) (Table 1). At Marietta, the rate was lowest for the earliest planting (M-1) and showed no systematic change for planting times from M-2 to M-5. At Portsmouth, the lowest rates corresponded to the mixed plantings indicating that the mixture slowed the rate of disease increase as well as reducing disease level. Of the single genotype plots, there were not enough planting dates to discern a systematic change in \( p \) with planting. Interestingly, the largest \( p \) at both Marietta and Portsmouth corresponded to plots planted between days 148 and 154 (M-2 and P-2).

The autocorrelation coefficient of the residuals was not greater

![Fig. 1. Disease incidence \( y \), patchiness \( m^* \), and standardized runs \( Z_t \) versus time \( t \) for three maize dwarf mosaic epidemics of maize at Wooster (W), OH. The drawn lines represent the predicted level of the three variables as a function of time, based on the results in Tables 1 and 2. In order to draw all values on a consistent axis of the ordinate the following point \( (t, m^* m) \) was omitted from W-7: \((31, 0.17)\). For patchiness and runs, open symbols represent randomness and closed (solid) symbols represent significant clustering \( (P = 0.05) \).](image)

![Fig. 2. Disease incidence \( y \), patchiness \( m^* \), and standardized runs \( Z_t \) versus time \( t \) for five maize dwarf mosaic epidemics of maize at Marietta (M), OH. The drawn lines represent the predicted level of the three variables as a function of time, based on results in Tables 1 and 2. In order to draw all values on a consistent axis of the ordinate the following point \( (t, m^* m) \) was omitted from M-2: \((14, 1.8)\). For patchiness and runs, open symbols represent randomness and closed (solid) symbols represent significant clustering \( (P = 0.05) \).](image)
than 0 for any of the epidemics ($P > 0.20$). No correction for autocorrelation such as first-difference regression (17,18), therefore, was needed. Standard errors of the estimated weighted mean rate parameters were calculated from ordinary least squares regression and are presented in Table 1.

The Weibull model provided good fits to all disease progress curves with five or more assessment times (Table 1). As predicted by theory (17,25,35), those plots best described by the monomolecular model had Weibull shape parameters ($c$) equal to

![Fig. 3. Disease incidence ($y$), patchiness ($m^* / m$), and standardized runs ($Z$) versus time ($t$) for five maize dwarf mosaic epidemics of maize at Portsmouth (P), OH. The drawn lines represent the predicted levels of the three variables as a function of time, based on results in Tables 1 and 2. In order to draw all values on a consistent axis of the ordinate, the following points ($t, m^* / m$) were omitted from P-6: (4, 0.49), (22, 0.64). The following points ($Z_t, Z_t$) were also omitted from P-1: (28, 49, 65) and from P-6: (14, 8, 41). Disease incidence reached 1.0 by the third assessment date at P-4 and P-7, and therefore, these epidemics are not presented. For patchiness and runs, open symbols represent randomness and closed (solid) symbols represent significant clustering ($P = 0.05$).

