Plant disease epidemiology: facing challenges of the 21st Century

Plant disease epidemiology: facing challenges of the 21st Century

Under the aegis of an International Plant Disease Epidemiology Workshop held at Landernau, France, 10–15th April, 2005

Edited by
S. Savary and B.M. Cooke

Reprinted from *European Journal of Plant Pathology*, Volume 115 Issue 1, 2006

Springer

A C.I.P catalogue record for this book is available from the library of Congress

ISBN 1-4020-5019-4 (HB)
ISBN 1-4020-5020-8 (e-book)

Published by Springer,
P.O. Box 17, 3300 AA, Dordrecht, The Netherlands

Printed on acid-free paper

Cover photos:
Patterns of change in multiple pathosystems over space: spatial distribution of four diseases in a groundnut plot, Côte d'Ivoire, France.
A- Groundnut rust, *Puccinia arachidis*;
B- Early leafspot, *Cercospora arachidicola*;
C- Late leafspot, *Cercosporidium personatum* (*Phaeoisariopsis personata*),
D- Web blight, *Rhizoctonia solani*.
Disease assessments were made at 90 days after sowing. Rust, early leaf spot, and late leaf spot: severity (% diseased leaf area) scales; web blight: incidence (% diseased plants) scale. From Lannou and Savary, 1991, modified.

Printed in the Netherlands

Contents

European Journal of Plant Pathology (2006) 115:1–2
DOI 10.1007/s10658-005-5946-6

Foreword

Plant Disease Epidemiology: Facing Challenges of the 21st Century

Plant disease epidemiology deals with diseases in plant populations. During the past century, it has become a vibrant field of science, achieving significant conceptual innovations with important impact on the management of plant diseases. Plant disease epidemiology mobilises concepts and methods from ecology, genetics, environmental physics, botany, and mathematics. It deals with cultivated and non-cultivated plants in environments where human activities have had large, or lesser, impact. As in many other fields of science, plant disease epidemiology faces important, sometimes new, questions. By and large, many of these questions emerge from changes in human societies and changes in the status of the planet on which we live.

Global climate is changing at a rapid rate: will it render plant diseases more, or less, harmful to man-made and spontaneous ecosystems? There is much debate on this issue, because global climate has varying, sometimes very large effects on the local environment of growing plant canopies, and because the physical micro-environment and its variation strongly influence plant diseases and their consequences on ecosystem functioning and performance; in addition, changes in global climate trigger many profound changes in the way ecosystems, cultivated or not, are managed. Interestingly, much of the early literature on botanical epidemiology dealt with climate-disease or climate-pathogen relationships – in fact these kinds of relationships have long been perceived as the bulk of epidemiological research by many. Plant disease epidemiologists thus have a strong scientific tradition in studying climate-pathogen-disease relationships. Can such an asset be mobilised by the epidemiological community to answer questions about the effect of climate change on plant diseases?

Global trade, and thus, trade of plant products, have increased at an unprecedented rate during the 20th century, and will continue to expand in the next century. Exchanges of plant materials at very different scales, local to global, have profound effects on plant diseases. Plant disease epidemiologists have become experts in assessing the risk of irruption of novel pathogens in plant communities, the consequences it may have on ecosystems, and ways to manage such perturbations. The concepts related to biological invasions or population displacements certainly are not new to plant pathologists: the epidemiological community in fact contributed to craft them in the past century. New threats may now also exist, whereby exotic or novel plant pathogens would intentionally be introduced: these threats must be dealt with. The consequences of plant pathogen transport are many: on local performances of spontaneous ecosystems and agricultural ecosystems; on farmers' livelihoods; on local, national, and regional economies; and perhaps more importantly, they can have adverse consequences on trade regulation. Will plant disease epidemiologists provide answers to such pressing questions?

Biodiversity, a buzzword of the past century, is also of global concern. The decline in global biodiversity that is currently taking place has been referred to as the sixth great extinction process our planet has experienced during its history, but this time, it is man-made. Generations of plant pathologists, and especially of plant disease epidemiologists, have been dealing with biodiversity. The huge diversity of life that resides in the rhizosphere and the phyllosphere are causes both of diseases in plants, and of their suppression. Much current research is addressing ways of harnessing such biodiversity not as enemies – of which pathogens are an inherent part – but rather as important biological allies to control disease epidemics. The diversity of plants is another facet of global biodiversity, and there are concerns about the decline in the genetic diversity of crop plants. It is from this diversity that possibly the most potent

instrument for disease management has been developed by plant pathologists: genetic host plant resistance. Will we run short of resistance genes against major plant pathogens? Host plant diversity, and the disease resistance genes it harbours, can be deployed over time and space, according to epidemiological principles. In-depth knowledge of the characteristics of individual pathogens causing specific diseases that must be controlled has been mobilised to develop appropriate strategies at the plant population, field, landscape, and sub-regional levels. Major successes have been achieved using such strategies, and the end of the past century has seen their recognition by the scientific community. Will epidemiologists succeed in the future in fully sharing these technologies with the farmer so that they are more fully utilised?

