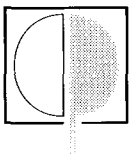
ELSEVIER

PII: S0261-2194(97)00055-0

Crop Protection Vol. 16, No. 7, pp. 603–618, 1997

0261-2194/97 $17.00+0.00



Quantifying the temporal dynamics of plant virus epidemics: a review

Forrest W. Nutter Jr
Department of Plant Pathology, Iowa State University, Ames, IA 50011, USA

The ongoing development of more sensitive and reliable plant virus detection methods offers new opportunities to accurately and reliably monitor the temporal dynamics of plant viruses within plant populations. This review provides operational definitions and examples for concepts pertaining to the sampling and assessment of host populations to quantify disease and/or pathogen incidence within host populations over time. The linear, monomolecular, exponential, logistic, and Gompertz population growth models are presented and discussed with regards to their use in modeling the temporal dynamics of plant viruses. Other quantitative measures of temporal disease (or virus pathogen) spread, such as the disease (pathogen) progress curve, area under the disease (or pathogen) progress curve, and time to reach 50% incidence (t_{50}), are also presented and discussed. © 1997 Elsevier Science Ltd

Keywords: disease and pathogen incidence; disease modeling; epidemiology; linear regression; sampling design and theory; statistical modeling

Quantitative comparisons of plant virus epidemics enable researchers to evaluate the effects of various disease management strategies and tactics on virus disease dynamics over time. Although the temporal dynamics of epidemics caused by fungi have been extensively studied, epidemics caused by plant viruses have not (Madden and Campbell, 1986). Unlike fungal diseases, that usually result in obvious symptoms and/or signs (leaf spots, blotches, cankers, wilts, etc.), it is often difficult to identify accurately virus-infected plants within host populations (crops) because obvious symptoms may be lacking or because multiple diseases, pests, and nutrient deficiencies may also cause virus-like symptoms. The confusion over virus symptoms (or lack thereof) has hampered efforts to model temporal spread in many virus pathosystems. However, modeling the temporal dynamics of plant viruses is becoming increasingly possible because of developments in virus detection methods that have enhanced sensitivity and specificity. These methods include, but are not limited to, enzyme-linked immunosorbent assay (ELISA), polymerase chain reaction (PCR), monoclonal antibodies, DNA fingerprinting and numerous other techniques (Hu *et al.*, 1995). Although the need is great, little fundamental research in population ecology has been directed towards plant virus pathosystems to meet today's agricultural needs. Moreover, there is an increasing need to quantify the epidemiological impacts that contemporary plant virus management practices will have on the temporal dynamics of plant virus epidemics, especially those involving genetically engineered organisms. For example, what will be the quantitative impacts of anti-sense or coat protein-mediated transgenic resistance strategies on the temporal dynamics of plant viruses? This information will be needed to assess the risk and benefits of new biotechnology-based resistance strategies.

This review focuses on assessing temporal changes in virus incidence (Nutter *et al.*, 1991) in plant virus pathosystems, how temporal change in virus incidence over time can be quantified, and how plant virus epidemics can be compared. The review also describes various sampling procedures, on how to select sampling units without bias, how to select the most appropriate population growth model to compare plant virus epidemics, and how mathematical modeling of plant virus epidemics can serve as an interpretive tool to quantify and compare the effects that plant virus disease management strategies and tactics have on the temporal patterns of virus spread.

Single point disease assessments

Surveys are often employed to determine whether a specific virus occurs within a defined geographical area (field, county, state, etc.), and surveys can provide important information concerning the identity, prevalence, incidence, and severity of plant viruses that affect crops at a specific time of sampling (Nutter *et al.*, 1991).

Because disease is the result of the interaction of a susceptible host, an aggressive pathogen, and a favorable environment, this interaction has inferred temporal and spatial requirements in that these three components of the disease triangle must interact within a defined spatial arena as well as during the

time of crop growth and development. Plant virus surveys are often conducted just once during a cropping season and, therefore, provide only a single 'snapshot' of the status of this dynamic interaction. Although a single assessment within a plant virus pathosystem can provide important information, a fuller understanding concerning the dynamics of plant virus epidemics normally requires multiple disease assessments. See Barnett (1986) for additional information regarding plant virus surveys.

Multiple point disease assessments

Plant virus epidemics are dynamic processes in that virus disease intensity changes in both time and space (Thresh, 1983). Thus, the quantitative study of virus epidemics requires that virus incidence within plant populations be assessed several times during the cropping season. Ideally, disease progress information should be obtained from several locations and over several years in order to define disease risk (Madden *et al.*, 1990). The power of multiple assessments, as opposed to the analysis of single assessments, is the ability to model temporal virus spread (disease progress). As will be shown later, a key to quantifying plant virus epidemics is to plot cumulative virus incidence (y) over time (t) and then to choose the most appropriate population growth model to allow the calculation and comparison of temporal rates of virus spread (Nutter and Parker, 1997). In addition to their use for quantifying the effects of virus disease management tactics on disease state and rate variables, multiple assessments are often required to model yield losses caused by plant viruses. This is because time of virus infection often has a significant impact on both crop yield and quality (Madden and Nutter, 1995; Nutter *et al.*, 1993).

Although plant virus infections are often systemic within an infected plant, the sampling unit that is often indexed or visually assessed is often a single leaf or other plant part. Thus, it is assumed that information obtained from a single plant part will accurately reflect the status of the entire plant; i.e. the entire plant is determined to be infected or not infected based on indexing or inspecting only part of the plant. This assumption may not be valid for plant genotypes that temporally restrict replication and movement of the virus within a host plant (Nutter, 1992), i.e. one leaf selected and indexed from a plant may test positive whereas other leaves selected and indexed from the same plant may test negative. This could lead to incorrect conclusions if the data are reported as 'incidence of infected plants' instead of 'incidence of infected leaves'. When the number of units sampled is large, the proportion of infected sampling units may closely represent the proportion of the infected plant tissue present in a field. It is also important to clearly define the population (disease versus pathogen) that is being assessed within a host population. This is particularly pertinent to plant virus pathosystems because the incubation period (time from inoculation to the time of symptom appearance) is usually longer than the latent period (time from inoculation to the time an infected host can be a source for new infections via insect, pollen, mechanical, or other means of virus transmission) in many plant virus pathosystems (Nutter, 1992).

Concepts and operational definitions for disease assessment terms

Before modeling the change in disease intensity (dy) with change in time (dt), we must first obtain accurate and precise measurements of disease intensity (Nutter *et al.*, 1991). *Disease intensity* is a general term for the amount of disease (injury) present in a host population (Nutter *et al.*, 1991), and the most common types of disease intensity measures are prevalence, severity, and/or incidence.

Prevalence is a term that is often used interchangeably (and mistakenly) with incidence (incidence is defined later in the section). *Disease prevalence* is more strictly defined as the number of fields (or other defined populations) within a specific geographic area (county, state, or region) where a disease has been visually observed (symptoms), divided by the total number of fields sampled and assessed (Campbell and Madden, 1990; Nutter *et al.*, 1991; Zadoks and Schein, 1979). To obtain information concerning *pathogen prevalence*, individual plants or plant parts are sampled from a host population (fields) and these samples are tested for the presence of the pathogen (virus) using a reliable method of indexing (infectivity, ELISA, PCR, presence of inclusion bodies, etc.) (Hu *et al.*, 1995; Kapa and Waterworth, 1981). Thus, pathogen prevalence may include quantitative information concerning the presence of a virus in asymptomatic as well as symptomatic fields. Prevalence data are often multiplied by 100 to give the percentage of fields (or populations) in which a disease or pathogen is present. Quite often, plant samples from a single geographical sampling unit (field, county, or state) are obtained, bulked, and processed as a single unit to ascertain whether the virus occurs within the geographic unit sampled. When plant samples are bulked, a single infected plant is sufficient to change the status of the geographic unit sampled (field, county, or state) from negative (−) to positive (+), provided the sensitivity of the indexing method is sufficient to detect the pathogen in a bulked (diluted) sample. Bulking samples is particularly useful when virus incidence is low because the number of bulked samples tested is often less than the number of individuals sampled and processed, thus reducing the cost of detection per sampling unit tested (Swallow, 1995). Clearly, prevalence data does not provide quantitative information about the relative amount of disease (or the virus) within the individual fields sampled. Within-field disease intensity measurements can be obtained by assessing plant populations within fields for disease incidence, pathogen incidence, and/or disease severity.

Compared to fungal and bacterial diseases, the concept of disease severity is usually problematical in virus pathosystems (see Campbell and Madden, 1990). The general definition for *disease severity* used here is the quantitative measure of the effect of a

plant virus on the health of an individual sampling unit. For example, disease severity might be assessed by measuring the degree of stunting of a plant that is infected with a plant virus (shoot length, plant height, or internode length), or by measuring the reduction in leaf area (L^2) relative to healthy plants. If a virus causes local lesions, disease severity may be defined operationally as the number of lesions per square centimeter (N/L^2) or the lesion area/total leaf area (L^2/L^2). Some researchers have developed disease rating scales in which a number is assigned to sampling units based on a written description of severity classes. Because in many plants viruses become systemic, Vanderplank (1963) considered virus-infected plants to be analogous (epidemiologically) to lesions in fungal pathosystems. Consequently, most virus epidemics are assessed using incidence data. Although the concepts of lesion size and lesion expansion of virus infections within whole plants is epidemiologically interesting and should be considered when sampling (Nutter, 1992), these concepts are not addressed further in this section.

A useful operational definition for *disease incidence* is the number of plant units visibly diseased by a virus divided by the total number of plant units sampled and assessed (see Madden and Hughes, 1995). As with the term prevalence, it is extremely important that researchers make a clear distinction as to whether incidence within a population of host plants is based on *pathogen incidence* (detection of a virus) or disease incidence (based on visual symptoms). Progress curves based on virus indexing methods, such as ELISA, may closely mirror progress curves based on disease symptoms such as the example shown in *Figure 1*; in other instances, however, the use of different assessment methodologies may result in progress curves with quite different shapes (*Figure* 2A). Consequently, researchers must decide which assessment methods provide the most meaningful information to achieve the objectives of their specific studies. This can be accomplished by regressing incidence values derived from one method (visual assessment, for example) with pathogen incidence data obtained by using a virus detection-based method (ELISA, for example). In the first example (*Figure 1*), regressing visual assessments for tobacco etch potyvirus (TEV) disease incidence in bell pepper (*Capsicum annuum* L.) (y) versus pathogen incidence based upon ELISA indexing (x), yielded the equation $y = 0.02 + 1.01x$ (coefficient of determinant, r^2 of 96.3%). Note that the slope of the regression line is very close to 1, the y-intercept is very close to zero, and the amount of variation in disease incidence (y) explained by pathogen incidence (x) is very high (96.8%). Thus, both incidence assessment methods are in close agreement. This is not so

