Letter to the Editor

Threshold Criteria for Model Plant Disease Epidemics. I. Asymptotic Results

M. J. Jeger and F. van den Bosch

First author: Natural Resources Institute, Central Avenue, Chatham Maritime, Chatham, ME4 4TB, UK; and second author: Department of Phytopathology, Agricultural University of Wageningen, P.O. Box 8025, 6700EE Wageningen, the Netherlands.
Current address of second author: Department of Mathematics, Agricultural University of Wageningen, Dreijenlaan 4, 6703HA Wageningen, the Netherlands.
The research of F. van den Bosch has been made possible by a fellowship of the Royal Netherlands Academy of Arts and Sciences.
Accepted for publication 15 September 1993.

There is much interest in and debate on the "threshold criteria" that determine whether a plant disease epidemic takes place. These criteria are derived from simplified mathematical models of an epidemic, rather than from actual observations. The justification for using simplified models is that a sufficient portion of an epidemic's characteristics are encapsulated to be relevant to real-life situations. For example, the threshold criterion generally accepted by plant disease epidemiologists states that for a polycyclic epidemic to occur the value of the dimensionless product, iR, must be greater than unity. The parameters i and R are defined in Vanderplank's (9) differential-delay equation:

$$dy_t/dt = R(y_{t-p} - y_{t-i-p})(1 - y_t) \quad (1)$$

in which y is a measure of disease scaled in a range from 0 to 1 (the number of infected units relative to the total number of units), R is a constant rate parameter (the number of new infected units per infectious unit per unit time), and p and i are the assumed constant lengths of the latent and infectious periods, respectively.

Although conjectured by Vanderplank (9), it was Waggoner (10) who provided a justification for the particular form of the threshold criterion, i.e., the intrinsic rate of disease increase, r, based on an exponential model of disease increase, is positive only if $iR > 1$, in which i and R are as defined in equation 1 (3).

Jeger (3), for the first time, provided a mathematical proof that the final size of a plant disease epidemic defined by equation 1 (the "asymptotic value") is given by the transcendental equation

$$L = 1 - A\exp(-iRL) \quad (2)$$

in which L is the final fraction of infected units strictly less than one and A is a constant dependant on initial disease (y_0). The threshold statement "$iR > 1$ for disease to increase from y_0" was then incorporated as an initial condition to determine a particular value for A. In the remainder of Jeger (3), only values of $iR > 1$ were considered in developing applications in comparative epidemiology, such as showing parallels with medical epidemiology, the difference between epidemic and endemic disease, and general population ecology. However, the condition used by Jeger (3) was not generally valid.

Hau (1) compared the two versions of equation 2 with $A = 1 - y_0$ and, as originally proposed by Jeger (3), $A = (1 - y_0)\exp(y_0)$. This showed that the latter criterion carries the implication that for values of $iR < 1$, the final amount of disease would be less than the initial disease. Hau (1) also pointed out that the assumption that disease would not increase for $0 < iR < 1$ was not valid. Hau argued that disease would increase according to a geometric series reaching a final size of $Y_\infty = Y_0/(1 - iR)$, in which Y_0 and Y_∞ are the initial and final numbers (not scaled) of infected sites, respectively (throughout, we use the capital Y notation to indicate absolute, or unscaled, units for the population). May (6), in a broad synthesis of the similarity between the plant and animal/medical literatures, derived equation 2 by considering the Vanderplank equation as a special case of the Kermack-McKendrick equation (5) and drew out a more general framework for epidemic analysis. May (6) also pointed out that if a quantity, R_0 (essentially the same as iR), is less than one then there will be a "decaying chain of infection" that corresponds to the convergent geometric series referred to by Hau (1).

It is clear from these contributions that where the amount of disease is initially small relative to the total population, the final amount of disease will be limited by values of $iR < 1$ and not by the availability of healthy tissue. Conversely, the threshold criterion of $iR > 1$ refers to a "runaway chain reaction, or epidemic" (6) during what would be analogous to exponential growth. Thus, Waggoner's original derivation of the result (10) is correct in stating there will be exponential disease increase only if $iR > 1$.

GENERAL ASYMPTOTIC RESULTS IN A FINITE PLANT POPULATION

An alternative to setting initial conditions to solve for the value of A in equation 2 is to explicitly introduce initial disease. This can be done either by introducing limits of integration directly into solving equation 1 (which requires assumptions concerning initial disease) or by introducing a more general function describing how an epidemic is initiated (Metz [7]). In either case, the usual solution to equation 1, $A = 1 - y_0$, can be obtained. In this paper, we are interested in analyzing the general case in more detail.