Food security was a central concern of the global agricultural research community in the middle of the 20th century, but apparently, not anymore. However, the world population still increases, and is expected to do so for several decades. One out of six human beings living on earth today suffers from lack of food. Many of today's poor live in cities, with no access to land and agriculture, and most of the projected increase in the world population will take place in the world's largest cities. Pests, including plant pathogens, cause losses in pre-harvest yield in the range of 20–40%; estimates of post-harvest losses are inadequate, but it is a fair assumption that they are often higher than 10 or 20%. Why are our estimates – the raison d'être of plant pathology – still so vague today? Seldom do economists currently address the issue of food security – why? Is it so that globalised exchanges, novel biological technologies, and the self-regulating mechanisms of trade, will be sufficient to fulfil the needs of future generations? Will these not have negative side-effects, and will they truly prevent the current over exploitation of natural resources, water and land in particular?

Sustainable production and crop protection systems need to be devised, which could exploit scarcer resources sparingly, and if possible enhance the resource base. Can these production and protection systems be designed so that they generate healthy, high-quality products that would find niche markets both locally and globally, and so provide farmers with the income they require, and offer consumers products that suit their needs and their incomes? Plant disease epidemiologists alone cannot provide answers to such questions, but certainly could significantly contribute to these new strategies.

The five-day International Plant Disease Epidemiology Workshop (held 10–15th April, 2005, in Landernau, France, the ninth of a series) reported in this special issue of the European Journal of Plant Pathology, obviously could not address all of these issues, and others, with all the depth good science demands. However it provided a unique opportunity for scientists interested in this field to meet and face challenging questions, contribute to animated debates, and reflect on the future development of the science of plant disease epidemiology.

Serge Savary
Mike Cooke

European Journal of Plant Pathology (2006) 115:3–23
DOI 10.1007/s10658-005-1229-5

Botanical epidemiology: some key advances and its continuing role in disease management

Laurence V. Madden
Department of Plant Pathology, The Ohio State University, Wooster, OH, 44691-4096, USA
(E-mail: madden.1@osu.edu)

Accepted 13 July 2005

Key words: basic reproduction number, Bayesian decision theory, epidemic, mixed models, model, population dynamics, receiver operating characteristic (ROC) curve

Abstract

Epidemiology involves the study of the temporal, spatial, and spatio-temporal dynamics of disease in populations, and the utilization of results of experiments and surveys to describe, understand, compare, and predict epidemics. Such understanding and description of epidemics can lead directly to the development and evaluation of efficient control strategies and tactics. Mathematical and statistical models are key tools of the epidemiologist. Recent advances in statistics, including linear and nonlinear mixed models, are allowing a more appropriate matching of data type and experimental (or survey) design to the statistical model used for analysis, in order to meet the objectives of the investigator. Coupled ordinary and partial differential equations, as well as simpler growth-curve equations, are especially useful deterministic models for representing plant disease development in fields in time and space over single seasons or many years, and their use can lead to appraisal of control strategies through metrics such as the basic reproduction number, a summary parameter that may be calculated for many general epidemic scenarios. Recently, compelling arguments have been made for the use of Bayesian decision theory in developing and evaluating real-time disease prediction rules, based on measured disease or weather conditions and either empirical or mechanistic models for disease or control intervention. Through some simple calculations of predictor accuracy and (prior) probability of an epidemic (or the need for control), the success of any predictor can be quantified in terms of the estimated probability of random observations being epidemics when predicted to be epidemics or not epidemics. Overall, despite the many contributions in epidemiology over the past four decades, more effort is still needed to convince those outside of epidemiology to more fully use epidemiological results and insights into the development and evaluation of disease controls.

Introduction

In 1963, van der Plank made a most compelling case for the importance of botanical epidemiology, both for understanding plant diseases at the population scale and for determining disease management strategies (van der Plank, 1963). He also made the bold statement at the time that 'epidemiology is here to stay.' Individual disciplines enjoy 'ups and downs' of popularity, of course, and epidemiology is no exception. The tremendous growth in the discipline within plant pathology during the 1960s, 1970s, and 1980s (e.g., Campbell and Madden, 1990; Kranz, 1990; Jones, 1998; Zadoks, 2001) has been eclipsed by growth in the larger field of molecular biology over the last two decades. Nevertheless, more than 40 years after van der Plank's book (1963), botanical epidemiology is still here, and still of utmost importance in giving a sound theoretical and practical basis for disease management. This view may not always be held outside of the discipline, however, and it

remains a challenge for epidemiologists to continue to make the compelling case that epidemiology matters.