| Field* | Planting date | Best model | $N$ | $y_0 \times 10^3$ | $r$ | $\rho$ | $R^2$ | $Y_t$ | $b$ | $c$ | $R^2$
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>W-5</td>
<td>198</td>
<td>M</td>
<td>6</td>
<td>0</td>
<td>0.008</td>
<td>0.004</td>
<td>.0004</td>
<td>0.97</td>
<td>0.26</td>
<td>98.2</td>
<td>1.2</td>
</tr>
<tr>
<td>W-6</td>
<td>198</td>
<td>M</td>
<td>5</td>
<td>0</td>
<td>0.001</td>
<td>0.001</td>
<td>.0001</td>
<td>0.96</td>
<td>0.02</td>
<td>958.9</td>
<td>1.2</td>
</tr>
<tr>
<td>W-7</td>
<td>198</td>
<td>M</td>
<td>6</td>
<td>0</td>
<td>0.002</td>
<td>0.001</td>
<td>.0001</td>
<td>0.97</td>
<td>0.06</td>
<td>521.7</td>
<td>1.0</td>
</tr>
<tr>
<td>M-1</td>
<td>120</td>
<td>G</td>
<td>5</td>
<td>0</td>
<td>0.058</td>
<td>0.014</td>
<td>.0008</td>
<td>0.99</td>
<td>0.38</td>
<td>37.8</td>
<td>3.1</td>
</tr>
<tr>
<td>M-2</td>
<td>148</td>
<td>L</td>
<td>6</td>
<td>0.03</td>
<td>0.362</td>
<td>0.060</td>
<td>.0003</td>
<td>0.99</td>
<td>1.00</td>
<td>21.5</td>
<td>4.5</td>
</tr>
<tr>
<td>M-3</td>
<td>183</td>
<td>L</td>
<td>8</td>
<td>0.09</td>
<td>0.226</td>
<td>0.039</td>
<td>.0015</td>
<td>0.99</td>
<td>1.00</td>
<td>26.3</td>
<td>3.0</td>
</tr>
<tr>
<td>M-4</td>
<td>204</td>
<td>L</td>
<td>5</td>
<td>0.06</td>
<td>0.338</td>
<td>0.056</td>
<td>.0015</td>
<td>0.98</td>
<td>1.00</td>
<td>16.2</td>
<td>2.7</td>
</tr>
<tr>
<td>M-5</td>
<td>232</td>
<td>L</td>
<td>6</td>
<td>0.4</td>
<td>0.231</td>
<td>0.038</td>
<td>.0004</td>
<td>0.95</td>
<td>1.00</td>
<td>27.5</td>
<td>4.4</td>
</tr>
<tr>
<td>P-1</td>
<td>126</td>
<td>L</td>
<td>8</td>
<td>0.01</td>
<td>0.242</td>
<td>0.040</td>
<td>.0036</td>
<td>0.95</td>
<td>1.00</td>
<td>42.3</td>
<td>8.5</td>
</tr>
<tr>
<td>P-2</td>
<td>154</td>
<td>G</td>
<td>6</td>
<td>0</td>
<td>0.250</td>
<td>0.062</td>
<td>.0070</td>
<td>0.95</td>
<td>1.00</td>
<td>16.6</td>
<td>4.3</td>
</tr>
<tr>
<td>P-3</td>
<td>182</td>
<td>L</td>
<td>6</td>
<td>2.8</td>
<td>0.206</td>
<td>0.034</td>
<td>.0029</td>
<td>0.97</td>
<td>1.00</td>
<td>29.2</td>
<td>2.5</td>
</tr>
<tr>
<td>P-4</td>
<td>210</td>
<td>...</td>
<td>3</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>1.00</td>
<td>1.00</td>
<td>...</td>
</tr>
<tr>
<td>P-5</td>
<td>182</td>
<td>L</td>
<td>6</td>
<td>1.4</td>
<td>0.092</td>
<td>0.015</td>
<td>.0008</td>
<td>0.98</td>
<td>0.36</td>
<td>51.0</td>
<td>5.0</td>
</tr>
<tr>
<td>P-6</td>
<td>238</td>
<td>G</td>
<td>7</td>
<td>0</td>
<td>0.124</td>
<td>0.031</td>
<td>.0007</td>
<td>0.98</td>
<td>0.49</td>
<td>54.6</td>
<td>4.5</td>
</tr>
</tbody>
</table>

*W = Wooster, M = Marietta; and P = Portsmouth.

* M = monomolecular; G = Gompertz; and L = logistic disease progression models.

*Number in parentheses = the standard error of the estimated weighted mean rate ($\rho$).

*Not possible to fit models to the disease progression data in these fields due to low number of times; data not shown in Fig. 3.
The differences in the rate of disease increase were also evident from the Weibull b parameter estimates (Table 1). Clearly, the epidemics at Wooster, in addition to having a different shape, progressed at a much slower rate.

**Disease patterns.** The spatial pattern of diseased plants varied considerably among plots and over time within plots (Figs. 1-3). Standardized runs (Z_t) declined in all Wooster and Marietta plots, as well as in two of the five plots in Portsmouth (P-1 and P-6). In P-2 there were no apparent trends, in P-3 there was a slight increasing trend, and in P-5 there was an initial decline in Z_t to a minimum and then an increase. In most plots, the standard normal test indicated nonclustering (open triangles in Figs. 1-3) at the first time and then changing to significant clustering at later times (closed triangles). W-5 and W-6, however, always exhibited clustering and nonclustering, respectively (Fig. 1). P-5 progressed from random to clustered and then back to random (Fig. 3).