We first wrote Vanderplank's equation (9) in the unscaled form:

$$dY_t/dt = (S_0 - Y_t)[(R/S_0)(Y_{t-p} - Y_{t-i-p}) + g(t)] \quad (3)$$

in which $y_t = S_0 y_t$ is the number of diseased units, S_0 is the initial number of healthy units in the population, $S_0 - Y(t)$ is the number of susceptible units in the population at time t, and $g(t)$ is a function giving the rate of increase due to initial infection or another (e.g., external) source of inoculum, with the property that ${}_0\int^\infty g(t)dt < \infty$.

Suppose, for example, that at $t = 0$, the epidemic is started from Z newly infected units (without specifying whether these are part of the total population but with $Z < S_0$), then

$$g(t) = \begin{cases} 0 & t < p \\ (R/S_0)Z & p \le t < i + p \\ 0 & t \ge i + p \end{cases} \quad (4).$$

The advantage of incorporating this function into the Vanderplank equation is that it enables the introduction of initial disease and solves directly for the constant, A, in equation 2.

Separating variables and integrating

$$\int dY_t/(S_0 - Y_t) = (R/S_0)\int(Y_{t-p} - Y_{t-i-p})\,dt + \int g(t)dt \quad (5)$$

from which, following the procedures in Jeger (3), can be found

$$Y_\infty = S_0 - A\exp[-iR(Y_\infty/S_0) - g_\infty] \quad (6)$$

in which g_∞ is the value of $\int g(t)dt$ as $t \to \infty$. The constant, A, follows from considering the situation in which no epidemic occurs, i.e., $g(t) = 0$ for all t and as $t \to \infty$, $Y_\infty = 0$.

Substitution in equation 6 shows that $A = S_0$ and, thus,

$$Y_\infty/S_0 = 1 - \exp[-iR(Y_\infty/S_0) - g_\infty]. \quad (7)$$

This is the most general form of the asymptotic value equation for the Vanderplank equation. Special cases can be obtained by making assumptions concerning 1) the form of $g(t)$ and 2) whether infected individuals form part of the total population.

For example, in the special case of equation 4, $g_\infty = iR(Z/S_0)$ gives

$$\begin{aligned} Y_\infty/S_0 &= 1 - \exp[-iR(Y_\infty/S_0) - iR(Z/S_0)] \\ &= 1 - \exp[-iR(Y_\infty + Z)/S_0]. \end{aligned} \quad (8)$$

We next considered two cases that depend on whether Z is part of the total population.

Case 1: Initial infections are part of the population. The total population, N, is the sum of the initial infected individuals, Z, plus the initial number of healthy individuals, S_0, i.e., $N = Z + S_0$. The asymptotic value, L, is defined by

$$L = (Y_\infty + Z)/N \text{ (with } L_0 = Z/N).$$

Substitution in equation 8 and rearrangement give the asymptotic value equation

$$L = 1 - (1 - L_0)\exp[-iR\,L/(1 - L_0)]. \quad (9)$$

The effect of iR and initial disease, $L_0(=Z/N)$, on the final size L, for this specification of total population, is shown in Figure 1.

Case 2: Initial infections are not part of the population. In this case, $N = S_0$, $L = Y_\infty/S_0$, and trivially $L_0 = 0$:

$$L = 1 - \exp[-iR(Z/S_0)]\exp(-iRL). \quad (10)$$

The effect of iR and the quantity, Z/S_0, on the final size, L, for this specification of total population is shown in Figure 2.

Clearly these two cases give quite different outcomes. In case 1, the epidemic is initiated within the population, whereas in case 2 infection can arise only from an external source, e.g., immigrant spores or vectors. The relationship between the two cases can be seen by considering the behavior as $iR \to 0$.

In case 1, as $\exp[-iRL/(1 - L_0)] \cong 1 - iRL/(1 - L_0)$, we find from equation 9 that $L \cong L_0 + iRL$, i.e., $L \cong L_0/(1 - iR)$, which is the sum of a convergent geometric series.