Until molecular biology or more traditional breeding results in durable resistance to all plant pathogens on all crops, coupled with the acceptance of the new cultivars by growers and the public, there will be plant disease epidemics, and many of these will result in substantial reductions in yield. There is certainly increasing use of crop GMOs around the world (James, 2003), but cultivars with very broad-acting and durable resistance have yet to be developed. Moreover, the public opinion against their use remains strong in many regions; thus, it would be naïve to expect 'super resistant' cultivars in the foreseeable future. Use of fungicides and other chemicals in a protectant or curative manner is only practical for some crops and some diseases, and there is increasing societal pressure to (drastically) reduce the use of these chemicals in many regions. Thus, a scientific basis for applying or not applying chemicals is needed, and the decision clearly involves knowledge (or prediction) of the disease dynamics under different environmental conditions. The development of resistance to fungicides and antibiotics continues, and new cultivars have a finite lifetime.

No control tactics are known that will totally eliminate epidemics in crops and forests where the pathogen is present over large areas. Biological and cultural controls may be very beneficial, depending on the pathosystem (Maloy, 1993), but variability of control efficacy may be high with the former, and grower acceptance may be low with the latter (e.g., unwillingness to rotate crops).

The public and the scientific community have been definitely reminded of the importance of epidemiology, and the research tools that epidemiologists can bring to a problem, in recent years. A few examples are given. With increasing world trade of agricultural commodities as well as international travel, the risk of pathogen invasion of new countries or regions is well recognized (NRC, 2002), and predictions of the risk of invasion involve many epidemiological characteristics of pathogens, such as survival probabilities and reproductive potential (Madden and Wheelis, 2003). Moreover, the decision to attempt to eradicate or not also involves knowledge of disease epidemiology. The cases of citrus canker in Florida, karnal bunt in Arizona, and plum pox in Pennsylvania, U.S., are three examples of disease invasions (Gildow et al., 2004; Gottwald et al., 2001; Rush et al., 2005).

New pathogens (or pathogens new to a given crop) continue to be discovered, as well as strains, races, or biotypes of previously known pathogens. The new very aggressive biotype of African cassava mosaic virus in Africa is an example of a newly evolved isolate (Legg, 1999; Strange and Scott, 2005) that is proving very difficult to control. Sudden oak death, caused by *Phytophthora ramorum*, is a newly identified disease of oak and several other plant species, which is spreading naturally and (unfortunately) with the assistance of man, in the U.S. and elsewhere (Rizzo et al., 2002).

For diseases such as sudden oak death or Asian soybean rust (newly introduced into the U.S.), there is a great need to know the extent of spread from current locations (e.g., from the point of introduction) to other locations. For any disease that is locally concentrated (e.g., around the point of a new introduction), or does not yet exist in a country or region, ethically one cannot deliberately introduce the pathogen where it does not occur in order to study spore movement and resulting disease intensity. Thus, modelling is a key research tool for understanding risks based on key epidemiological characteristics or traits of a disease (Madden and van den Bosch, 2002; Madden and Wheelis, 2003). Epidemiology as a discipline depends heavily on the tools of mathematical and statistical modelling (Campbell and Madden, 1990), so epidemiologists are, in general, quite prepared to tackle the problem of disease spread through modelling. Model parameters for these types of situations can be obtained from observations where the disease of interest does occur naturally.

Most practicing epidemiologists would strongly support van der Plank's (1963) statement that epidemiology sets the strategy for disease control, and numerous examples can be given where epidemic knowledge leads to better control (Zadoks and Schein, 1979; Fry, 1982; Maloy, 1993). Furthermore, epidemiological principles and results can also lead to *specific* control recommendations, through the process of disease forecasting or risk prediction (Hardwick, 1998; Hughes et al., 1999), as demonstrated 45 years ago (Waggoner, 1960).

However, as pointed out recently by Jeger (2004), many controls are utilized and evaluated without explicit consideration of disease dynamics in fields. Although there is great danger in basing conclusions on disease intensity measured at one time in an epidemic (especially for polycyclic diseases; see Campbell and Madden, 1990), this unfortunately happens too often. Thus, epidemiologists still need to be pro-active in working with others in developing and evaluating disease control measures.

In the remainder of this article, I discuss a few developments that I consider to be very important in the development of plant disease epidemiology. Many more topics could have been covered. I have two major themes. One deals with the advancement in our theoretical understanding of the population-dynamic processes of disease spread in space and increase in time, coupled with the improvements in relating certain models (or their parameters) to empirical results (i.e., model fitting). The other theme deals with the prediction of plant disease on a real-time basis, or prediction of the need to impose a control measure, based on principles from probability theory. Citations are deliberately sparse, and are mainly to major reviews of topics rather than to all the (many) important original papers published over the last few decades. I assume throughout that modelling and statistical data analysis are methodological foundations for understanding epidemics and utilizing any gained knowledge in disease control.

