The Study of Plant Disease Epidemics
Laurence V. Madden, Gareth Hughes, and Frank van den Bosch

ISBN: 978-0-89054-505-8

7

Spatial Aspects of Epidemics—I: Pathogen Dispersal and Disease Gradients

Everything is related to everything else, but near things are more related than distant things.

W. R. Tobler (*Tobler's Law of Geography*)

7.1 Introduction

Epidemics are population processes that occur in time and space (see Chapter 1). For the previous three chapters, however, we have mostly ignored the spatial component of epidemics so that a general framework could be developed to characterize and analyze disease development over time. Plant diseases generally do spread in space as they increase in time, and an understanding of plant disease epidemics would be incomplete without consideration of this aspect of disease development.

Disease spread is the result of inoculum dispersal. *Dispersal* is the movement of infectious units (e.g., spores) of a pathogen from one place to another or the movement of infectious units from the place they were formed to other locations (Campbell and Madden, 1990). We consider the terms propagules, propagative units, and units of inoculum as being equivalent to infectious units in this and other chapters. Because propagative units mostly move short distances, new infections primarily occur near other infections and, therefore, there usually is higher disease intensity in some locations than others. If the inoculum initially is concentrated in one location, one can study dispersal by measuring the spore deposition on surfaces (e.g., leaves, flowers) at a range of distances from the inoculum source. Likewise, if diseased individuals are initially concentrated in a focus in one area, disease spread can be determined by measuring disease intensity at a range of distances from the initial disease source. These types of situations are considered in this and the next chapter. However, if there are many inoculum sources, or diseased individuals are located at multiple locations at the start of an epidemic, then disease spread may not be easily measured (although the same biological process is occurring). This is because it will not be known where a deposited spore originated or which inoculum source produced the spore that resulted in a new diseased individual at a given location. In this situation, different spatial methods typically are used for characterizing epidemics (see Chapter 9).

By re-addressing the models and analysis of the previous chapters, one can easily see the importance of the spatial component of epidemics. For instance, for monocyclic diseases, the probability per unit time that a unit of inoculum will come in contact with a host individual and cause an infection (see section 4.4.2) is a function of the distance between the inoculum and the host. Also, the probability that a spore produced on a lesion during the epidemic of a polycyclic disease will reach another host individual is a function of distance between the diseased individual and the "other" host individual. By treating these terms as constants, one is assuming that the epidemic is homogeneous in space, in that there are many inoculum sources and/or infected individuals evenly positioned throughout the host population. Such an assumption is not realistic when the epidemic starts with a single (or a few) spatially-restricted inoculum sources.

We now explicitly consider the spatial component of epidemics. In particular, this chapter deals directly with characterizing disease spread from a single (original) inoculum source. Relatively simple models are presented and their properties discussed. It is shown how to estimate model parameters, and how to compare the spread of two or more epidemics based on the fitted models. It is assumed here that the reader has already digested the information in Chapter 4 on the temporal characterization of epidemics, especially the concepts related to statistical model fitting and comparisons. Thus, where the concepts are the same, the detailed statistical methods presented in Chapter 4 will not be repeated here.

In the next chapter, a much more mechanistic approach is taken to understand the spread of disease through the movement of inoculum. There, we expand on the rate equations (dy/dt) for epidemics in Chapter 5, by explicitly incorporating distance functions into terms for the probability of a spore starting at one location coming in contact with a host at another location. Then, in Chapter 9, we take a different approach and show how to characterize spatial patterns of epidemics without explicit reference to the source(s) of inoculum.

7.2 Dispersal Gradients, Disease Gradients, and Disease Spread

7.2.1 Concepts

The study of spore dispersal is very important historically in plant disease epidemiology. In fact, before the concepts of Vanderplank (1963) started to take hold, "epidemiology" often was used to describe the study of pathogen spores in the atmosphere, including the study of spore movement to new infection sites, and how the environment affected the production, movement, and germination of spores. Gregory (1968, 1973) was a pioneer in the study of spore dispersal, and many of the concepts and principles he developed are still valid today and are reflected in this chapter. Gregory's work was even instrumental in the ultimate formation of an entire scientific discipline—*aerobiology*—that focuses on all aspects of spores, pollen, and other organisms in the atmosphere. Although plant disease epidemiology is now considered to involve a lot more than dispersal, this process is clearly a key component of disease development.

Dispersal of propagative or infectious units (e.g., spores) involves liberation (takeoff), transport (flight), and deposition (landing). The terminology is heavily influenced by the study of spore movement in the atmosphere, but dispersal certainly occurs in the soil as well, although the distances involved are generally much shorter. Inoculum also needs to be considered more broadly than just spores in order to account for the movement of bacterial cells and pathogens in arthropods or other vectors. As discussed in Chapter 4, we deliberately use a loose definition of "inoculum" in this book that can accommodate the movement of vectors that have acquired plant viruses, spiroplasmas, phytoplasmas, and other pathogens.

Liberation of inoculum involves removal of spores or cells from diseased individuals and their release into the transport medium (i.e., air, water, or soil). Transport involves movement of inoculum through a turbulent medium such as air or water. Deposition involves the removal of the inoculum from the transport medium and placement on a surface (e.g., leaf). Overall, dispersal is both a physical and biological process, and research on the topic typically is interdisciplinary in nature. We will not cover the physics of inoculum dispersal in this book, but refer the interested reader to several thorough references (Aylor 1999, and references within; Ferrandino, 1993; Madden, 1992; McCartney and Fitt, 1985).