Several different regression equation forms were necessary to relate standardized runs to time (Table 2). Several of the present regression equations did not fit the data at P < 0.05. The best results obtained are presented with the achieved significance level. With P-2, it was not even possible to obtain an equation with P < 0.50. The low numbers of observations hindered obtaining highly significant results.

Patchiness declined in nine of the 13 plots. There was a trend for patchiness to increase to a maximum and then decline in two plots, M-3 and P-6. In W-6, patchiness was virtually constant after an initial lower value. Only in W-7 was there an increasing trend throughout the epidemic.

Few plots exhibited an exclusively clustered disease pattern (Figs. 1 and 2). Some plots started with a random pattern and, at least for one or more times, then exhibited clustered patterns of infected plants (W-6, W-7, M-2 to M-3, P-1, P-5, and P-6). Other plots started with a clustered pattern and then later displayed a random pattern (M-4, P-2, and P-3). Seven of the plots had clustered disease patterns at the last assessment date when y_0 < 0.95.

Different regression equation forms were needed to relate patchiness to time adequately (Table 2). For W-6, it was not even possible to obtain an equation with P < 0.50. Some of the regression equations fit the data quite well with values of R² > 0.80.

Regression analysis of the variances and means, based on Taylor's power relation (Eq. 1) resulted in a wide range of aggregation indices (b_t) (Table 2). The relation was not significant at P < 0.05 for three of the plots. Six of the plots had values of b_t significantly greater than 1, indicating clustering of diseased plants. Some of these results appeared to be in conflict with the patchiness (or individual V/m) results. For instance, Taylor’s analysis indicated clustering for plot M-3, yet this plot exhibited a random pattern on several of the assessment dates (Fig. 2, Table 2). M-5 exhibited clustering on most of the assessment dates, however, Taylor’s b_t was less than one.

**Relation between temporal and spatial epidemic characteristics.** There were few significant (P < 0.05) correlations between disease progression characteristics and initial and final levels of aggregation. There was no relation between y_0 and the initial or final standardized runs (Fig. 3). Likewise, there was no relation between y_0 and initial patchiness. However, there was a nonlinear, negative relation between y_0 and the final level of patchiness (Fig. 3). Lower levels of y_0 were associated with higher levels of final patchiness.

There was no significant (P < 0.05) association between b_t and initial or final standardized runs, nor between b_t and initial patchiness (Fig. 5). There was a strong negative relationship, however, between b_t and final patchiness (Fig. 5); lower rates of disease increase resulted in higher final patchiness.

There was no significant association between the shape of the

![Initial patchiness vs Final patchiness](image)

**Fig. 4.** Estimated initial incidence (y_0) of maize dwarf mosaic of maize in plots at three locations in Ohio versus initial and final patchiness (m^t/m) and standardized runs (Z_t). The rank correlation between the variables are presented in the figure; an absolute value of the correlation ≥ 0.56 is necessary for a significant association at P = 0.05.

<table>
<thead>
<tr>
<th>Field</th>
<th>N</th>
<th>Standardized runs</th>
<th>Patchiness</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Model</td>
<td>b_0</td>
</tr>
<tr>
<td>W-5</td>
<td>6</td>
<td>t</td>
<td>27.0</td>
</tr>
<tr>
<td>W-6</td>
<td>5</td>
<td>t</td>
<td>11.0</td>
</tr>
<tr>
<td>W-7</td>
<td>5</td>
<td>t</td>
<td>16.5</td>
</tr>
<tr>
<td>M-2</td>
<td>4</td>
<td>t</td>
<td>4.4</td>
</tr>
<tr>
<td>M-3</td>
<td>5</td>
<td>t</td>
<td>12.0</td>
</tr>
<tr>
<td>M-4</td>
<td>3</td>
<td>t</td>
<td>-3.0</td>
</tr>
<tr>
<td>M-5</td>
<td>6</td>
<td>t</td>
<td>0.6</td>
</tr>
<tr>
<td>P-1</td>
<td>4</td>
<td>t</td>
<td>187.0</td>
</tr>
<tr>
<td>P-2</td>
<td>3</td>
<td>t</td>
<td>2.36</td>
</tr>
<tr>
<td>P-3</td>
<td>5</td>
<td>t</td>
<td>5.9</td>
</tr>
<tr>
<td>P-4</td>
<td>6</td>
<td>t</td>
<td>5.3</td>
</tr>
</tbody>
</table>

* Determined as the slope of regressing log(ν) on log(m), in which ν and m are the variance and mean, respectively, for the number of infected plants per quadrat.