Temporal and spatial dynamics of disease

Growth curve modelling and analysis

Van der Plank (1963) used the monomolecular and logistic equations as heuristic models of monocyclic (simple interest) and polycyclic (compound interest) disease epidemics. These models continue to be the benchmarks for quantification of epidemics, especially over single growing seasons. However, plant pathologists discovered in the 1960s and 1970s that these two models did not necessarily provide an adequate description (based on statistical principles of model fitting) for many disease progress curves (Campbell and Madden, 1990). Several alternative models were proposed or developed, some of them flexible in the sense that different degrees of skewness could be represented with the same model (depending on a realized value of a shape parameter). A feature of these models is that they are all based on a single response variable (disease intensity, y) in relation to continuous time, which can be obtained as a solution for the rate of change of y with time, dy/dt [e.g., $dy/dt = r_L y(1-y)$ for the logistic model]. In some cases, the solution can be expressed as a linear model, e.g., $\text{logit}(y) = a + r_L t$, where a is a transformation of disease intensity at time 0, r_L is the per capita rate parameter, and logit(y) is a linearizing transformation of y.

A good fit of an empirical model, or even a perfect fit, to data collected over time, is not proof of any mechanism for population growth (Campbell and Madden, 1990; Zadoks, 2001). But a good fit of a particular model for several disease progress curves could lead one to hypothesize about mechanisms, and then test the hypothesis with additional data or experiments. Moreover, using a model that provides a (reasonably) good fit to data is extremely important to accurately compare epidemics; among other things, using an inappropriate model will lead to biased estimates of the rate parameter and its standard error (Neter et al., 1983).

One clear trend in botanical epidemiology is the dramatically increasing complexity of statistical models and methods that have been applied to all epidemiological data over the last few decades (e.g., Gilligan, 2002; van Maanen and Xu, 2003). This is a natural development given the fact that epidemiology is a science of populations, and populations can only be adequately characterized and compared using the methodology of statistics. Although I am sure there are some who feel that the emphasis on mathematics and statistics obscures the understanding of the biology of epidemics, I would make the opposite claim, and declare emphatically that mathematical and statistical modelling are foundations for understanding epidemics. I further believe that, with some exceptions, the use of statistical analysis is actually still inadequate in most of epidemiological research, and certainly in most of plant pathology research! Many investigators still only: measure disease at a single time, do not match the chosen form of data analysis to the type of disease intensity variable (discrete for incidence, continuous but unequal variance for severity, ordinal for many disease rating scales); do not base their analysis on the

chosen experimental design; or perform inefficient (and sometimes uninformative) analyses. An example of the latter is the still common practice of performing a separate data analysis for each assessment time during an epidemic rather than simultaneously analyzing treatments (between-subject factors) and time (within-subject factors), and their interactions. Garrett et al. (2004) and citations therein can lead the reader to some of the important recent advances in statistical data analysis of relevance in plant pathology.

It has been known for many years (Madden, 1986) that disease values collected over time in the same experimental or sampling unit (e.g., plot) are serially correlated and that the variation in disease over time within plots is different from the variation between plots. This may be in part due to the cumulative nature of disease progress curves (see pp. 521–522 in Schabenberger and Pierce, 2002, for general discussion of cumulative processes over time). Serial correlations, sometimes called temporal autocorrelations, are especially troublesome in the comparison of treatments. My recent studies now show, however, that fitting of appropriate population-growth models to disease progress data often reduces the correlation of residuals to near zero for *individual* disease progress curves, reducing the need to directly utilize cumbersome adjustments to standard errors for calculated rates (unpublished). However, in the larger setting of multiple disease progress curves, corresponding to multiple treatment factors and blocks, there will always be non-zero correlations of observations *within* the plots by the nature of the experimental design (Schabenberger and Pierce, 2002). However, the structure of the correlations and variances may be quite complex, due to the cumulative process of disease development, but simple variance-covariance models can adjust for this property. For disease progress models that can be expressed in linear form through the use of a transformation of y [e.g., logit(y)], linear mixed models provide a tremendous (and still underutilized) tool for a thorough analysis of the epidemics (Garrett et al., 2004). Most plant pathologists (including epidemiologists) are not aware of the major advances made in mixed model analysis in statistics, a field that encompasses classical ANOVA and regression, and many other topics in a unified manner (Schabenberger and Pierce, 2002; Garrett et al., 2004). Instead of estimating disease progress model parameters for each epidemic, with a follow-up analysis of variance, through mixed models one can simultaneously estimate the disease progress parameters and their appropriate standard errors based on the explicit features of the design. The former approach (e.g., estimated slope for each plot, and then an ANOVA of these slopes), still common with researchers, is known to be the least powerful approach to detect differences in treatments (Wolfinger, 1996). Through these mixed-model methods, random effects (such as locations, blocks, and possibly genotypes), and their interactions with fixed effects (treatments) can be appropriately estimated and realistic inferences made.