Instead of focusing on the biophysical details of inoculum movement, we summarize the dispersal processes by the *disease gradient*—the change (generally, decline) in disease intensity or number of infections over distance from an inoculum source. A note of caution should be mentioned here. Gradients of disease can be due to multiple causes. For instance, there can be a gradient of root disease incidence from one side of a field to the other due to an *environmental gradient* such as decreasing soil moisture. We do not deal with such environmental gradients here, but limit our discussion (for the most part) on gradients of disease that are caused by dispersal. Often, the term *dispersal gradient* is used to describe a disease gradient that is caused by variation in deposition of inoculum in relation to distance from an inoculum source. By moving a little earlier in the chain of events, one can deal directly with a *deposition gradient*, which is a measure of spores per unit area or length in relation to distance from an inoculum source. Multiplying this spore (or infectious unit) density by an infection-efficiency factor produces number of infections in relation to distance (i.e., the disease gradient).

Typically, observed gradients may be caused by a mixture of inoculum dispersal and environmental gradients, and researchers should be careful in over-interpreting disease gradient data if they do not have some evidence that dispersal is primarily responsible for the observed systematic change in disease over space. Because this is a book about diseases in plant populations, we focus mainly on disease gradients, and assume in this chapter that they are the result of dispersal. That is, we assume that the disease gradient is a dispersal gradient. One can categorize the dispersal gradient as either primary or secondary. In a *primary gradient*, all infections are due to spores (or other infectious units) originating at the inoculum source. Essentially, primary gradients are due to primary spread of disease (from initial inoculum, which, in this situation, is located at a well-defined inoculum source). In a *secondary gradient*, infections are due to inoculum produced beyond the original source. Essentially, secondary gradients occur when there is secondary spread of disease during an epidemic, which occurs for polycylic diseases (see Chapter 4).

Typically, diseased plants do not move. Thus, plant disease spread is not exactly the same as, for example, the spread of insects from one field to another. The inoculum does move, and as mentioned in the Introduction (section 7.1), spread can be considered the result of inoculum dispersal. In other words, the area occupied by diseased individuals, and the intensity of disease at a given distance from the source, increase due (at least in part) to the dispersal process. In this sense, disease spread is analogous to the population expansion of animals where the parents do not move but the offspring move out to new areas.

Example gradients are shown in Fig. 7.1. These were chosen to show a wide range of pathogen and host systems, and to show that distances of disease spread can be short or long. In most studies, disease gradients are measured over a few meters; thus, most disease gradients are quantified for distances less than 10 m. However, in some cases, gradients are determined over greater distances (>1 km), but this is most easily done when there is a known single inoculum source. In Fig. 7.1, disease intensity is measured or expressed in various ways, such