In case 2, as $\exp[-iR(Z/S_0)] \cong [1 - iR(Z/S_0)]$ and $\exp(iRL) \cong (1 - iRL)$, we find from equation 10 that $L \cong iR(Z/S_0) + iRL$ (ignoring the quadratic term), i.e., $L \cong iR(Z/S_0)/(1 - iR)$. The term iRZ/S_0 is simply the fraction of the population infected due to the (external) source of infection and is directly comparable to L_0 and again is related to the convergent geometric series.

In Figure 3, the normal version of the asymptotic value equation is plotted with $A = 1 - L_0$, which for small values of L_0 is almost identical to Figure 1.

ASYMPTOTIC RESULTS FOR $iR < 1$

Both Hau (1) and May (6) gave intuitive arguments for the asymptotic result when $iR < 1$. We have demonstrated that in a finite plant population, in which the availability of healthy tissue is limiting to disease increase, there is a close relationship between the asymptotic result and a convergent geometric series. We now demonstrate the convergent geometric series result formally and how it can be obtained with the techniques described by Jeger (3). To show that the final level of disease is constrained by values of $iR < 1$, we assume there is an infinite plant population.

Convergent geometric series. Suppose that in an infinite plant population there is an initial disease population (which may be

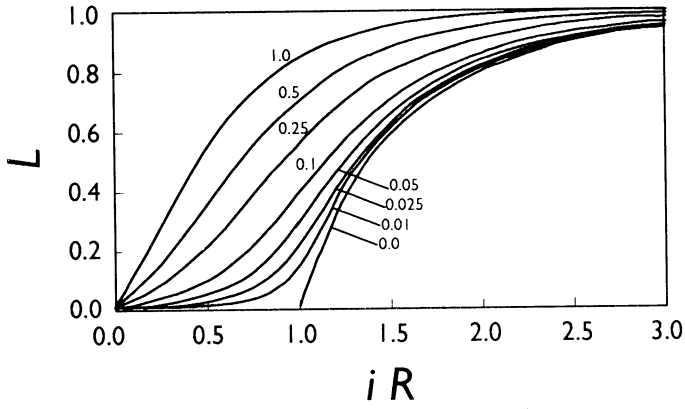

Fig. 2. The final size of an epidemic (L) in relation to the value of iR for different values of the quantity, Z/S_0 (denoted on curves), according to equation 10.

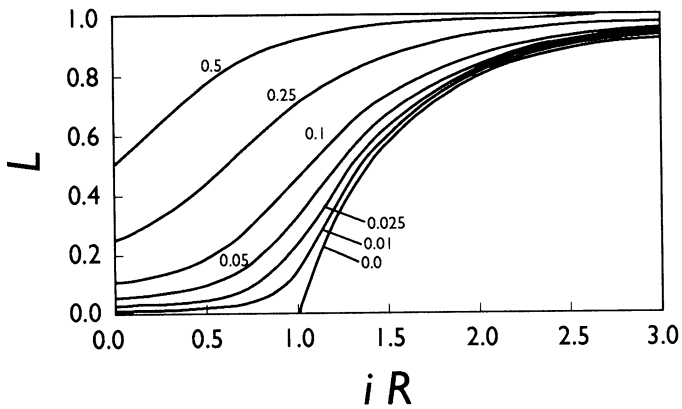

Fig. 1. The final size of an epidemic (L) in relation to the value of iR for different values of initial disease, L_0 (denoted on curves), according to equation 9.

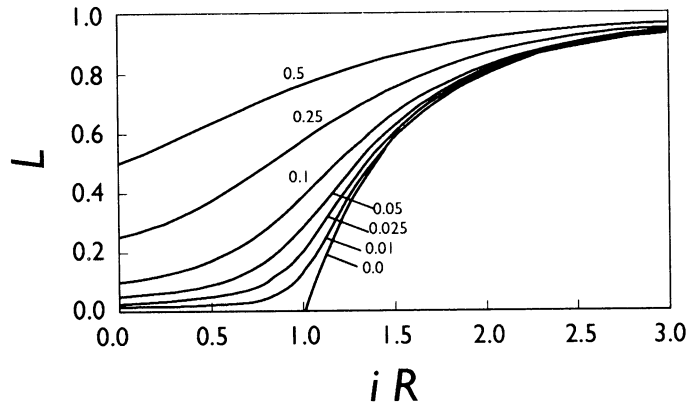

Fig. 3. The final size of an epidemic (L) in relation to the value of iR for different values of initial disease, L_0 (denoted on curves), according to equation 2 with $A = 1 - L_0$.

lesions, plant parts, or plants) at time $t = 0$, which after a time, p, produce infectious propagules (e.g., spores) and are, thus, infectious lesions. After a further time, i, these lesions cease producing spores and are postinfectious. Suppose that each lesion leads to R, new lesions, for each unit of time the lesion is infectious; then during the lifetime of that lesion, a total of iR lesions will have resulted from each original lesion. These are the standard characteristics of a polycyclic disease, with, in this case, the plant population not limiting the increase in new lesions.