* With y representing either Z_t or m^t/m, models were of the following forms: t: Y = b_0 + b_1 t; t': Y = b_0 + b_1 t'; t²: Y = b_0 + b_1 t²; t: Y = b_0 + b_1 t + b_2 t².

* W = Wooster; M = Marietta; P = Portsmouth.

* Indicates that there was a significant (P = 0.05) relation between log(ν) and log(m). Italicized numbers indicate that b_t was significantly greater than 1.0

* Not possible to find an acceptable model with P < 0.50.

152 PHYTOPATHOLOGY
disease progress curves, as measured by the Weibull c, and the
measures of aggregation. There was a significant correlation
between the Weibull b and final patchiness, but this result parallels
the significant correlation between ρ and final patchiness.

There was a significant negative correlation between both p₀ and
ρ with Taylor’s b₁ (Fig. 6). Lower initial levels of disease and slower
rates of increase were both associated with higher values of the
aggregation index b₁.

**Disease incidence in relation to aphid numbers.** The relationship
between y and the cumulative number of aphids was well described
by equation 2 (Fig. 7). All regressions were significant at P ≤ 0.01.
The maximum level of disease (yₘₐₓ) was assigned equal to 1 for all
epidemics. Coefficients of determination for the transformed
model, except for P-5, were greater than 0.80 (Table 3). Eight of
the plots, including all at Wooster, were described by the two
parameter form. Only M-1, M-3, P-1, and P-2 required a quadratic
term (k₁t²) in the model.

For the epidemics without the k₂ term, one can use the estimated
k₁ parameter to quantify the rate of disease increase per unit
increase in aphids, or equivalently, the “field transmission
efficiency” of the aphids. This can best be demonstrated with the
Wooster plots, all three of which were planted on the same day.
W-5 had the largest k₁, and W-6 the lowest. For unknown reasons,
a unit increase in aphids at W-5 resulted in a greater disease
increase than at W-6. The values of k₁ for the Wooster plots were
considerably less than for the other locations. It is not known if
aphid activity (behavior) or the lower percentage of viruliferous
aphids accounted for this result. For epidemics with the k₁ and k₂
terms, the change in disease with change in aphids depends on the
time as well as the number of aphids.

**DISCUSSION**

All of the MDM disease progress curves with five or more
nonzero data points were well described by one of the previously
proposed epidemic models (17). No conclusion about model form
for the two epidemics with only three nonzero values <1 can be
made. No single model, however, was adequate to describe the
remaining 13 epidemics. Location was the only clear determinant
for the form of disease progression. The monomolecular, or so-
called “simple-interest disease” model, was appropriate only in
Wooster. Using no other information, one might then conclude
that at Wooster, MDM behaved as a simple-interest disease, i.e.,
disease increased without spread from plant to plant. Simple
interest diseases, as proposed by Vanderplank (41), increase either
from inoculum in the soil or seed, or from inoculum entering the
fields from without. The source of inoculum was provided at
Wooster in the form of nearby plots with MDMV-inoculated
maize.

Epidemics at the other two locations were all accurately
described by either the logistic or Gompertz models. The Weibull c
was always greater than 1. Both models are appropriate for
“compound-interest diseases” in which there is spread from plant
to plant, or equivalently, diseased plants within the plots serve as
inoculum for further disease increase. The differences between
the logistic and Gompertz models are not always obvious from simply
looking at y versus t (P-1 and P-2 in Fig. 3). The absolute rate of
disease increase dy/dt of the Gompertz model is greater than the
logistic’s during the early part of an epidemic. Additionally, dy/dt of
the Gompertz reaches a maximum earlier than does dy/dt of the
logistic (2,35). The three epidemics better described by the
Gompertz model did not have any characteristics in common to
separate them from the logistic-type of disease epidemics.