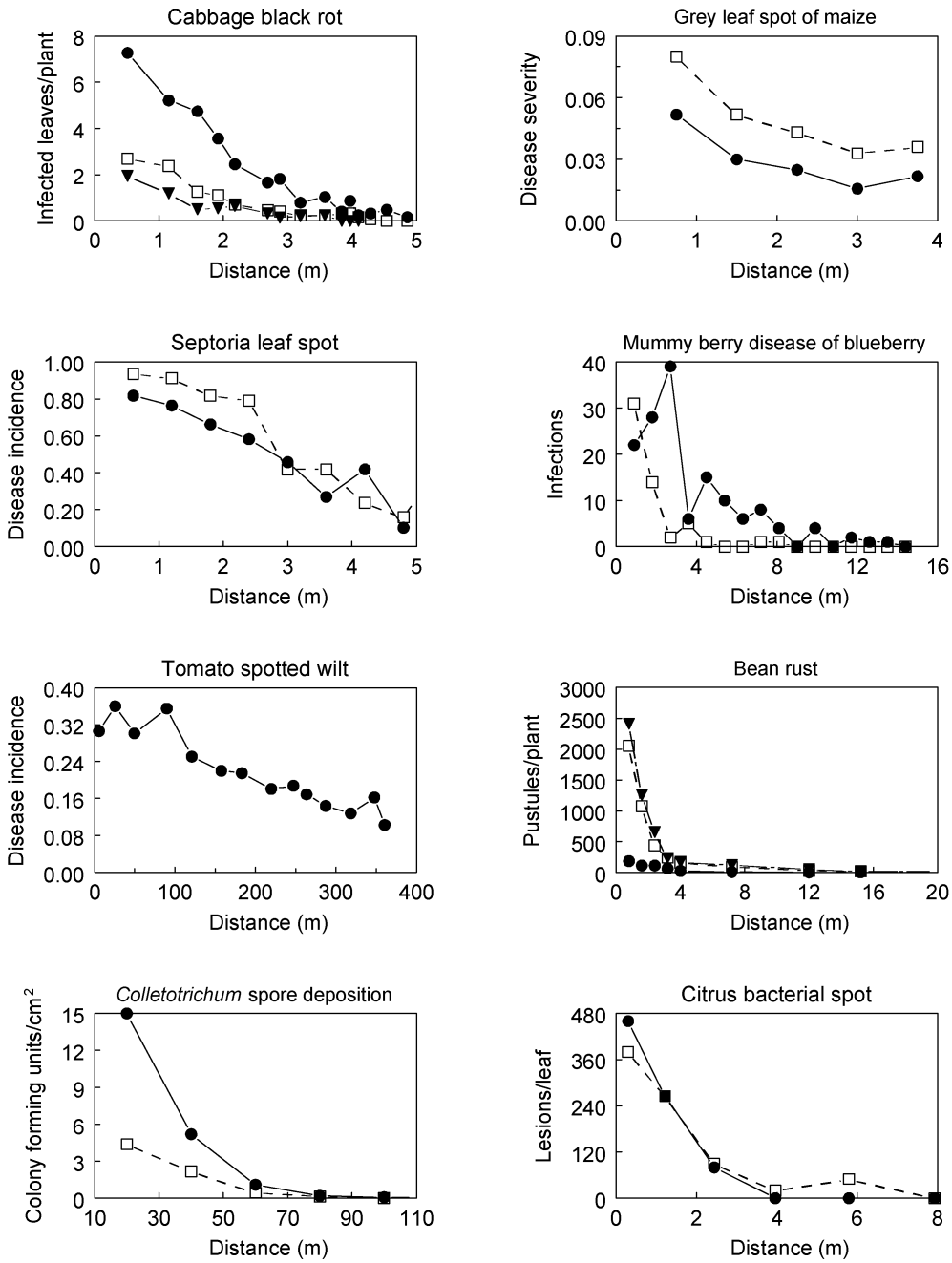

FIG. 7.1. Exampled dispersal and disease gradients. *Left-hand column*: Number of cabbage leaves infected by *Xanthomonas campestris* pv. *campestris*, cause of black rot, from a point source (Kocks and Ruissen, 1996). Incidence of Septoria leaf spot of tomato, caused by *Septoria lycopersici*, in relation to distance from a point source (Parker et al., 1997). Lines are for two different years. Incidence of tomato spotted wilt of peppers, caused by tomato spotted wilt virus, from an area source in Georgia (Gitaitis et al., 1998). Number of spores of *Colletotrichum acutatum* deposited over 1 min at 5 and 15 min into a generated rain over a plastic mulch surface (Yang et al., 1991). *Right-hand column*: Severity of grey leaf spot of maize, caused by *Cercospora zeae-maydis*, in relationship to an area source (Asea et al., 2002). Lines are for two different seasons. Number of shoots of blueberry infected by *Monilinia vaccinii-corymbosi*, cause of mummy berry disease, from a point source (Cox and Scherm, 2001). Lines are for up- and down-wind from the source. Number of bean rust pustules, caused by *Uromyces appendiculatus*, in relation to distance from a line source (Aylor and Ferrandino, 1989). Lines are for different times. Number of lesions/leaf of citrus bacterial spot, caused by *Xanthomonas axonopodis* pv. *citrumelo*. Lines are for two treatments at one assessment time (Gottwald et al., 1997).

as number of infected or diseased leaves, rust pustules per plant, proportion of individuals infected (incidence), and proportion of area with lesions (severity). An example of a spore deposition gradient is also shown for comparison purposes.

The disease gradients have, for the most part, a characteristic shape that is determined largely by the physical process of dispersal. Disease intensity is high at and near the inoculum source (especially if the inoculum source is at or near the soil surface), and generally declines with

increasing distance from the source. The rate of decline is usually large at small distances from the source, especially early in the epidemic, and small at large distances from the source. When disease intensity is moderately high near the source (e.g., considerably more than 10% of the maximum possible), intensity only declines slowly at small distances, and then a steeper decline occurs before the decline becomes small again at large distances (see *Septoria* example). In some cases, there is an increase in disease intensity before the general pattern of decline in disease with distance is observed (see mummy berry disease example). This is more typical when the inoculum source is elevated (i.e., in the upper parts of the host plants).

The disease/dispersal gradients shown in Fig. 7.1 may be quantified and compared using the models presented in this chapter.

7.2.2 Inoculum sources

We have used the term "inoculum source" so far without really defining it. For a polycyclic disease, any infected (and infectious; see Chapter 5) individual is an inoculum source for other healthy individuals. For a monocyclic disease, an inoculum source can be infected individuals from another location or a previous year, or the source can be overwintering spores in the soil. For the purposes of characterizing disease spread, however, we tend to think of an *inoculum source* as a spatially-restricted concentration of infectious units, such as spores, bacterial cells, or even diseased individuals that can be producing the infectious units. That is, the source is separate from, or on the "edge" of, the area over which disease spread is measured, at least at the start of the epidemic. The inoculum source can also be called a disease *focus*, but we mostly use this term in the following chapters. Generally, the source covers an area far less than the area where disease is measured, although this is not a requirement.

It is important to note that this spatially-restricted concept of an inoculum source is, in part, a matter of convenience. That is, in many (most) epidemics, the inoculum source is not separate from the host individuals, but is part of the system where disease is developing. With polycyclic diseases, every lesion (for example) can be an inoculum source. Likewise, with monocyclic diseases, overwintering spores may be found throughout the soil where the host crop is grown. To *study* disease gradients and spread, in practice, one often must use systems where the (original) inoculum source is well-defined in space. Exciting new statistical approaches have been developed in recent years, where such as restriction is not needed, as will be discussed in section 9.10 (Gibson, 1997a, b).