Many population dynamic processes can be expressed only in nonlinear form (e.g., $y = f(t; a,b)$, where $f(\bullet)$ is a nonlinear function). The recent advances in nonlinear mixed models (Garrett et al., 2004) can be applied to these situations, but the range of experimental designs is much more limited (currently), and considerably larger data sets are required to estimate and compare parameters. Nevertheless, statistically savvy and motivated epidemiologists can make considerable progress here.

Mechanistic modelling (linked differential equations)

Van der Plank (1963) clearly realized that models such as the logistic were inadequate for a biologically meaningful characterization of disease progress in time. His approach was to use a so-called differential-delay equation in order to represent polycyclic disease development. This model relates dy/dt to the *infectious* disease intensity rather than to total disease intensity, with infectious disease estimated based on assumed fixed-duration latent and infectious periods. Although the use of differential-delay equations serve as a good foundation for developing computer simulation models with fixed time steps, such equations are extremely cumbersome for mathematical analysis, making it difficult to explore implications of different biological properties of hosts and pathogens, or of different control strategies, on long-term disease development. Eventually, plant pathologists discovered the mathematical elegance of linked or coupled differential equations for characterizing disease progress (Jeger, 1986a, b; van Maanen and

Xu, 2003). The approach – which was utilized as long ago as 1911 for representing malaria epidemics (Ross, 1911) – is to use two-to-several differential equations, with some variables of interest and parameters appearing in more than one of the equations. The beauty of this approach is that new terms can be easily added, as needed, to meet the objectives of the investigator and the details of the pathosystem, and asymptotic and steady state results (such as disease persistence) can be explored quantitatively. Furthermore, even though analytical solutions cannot generally be obtained (i.e., one cannot write out y as a function of parameters and time without the use of the integral symbol), numerical solutions are now easy to obtain with many mathematical programmes such as MATHCAD and MATHEMATICA.

Statistical software such as PROC MODEL of the SAS/EST system allows direct parameter estimation of one or more parameters for these types of models (Madden et al., 1987). The approach is iterative and computationally intensive, but readily accomplished by those who have a good understanding of nonlinear models and statistics. However, unlike the case for models with analytical solutions (linear or nonlinear; see previous sub-section above), one cannot easily incorporate the features of the experimental design (e.g., split plot, etc.) into the model fitting. Rather, one generally needs to estimate parameters for each individual epidemic (e.g., each field or plot) and then perform t-tests or analysis of variance on the estimated parameters (depending on the experimental design).

A relatively simple coupled differential equation model for a polycyclic disease with no plant mortality is given by:

$$\begin{aligned}\frac{dH}{dt} &= -\beta HI \\ \frac{dL}{dt} &= \beta HI - \omega L \\ \frac{dI}{dt} &= \omega L - \mu I \\ \frac{dR}{dt} &= \mu I\end{aligned} \qquad (1)$$

where H, L, I and R are the densities of disease-free (healthy), latently infected, infectious, and post-infectious (removed) individuals (e.g., plants, leaves, roots, or even sites on leaves), $1/\omega$ is the mean latent period, $1/\mu$ is the mean infectious period, and β is the per capita transmission rate (new diseased individuals per diseased individual per healthy individual per unit time). For fungal (or oomycetes) diseases, β is the product of spore production per time unit per infectious individual, the probability that a spore comes in contact with a healthy individual, and the probability that a spore in contact with a healthy host individual causes an infection. Total disease at any time is determined as $Y = L + I + R$, and disease intensity as a proportion is given by $y = Y/(H + L + I + R)$. If initial disease intensity is very low, then at $t = 0$, initial total host density is virtually the same as initial healthy host density, H_0. The product βH_0 is analogous to van der Plank's (1963) corrected basic infection rate (new diseased individuals per diseased individual per unit time).

A fundamental result with this model is that disease will increase (i.e., an epidemic will occur) only if $\beta H_0/\mu > 1$. The expression to the left of the inequality is known as the basic reproduction number, R_0 (Diekmann and Heesterbeek, 2000). This composite parameter also indicates the final intensity of disease (after a long time) and the initial exponential rate of increase (see Segarra et al., 2001, for details). An example realization of the model in equation 1 is shown in Figure 1 for the situation with $R_0 = 2.5$. Final disease is less than 100%, and is estimated by iteratively solving $y_\infty = 1\text{-}\exp(-R_0 y_\infty)$. Control strategies are developed or evaluated by finding combinations of β, ω, and μ that give $R_0 < 1$; specific control tactics (e.g., host resistance, protectant fungicide, curative fungicide) can then be directed at reducing β, etc.