Differences between these two types of epidemics (or types of
models) are small compared with those between the monomolecular
and the compound-interest type of epidemics. All of the epidemics
at Marietta and Portsmouth were not adequately described by the
monomolecular model, suggesting that, using no other information,
there was spread from plant to plant. Aphids were present at both
locations for spread among plants; there were also outside sources of
MDMV inoculum in the form of naturally infected johnsongrass.

It is now generally accepted in epidemiology that one cannot
simply use the most appropriate disease progression model to infer
the mechanism of disease increase (22,26). This is because many
assumptions are used in developing theoretical models, such as the
monomolecular or logistic. Failure of an epidemic to satisfy all of
the assumptions could alter the shape or form of a disease progress
curve. It is less important to fix exact biological meaning to the
the type of model that describes an epidemic than it is to know which

---

**Fig. 5.** Estimated weighted mean rate of increase (p) for maize dwarf mosaic
of maize in plots at three locations in Ohio versus initial and final patchiness
and standardized runs (ζ). The rank correlation between variables are presented in the figures; an absolute value of the correlation ≥0.56 is necessary for a significant association at P = 0.05. The relation between ρ and final patchiness was described by: \( \rho = 1.54 - 2.58(m^* / m) + 1.08 (m^* / m)^2 \);

**Fig. 6.** Estimated initial incidence (p₀) and weighted mean rate of increase
(p) of maize dwarf mosaic of maize in plots at three locations in Ohio versus
Taylor’s index of aggregation (b₁) estimated with equation 1 of the text.
The rank correlation between the variables are represented in the figures; an
absolute value of the correlation ≥0.56 is necessary for a significant
association at P = 0.05. Lines drawn correspond to a linear relation
between the variables. The closed triangle represents three data points that
lie on the same position.
models are appropriate for analysis and comparisons, and ultimately to discern why different models describe different epidemics.

Two assumptions of the epidemic models used in this study are a constant environment and a random distribution of diseased plants. Of particular relevance is that the biotic environment, as represented by the aphids in the plots, is assumed to be constant over time. When \( y \) was modeled as a function of aphids (Eq. 2), instead of assuming an invariable number, it was possible to describe accurately disease progression in terms of cumulative numbers of aphids. The development of models for virus disease epidemics as a function of vector numbers is still in its early, formative stages of development, and there is yet no commonly accepted approach to this research area. The goodness-of-fit of equation 2 is better than we expected, given the following limitations to the aphid data: A comprises many species that vary in their behavior on maize, and their transmission efficiency (15,20); the proportion of viruliferous aphids trapped, likely, varies over time; and aphids were collected in yellow pans that are biased estimators of species composition and landing rates (11). The

Fig. 7. Disease incidence \( (y) \) of maize dwarf mosaic versus the cumulative number of aphids \( (A) \) caught in yellow pan traps within 12 plots at three locations in Ohio. The drawn lines represent the predicted level of \( y \) as function of \( A \), based on the results in Table 3. There were not enough aphid data for plot M-5 to analyze with regression.
model, also, does not account for a latent period for MDMV development in maize that varies with the environment (9). Interestingly, modeling disease incidence as a function of \( A \) 1 wk earlier (approximate latent period) consistently gave lower coefficients of determination than the models presented here (Madden, unpublished). This may be due to totalizing aphid numbers per week when the latent period varies by a few days. For instance, aphids caught at the beginning of the current week may have more impact on disease than aphids trapped at the beginning of the previous week. Simultaneously incorporating species composition, variable proportions of viruliferous aphid individuals, and latent period would require computer simulation, as performed by Ruesink and Irwin (29). With the data used here, MDM disease progression was, nevertheless, highly correlated with cumulative aphid numbers.