There are three general types of inoculum sources based on geometry and relative size: point, line, and area sources. The distinction is important for the characterization of physical movement of propagules across a plane. Even if the physical process is not of direct concern, parameters in gradient models will depend on source type; thus, one must be cognizant of the source type when comparing results among systems. Although a point is mathematically dimensionless, in the context of disease spread, the term *point source* is used for a very small source relative to the area where spread is being studied. A single infected plant with sporulating lesions in the center or at one corner of a field is an example of a point source. For practical purposes, the diameter (or, more generally, width) of a point source is about 1% of the distance over which spread is being measured. One can think of a *line source* as a point source stretched out across the entire width of a field or plot, perpendicular to the direction of measured disease spread. An example could be a row of infected plants along one edge of a field, or across the center of a field. Fig. 7.2 shows idealized examples of a point and line source, but these are placed at difference locations for clarity of presentation.

An *area source* is bigger than a point or line source, but the meaning is relative to the size of the area over which dispersal or spread is measured. A planting of cedars infected by the cedar apple rust pathogen, *Gymnosporangium juniperi-virginianae*, is an area source for a neighboring apple orchard. One can think of an area source as a line source with depth, so that there is, for example, more than one row of infected (and infectious) plants. One can also think of an area source as an expanded point source, such that its

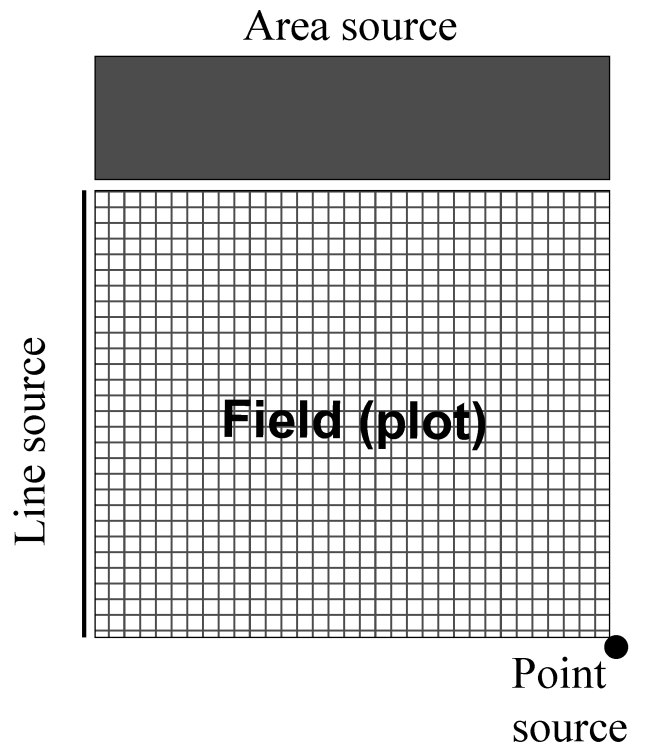

FIG. 7.2. Schematic of a field of host plants with a point, line, and area inoculum source.

diameter is considerably more than 1% of the distance over which spread is quantified. A potato field with late blight may be considered an area source for an adjacent field but a point source for a region, county, or province. Fig. 7.2 shows diagrammatically one version of an area source.

7.3 Models

There are many ways of modeling dispersal and disease spread. Although not presented here, spore dispersal can be characterized by various types of models based on the physics of particle movement in different types of media (Aylor, 1999; Ferrandino, 1993; McCartney and Fitt, 1985). There are also many relatively simple nonlinear models for representing spore deposition, infections, or disease intensity in relation to distance from a source (Jeger, 1983; Maffia and Berger, 1998; Minogue, 1986). These typically are correlative or descriptive (see Chapter 3), but they can be consistent with known mechanisms of dispersal. There are also quite elaborate theoretical mathematical representations of dispersal (and movement of individuals, in general) that account for many population attributes of the biological processes involved without explicitly dealing with physics of particle transport (Othmer et al., 1988; Pielaat and van den Bosch, 1998; van den Bosch et al., 1990a).

For the most part, we focus our attention on some fairly simple nonlinear models of disease gradients that have been proposed. Being consistent with the models used in previous chapters, the ones used here throughout most of the chapter are deterministic and treat distance (s, as in "*s*pace") as a continuous variable. Although descriptive in nature, they are also generally consistent with more mechanistic models of dispersal and can accurately characterize observed gradients. These models also have stochastic versions that are useful for understanding certain aspects of the dynamics of disease development over time and space.

As before, we use Y to represent disease in absolute units, such as number of lesions, number of infected leaves, or area of lesions, and y for disease intensity as a proportion ($y = Y/M$, where M is number of individuals or area of host that can be infected). For convenience, we can refer to Y as number of infections, because if intensity is not too high, one can count lesions and consider these to represent individual infections. Assuming that a fixed proportion (ψ) of deposited spores (or other infectious units) cause infections, then Y/ψ represents the number of deposited spores when Y is not high. For completeness, we can write $Y(s)$ for Y at distance s from the inoculum source, but typically we do not need the extra notation in this chapter. Moreover, because space is explicitly dealt with here, Y represents infections per unit area (such as infections per square centimeter of the surface where infections are counted or measured), or infections per unit length (such as lesions per centimeter distance along a radius of the circle centered on a point source).