Initially suppose the number of lesions is given by Y_0. At the first generation, we have $Y_1 = Y_0 + iRY_0 = Y_0(1 + iR)$ lesions. The second generation of lesions is produced by the infectious lesions of the first generation (iRY_0):

$$\begin{aligned} Y_2 &= Y_1 + iR(iRY_0) \\ &= Y_0 + iRY_0 + (iR)^2 Y_0 \\ &= Y_0[1 + iR + (iR)^2] \end{aligned}$$

and in general at the nth generation

$$Y_n = Y_0[1 + iR + (iR)^2 + \ldots + (iR)^n].$$

As n approaches ∞, we have the infinite sum

$$Y_\infty = Y_0[1 + iR + (iR)^2 + \ldots + (iR)^n + \ldots] \qquad (11)$$

This infinite sum has two outcomes depending on whether $iR < 1$ or $iR \geq 1$. If it is the latter, then Y_∞ approaches ∞. Of course, the assumption of an infinite population becomes unsatisfactory when $iR > 1$. When $iR < 1$, the sum approaches the value $Y_\infty = Y_0/(1 - iR)$. The case when $iR = 1$ is a special one in that, from equation 6, $Y_n = (n + 1)Y_0$, i.e., Y increases linearly. Also, the term denoting infectious lesions at the first generation is iRY_0, at the second $(iR)^2 Y_0$, ..., and so on to $(iR)^n Y_0$ at the nth generation. If $iR < 1$, then the number of infectious lesions decreases from the original number to approach zero as n becomes large, confirming the observation of Jeger (2) with respect to linked differential equations specifying disease dynamics, i.e., that the equivalent to iR in such a system must be strictly greater than one for the amount of infectious disease to increase.

Solution of differential-delay equation. Suppose that disease increases in an infinite population of plants, such that susceptible plant tissue is not limiting. We denote by Y the number of infected sites, i.e., an absolute rather than scaled measure of disease. The differential-delay equation (equation 1) becomes

$$dY_t/dt = R(Y_{t-p} - Y_{t-i-p}). \qquad (12)$$

On the assumption that $Y_t = Y_0$ at $t = 0$, this equation can be integrated directly to give

$$Y_t - Y_0 = R\int[Y(t-p) - Y(t-i-p)]dt.$$

For $iR < 1$, and as $t \to \infty$, it follows, using Jeger's method to solve the integral on the right side, that

$$Y_\infty - Y_0 = iRY_\infty.$$

From which, by rearranging,

$$Y_\infty = Y_0/(1 - iR), \qquad (13)$$

which is the sum of the convergent geometric series.

Suppose now that we wish to consider the more general case by introducing the function $g(t)$ as for the finite population. Then

$$dY_t/dt = R(Y_{t-p} - Y_{t-i-p}) + g(t). \qquad (14)$$

Integrating directly and allowing $t \to \infty$,

$$Y_\infty = A + iR(Y_\infty + g_\infty) \qquad (15)$$

in which $A = 0$, if $Y_\infty = 0$ for $g_\infty = 0$.

For the case in which $g_\infty = iRZ$ (there is clearly no S_0),

$$Y_\infty = iR\,Y_\infty + iRZ = iRZ/(1 - iR), \qquad (16)$$

which is equivalent to the expression for case 2 in the finite population.

DISCUSSION

It has been shown that in a plant population, effectively of infinite size, in which the availability of host tissue is not limiting, two quite distinct patterns of epidemic behavior for polycyclic diseases are possible depending on whether $iR < 1$ or $iR \geq 1$. In the former case, disease will approach an asymptotic value, given by the sum of the geometric series $Y_0/(1 - iR)$. Because the population size is infinite, the actual proportion of disease is zero; thus, for $iR < 1$, there is no epidemic. In the latter case, disease will increase to infinity. However, in the former case, it is clear that considerable multiplication of disease can still take place. For example, for a given Y_0, the asymptotic value, Y_∞, can be made as large as possible simply by choosing values of iR close to but less than unity. Nevertheless, the two qualitative outcomes for such a model population yield a clear interpretation of the threshold criterion. Again the difference is the finite increase in the case in which $iR < 1$ to the runaway chain reaction in the case in which $iR > 1$.