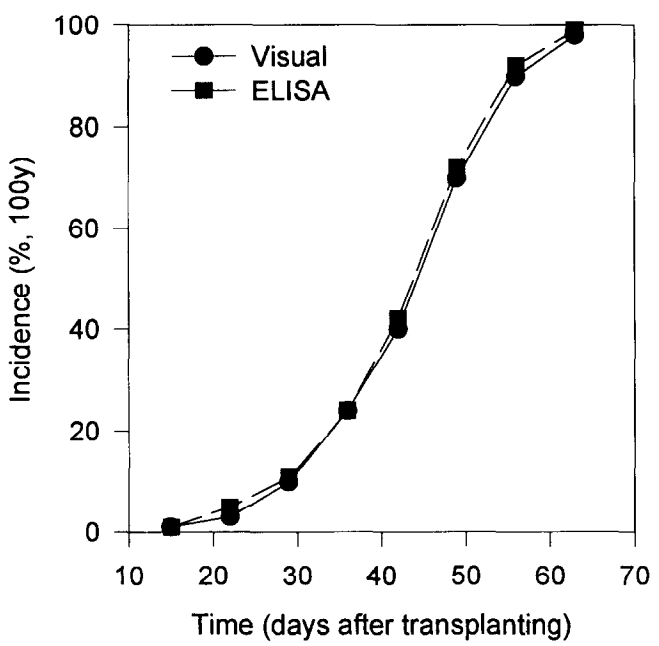


Figure 1. Comparison of disease progress curves for tobacco etch virus incidence in bell pepper based on visual inspection of plants versus indexing plants for the presence of TEV using an enzyme-linked immunosorbent assay (ELISA). The relationship between the two methods was highly significant ($r^2 = 0.963$, $P \leq 0.001$). Data from Padgett *et al.* (1990)

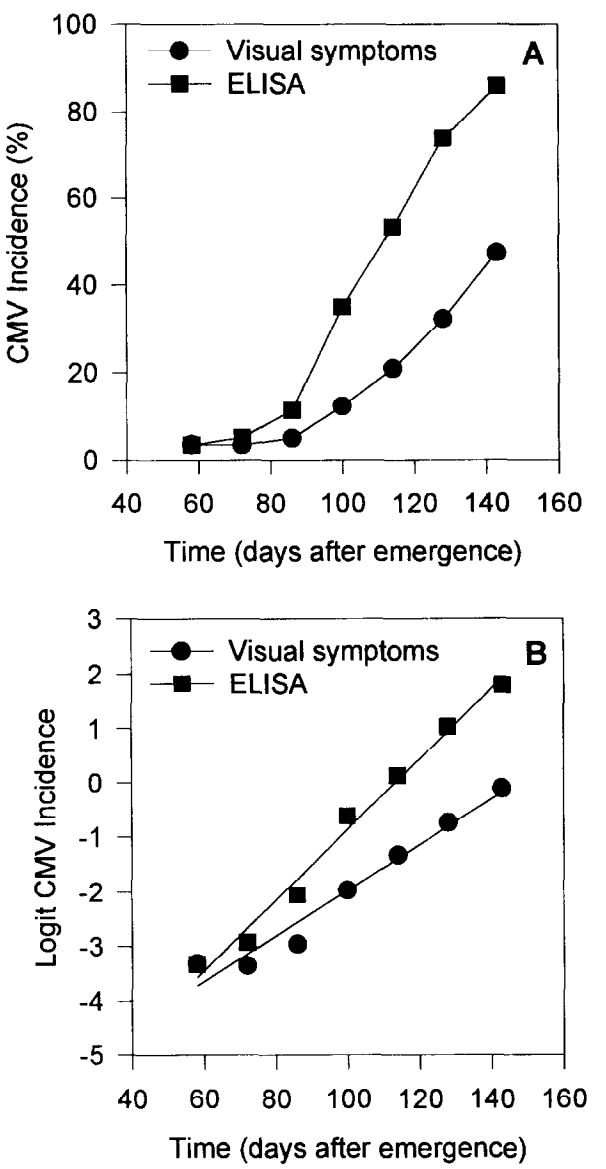


Figure 2. (A) Progress curves for disease caused by cucumber mosaic virus in narrow-leafed lupin based on visual symptoms as opposed to indexing by enzyme linked immunosorbent assay (ELISA). (B) Rate of CMV infection as a measure of epidemic risk based on visual symptoms (0.03 logits/day) versus ELISA (0.06 logits/day)

for the second example, however, as there is not good agreement between visual assessments for disease incidence caused by cucumber mosaic cucumovirus (CMV) in lupin (*Lupin angustifolius* L.) versus CMV incidence based on ELISA (Geering, 1992). See Nutter and Schultz (1995) regarding techniques to evaluate the accuracy and precision of assessment methods. Although there appears to be a very good relationship between visual and ELISA assessments for CMV disease and pathogen incidence, respectively ($r^2 = 93.5\%$), there is clear evidence that systematic bias is present as the slope of the regression line relating CMV pathogen incidence based on ELISA (x) to CMV disease incidence based on visual symptoms (y) was 0.49 (*Figure 3*). Because this slope deviated significantly from a slope of 1.0, the two assessment methods were not in agreement. The slope should have been close to 1.0 if the two methods were in agreement, and a slope of 0.49 means that for each 1% increase in pathogen incidence, disease incidence increased by 0.49% (about half) (Nutter and Schultz, 1995). There was no constant bias present between the two assessment methods as the y-intercept (-0.80) was not significantly different from zero. The methodology used in virus disease assessment is particularly important if the rate of infection (r) is taken as a measure of epidemic risk, i.e. the higher the infection rate r, the higher the risk. The infection rate (slope) based on ELISA in *Figure* 2B is nearly two times higher than the infection rate based on visual virus symptoms. Thus, when visual assessments are used to estimate the infection rate in this pathosystem, risk may be greatly underestimated. The effect of assessment method on estimating r underscores the need to utilize the *same* assessment methodologies when comparing infection rates for different epidemics (years, virus management tactics, etc.). Moreover, the assessment methods should be fully described when comparing and reporting virus rates of infection.

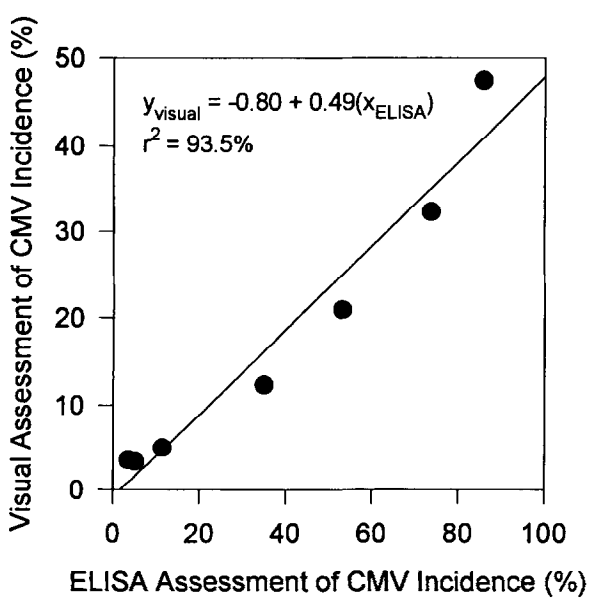


Figure 3. Linear relationship between CMV incidence based on ELISA (x) versus visual symptoms (y). The slope (0.49) was significantly different from 1.0, indicating the presence of systematic bias between the two methods

Moreover, information concerning the accuracy and precision of the assessment methods should also be reported (Nutter and Schultz, 1995).

The disease progress curve

Once a method for virus disease assessment and/or pathogen detection has been developed and verified as being accurate and precise, the next step is to decide the number of times during the period of crop growth and development that incidence needs to be assessed in order to accurately characterize the change in disease intensity over time. Because incidence is by far the most common measure of intensity employed in plant virus pathosystems, all the examples cited here will involve incidence data. Once incidence has been assessed several times during the cropping season, *disease (or pathogen) progress curves* can be constructed by plotting incidence on the ordinate axis versus time on the abscissa axis. Disease and pathogen progress curves provide a visual representation of the stimulus–response relationship between incidence and time (Zadoks and Schein, 1979). A disease progress curve is the 'signature' of an epidemic and represents the integration of all host, pathogen, and environmental effects (including virus vectors, alternative host reservoirs, and disease management tactics) that occur during the period of interaction (Campbell and Madden, 1990). If sampling is done properly and with sufficient care, accurate and precise estimates of plant virus epidemics can be obtained.

Defining a sampling unit

A sampling unit is defined as the smallest unit on which a disease assessment is made, whereas the sample population is the aggregate from which sample units for disease assessment are selected. For example, Nutter *et al.* (1994) defined a sampling unit as the newest, fully expanded soybean leaf selected from single plants in order to assess the incidence of soybean mosaic virus-infected plants using ELISA. In the tobacco etch virus (TEV) bell pepper (*Capsicum annuum* L.) pathosystem, Padgett *et al.* (1990) defined a sampling unit to be a whole pepper plant and individual plants were inspected for the presence of TEV symptoms to determine TEV disease incidence in the crop.

Sampling frequency and sampling interval

The time scale of multiple assessments can be relatively short (days) or long (years), depending on the type of crop. As a general rule, the minimum number of assessments required to define the shape of a disease (or pathogen) incidence curve and calculate temporal rates of virus spread is five, excluding 0 and 100% incidence values. The first assessment, which is a measure of the initial incidence of the disease or pathogen (y_0), should be made when incidence is near zero so that the rate of disease increase early in the epidemic can be quantified. Berger (1989) has suggested that the interval

between assessments should be half the approximate latent period, although time and resources will also impact on the time interval between assessment periods. Moreover, for some plant virus pathosystems, the timing and frequency of disease (or pathogen) incidence assessments may depend less on the latent period of the virus in the host population and more on the size, activity, and propensity of the viruliferous vector population (Madden and Campbell, 1986; Madden *et al.*, 1990). Ideally, assessments should continue at appropriate intervals until incidence values approach a level where no increase in incidence is discernible or until the crop is physiologically mature or harvested.

Determining sample size

Once the population of sampling units has been defined, the next step is to determine the number of sampling units that will be needed in order to achieve a specified level of precision. Sampling constitutes one of the most important activities in quantifying plant virus epidemics (Campbell and Madden, 1990; Nutter and Gaunt, 1996). The selection of sampling units from a population is often undertaken because generally one cannot take a complete census of all the plants in an entire field or even within an experimental plot. A census is a complete enumeration of all the sampling units in a population. Because of both logistics and costs, a complete census of the disease status of each individual in a host population is rarely undertaken; this is particularly true of field crops where the number of plants or sampling units per unit area can be extremely large. However, when the number of sampling units per unit area is relatively small, such as in orchards, plantations, and vineyards, the enumeration of all sampling units in a host population may be possible. For example, Habili and Nutter (1997) used several methods of virus indexing to determine the incidence of grapevine leaf-roll-associated virus 3 in Pinot Noir grapevines over an 11-year period for all grapevines ($N = 104$) in an experimental plot located in South Australia. When sampling is used to estimate incidence in a host population, the accuracy (closeness to the truth) and precision (reliability) of an estimate of disease and/or pathogen incidence in a host population is related directly to sample size (Madden *et al.*, 1996). It is often useful to have some preliminary data available to help determine sample size, i.e. the number of sampling units required to conduct a valid study (Campbell and Madden, 1990). An approximate determination of the number of samples (n) is needed to help ensure the acquisition of reliable disease assessment data. Too few samples will result in unreliable and/or unrepresentative data, but an excess number of samples is not cost-efficient. Often the number of samples must be optimized with the resources (labor and/or dollars) available to conduct the study. An approximate indication of the number of samples needed can be obtained by plotting standard deviation as a function of increasing sample size. Campbell and Madden (1990) state that these curves will often flatten after a certain number of samples are assessed, and the point of relative stability of these curves may be chosen as an approximate measure of sample size.