For the physical purists, our presentation below is strictly valid for a line source of inoculum, with spread perpendicular to that source. In other words, we present model just for a one-dimensional spatial system. In Chapter 8, we discuss some of the issues regarding spread in one versus two dimensions.

7.3.1 Exponential

The first model we consider is the exponential, sometimes called the negative exponential, which has been used to represent gradients since at least 1940's (Frampton et al., 1942; Gregory and Read, 1949). Here, Y is proportional to an exponential function of distance from the source:

$$Y = a_E \exp(-b_E s) \tag{7.1}$$

in which a_E and b_E are parameters. The parameter a_E represents the (theoretical) level of Y at the inoculum source ($s = 0$), because at the inoculum source one writes equation 7.1 as:

$$Y = a_E \cdot e^0 = a_E 1 = a_E$$

This parameter is thus an indication of *source strength*—the amount of inoculum or inoculum-producing disease intensity at $s = 0$. The parameter b_E, sometimes known as the *spread parameter*, measures the steepness of the gradient and has units of 1/distance. The inverse of b_E (i.e., $1/b_E$) has units of distance and is an indicator of the spatial scale over which spread (or dispersal) is occurring. More will be said about this later. The lower limit for each of the two parameters is 0 (when Y is non-increasing with distance) and there is no upper limit (at least in the model). Obviously, if $a_E = 0$, there are no diseased individuals (infections), and no dispersal is being assessed ($Y = 0$ at all s). When $b_E = 0$, there is no decline in Y with increasing distance. This would be an indication that there was no true inoculum source at $s = 0$; it also could be because disease intensity was measured after a long time and that all plant individuals were diseased. Example realizations of equation 7.1 are shown in Fig. 7.3. As expected, increasing b_E results in a steeper decline in Y over s, and increasing a_E results in an increase in the height of the curves. Equation 7.1 can also be used for disease intensity as a proportion (y). If M is the number of sites that can be infected at a given distance (e.g., maximum possible number of lesions), then $y = Y/M$. Dividing both sides of equation 7.1 by M gives an equation for y instead of one for Y.

The exponential model can be linearized by taking natural logs of both sides of the equation, producing:

$$\ln(Y) = \ln(a_E) - b_E s \tag{7.2}$$

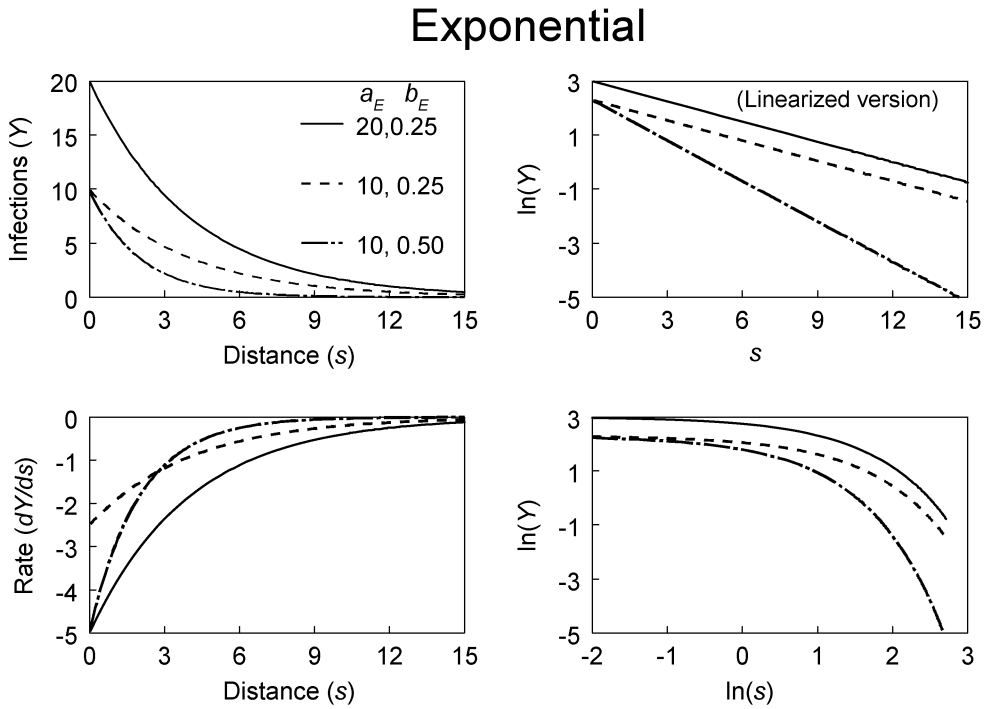

FIG. 7.3. Dispersal gradients (Y versus s) characterized by the exponential model (equation 7.1).

Equation 7.2 is a straight-line relationship between $\ln(Y)$ and s, with slope of $-b_E$ and intercept $\ln(a_E)$ (linearized version frame in Fig. 7.3). In this transformed scale, it can be easier to see the effects of a_E and b_E on the gradients. For instance, two gradients with the same b_E but different a_E values are seen as parallel lines when $\ln(Y)$ is plotted versus distance. With different b_E values, two gradient lines would not be parallel, and the lines could cross depending on the values of a_E. On an untrasformed scale (i.e., Y versus s), it can be very hard to see the effects of different spread parameters on the gradients. If one only had observations of Y at two distances, $Y(s_1)$ and $Y(s_2)$, one could determine b_E using: $b_E = [\ln\{Y(s_2)\} - \ln\{Y(s_1)\}]/(s_2 - s_1)$.