An advantage of the equation 1 formulation is the easy expansion for other situations. For instance, a simple-interest disease component (infections from resident inoculum throughout the epidemic, rather than just at the start) can be incorporated by using the πxH term, where x is the density of inoculum and π is a simple-interest rate parameter. One can consider x to be constant or to change (typically, decline over time), so that $dx/dt = \mathrm{K}x$. When x does not change, then πx is equivalent to the monocyclic rate parameter (r_M) of the monomolecular model. The πxH term is subtracted from dH/dt and added to dL/dt in equation 1. A pure simple-interest epidemic results if $\beta = 0$; otherwise, a composite of polycyclic and

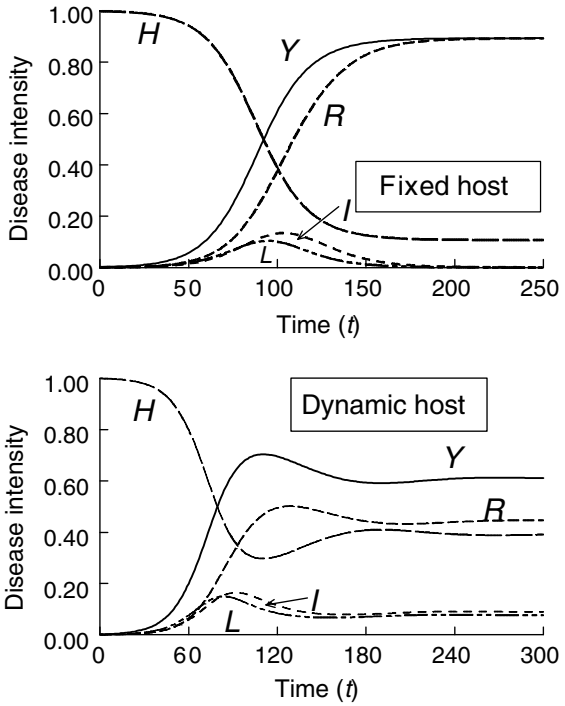

Figure 1. Density of healthy (H), latently infected (L), infectious (I), and post-infectious (R) individuals (on a proportion scale), together with total disease ($Y = L + I + R$), based on equation 1 (upper frame) and equation 2 (lower frame). Mean latent period ($1/\omega$) was 7, and mean infectious period ($1/\mu$) was 10 time units. Upper frame: $\beta H_0 = 0.25$ per time unit. Lower frame: $\beta H_0 = 0.35$ per time unit, $\eta = 0.02$, and $\pi = 0$ (no simple-interest component). Because of proportion scale, y and Y are the same here.

monocyclic processes occurs over time, very typical for root diseases (Gilligan, 2002). Host mortality can be incorporated by using a death-rate parameter η. Then ηH, ηL, ηI, and ηR are subtracted from the right hand sides of the equations for dH/dt, dL/dt, dI/dt, and dR/dt, respectively. Host growth can be incorporated in various ways. One approach is to consider just a single per capita growth rate (Ω) for disease-free individuals, and add the term Ω to the right hand side of the dH/dt equation. Suppose, further, that host size (e.g., number of citrus trees in a region) is fixed (say, at $H_{\max}$), and that new trees are only planted if others die. Then, the growth rate is also the mortality rate, and new host individuals can be expressed as $\Omega = \eta H_{\max}$; the combined growth/mortality for H can then be written as $\eta(H_{\max} - H)$.

A more general epidemic model can be written as

$$\begin{aligned}
\frac{dH}{dt} &= -\beta HI - \pi xH + \eta(H_{\max} - H) \\
\frac{dL}{dt} &= \beta HI + \pi xH - \omega L - \eta L \\
\frac{dI}{dt} &= \omega L - \mu I - \eta I \\
\frac{dR}{dt} &= \mu I - \eta R \\
\frac{dx}{dt} &= -\vartheta x
\end{aligned} \tag{2}$$

Note that in this example, total host size ($H + L + I + R$) does not change, even though there is continuous loss and addition of the host individuals (with a balance between the additions and losses). This can be seen by noting that $H_{\max} = H + L + I + R$ and adding the rates: $dH/dt + dL/dt + dI/dt + dR/dt = 0$. The model can be written in different ways to unlink the growth and mortality, to incorporate more complicated linkages, and to account for more than one disease or more than one host genotype at a time, but the example is useful to show one model formulation. When $\pi = 0$ (no simple interest component), an R_0 can be defined for many host-growth/mortality model situations. For instance, with $\pi = 0$ (no simple-interest component), $R_0 = [\beta H_{\max}/(\mu + \eta)] \cdot [\omega/(\omega + \eta)]$. An example realization of this model is shown in the lower frame of Figure 1. Note that Y ($= L + I + R$) and H oscillate a little before settling down to the steady states. The steady-state level of disease at a given R_0 is lower for the dynamic host than the fixed-host situation (equation 1); without the simple-interest component, the steady state Y is $1 - (1/R_0)$.