Sampling designs

The purpose of a sampling design is to optimize the accuracy, precision, speed, and scope of obtaining a reliable sample relative to costs. Basically, there are three types of sampling designs employed to obtain disease intensity estimates from field studies; these are random, arbitrary, and systematic designs. Although many scientists state that they have employed a simple random sample to select sampling units, this is often not the case. Each sampling unit from the population of sampling units must have an equal chance of being selected in a truly random sampling design. The primary reason for taking a random sample is to remove any possibility of bias, unconscious or otherwise, on the part of the person collecting the samples for disease assessment. Biased sampling, by definition, leads to an overestimation or underestimation of the 'true' level of disease intensity. For a truly random sample, each sampling unit in the population (leaf, stem, plant, unit length of row or ground area, etc.) would be numbered from 1 to n and a random number generator (or some other accepted method) would be used to select the sampling units to be assessed. For example, a random sample might be employed when the number of sampling units per unit area is relatively small (such as an orchard or plantation), yet the number of sampling units is too large to conduct a complete census. More often, scientists working in plant virus pathosystems employ arbitrary or systematic sampling schemes. Arbitrary sampling, for example, is done when a person walks into a field and stops on a whim to select a sampling unit. Campbell and Madden (1990) stated 'Such sampling is not necessarily inappropriate in many applications, but it (arbitrary sampling) should be identified for what it is and not referred to erroneously as random sampling.' From my own practical experience, arbitrary sampling schemes have the potential to introduce bias if the sampling units are visually selected. This is because, unconsciously, the eye may be preferentially directed to select diseased versus healthy sampling units and this will lead to an overestimation of the true disease incidence in the host population. One way to reduce this source of bias would be to reach for and select sampling units by touch, rather than by eyesight. One should be aware, however, that when selecting sampling units by touch, another source of bias may be introduced. This form of bias may occur in virus pathosystems in which the infecting virus causes stunting of the host and thus taller, virus-free plants would have a higher probability of being selected for assessment compared to shorter, virus-infected plants. This form of bias would result in the underestimation of the actual virus incidence in a host population. To avoid such problems of bias, Benner *et al.* (1995) devised a sampling scheme in which transects of 25 consecutive pepper plants were individually assessed

for disease incidence (symptoms) beginning at arbitrary starting points.

Because it is often difficult to identify and number sampling units as required for completely random sampling, many researchers employ a systematic sampling design. In systematic sampling, the person collecting samples does so in a predetermined pattern, starting at an arbitrarily (or randomly) chosen point and then obtaining a sampling unit every hth unit, where h is a prespecified distance and direction away from the previous sample.

Any of these sampling designs are sometimes employed after dividing a field or plot into non-overlapping subpopulations called strata. Strata are often quadrats or strips of a field and are numbered N_1, N_2, N_3, to N_n. After strata are established, random, arbitrary, or systematic samples are taken from within each stratum. Stratification prior to sampling has the benefit of ensuring that a specified number of samples will be selected from each section of a field. The gain in precision due to stratification can be estimated by comparing the variance in the stratified sample with an estimate of what it would be if a simple random sample had been used. In addition to the above sampling designs, the use of cluster sampling, inverse sampling, and sequential sampling to estimate mean disease incidence with a prespecified degree of precision have recently been discussed by Hughes *et al.* (1996) and Madden *et al.* (1996).

Sampling patterns

Several sampling patterns can be used in conjunction with the above sampling designs to estimate disease intensity. Examples of sampling patterns include the V, X, W, diamond, and the diagonal (line) transect. For example, using arbitrary sampling with a W sampling pattern, a researcher would arbitrarily select n number of samples along each arm of the W (i.e. four arms). A systematic W pattern would specify the interval (distance or plant number) between samples selected within an arm. When disease is random within a field, increasing the number of samples (n) is more beneficial in terms of accuracy and precision than increasing the number of sampling arms. Conversely, if disease is aggregated in a field or plot, more benefit is obtained by increasing the number of sampling arms.

Statistical analyses of disease progress data

Simple linear regression

There are two common approaches to quantify and compare the temporal dynamics of epidemics, and these approaches are not mutually exclusive. The first is to use population growth models coupled with *linear regression* to calculate descriptive parameters such as the rate of disease (or pathogen) progress (r) and the initial incidence of disease (y_0). These two parameters may be obtained with linear or nonlinear regression analysis (Madden and Campbell, 1990). With linear regression, y is transformed based on the appropriate model (e.g. $y^* = \ln(y)$), and the estimated slope parameter is one estimate of r. The parameter for the intercept with the ordinate is a measure of transformed initial disease (or pathogen) incidence (y_0^*). Temporal rates of progress can then be compared statistically, but these comparisons are valid *only* when the same population growth model is used to calculate rate parameters for each of the disease or pathogen progress data sets being compared. When different models are required, one must use either a flexible nonlinear model or calculate the mean weighted absolute rate of change in each progress curve (Campbell and Madden, 1990).

Coefficient of determination

In addition to the slope and intercept parameters, statistics that can be obtained include the coefficient of determination (r^2), the coefficient of variation (cv), and the standard error of the estimate for y (SE_y). The r^2 indicates the proportion of the variation in the response (y) that can be explained by the stimulus (e.g. t) (Madden *et al.*, 1990). However, r^2 values should be interpreted with caution. High r^2 values may be obtained for more than one model simply because the amount of y present increases over time and higher slopes automatically give higher r^2 values. Moreover, there is often a high degree of autocorrelation among disease intensity values assessed over time (Campbell and Madden, 1990).

Coefficient of variation

The coefficient of variation (cv) provides a good overall index as to the precision with which epidemics are quantified. It is calculated by expressing the error mean square (variance) as a percentage of the overall mean, and thus has the advantage of standardizing the units for inherently different methods of measurement (e.g. disease incidence using visual assessment versus pathogen incidence based on ELISA indexing). To compare single point assessments of virus incidence at one time, for instance, the cv is given as:

$$cv = \frac{\text{Error Mean Square}}{\text{Grand Mean}} \times 100$$

When linear regression is used, the following equation can be used to calculate the cv:

$$cv = [\sqrt{(\text{Error Adjusted Mean Square of } y}/\text{Grand Mean of } y)] \times 100$$

This equation can be used to compare regression results for virus incidence based on visual assessments versus serological indexing. The higher the cv, the lower the precision of the assessment method; therefore, the cv provides a dimensionless numerical value to compare assessment methods or raters (Nutter and Schultz, 1995). For example, the cv may be affected by the experience of the rater when visually assessing virus disease incidence in a crop. The sensitivity and/or specificity of the assessment method (e.g. use of monoclonal versus polyclonal antibodies) may also impact the cv as it is quite possible that more specific virus detection methods such as PCR may actually

Table 1. Five population growth models commonly used to describe temporal disease progress. The EPIMODEL computer program reference uses the linearized forms of each equation to transform disease severity values and compute regression statistics[a]

Model	Integrated expression	Absolute rate equation	Linearized equation
Monomolecular	$y = 1-(1-y_0)\exp(-r_M t)$	$dy/dt = r_M(1-y)$	$\ln[1/(1-y)] = \ln[1/(1-y_0)]+r_M t$
Exponential	$y = (y_0)\exp(r_E t)$	$dy/dt = r_E y$	$\ln(y) = \ln(y_0)+r_E t$
Logistic	$y = 1/\{1+[(1-y_0)/y_0]\exp(-r_L t)\}$	$dy/dt = r_L y(1-y)$	$\ln[y/(1-y)] = \ln[y/(1-y_0)]+r_L t$
Gompertz	$y = \exp\{[\ln(y_0)]\exp(-r_G t)\}$	$dy/dt = r_G[-\ln(y)]$	$-\ln[-\ln(y)] = -\ln[-\ln(y_0)]+r_G t$
Linear	$y = b_0+r_l t$	$dy/dt = r_l$	$y = b_0+r_l t$

[a]y, disease (or pathogen) incidence; t, time; r, rate parameter; y_0, constant of integration corresponding to y at $t = 0$

have a higher *cv* than less specific methods such as ELISA due to the specificity of strain detection.

Standard error of the estimate for y (SE_y)

The standard error of the estimate for (SE_y) is another way to evaluate the precision of disease assessments obtained by different raters or methods. The precision of disease assessments is indicated by the standard error of the raters' disease intensity estimates around the predicted values obtained by regression. In general, the smaller the SE_y, the better the fit of the data to the predicted line. However, SE_y values obtained using different models or assessment methods cannot be directly compared because SE_y values are influenced by the scale used to describe disease intensity. For example, consider the range of numerical values encountered when disease incidence is expressed as a proportion (0.0 to 1.0) or on a logistic scale (−9.2 [or lower] to +9.2 [or higher]). See chapter 8 in Campbell and Madden (1990) on comparing statistics obtained from different models. The standard error of estimate for y is computed as follows:

$$SE_y = S_y\sqrt{(1-r^2)}$$

where S_y is the standard deviation for y.

Plot of residuals

Finally, it is often helpful to examine subjectively the plot of residuals ($y_{actual}-y_{predicted}$) versus t or $y_{predicted}$ to help determine how well the regression line fits actual disease data. An obvious pattern in the plot of residuals indicates that the data set is not adequately explained by the model under consideration (Nutter and Parker, 1997).

Area under the disease progress curve

A second common approach to quantify disease progress curves (epidemics) is to calculate the area under the disease progress curve (*AUDPC*). This method is particularly useful when there is no single population growth model that best fits all of the disease progress curves among the set of curves to be compared, or if there are too few assessments. This 'synoptic' approach is the numerical integration over time of the proportion or percentage of the host population that is diseased. Bear in mind that *AUDPC* values are based on disease incidence data and that area under the pathogen (incidence) progress curve (*AUPPC*) values can also be calculated. *AUDPC* and *AUPPC* are usually calculated by trapezoidal integration (Berger, 1989) and, therefore, the accuracy of these values as a measure of the virus epidemic depends on the time interval between assessments. *AUDPC* and *AUPPC* values calculated from progress curves of different durations can be normalized by dividing each *AUDPC* (or *AUPPC*) value by the time duration from the first to last assessment of the epidemic.

Other measurements that are derived from progress curves and can be used to compare epidemics include the date of disease onset (from time when $y > 0$), the time to reach a certain level of y (e.g. 50% incidence), and the maximum disease (pathogen) incidence (y_{max}). When experimental treatments are replicated, temporal rates, *AUDPC* (or *AUPPC*), y_{max}, and time to, say, 50% incidence values can be calculated for each replication and then treatment effects for these estimated parameters can be evaluated with analysis of variance (Padgett *et al.*, 1990). The above discussion on statistics is relevant for *AUDPC* (*AUPPC*) and other variables as well.