Equation 7.2 is useful for visualizing disease gradients and for parameter estimation (see below). However, the actual exponential gradient of disease is a curve and not a straight line (equation 7.1), meaning that the steepness varies continuously with distance from the source. This can be seen by writing the exponential gradient model in differential-equation form:

$$\frac{dY}{ds} = -b_E Y \tag{7.3}$$

in which dY/ds is the rate of change in Y over distance from the source. We could have introduced the exponential model with equation 7.3 instead of equation 7.1 to emphasize the dynamics of disease spread in space. However, because these gradient models often are justified based on their ability to describe observed gradients and not based on dynamics arguments, we have elected to present the models first in their integrated form. Readers should refer to Chapter 4 for detailed explanations of differential equations and rates of change.

The rate dY/ds is a direct measure of the gradient steepness. As shown in the bottom-left frame of Fig. 7.3, the rate is a large negative value close to the source (near $s = 0$), and approaches zero at large s (depending on the value of the spread parameter). It can be shown that $dY/dt = -a_E b_E$ at $s = 0$ (because Y at $s = 0$ is given by a_E). The steepness of the gradient is a function of only one variable, Y, and one parameter, b_E. The spread parameter can, therefore, be thought of as a summary of all the conditions and factors (e.g., environment) that influence disease gradients other than the actual level of disease intensity. Two exponential gradients that have the same b_E can be said to have equal gradients, no matter what average disease intensity is found for a data set, because they will have the same dY/ds at any common Y.

7.3.2 Power model

The second model considered in this chapter for representing disease gradients is known as the power, power law, or inverse-power model. It has also been called the *Gregory model*, because he was an early proponent of its use to describe gradients and he utilized it extensively to summarize dispersal results (Gregory, 1968, 1973). For the power model, Y is proportional to a power of distance:

$$Y = a_P s^{-b_P} \tag{7.4}$$

in which a_P and b_P are parameters. These parameters have some similarity to the parameters of the exponential

model, but there are also some fundamental differences. Note that when $s = 1$, equation 7.4 reduces to $Y = a_P 1^{-b_P} = a_P 1 = a_P$. Thus, a_P represents theoretical Y at a distance of 1 unit from the source [i.e., $a_P = Y(1)$], not at the source. In fact, Y becomes indefinitely large as s approaches 0 (if b_P is positive), which is biologically meaningless. However, just because the model gives a meaningless result at one particular distance does not mean that it is not useful at all (or most) other distances. Moreover, if the distance between plants (or between the source and the nearest plant) is, say, x cm, it does not matter if the model generates unrealistic values at distances less than x cm. What matters is that the model generates realistic values at distances of x, $2x$, and any actual distance between the source and a plant.

Because the meaning of a_P is linked to a non-zero distance, its physical meaning depends on the units for recording s. For instance, if equation 7.4 was appropriate for representing the gradients in Fig. 7.1, and distance was recorded in meters, then a_P would be the (theoretical) value of Y at 1 m from the source. However, if the distances were recorded in centimeters, then a_P would be the value of Y at 1 cm from the source. Thus, one must always keep the distance units in mind when interpreting the specific value of a_P. If one wants to use a_P as an approximate measure of Y at (more precisely, "very near") the source, then the distance units should be chosen so that $s = 1$ is a lot closer to the source than to the first observed Y value. For example, if the closest distance to the source with observed Y was at 0.5 m, then recording distance in centimeters (and not meters or kilometers) would be appropriate.

In equation 7.4, b_P is also known as a spread parameter. However, it is a unitless term, unlike b_E, which has units of 1/distance. Thus, there is a fundamental difference in the physical interpretation of the exponent parameter of the two models. A rescaling of distance (say, from kilometers to centimeters) will change the value of b_E but not of b_P. (In contrast, remember that a rescaling of distance will change the value of a_P but not of a_E).

Because the value of Y at $s = 0$ is not defined with the power gradient model, the model often is modified by adding a small positive constant (λ) to s:

$$Y = a_P(s + \lambda)^{-b_P} \tag{7.5}$$

We call this the *modified power model* or *modified power-law model*. Here, λ has the same units as s (i.e., units of distance). When $s = 0$, equation 7.5 reduces to $Y = a_P\lambda^{-b_P} = Y(0) = Y_0$, which is a positive constant for positive values of a_P and b_P. In other words, $a_P\lambda^{-b_P}$ is the source-strength term for the power model. The interpretation of a_P can be determined by inserting $1-\lambda$ for s in equation 7.5. This gives $Y = a_P 1^{-b_P} = a_P$. In other words, a_P is the theoretical Y at a distance of $1-\lambda$ from the center of the inoculum source. Mundt (1989) made the interesting argument that λ can be considered a measure of the physical "size" (diameter, width) of the inoculum source.

Of course, when $\lambda = 0$, equation 7.5 reduces to equation 7.4. For generality, we refer to equation 7.5 as the power model for gradients, with the understanding that often it is used with an assumed value of 0 for λ. A plot of Y versus s (Fig. 7.4) for this model has some similarity to the plot for the exponential model (Fig. 7.3). However, for the power model, the curve is steeper close to the inoculum source and shallower a large distances.