Most of the MDM epidemics could be related to \( A \) without the empirically derived quadratic term. The Wooster epidemics, all described by the monomolecular model, were also all described by the simple form of equation 2. The cumulative number of aphids should increase with a slope of \( r/k_1 \) if \( y \) increases in a monomolecular fashion and the simple form of equation 2 also represents the epidemic (12). At Wooster, the cumulative number of aphids did increase in an approximately linear manner (\( R^2 = 0.91 \)) with a slope of 140 per day (S.E. = 22.2). Based on calculations with more significant digits than shown in Tables 1 and 2, \( r/k_1 \) for the three Wooster plots equaled 157, 171, and 165 per day. The confidence interval for 140 included these three values (\( P > 0.05 \)).

No specific reason can be given for the four epidemics that required a quadratic term to relate \( y \) to \( A \). Two were better described by the Gompertz model (M-1 and P-2) but the others by the logistic (M-3 and P-1). There was no association between the increase in \( A \) (linear or not) and the type of disease progression model or the form of equation 2. We attempted to fit the four epidemics that required a quadratic term with a logistic function, i.e., \( \ln(y/(1-y)) = k_0 + k_1 A \). However, only M-1 was adequately described by this model (Madden, unpublished). Further investigations to relate virus disease progress in relation to \( t \) and \( A \) are needed.

Appraisal of the pattern of MDMV-diseased plants depended on the statistical method used. Ordinary runs analysis generally indicated a progression from a random pattern early in the epidemic to a highly clustered pattern later. If MDM epidemics were initiated by seedborne infection or transmission by viruliferous aphids arriving from many distances, one would indeed expect to observe a random sequence of diseased plants. As the epidemic continues, one would then expect a clustered sequence of diseased plants if there was spread from plant to plant within the plots. This is supported by the data in M-1 to M-5 and also P-1 and P-6. All of these epidemics also are accurately described by compound-interest type disease progression models. Nonclustered patterns were always observed at W-6, a plot that had an epidemic described by the monomolecular model. W-5, on the other hand, always exhibited a clustered pattern, indicating plant-to-plant spread. Disease progress as a function of \( r \) or \( A \) in W-5, however, suggested a simple-interest type of increase with no spread from plant to plant. A side of W-5 and W-7 was adjacent to MDMV-inoculated maize, which violated the additional assumption of runs analysis that the plots need to be homogeneous. Caution should be given in interpreting runs analysis when there is a nearby outside source of inoculum.

Variability in the standardized runs of plots P-2 to P-5 present no immediate explanation. These epidemics quite possibly are due to a combination of within-field spread and spread from outside sources. At various times during the epidemics, either form of disease development may be dominant, as indicated by runs analysis. Difficulties with the way in which the analysis was conducted could also be responsible for the results. When patterns are assessed one up and down the "rows," the information derived from the analysis is limited. Two-dimensional approaches to analyze data of this type are possible (7), but the complexity of the analysis may discourage most researchers.

As with runs analysis, there were variable results when patterns of diseased plants were assessed with the techniques based on the variance and mean of contiguous quadrats. These spatial point pattern techniques can be used to detect randomness or clustering. Here, a random pattern implies that every plant in a field has an equal probability of becoming infected. Based on the \( V/m \) ratio or patchiness, most plots had a clustered pattern for one or more times; only two plots showed clustering for the entire epidemic. Patchiness declined with most of the MDM epidemics, indicating a tendency towards randomness. No decline occurred in plots that did not approach the maximum \( y \) (e.g., M-1). As with runs analysis, it was not possible to explain the epidemics in which patchiness rose to a maximum then declined (M-3 and P-6). As with ordinary runs, there also are problems with this type of analysis. The main difficulty with analysis of spatial point patterns is that the positions of the sample units (quadrats) are not used in the analysis. The size of the quadrats additionally can influence the results. Techniques of this type may not be able to distinguish between certain spatial patterns (24).

Analysis of aggregation with Taylor's technique was only partly successful. Based on \( V \) and \( m \), one would expect a basic agreement among the procedures when these statistics are used. Yet, several of the Taylor regressions were not significant and, with some plots, there was a disagreement between the results. For instance, \( b_t \) indicated clustering when tests at individual times indicated randomness, or \( b_t \) indicated randomness when the majority of the tests at individual indicated clustering. Thus, despite the success of Taylor's index to assess aggregation of a large number of insect species and other organisms (33), the technique was not very useful for MDM epidemics.