Population growth models

Model selection

The selection of a temporal model that best describes the progress data of plant virus epidemics involves consideration of both subjective and objective criteria (Madden, 1980, 1986; Waggoner, 1986). The temporal population growth models commonly used to analyze plant disease epidemics are the linear, monomolecular, exponential, logistic, and Gompertz population models (Nutter and Parker, 1997). Expressions for the absolute rate of change in y with t (dy/dt) and the integrated and linearized forms of these models are given in *Table 1*.

To select the most appropriate population growth model for virus progress data sets, several steps are involved (Campbell and Madden, 1990). One of the first steps is to graph y (usually disease or pathogen incidence) versus time t. As will be shown later with specific examples, the shape of the incidence progress curve is very helpful in identifying the best model. The next step is to graph the estimated dy/dt versus t. The shape of these curves will also help to identify the model that will best fit virus progress data (see below). If the appropriate model is chosen, the plot of the y (y^*) versus (t) should approximate a straight line.

After viewing the above graphs, objective criteria for the acceptance or rejection of a model are provided by the estimated regression parameters and statistics described earlier in this section such as the r^2, cv, SE_y and standard deviations of parameter estimates (Campbell and Madden, 1990). Finally, the inspection of residual plots reveals if the model is satisfactory. For additional information concerning these and other temporal models used to quantify disease dynamics, the reader is referred to Berger (1989), Campbell and Madden (1990), Madden and Campbell (1990), Nutter and Parker (1997), Vanderplank (1963), and Waggoner (1986).

The simple linear model

The simple linear model is appropriate to quantify and compare epidemics for which dy/dt appears constant throughout the epidemic, resulting in a straight line (linear) progress curve. Unlike the other models discussed here, the linear model is too simplistic to predict or explain the biological characteristics of an epidemic. The linear model, however, can be used to provide simple descriptions of epidemics and to give information concerning the epidemiological benefits of various disease control tactics (Gray *et al.*, 1994). A good example is the disease progress curve for a virus disease of chili (Roff and Ho, 1991) (*Figure 4*A). Note that the rate curve (dy/dt vs t) for the linear model is approximately a horizontal line (*Figure 4*B) and that the simple linear model provides a very good fit to this epidemic (*Figure 4*C).

The monomolecular model

The monomolecular model assumes that dy/dt is greatest at the beginning of the epidemic and that dy/dt slows in direct relation to the remaining amount of disease-free (or pathogen-free) plant tissue $(1-y)$ (*Table 1*). For simplicity, it is assumed here that y_{max} is 1 (100%). The expression $1-y$ accounts for the constraints to further disease increase caused by the lack of healthy plants. At low levels of disease y, the expression is approximately $dy/dt \cong r_M$, and therefore dy/dt is not dependent on y and is similar to the linear model. This model is also called the negative exponential model (Campbell and Madden, 1990) or the 'simple interest' model (Vanderplank, 1963). The monomolecular population growth model is applied to epidemics for which there is no spread from plant-to-plant, i.e. there is no secondary spread within a growing season. The monomolecular model can also be used to quantify changes in virus incidence over years. For example, Jones (1979) annually recorded the incidence of raspberry virus in raspberries based on visual symptoms beginning at planting (May 1971) through 1978. The disease progress curve for incidence of raspberry virus over this 8-year period is shown in *Figure 5*A. Note that the rate dy/dt decreases with time for this model (*Figure 5*B) and that the monomolecular model provides a good fit to this epidemic as indicated by the regression line and equation in *Figure 5*C.

The exponential model

The exponential model (also known as the logarithmic, geometric, or Malthusian model) is the simplest of the two 'compound interest' models of Vanderplank (1963). It is appropriate for when there is virus spread from plant-to-plant over time; that is, new diseased (infected) individuals lead to more diseased (infected) individuals. Spread may be from plant-to-plant within a crop and/or be the result of virus spread from infected alternative virus hosts (reservoirs) located within or outside the crop to crop plants within a field. The exponential model dates back at least to Malthus who, in 1798, used it to predict future increase in the human population. This model assumes that dy/dt increases throughout the epidemic and that the absolute rate of virus increase is directly proportional to the present level of virus incidence as well as the rate of infection r. This

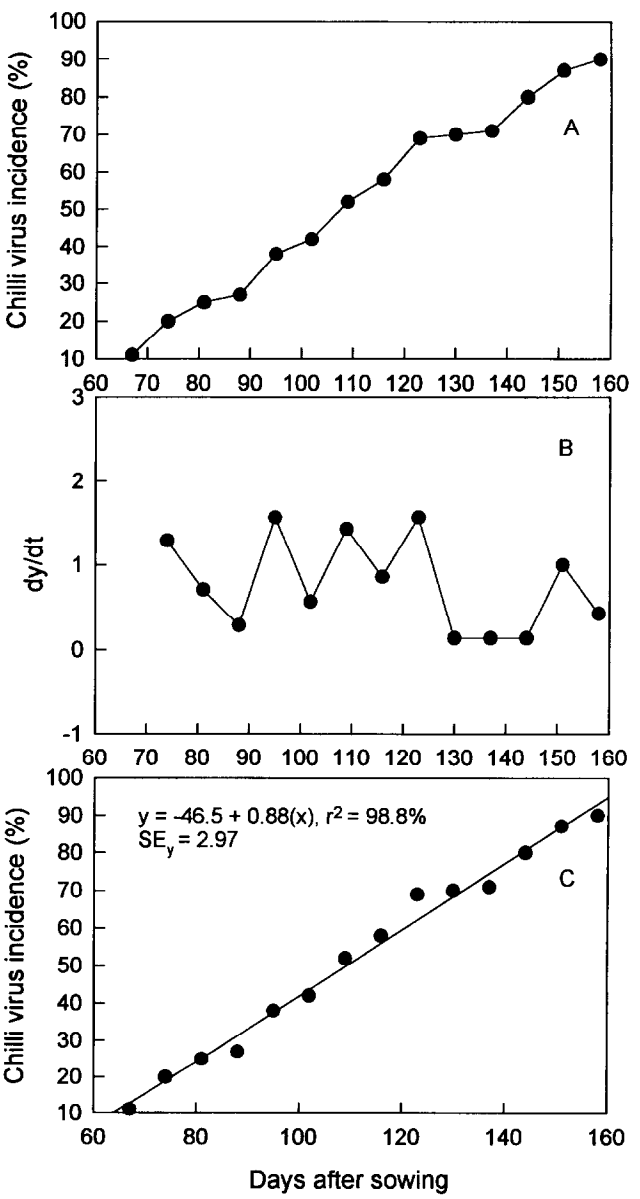


Figure 4. (A) Disease progress curve for chili virus disease incidence (y) in chili pepper. (B) Rate of change of chili virus incidence (dy/dt) versus time for the same epidemic. (C) Linear regression line and equation for the increase in chili virus incidence versus time; $y = 46.5+0.88(t)$, $r^2 = 98.8\%$, $SE_y = 2.97$

model also implies that factors such as the carrying capacity of the host crop do not limit the rate of disease increase and, therefore, virus progress curves are typically J-shaped and lack an inflection point. A plot of the rate curve, dy/dt vs t, also has no inflection point and dy/dt increases with time. The exponential model can be used to model changes in disease prevalence on a geographic scale and can be applied to describe the very early phases of most polycyclic (compound interest) epidemics. This is because, when y is very small, there is little effect of healthy host level on dy/dt ($1-y$ is close to 1).

*Figure 6*A shows a pathogen progress curve for a soybean mosaic virus epidemic in soybean (Nutter *et al.*, 1994) in which the exponential model would be the most appropriate model to obtain a linear relationship between transformed pathogen incidence ($\ln y$) and t. The date of sowing for this experiment was 17 June 1992 (day of year 168). The sampling unit defined in this study was unique in that individual leaves or plants were not assessed. The sampling unit was defined as a quadrat and each quadrat was 30 cm long and one row wide. Each plot consisted of 25 quadrats per row by six rows (i.e. there were 150 quadrats per plot). A single leaf was sampled from each plant within a quadrat (three to four plants) and the leaves from each quadrat were bulked and then tested for the presence of SMV using ELISA. The incidence of quadrats testing positive for SMV was then plotted with respect to time. Note that there is no inflection point in the pathogen incidence curve shown in *Figure 6*A and that dy/dt also increases with time (*Figure 6*B). The linearized model provides a good fit to the data as indicated by a high r^2 and low SE_y (*Figure 6*C). The slope of the regression equation

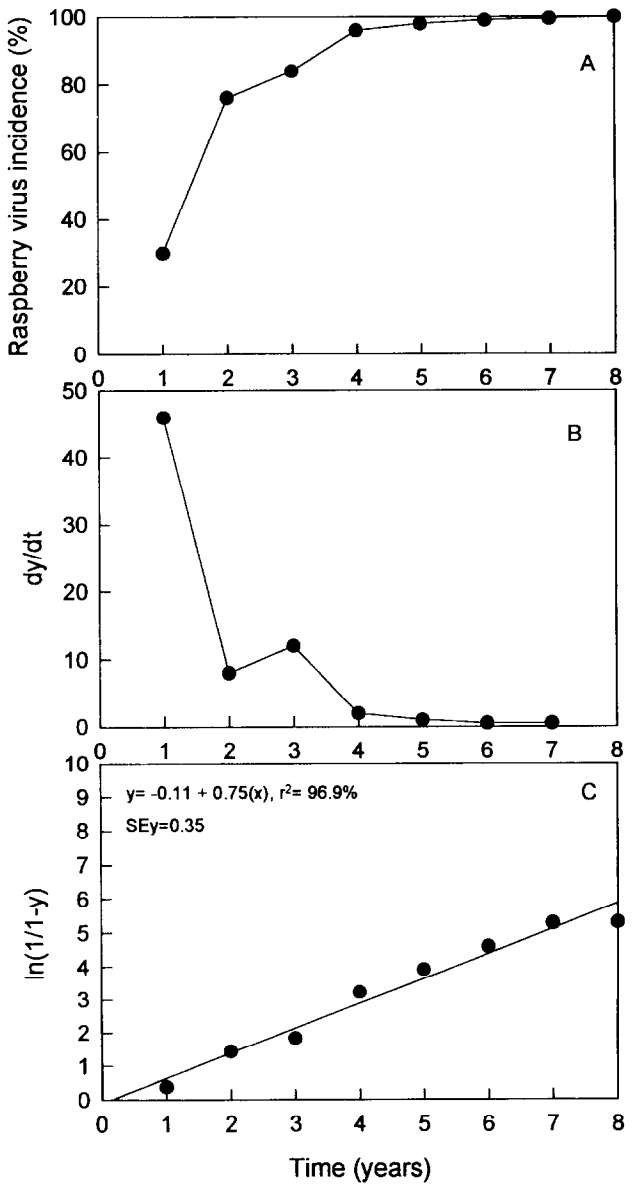


Figure 5. (A) Disease progress curve for raspberry virus incidence in raspberry. (B) Rate of change of raspberry virus incidence (dy/dt) versus time for the same epidemic. (C) Linear regression line and equation using the monomolecular model $\ln[1/(1-y)]$ versus time to transform incidence data; $\ln[1/(1-y)] = -0.11+0.75(t)$, $r^2 = 96.9\%$, $SE_y = 0.35$