Equation 7.5 can be linearized (at least in terms of the a_P and b_P parameters) by taking logs of both sides:

$$\ln(Y) = \ln(a_P) - b_P \ln(s + \lambda) \tag{7.6}$$

This equation describes a straight-line relation between $\ln(Y)$ and $\ln(s + \lambda)$ (see insert of lower right frame of Fig. 7.4). In fact, for this model, a plot of $\ln(Y)$ versus $\ln(s)$ is very close to a straight line except close to the source (i.e., where s is approaching λ). This kind of situation could arise if the true relation was equation 7.5 or 7.6 with $\lambda > 0$, but one fitted equation 7.4 (i.e., $\lambda = 0$) to the data. On the other hand, a plot of $\ln(Y)$ versus s (no transformation), which is the standard scaling for the exponential gradient model, exhibits a large amount of curvature (Fig. 7.4).

The (modified) power model can be expressed in differential-equation form:

$$\frac{dY}{ds} = \frac{-b_P Y}{s + \lambda} \tag{7.7}$$

As with the exponential model, dY/ds is proportional to Y, meaning that the gradient is most steep (most negative) at the highest values of disease (which are typically near the source). However, dY/ds is also indirectly proportional to s (or $s + \lambda$, if λ is non-zero). Thus, at a given level of Y, the gradient is steeper close to the source than farther away. The lower left frame of Fig. 7.4 shows some typical dY/ds curves for the power model. These curves are similar to those for the exponential model, except the steepness is greater near the source for the power model. In fact, if $\lambda = 0$, dY/ds becomes indefinitely large and negative as s approaches 0 (or, for positive λ, dY/ds becomes indefinitely large and negative as $s + \lambda$ approaches 0). For positive λ, dY/ds has a finite value at all s values, and equals $-a_P b_P \lambda^{-(b_P+1)}$ at $s = 0$. This can be seen in the inset of the lower left frame of Fig. 7.4.

In some ways it is more difficult to compare power-model gradients than exponential gradients. This is because the steepness of the gradient in the former case depends on both Y *and* s (see equation 7.7) (see Minogue (1986) for a valuable discussion). Thus, two gradients with the same b_P values do not necessarily have the same gradient steepness (dY/ds) at a given (chosen) Y. Nevertheless, it is still common to compare gradients

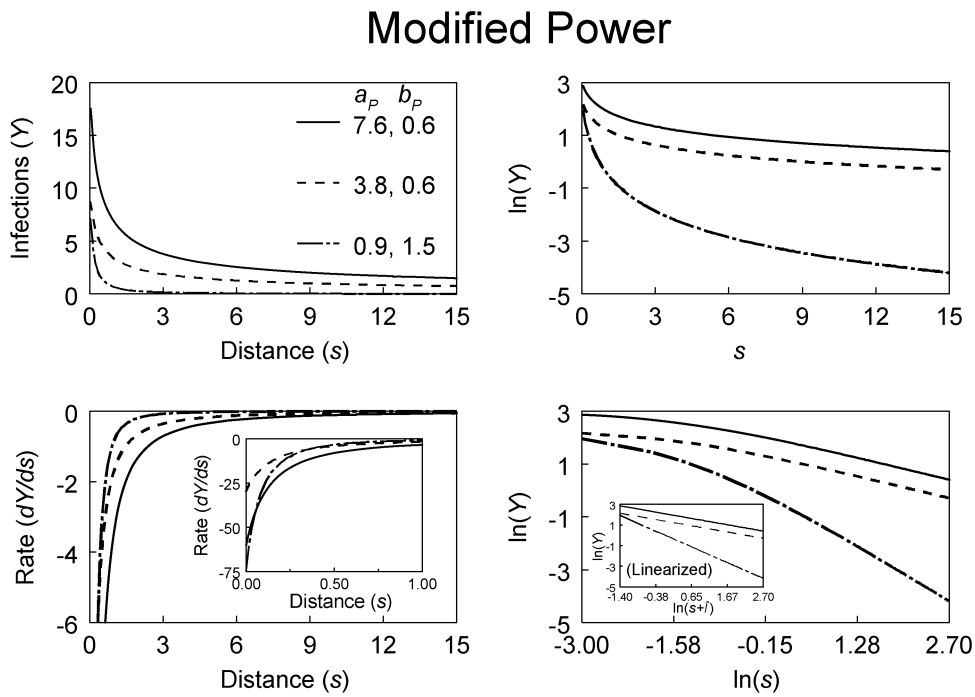

FIG. 7.4. Dispersal gradients (Y versus s) characterized by the (modified) power model (equation 7.5), with $\lambda = 0.2$. Insert in dY/ds graph shows the rate at very close to the source ($s = 0$). Insert in ln(Y):ln(s) graph shows ln(Y) versus ln($s + \lambda$).

through their (estimated) b_P values. In comparative studies, investigators must be aware that two gradients with identical b_P values will have equal changes in ln(Y) with unit changes in ln(s) [i.e., $d\ln(Y)/d\ln(s) = b_P$ (on average); see equation 7.6], but not equal changes in actual measures of disease intensity with unit changes in distance at a given Y.

7.3.3 Power versus exponential model

The exponential and (modified) power models of gradients are, by far, the most common for characterizing disease (dispersal) gradients. In fact, these two models, or similar ones, do provide a very good representation of many observed gradients (e.g., Fitt et al., 1987). Moreover, these two models serve as the basis for further model expansions that are used to represent spatio-temporal dynamics of plant disease development (Gibson and Austin, 1996; Xu and Ridout, 1998). Thus, it is beneficial to discuss these two models, and the gradients they characterize, in a little more detail. The differences in the concept (or meaning) of gradient equality for the two models was already discussed at the end of the previous sub-section. This discussion here expands on some points already made, and covers some of the similarities and differences in the model behavior and interpretation.