This approach of using a dynamic (but fixed total) host population size has been used in plant disease epidemiology (e.g., Madden et al., 2000), and even more so in medical epidemiology (Anderson and May, 1991) to determine whether or not an epidemic can occur (i.e., a disease invasion) as well as the persistence (or not) of disease long term. With primary infections occurring throughout the epidemic ($\pi > 0$), the concepts become a little more complicated, but there may still be a threshold (combination of parameters) that must be met for disease to persist (see review in Gilligan, 2002, and references cited therein).

Many other biological features can be incorporated in the model of equation 2. For instance,

since most plant viruses are transmitted by arthropod vectors, the rate of change in H and L does not directly depend on infectious plant individuals (I) but on infective vectors per plant (Z). Thus, the contact rate term, βHI in the first two equations of the model must be replaced by βHZ, where Z is the density of infective vectors per plant. Other components would be unchanged. There is also a need to add equations for the dynamics of the vector population, including virus-free and infective vectors. Details are given in Madden et al. (2000) and Jeger et al. (2004). Other expansions can incorporate disruptions caused by harvesting and/or planting for a multi-season time scale, as well as host responses to infection (e.g., Gilligan, 2002; Madden and van den Bosch, 2002).

The models shown so far are all deterministic. These can all be expressed in stochastic form, which is useful if one is specifically interested in heterogeneity of epidemics, small population sizes, or the epidemic outcome for individual plants or plant units. Gilligan (2002) and Gibson et al. (1999) provide more details. The mathematics definitely becomes more difficult with stochastic models.

Some spatial aspects of epidemics

There are two different threads to the characterization of the spatial component of plant disease epidemics. One thread deals with dispersal and resulting disease gradients, and the use of observed gradients to elucidate the form of the contact distribution (Campbell and Madden, 1990), the probability of a unit of inoculum at one location (ξ) coming in contact with a host individual at location s. This approach has been especially valuable for determining the rate of disease expansion from a focus, both within fields and higher spatial scales (e.g., continents) (van den Bosch et al., 1999). The contributions of van den Bosch and Zadoks (see Zadoks, 2001), Ferrandino (1993), and Aylor (1999) are especially noteworthy for aerial pathogens, and of Gilligan and colleagues (2002) for root diseases.

One of the advantages of the coupled differential equation approach of the previous section is that it can be directly expanded to account for disease at any location as well as any time. With two physical dimensions, it is now necessary to be explicit in notation about time t and location s. With two dimensions, we need to use partial derivatives rather than ordinary derivatives. Expanding equation 1, we can write the spatio-temporal model as:

$$\begin{aligned}
\frac{\partial H(t,s)}{\partial t} &= -\beta H(t,s) \int_{-\infty}^{\infty} I(t,\xi) D(s-\xi) d\xi \\
\frac{\partial L(t,s)}{\partial t} &= \beta H(t,s) \\
&\quad \times \int_{-\infty}^{\infty} I(t,\xi) D(s-\xi) d\xi - \omega L(t,s) \\
\frac{\partial I(t,s)}{\partial t} &= \omega L(t,s) - \mu I(t,s) \qquad (3) \\
\frac{\partial R(t,s)}{\partial t} &= \mu I(t,s)
\end{aligned}$$

where all parameters are as defined before, and $D(s-\xi)$ is the contact distribution, which is simply a scaled version of a disease gradient. Example contact distributions include the exponential, Pareto, Cauchy, and normal. Unlike with the simpler purely temporal model(s), the rate of decline in healthy host individuals at location s (and the rate of increase in latently infected host individuals at s) is explicitly based on the integration of the contributions of infectious individuals at all locations (all ξ values). The specific contribution at ξ to disease at s is the product of magnitude of infectious individuals at ξ multiplied by the probability that a unit of inoculum (say, spore) at ξ reaches location s (based on the contact distribution).