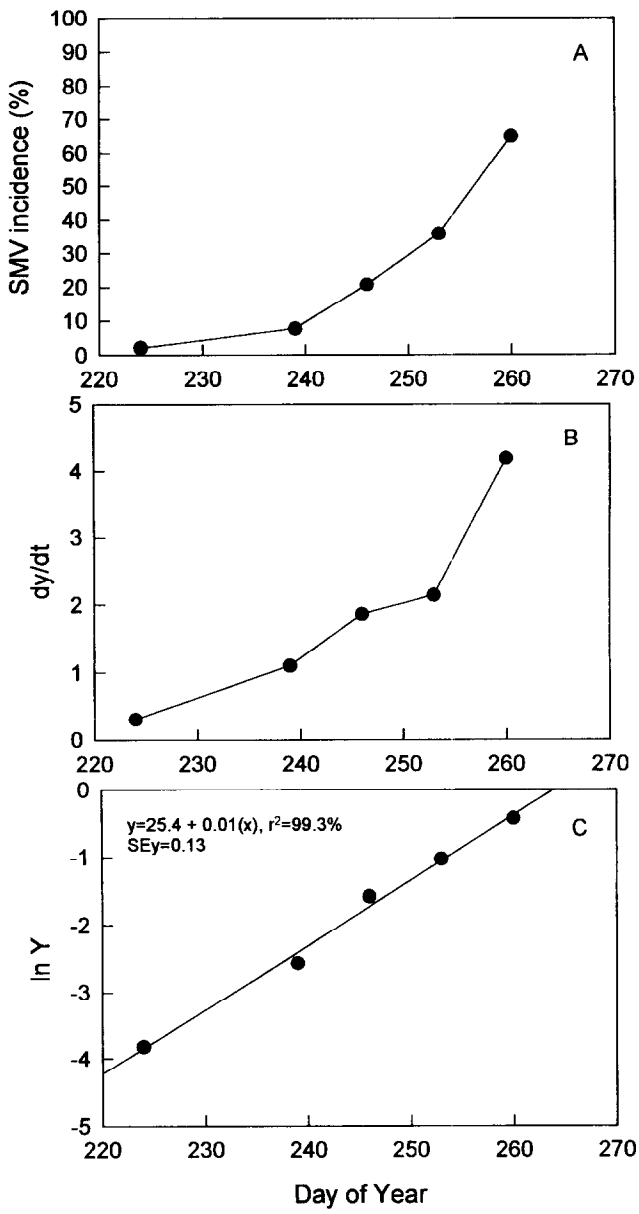


Figure 6. (A) Progress curve for incidence of soybean mosaic virus (SMV) in soybean cv. 'Corsoy'. (B) Change in rate of SMV incidence (dy/dt) versus time for the same epidemic. (C) Linear regression line and equation using the exponential model [$\ln(y)$] versus time to transform incidence data; $\ln(y) = 25.4+0.10(t)$, $r^2 = 99.3\%$, $SE_y = 0.13$

(0.10/day) is a measure of the exponential rate of pathogen increase per day.

The logistic model

The *logistic model* was proposed by Verhulst in 1838 to represent human population growth and is probably the model most frequently used to describe viral epidemics. It is the second type of compound interest model of Vanderplank (1963) and may be appropriate for entire epidemics ($0<y<1$) of viral diseases where there is spread from plant-to-plant. Disease and pathogen progress curves typically have a characteristic sigmoid (S-shaped) form, with an inflection point at the time when virus incidence reaches a proportion of 0.5 (50%). The absolute rate curve is symmetrical, with the highest rate occurring when $y = 0.5$. A biological interpretation is that, early in the epidemic, dy/dt accelerates as y approaches 0.5 because an increasing number of newly diseased plants become infectious and contribute to more diseased plants. At later stages, the diminishing incidence of healthy plants remaining $(1-y)$ limits the rate of increase. This relationship can be expressed in a differential equation for the rate of increase as $dy/dt = r_L y(1-y)$. Thus, the absolute rate of virus disease increase is proportional to the incidence of noninfected plants late in the epidemic $(1-y)$, the level of virus incidence early in the epidemic, and the rate of infection (r_L).

A typical S-shaped disease progress curve is demonstrated in *Figure 7*A for the disease caused by tobacco etch potyvirus (TEV) in bell pepper. Pepper plants were transplanted to the field on 6 June 1986 (day of year 157) and all pepper plants (census) in plots of six rows by 24 plants per row ($n = 144$ plants) were visually assessed each week for symptoms typical of those caused by TEV. The first diseased plants were observed on 3 July (day of year 184) and TEV disease incidence was 99% by 20 August (day of year 232). The absolute rate curve is bell-shaped (*Figure 7*B), which further indicates that the logistic model would probably provide a good fit to linearize disease progress data (*Figure 7*C). The slope of the regression line for transformed y versus t (r_L) is 0.21/day. It should be noted that r_L is an overall measure of the host, pathogen, and environment interaction, including vector population dynamics. Using this model, the independent variable (time) explained 96% of the variation in the transformed disease incidence and the SE_y was quite small, indicating that the logistic model provided a good fit to the data.

The Gompertz model

The Gompertz model is borrowed from an animal growth study and was originally proposed in 1825 (Madden, 1980). As with the logistic model, the progress curve has an inflection point, but it is located when $y = 0.37(1/e)$, and a large portion of the area under the rate curve is located to the right of the inflection point. The absolute rate curve reaches a maximum more quickly and then declines more gradually than with the logistic model. The model could be appropriate for polycyclic (compound interest) diseases as an alternative to the logistic model. The 'correction factor' for decreasing healthy host plants of the logistic model $(1-y)$ is replaced by $-\ln[-\ln(y)]$ in the Gompertz model.

An example of a disease progress curve that is best fitted by this model comes from Madden *et al.* (1987) for a TEV epidemic in tobacco (*Figure 8*A). The tobacco cultivar Kentucky 14 was transplanted to the field during 23–28 May 1985. An experimental plot consisted of 22 rows of 150 plants each and tobacco plants were assessed for the presence of tobacco etch virus symptoms each week in the experimental plot. The disease progress and the absolute rate curves are both skewed (*Figure 8*B). Because of the logarithmic function in the equation for dy/dt, the appropriateness of this model implies that equal proportions of

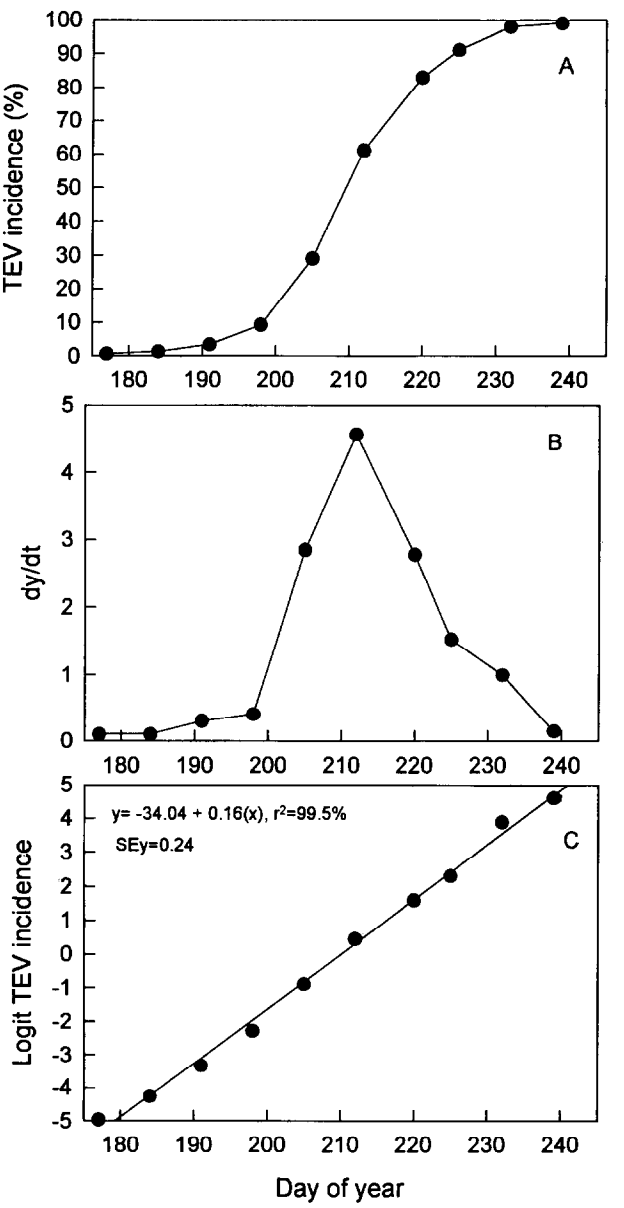


Figure 7. (A) Progress curve for incidence of tobacco etch potyvirus in bell pepper cv. 'Yolo Wonder B' in Georgia. (B) Change in rate of TEV incidence (dy/dt) versus time for the same epidemic. (C) Linear regression line and equation using the logistic model $[\ln y/(1-y)]$ versus time to transform incidence data; $[\ln y/(1-y)] = -34.04+0.16(t)$, $r^2 = 99.5\%$, $SE_y = 0.24$

the pathogen's ability to increase are reduced with an increase in time. This may possibly be due to the temporal dynamics of the aphid vector population and/or a trend towards increased host resistance as the crop matures (Webb and Nutter, 1997). *Figure 8*C shows that the Gompertz model provided an excellent fit to the transformed data, with an r^2 value of 98.9% and a very small SE_y.

Analysis and interpretation of disease progress curves in selected virus pathosystems

Example one

The first example involves three disease progress curves obtained from the TEV/bell pepper pathosystem (*Figure 9*A). The experiment was planted on 1 June 1987 (day of year 152) and each treatment was replicated four times. The first curve (closed circles) shows the effect of having a single TEV-infected plant present in a $8.3 \times 8.0\ m^2$ plot at the time peppers were transplanted to the field. The second curve (open triangles) depicts disease progress typical of that found in grower's fields whereby no weeds are present within the crop at transplanting and all pepper transplants are initially disease-free. The third disease progress curve (closed triangles) shows the effect of aluminum plastic mulch in delaying both the onset and rate of TEV spread with time. The latter control tactic is thought to repel aphids that may already have acquired TEV from alternative weed