In a plant population of finite size, the literal meaning of the threshold criterion is not obvious. When plant tissue is constraining disease increase, the difference in outcome between an iR value of 0.99 is not qualitatively different from the outcome when $iR = 1.01$ (Figs. 1–3). This situation is most apparent when the amount of initial disease is relatively high. May (6) questioned whether in these circumstances an epidemic had not already occurred. This observation potentially denotes a distinction in perception between plant disease and animal/human epidemiologists. Crops range from short-lived annuals to long-lived perennials and are grown in environments in which seasonality may or may not be present. Moreover, stratification occurs, in which epidemics are occurring at different hierarchical levels. For example, if a forest or a fruit tree is considered the unit of population, then the appropriate time scale for epidemiological events may be over 10–20 yr. On the other hand, if the population of vegetative shoots within a single tree is considered, then an annual epidemic of a foliar fungal disease may occur over a time scale of a few months. How an epidemic is defined depends very much on the scale of resolution considered appropriate (8); expressions such as "an endemic epidemic" are not self-contradictory but, in fact, point out the futility, in some cases, of trying to distinguish between the two categories of disease.

Threshold criteria derive naturally from simple models of plant disease epidemics. That similar criteria can arise from both a single differential-delay equation and from systems of differential equations (4) attests to their robustness; in fact, both types of model are special cases of the general Kermack and McKendrick (5) epidemic model (6; F. van den Bosch and M. J. Jeger, *unpublished data*). The criteria are useful in determining whether (under idealized model conditions) an epidemic will result given a range of assumptions concerning initial disease, etc., and, if so, an estimate of the likely final size of the epidemic. The question may legitimately be asked whether such criteria are relevant for actual epidemics constrained by weather, seasonality, host growth and immigration of fungal spores or virus vectors. In our view, the answer is yes. For example, in this paper we have shown how the introduction of a "source" function $g(t)$ enables analysis of the effects of the ways in which epidemics are initiated.

There are few key indicators that influence the ways in which plant disease epidemiologists view epidemics. Threshold criteria (and their analogues in wider population theory) are one such set of indicators and rather than reject them as oversimplistic or lacking realism, the effects of the above constraints should be investigated, both experimentally and theoretically.

LITERATURE CITED

1. Hau, B. 1988. Anmerkungen zum asymptotischen Verhalten und dem Schwellentheorem bei Modellen für Epidemien von Pflanzen krankheiton. Z. Pflanzenkr. Pflanzenschutz 95:372-376.
2. Jeger, M. J. 1982. The relation between total, infectious, and postinfectious diseased plant tissue. Phytopathology 72:1185-1189.
3. Jeger, M. J. 1986. Asymptotic behaviour and threshold criteria in model plant disease epidemics. Plant Pathol. 35:355-361.
4. Jeger, M. J., and van den Bosch, F. 1994. Threshold criteria for model plant disease epidemics. II. Persistence and endemicity. Phytopathology 84:28-30.
5. Kermack, W. O., and McKendrick, A. G. 1927. A contribution to the mathematical theory of epidemics. Proc. Roy. Soc. Lond. Ser. A 115:700-721.
6. May, R. M. 1990. Population biology and population genetics of plant-pathogen associations. Pages 309-325 in: Pests, Pathogens and Plant Communities. J. J. Burdon and S. R Leather, eds. Blackwell Scientific Publications, Oxford.
7. Metz, J. A. J. 1978. The epidemic in a closed population with all susceptible equally vulnerable; some results for large susceptible populations and small initial infections. Acta Biotheor. 27:75-123.
8. Onstad, D. W. 1992. Temporal and spatial scales in epidemiological concepts. J. Theor. Biol. 158:495-515.
9. Vanderplank, J. E. 1963. Plant Diseases: Epidemics and Control. Academic Press, New York.
10. Waggoner, P. E. 1977. Contributions of mathematical models to epidemiology. Ann. N. Y. Acad. Sci. 287:191-206.