Both so-called wave-like and non-wave-like disease expansion is documented, where the velocity of disease expansion into new areas is constant or increases with time, and supported by the theory summarized in equation 3. The velocity of expansion (or the acceleration of expansion) is generally proportional to $\ln(R_0)$, so that there is no spread if $R_0 \leq 1$. The form of the contact distribution makes the difference in type of expansion. An example realization is shown in Figure 2 for non-wave-like expansion. The linkage of temporal population dynamics of disease and focus expansion rates is of fundamental importance because it shows (qualitatively and quantitatively) how reproduction (infection) and contact probabilities (dispersal) fully determine spatio-temporal

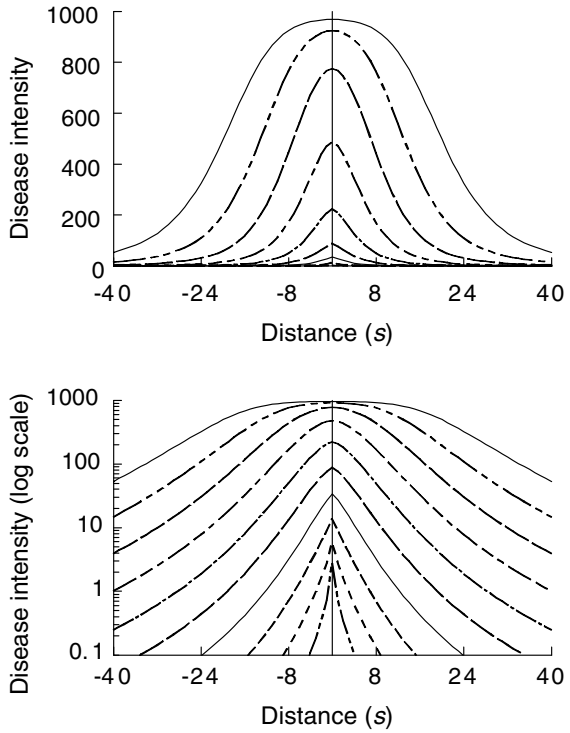

Figure 2. Density of diseased individuals ($Y = L + I + R$) vs. distance from a line source at 10-day time increments based on the numerical solution of equation 3. $H_0 = 1000$. Mean latent period ($1/\omega$) was 7, and mean infectious period ($1/\mu$) was 10 time units. $\beta H_0 = 0.4$ per time unit. A Pareto distribution was used for the contact distribution. The horizontal distance between pairs of successive curves at a single Y value (e.g., 0.1), divided by 10 gives the velocity of disease expansion.

outcomes, given a set of initial conditions. Control strategies are based, once again, on reducing R_0 to below 1, as well as reducing the scale of the contact distribution (spread parameter of the dispersal gradient) to a low value.

Equation 3 can be expanded for host growth, simple-interest dynamics, and so on, just as equation 1 was expanded to equation 2. It is (much) more difficult to work with partial differential equations than with ordinary ones, and finding numerical solutions can even be tedious. When the epidemic starts with a single focus (say, at the edge or centre of a region), then mathematical progress can be made, usually with additional assumptions (van den Bosch et al., 1990).

When there are several initial foci of infections, or unknown number and locations of initial inoculum, spatio-temporal differential-equation models, such as equation 3, are much less useful for studying epidemics because there is no single spatial starting point. With many original starting points (foci with disease at time 0), numerical solutions to equation 3 – or solutions to stochastic analogues of equation 3 (Xu and Ridout, 1998; van Maanen and Xu, 2003), – can be used to describe epidemics and explore implications of biological and physical features on disease progress, but it is more difficult to develop general principles or characterize expansion rates. Moreover, fitting a model such as equation 3 to data is generally impractical with standard statistical programmes. Thus, in epidemiology – as in ecology (Pielou, 1977) for that matter – more statistical (rather than mathematical) approaches have been generally followed to study spatial aspects of epidemics (Madden and Hughes, 1995, 2002; Hughes et al., 1997). This is the second thread of spatial characterization of epidemics. Concepts of clustering, aggregation, and regularity are utilized in terms of many different (but interrelated) statistical methods such as indices of dispersion, correlation, semivariograms, and distance statistics. This conceptual approach goes back to Cochran (1936) and Bald (1937) in plant pathology. A further advantage of the statistical approaches is that results (or concepts) are often directly useful for developing sampling plans, for either estimating disease intensity or making a decision regarding a control intervention (Madden and Hughes, 1999; Hughes et al., 2002).

The interrelationships between spatial aggregation of disease and temporal dynamics is gradually becoming more apparent. Using stochastic simulation, Xu and Ridout (1998) nicely showed how initial conditions, reproduction, and spatial contact distribution affect disease dynamics. A more theoretical approach has been to incorporate spatial properties of epidemics without explicitly using a spatial dimension (i.e., using models similar to equation 1). Models of this type are sometimes called spatially implicit, in contrast to the spatially explicit ones such as equation 3. The approach generally involves using a nonlinear function of I and/or H in the contact term, where the function depends on degree of aggregation (Zhang et al., 2000).

In recent years there has been considerable progress in bringing the two threads together (Gibson, 1997; Keeling et al., 2004), through the ingenious use of stochastic models and parameter