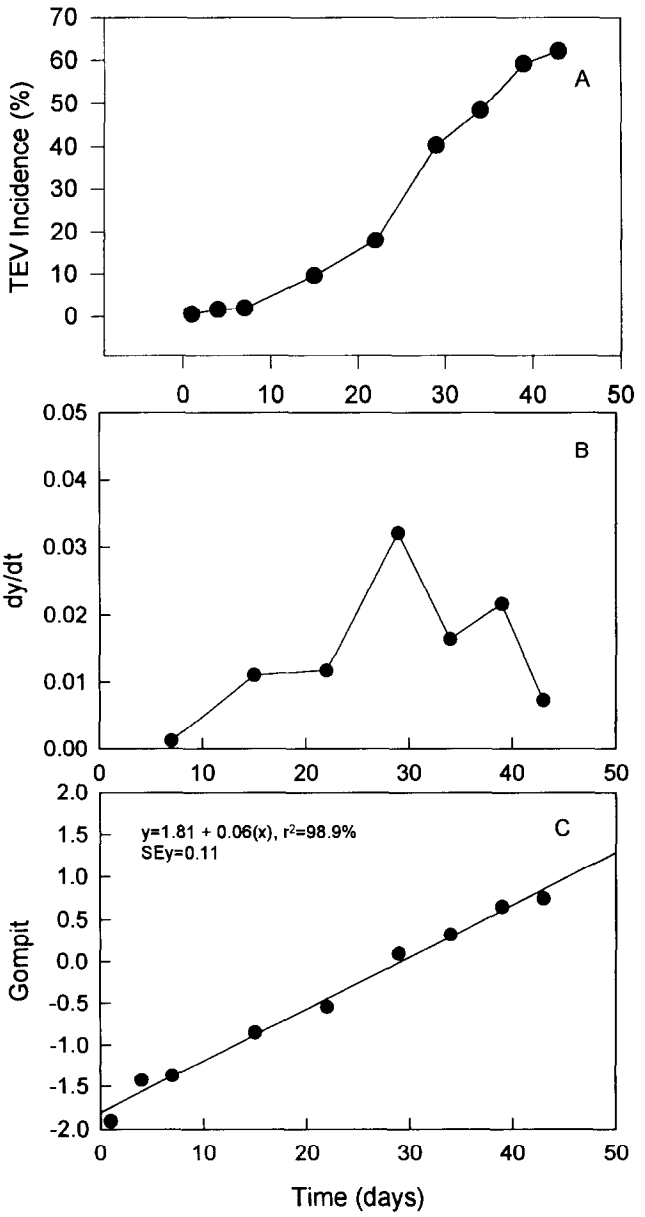


Figure 8. (A) Progress curve for incidence of tobacco etch potyvirus in tobacco in Kentucky. (B) Change in rate of TEV incidence (d*y*/d*t*) versus time for the same epidemic. (C) Linear regression equation using the Gompertz model $-\ln[-\ln(y)]$ versus time to transform incidence data; $-\ln[-\ln(y)] = 1.81+0.06(t)$, $r^2 = 98.9\%$, $SE_y = 0.11$

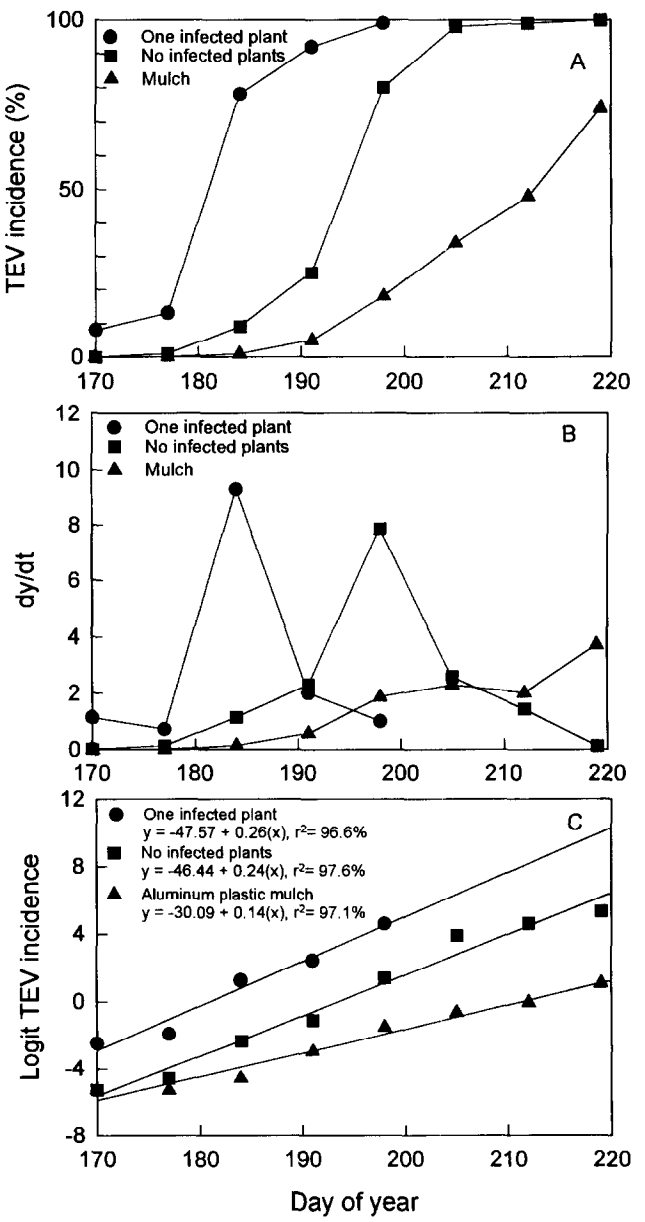


Figure 9. (A) Disease progress curves for three epidemics of tobacco etch potyvirus (TEV) incidence in bell pepper as affected by (i) a single TEV-infected plant present in the plot at the time peppers were transplanted, (ii) no infected plants at time of transplanting, and (iii) use of aluminum plastic mulch. (B) Rate of change in TEV virus incidence versus time for the same three epidemics. (C) Linear regression lines and equations using the logistic model $[\ln y/(1-y)]$ versus time to transform incidence data; ● $-47.57+0.26(t)$, $r^2 = 96.6\%$; ■ $-46.44+0.24(t)$, $r^2 = 97.6\%$; ▲ $-30.09+0.14(t)$, $r^2 = 97.1\%$

hosts found in bordering pepper fields (Summers *et al.*, 1995). Thus, the epidemiological effect of this tactic is to reduce the amount of virus inoculum entering the field as well as the number of aphids that could potentially spread TEV within the field.

Obviously, the presence of a single TEV-infected plant at transplanting enhances disease progress and the use of aluminum plastic mulch restricts disease progress, compared with the center disease progress curve which represents standard cropping practices. Note also that the aluminum mulch treatment changes the shape of the disease progress curve from sigmoidal to exponential, although there are no y values above 70% in this curve. The absolute rate curves indicate that the logistic model provides a good fit to the other two progress curves but may not be the best transformation for the aluminum mulch treatment because of the continually increasing dy/dt (*Figure* 9B). This curve indicates that the exponential model would provide a better fit to this epidemic. The logical question, therefore, is which model should be selected to make comparisons among all three curves? There are several possible approaches to this problem. First, one could argue that in the aluminum mulch treatment, the disease progress curve would have become sigmoidal if the onset and rate of the epidemic were not slowed to the point where the crop was mature before high disease levels of incidence ($> 70\%$) could be reached. If so, then the logistic model would be a realistic choice. To test this, all three disease progress curves were transformed based on the logistic model and regression parameters for each treatment were calculated. *Figure* 7C shows that the logistic model could, indeed, be used to adequately describe and compare all three curves as this model provided a linear relationship between transformed y with time. Moreover, the use of this model accounted for 96.6 to 97.1 of the variation in logit TEV incidence (high r^2 values).

For these three epidemics, the effect of aluminum mulch compared to no infected plants at time of transplanting can be seen as a reduction in r_L from 0.24 to 0.14/day, with no change in transformed initial disease, $\ln(y_0/[1-y_0])$. Adding an infected plant increased $\ln(y_0/[1-y_0])$, which is the initial disease incidence, and the rate of TEV infection was increased from 0.24 (no infected plants at transplanting) to 0.26 logits/day when one infected plant was present at transplanting.

Another possible approach to quantify and compare these three epidemics is to calculate the time to 50% disease incidence (t_{50}) for each of the four replicate plots and then use analysis of variance and mean separation tests to compare the three treatments. The t_{50} values for one infected plant at transplanting, no infected plants at time of transplanting, and aluminum mulch treatments were days of year 181 (30 June), 196 (15 July), and 219 (7 August), respectively. Note in this example that final disease incidence (y_{max}) does not distinguish between the first two epidemics, yet *AUDPC* values for all three epidemics were significantly different ($P \leq 0.05$) (*AUDPC* values were 3680, 2534, and 1001 for the same three treatments, respectively). Thus, the comparison of r_L, *AUDPC*, and/or t_{50} values appear to be the best methods to obtain quantitative information to compare these three disease progress curves (*Figure* 7A).

Example two

A study by Lecoq and Pitrat (1983) quantified the effects of weed control measures and host resistance on cucumber mosaic cucumovirus (CMV) in muskmelon. The form of resistance evaluated was known to be effective against some CMV strains but completely ineffective against others. Each plot was approximately 600 m^2 (260 plants) planted to either the CMV-susceptible cultivar Vedrantais or the resistant cultivar (line) Songwhan charmi (PI 161375) in the summer of 1979. Four disease progress curves for the 1979 summer planting are shown in *Figure 10*A. Based on the sigmoid shape of these four curves and the absolute rate curves (not shown), it appeared that the logistic model was the model most appropriate to describe all four curves (*Figure 10*B). Vanderplank (1963) stated that resistance effective against some strains (races) of a pathogen may have

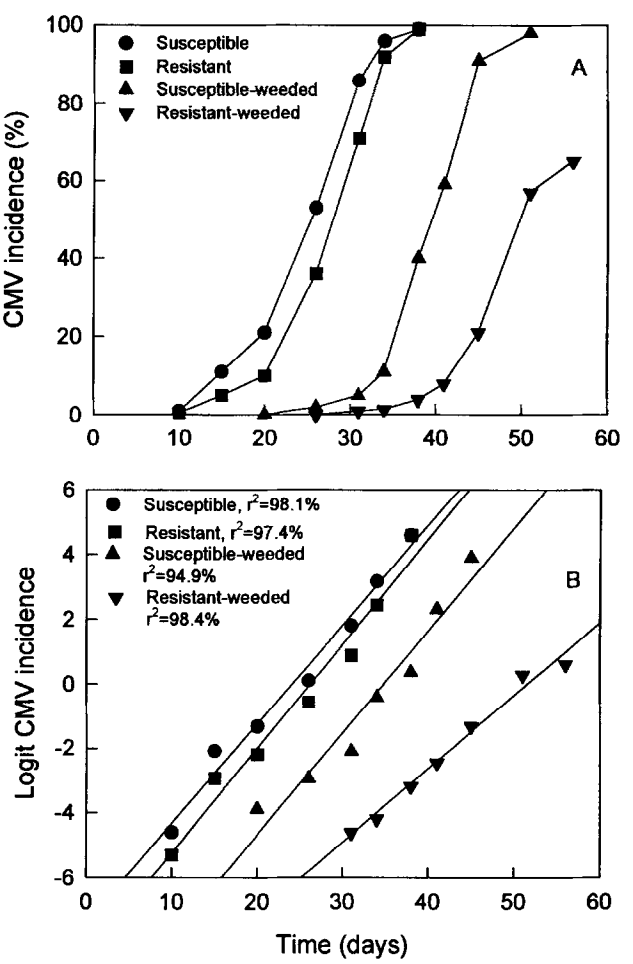


Figure 10. (A) Progress curves for incidence of cucumber mosaic virus in resistant and susceptible muskmelons and in plots weeded or not weeded. (B) Linear regression lines and equations using the logistic model $[\ln y/(1-y)]$ versus time to transform incidence data; ● $-7.66+0.32(t)$, $r^2 = 98.1\%$; ■ $-8.47+0.32(t)$, $r^2 = 97.4\%$; ▲ $-11.02+0.32(t)$, $r^2 = 94.9\%$; ▼ $-11.72+0.23(t)$, $r^2 = 98.4\%$

the effect of delaying the epidemic, that is, y_0 will be reduced, but not r_L. The estimates of r_L for the susceptible and resistant cultivars were 0.32 and 0.32/day (with no weeding), respectively. The lack of a difference in r_L values between cultivars can be attributed to the fact that this form of resistance was only effective against some virus strains, and therefore transformed y_0 was somewhat reduced but r_L was unaffected. As predicted by Vanderplank (1963), the epidemiological effect of this form of resistance was to delay the epidemic, but only for a few days because r_L was very fast. The initial amount of virus infection in the field was reduced because only the compatible strain(s) of the virus continued to spread within the crop. The epidemiological effects of this form of resistance can best be measured by determining and comparing the time to reach 50% disease incidence (t_{50}). Although t_{50} can be estimated by finding the time of $y = 50\%$ for each curve, one could also use the appropriate linearized version of the population growth model (*Table 1*) that was selected and solving the equation for t (with $y = 50\%$) to obtain t_{50} values (Padgett *et al.*, 1990). In this example, the benefit of the resistant cultivar was to delay the epidemic by only 2 days, that is t_{50} was increased from 24 to 26 days after transplanting; thus, this form of resistance alone is probably not cost-effective.