Two sources of difficulty can be identified when using Taylor's approach with epidemic data as conducted in this study. As originally proposed, the data used to determine different levels of \( m \) and \( V \) should be independent and have no upper limit. Because the same plants were observed over time with the MDM epidemics, the estimates of \( V \) and \( m \) at each time were not independent. When \( y \) and \( t \) are the dependent and independent variables, respectively, there are well-established procedures to "eliminate" the autocorrelation (18); it is not known how best to handle the case where \( V \) and \( m \) are the dependent and independent variables. The results in Table 2 for \( b_t \) should, therefore, only be considered approximate; the significance levels, likely, should be higher than those calculated. The second difficulty is due to the saturation of diseased plants as \( y \) approaches a maximum. Even though cases where \( y > 0.95 \) were not used, there may still be a significant effect on the limit to \( y \) on the results of Taylor's analysis.

**TABLE 3. Results from modeling maize dwarf mosaic disease incidence as a function of the cumulative number of aphids trapped in the plots**

<table>
<thead>
<tr>
<th>Field</th>
<th>Model parametersa ( k_0 ) ( k_1 \times 10^2 ) ( k_2 \times 10^3 )</th>
<th>( R^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>W-5</td>
<td>0.02</td>
<td>0.54</td>
</tr>
<tr>
<td>W-6</td>
<td>0.00</td>
<td>0.04</td>
</tr>
<tr>
<td>W-7</td>
<td>0.00</td>
<td>0.11</td>
</tr>
<tr>
<td>M-1</td>
<td>1.04</td>
<td>-27.33</td>
</tr>
<tr>
<td>M-2</td>
<td>-3.61</td>
<td>45.24</td>
</tr>
<tr>
<td>M-3</td>
<td>297.54</td>
<td>-952.0</td>
</tr>
<tr>
<td>M-4</td>
<td>-17.22</td>
<td>648.5</td>
</tr>
<tr>
<td>P-1</td>
<td>21.87</td>
<td>187.1</td>
</tr>
<tr>
<td>P-2</td>
<td>68.29</td>
<td>790.7</td>
</tr>
<tr>
<td>P-3</td>
<td>-0.09</td>
<td>31.80</td>
</tr>
<tr>
<td>P-5</td>
<td>-0.05</td>
<td>9.94</td>
</tr>
<tr>
<td>P-6</td>
<td>-0.03</td>
<td>2.54</td>
</tr>
</tbody>
</table>

aModels were of the following forms: ln(1/(1-y)) = k_0 + k_1 A + k_2 A^2, in which \( y \) is disease incidence, \( A \) is the cumulative number of aphids, and the \( k_x \) are parameters. The coefficient of determination (\( R^2 \)) is based on the agreement between the observed and predicted transformed \( y_x \).

bW = Wooster; M = Marietta; P = Portsmouth. In addition to P-4 and P-7, which had only three assessment times, aphids were not trapped for enough times in M-5 to perform a regression analysis. Data not shown for M-5 in Fig. 7.
The degree of aggregation at the beginning of the MDM epidemic was not related to the epidemic characteristics of initial disease, rate of increase or shape of the disease progress curve. Final aggregation (as measured by patchiness) was significantly related to the estimates of both $q_0$ and $q_e$. With MDMV, we therefore have evidence that disease progression determines final aggregation but that initial aggregation does not determine progression.

The dynamic aspects of spatial patterns of infected plants or of plant pathogens has been little studied. Even if spatial patterns were assessed at several times, investigators usually were not interested in how the pattern changed with time. We have found that the spatial pattern of MDMV-infected plants generally changes over time in an epidemic. In fact, changing pattern was the exception rather than the rule. Taylor et al (33, 34) have postulated that aggregation of organisms varies with population level (i.e., the mean). Because in plant disease epidemics, the mean generally changes, aggregation does also. Taylor et al (34) presented a functional form for the change, but the function assumes that equation 1 is always appropriate. With MDMV epidemics, no theoretical framework for the change is yet possible. Various empirical regression models were found at least partly acceptable for describing aggregation (patchiness or runs). Because aggregation is a fundamental characteristic of organisms (32), epidemiologists need to study further aggregation/time relations, both experimentally and theoretically.

LITERATURE CITED