There was no effect of weed control (use of herbicides and mechanical cultivation) on the rate of the epidemic (r_L) when the susceptible cultivar was grown ($r_L = 0.32$ for both weeded and unweeded epidemics), but there was an 11-day delay in the time to reach 50% TEV incidence when the susceptible cultivar was weeded. When resistance was combined with weed control, however, t_{50} was further delayed from day 24 (susceptible cultivar, no weed control) to day 51 (resistant cultivar plus weed control). The two tactics combined reduced r_L from 0.32 to 0.23 logits/day. The biological interpretation of these results is that resistance reduces initial disease, and together with weed control, alloinfection among plants is reduced throughout the epidemic by reducing the number of foci from which aphids could potentially acquire CMV. This study clearly demonstrates that when disease progress (or pathogen) curves are similar in shape, t_{50} values can provide an excellent alternative means to quantify the effects of virus control tactics that delay epidemics. Furthermore, this study also provides an excellent example of how the benefits of integrated virus management tactics can be quantified.

Example three

The third example again comes from the TEV/bell pepper pathosystem (Padgett *et al.*, 1990). The objectives of this study were to identify pepper genotypes that reduce the temporal rate of TEV epidemics and to quantify the effect of this form of resistance on epidemiological characteristics. *Figure 11*A shows disease progress curves for a susceptible cultivar 'Yolo Wonder' and two pepper genotypes with rate-reducing resistance: 'Tambel 2' and an 'Asgrow' line. Plot sizes were 7.2 m long (24 plants) and six rows wide and each treatment (pepper genotype) was replicated four times. All pepper plants within a plot were assessed each week for the presence of TEV symptoms. The disease progress curve for TEV in 'Yolo Wonder' is sigmoidal, but the disease progress curves for the other two pepper lines are J-shaped. It was assumed that these two progress curves would have become sigmoidal had the annual growing season been extended; indeed, the logistic model provided a good fit to all three curves for purposes of comparison (*Figure 11*B). Using this model, the estimates of r_L for Asgrow and Tambel 2 lines were about half of the r_L compared with 'Yolo Wonder' (*Table 2*). A reduction in r_L delayed the t_{50} by 25 to 34 days, and by the end of the growing season, TEV incidence levels (y_{max}) were substantially lower for the two rate-reducing genotypes (*Table 2*). Had the logistic model not adequately described all

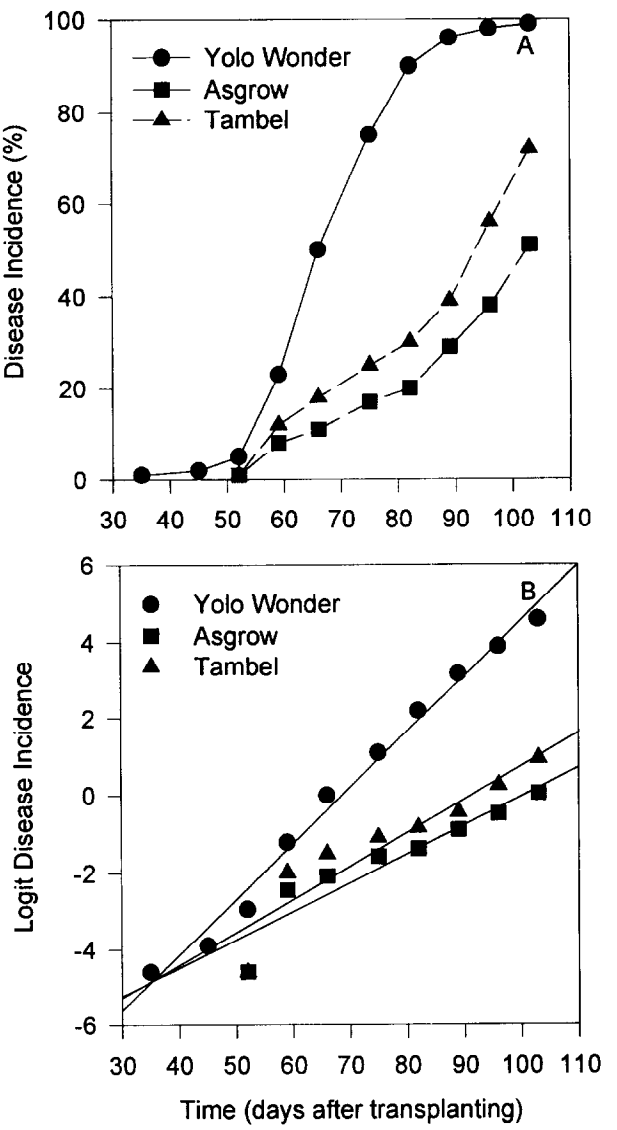


Figure 11. (A) Progress curves for incidence of tobacco etch potyvirus in three pepper genotypes. (B) Linear regression lines using the logistic model $[\ln y/(1-y)]$ versus time to transform incidence data. Regression equations for the transformed y values versus time (t) are: Yolo Wonder $-10.18+0.15(t)$, $r^2 = 99.1\%$; Asgrow $-7.48+0.07(t)$, $r^2 = 87.7\%$; Tambel 2 $-7.63+0.08(t)$, $r^2 = 85.4\%$

Table 2. Disease progress measurements[a]: effect of bell pepper genotype on final incidence of tobacco etch virus (TEV) incidence at time of harvest, relative area under the disease progress curve (RAUDPC), time for TEV incidence to reach 50% (t_{50}), and apparent infection rate (r_L) in Blarisville, GA, 1986

Genotype	Final incidence (%)	RAUDPC (%)	Apparent infection rate (r_L/day)	t_{50} (days)
Yolo Wonder B	97.2a	100.0a	0.15a	205a
Tambel 2	73.4b	46.4b	0.08b	246b
Asgrow XPG-5021	51.4c	32.3c	0.07c	255a
SED_2	3.7	4.8	0.02	3.2

[a]Within each column, means followed by a different letter are significantly different using the Waller–Duncan *K*-ratio test ($P \leq 0.05$)

three TEV disease progress curves, the effect of host genotype on disease progress could still have been quantified by calculating the *AUDPC* for each replicate plot and subjecting the data to analysis of variance and a means separation test. This was done and the results are shown in *Table 2*. In addition to standardizing *AUDPC* by dividing the time from the first to the last assessment, another useful standardization is to divide *AUDPC* values by the highest *AUDPC* value obtained in the study and multiply by 100. Using this approach, the relative area under the disease progress curve values for Tambel 2 and Asgrow were 46 and 32, compared with 100 for 'Yolo Wonder' (*Table 2*).

Example four

Seasonal effects on alfalfa mosaic virus alfamovirus (AMV) epidemics in annual medics (*Medicago* spp.) were quantified and compared by obtaining estimates of the rate parameter for AMV spread in each of four seasons. The original data were collected by R.A.C. Jones and D.G. Ferris, Plant Pathology Branch, Department of Agriculture, Perth, Western Australia (Jones and Ferris, 1995). The rate of spread of AMV in annual medics was determined by plotting cumulative AMV incidence versus time to obtain disease progress curves for the years 1991–1994 (*Figure 12*A). In this study, AMV epidemics began as early as 18 August (1992) or as late as 13 September (1994). The computer program EPIMODEL (Nutter and Parker, 1997) was used to determine which population growth model best described AMV disease progress in annual medic pastures. Model evaluation included the following criteria: (i) the *F*-statistic for the overall model which indicates that there is a significant linear relationship between transformed AMV incidence (y) and time (x); (ii) the coefficient of determination (r^2) which indicates the amount of variation in y that is explained by time; (iii) the standard error of the estimate for a predicted y value (SE_y); and (iv) visual inspection of the plot of the residuals versus time (a model is considered inadequate if residuals plotted versus x appear in a defined pattern).

Based on the above criteria, the logistic model best explained the temporal spread of AMV in annual medics of the study and, therefore, AMV incidence was transformed to logits $[\ln y/(1-y)]$ and regressed against time to obtain slope values which are estimates of the rate of AMV spread (*Figure 12*B). AMV spread in annual medic pastures was moderately fast in 1991 (0.12 logits/day) and moderately slow in 1992 (0.06 logits/day) (*Figure 2*). CMV spread in 1993 and 1994 was intermediate (0.09 logits/day). Using the highest infection rate recorded during the 4-year study (0.12 logits/day), AMV has the potential to increase from 1 to 100% incidence in just 96 days. The logistic model can also be used to calculate the time to 50% AMV incidence and this variable may also be used to compare AMV epidemics among

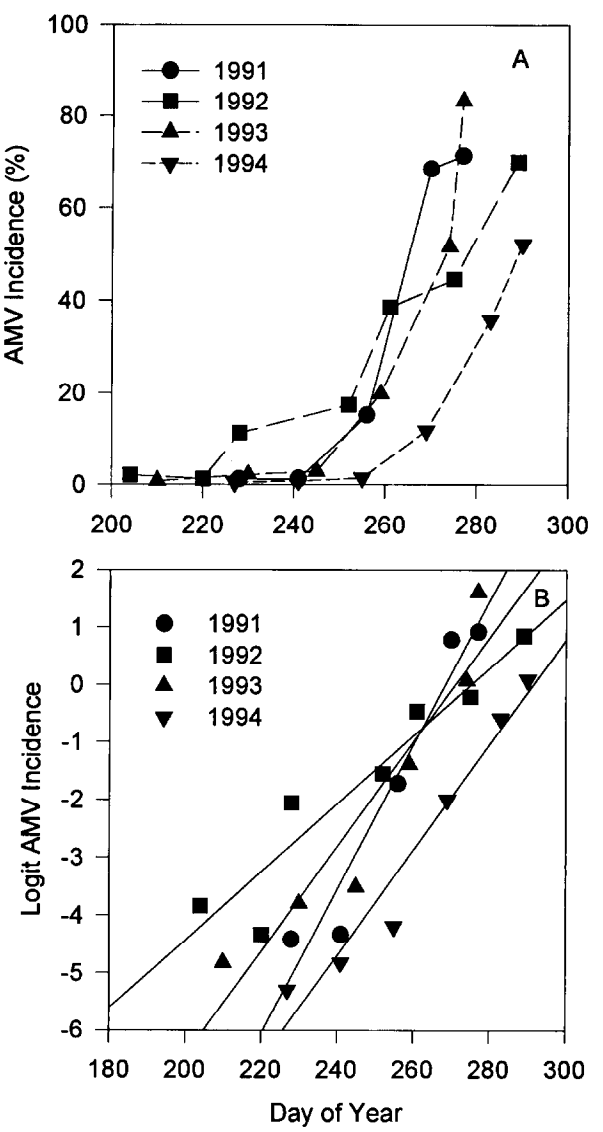


Figure 12. (A) Progress curves for incidence of alfalfa mosaic virus (AMV) in annual medic during the years 1991–1994. Disease incidence reached 53 to 83% within the first growing season with epidemics starting as early as day of year 220 (August 18) and as late as day of year 256 (September 13). (B) Rate of spread of alfalfa mosaic virus in annual medic: 1991, 0.12 logits/day; 1992, 0.06 logits/day; 1993 and 1994, 0.09 logits/day

years. For three of the four years, time to 50% AMV incidence occurred within a 6-day window (day of year 266, 269, and 272 for the years 1991, 1992, and 1993, respectively). For some as yet unknown reason, time to 50% AMV incidence occurred much later in 1994 (day of year 290).

Concluding remarks

Although the temporal analysis of plant virus disease (or pathogen) progress curves can provide much information on plant virus epidemics, our understanding of virus epidemics is still sparse compared to pathosystems involving fungal and bacterial pathogens. The development of virus detection tools that are low cost, fast, and accurate in terms of specificity and sensitivity will facilitate the undertaking of more extensive and quantitative epidemiological studies of plant virus pathosystems. The ability not only to detect the presence of specific viruses in time and space, but to differentiate virus strains, will lead to the generation of important new epidemiological data that can be quantitatively analyzed and interpreted by using the methods discussed here. These new results will help provide critical knowledge concerning the risks and benefits of virus disease management strategies and tactics, especially those involving crop plants genetically engineered for resistance to plant viruses. Moreover, quantitative information concerning the epidemiological impacts of specific virus management tactics on initial disease (or pathogen) incidence (y_0) and the increase in virus incidence with respect to time (r) will facilitate the integration and evaluation of more effective plant virus management programs.

Acknowledgements

This work was partly supported by the South Australian Research and Development Institute while the senior author was on Faculty Improvement (Sabbatical) leave and by Hatch Act and State of Iowa funds, Iowa Agriculture and Home Economics Experiment Station, Ames. Journal Paper No. J-16315, Project No. 3394. Appreciation is expressed to L. V. Madden, Department of Pathology, Ohio State University for reviewing and editing the manuscript prior to submission.

References

Barnett, O. W. (1986) Surveying for plant viruses: design and considerations. In *Plant Virus Epidemics: Monitoring, Modeling and Predicting Outbreaks*, eds G. D. Mclean, R. G. Garrett and W. G. Ruesnik, pp. 147–164. Academic Press, New York

Benner, C. P., Kuhn, C. W., Demski, J. W., Dobson, J. W., Colditz, P. and Nutter, F. W. Jr (1995) Identification and incidence of pepper viruses in northeastern Georgia. *Plant Dis.* **69**, 999–1001

Berger, R. D. (1989) Description and application of some general models for plant disease epidemics. *Plant Dis. Epidemiol.: Genet., Resist. Manag.* **2**, 125–149

Campbell, C. L. and Madden, L. V. (1990) *Introduction to Plant Disease Epidemiology*. John Wiley, New York, 532 pp

Geering, A. D. W. (1992) The epidemiology of CMV in narrow-leafed lupins (*L. angustifolius*) in South Australia. Ph.D. Dissertation, Department of Crop Protection, Waite Agricultural Research Institute, The University of Adelaide, South Australia

Gray, S. M., Smith, D. and Sorella, M. (1994) Reduction of disease incidence in small field plots by isolate-specific resistance to barley yellow dwarf virus. *Phytopathology* **84**, 713–718

Habili, N. and Nutter, F. W. Jr (1997) Temporal and spatial analysis of grapevine leafroll-associated virus 3 in Pinot Noir grapevines in Australia. *Plant Dis.* **81**, 625–628

Hu, J. S., Li, H. P., Barry, K., Wang, M. and Jordan, R. (1995) Comparison of dot blot, ELISA, and RT-PCR, assays for detection of two cucumber mosaic virus isolates infecting banana in Hawaii. *Plant Dis.* **79**, 902–906

Hughes, G., Madden, L. V. and Munkvold, G. P. (1996) Cluster sampling for disease incidence data. *Phytopathology* **86**, 132–137

Jones, A. T. (1979) Further studies on the effect of resistance to *Amphorophora idaei* in raspberry (*Rubus idaeus*) on the spread of aphid-borne viruses. *Ann. Appl. Biol.* **92**, 119–123

Jones, R. A. C. and Ferris, D. G. (1995) Application of insecticides to limit spread of alfalfa mosaic virus in grazed annual medic swards. In *Epidemiological Aspects of Plant Virus Control*, eds Plant Virus Epidemiology Committee, International Society of Plant Pathologists, P 7. 6th International Plant Virus Epidemiology Symposium, 23–28 April 1995, Jerusalem, Israel

Kapa, J. M. and Waterworth, H. E. (1981) *Handbook of Plant Virus Infections and Comparative Diagnosis*, ed. E. Kursta, pp. 257–332. Elsevier/North Holland, New York

Lecoq, H. and Pitrat, M. (1983) Field experiments on the integrated control of aphid-borne viruses in muskmelon. In *Plant Virus Epidemiology*, eds R. T. Plumb and J. M. Thresh, pp. 169–183. Blackwell Scientific, Oxford

Madden, L. V. (1980) Quantification of disease progression. *Protect. Ecol.* **2**, 159–176

Madden, L. V. (1986) Statistical analysis and comparison of disease progress. In *Plant Disease Epidemiology Population Dynamics and Management*, eds K. J. Leonard and W. E. Fry, Vol. 1, pp. 55–84. Macmillan Publishing, New York

Madden, L. V. and Campbell, C. L. (1986) Descriptions of virus epidemics in time and space. In *Plant Virus Epidemics: Monitoring, Modeling, and Predicting Outbreaks*, eds G. D. McLean, R. G. Garrett and W. G. Ruesnik, pp. 273–293. Academic Press, New York

Madden, L. V. and Campbell, C. L. (1990) Nonlinear disease progress curves. In *Epidemics of Plant Disease*, ed. J. Krantz, 2nd edn, pp. 181–229. Springer-Verlag, Berlin

Madden, L. V. and Hughes, G. (1995) Plant disease incidence: distributions, heterogeneity, and temporal analysis. *Ann. Rev. Phytopathol.* **33**, 529–564

Madden, L. V. and Nutter, F. W. Jr (1995) Modeling crop losses at the field scale. *Can. J. Plant Path.* **17**, 174–185

Madden, L. V., Pirone, T. P. and Raccah, B. (1987) Temporal analysis of two viruses increasing in the same tobacco fields. *Phytopathology* **77**, 974–980

Madden, L. V., Raccah, B. and Pirone, T. P. (1990) Modeling plant disease increase as a function of vector numbers: non persistent viruses. *Res. Popul. Ecol.* **32**, 47–65

Madden, L. V., Hughes, G. and Munkvold, G. P. (1996) Plant disease incidence: inverse sampling, sequential sampling, and confidence intervals when observed mean incidence is zero. *Crop Protect.* **15**, 621–632

Nutter, F. W. Jr (1992) Quantification of components contributing to rate-reducing resistance in a plant-virus pathosystem. In *Systems Approaches to Agricultural Development*, eds F. Penning de Vries, P. S. Teng and K. Metselaar, pp. 299–310. Kluwer Academic, The Netherlands

Nutter, F. W. Jr and Gaunt, R. E. (1996) Recent development in methods for assessing disease losses in forage/pasture crops. In *Pasture and Forage Crops*. Society of Agronomy, Inc., Madison, WI

Nutter, F. W. Jr and Parker, S. K. (1997) Fitting disease progress curves using EPIMODEL. In *Exercises in Plant Disease Epidemiology*, eds L. J. Francl and D. A. Neher, pp. 24–28. APS Press, St. Paul, MN

Nutter, F. W. Jr and Schultz, P. M. (1995) Improving the accuracy and precision of disease assessments: selection of methods and use of computer-aided training programs. *Can. J. Plant Pathol.* **17**, 174–184

Nutter, F. W. Jr, Teng, P. S. and Shokes, F. M. (1991) Disease assessment terms and concepts. *Plant Dis.* **75**, 1187–1188

Nutter, F. W. Jr, Teng, P. S. and Royer, M. H. (1993) Terms and concepts for yield, crop loss, and disease thresholds. *Plant Dis.* **77**, 211–215

Nutter, F. W. Jr, Hill, J. H. and Schultz, P. M. (1994) Modeling within-field spread of soybean mosaic virus with strain-specific monoclonal antibodies. *Phytopathology* **84**, 1159

Padgett, G. B., Nutter, F. W. Jr and Kuhn, C. W. (1990) Quantification of disease resistance that reduces the rate of tobacco etch virus epidemics in bell pepper. *Phytopathology* **80**, 451–455

Roff, M. N. M. and Ho, B. L. (1991) Maize as a barrier crop in reducing aphids, the virus vector of chili. *MARDI Res. J.* **19**, 251–258

Summers, C. G., Stapleton, J. J., Newton, A. S., Duncan, R. A. and Hart, D. (1995) Comparison of sprayable and film mulches in delaying the onset of aphid-transmitted virus diseases in zucchini squash. *Plant Dis.* **79**, 1126–1131

Swallow, W. H. (1995) Group testing for estimating infection rates and probabilities of disease transmission. *Phytopathology* **75**, 882–889

Thresh, J. M. (1983) Progress curves of plant virus disease. *Adv. Appl. Biol.* **8**, 1–85

Vanderplank, J. E. (1963) *Plant Diseases: Epidemics and Control*. Academic Press, New York, 344 pp

Waggoner, P. E. (1986) Progress curves of foliar diseases: their interpretation and use. In *Plant Disease Epidemiology Population Dynamics and Management*, eds K. J. Leonard and W. E. Fry, Vol. 1, pp. 3–37. Macmillan Publishing, New York

Webb, D. H. and Nutter, F. W. Jr (1997) Effects of leaf wetness duration and temperature on infection efficiency, latent period, and rate of pustule appearance of rust of alfalfa. *Phytopathology* **87** (in press)

Zadoks, J. C. and Schein, R. D. (1979) *Epidemiology and Plant Disease Management*. Oxford University Press, New York, 427 pp

Received 8 November 1996
Revised 3 June 1997
Accepted 4 June 1997