Both models represent a situation with a nonlinear decline in Y with s (Fig. 7.5, top frame). For the power model, the gradient is steeper near the source and shallower far from the source, compared with the exponential model. The model difference is even more striking when plotting ln(Y) versus s (Fig. 7.5): the exponential model results in a straight line but the power model can show considerable curvature. This graph is especially useful to keep in mind in conceptualizing the meaning of the two models. For the power model, at large distances from the source, ln(Y) changes very little with unit change in distance (relative to the exponential model), and there remains a small (but non-trivial) level of disease even at very large distances. In terms of dispersal, this means there is a small (but non-trivial) chance of spore deposition even at very large distances.

As discussed in several articles (Aylor, 1999; McCartney and Fitt, 1998), the two models are representative of different physical aspects of spore dispersal in the air. Specifically, propagule dispersal is characterized by:

- the rapid decrease in the aerial spore concentration, and
- the rate of deposition of spores on plants with distance away from the source.

The rapid decrease in spore concentration is due to *filtration* of spores by plants (or other structures) and to *dilution* by mixing of the air, which leads to *escape* of spores from the canopy (Aylor, 1999). When air turbulence is low, the tendency for spores to be deposited is much greater than the tendency to be diluted and escape, and spore density decreases exponentially, on average, with distance (corresponding to equation 7.1). This situation also would occur with splash-dispersed pathogens where air turbulence is less of a factor (Pielaat et al., 1998; Yang et al., 1991). Conversely, when air turbulence

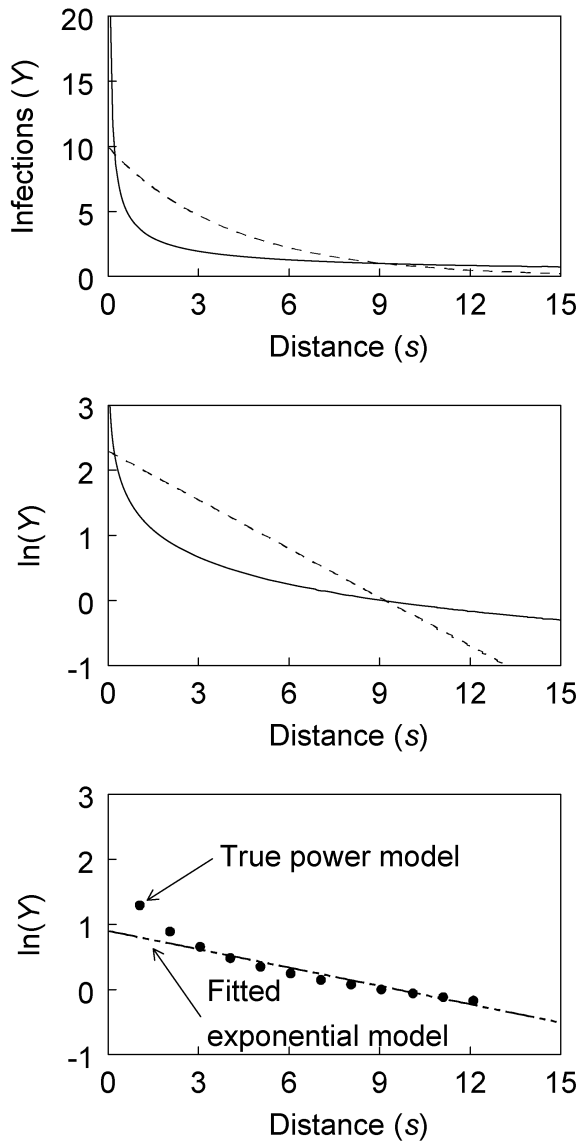

FIG. 7.5. Direct comparison of exponential (equation 7.1; dashed line) and (modified) power (equation 7.5; solid line) dispersal functions, on linear and log scales. In bottom frame, exponential model is fitted to true power-model values.

is high, the tendency for spores to be diluted and escape the canopy is much greater than the tendency to be deposited. In this case, spore density decreases, on average, according to a power model (corresponding to equation 7.4).

Of course, it is quite possible for neither dilution/escape nor deposition to dominate in an actual dispersal event; then, more complex models (see section 7.3.5 and Aylor, 1999) may be needed to adequately characterize the process. In fact, Aylor (1999) points out that deposition and dilution/escape may dominate over different distances. So, in principle, if one could observe distances at a very fine spatial resolution (say, every centimeter), starting at the source and extending for several kilometers, one could observe power-type and exponential-type gradient curves over different distance segments. Fortunately, such refinement in the modeling is not needed for many purposes, and models of the simple exponential or (modified) power types are sufficient for representing a high percentage of observed gradients.

7.3.4 Contact distributions

So far, we have taken a mostly descriptive approach to characterize disease or dispersal gradients (Fig. 7.1), and have indicated how the power and exponential models are consistent with known physical processes of spore dissemination. It is sometimes helpful to characterize dispersal with probabilistic (i.e., statistical) models rather than deterministic ones. Assume that all spores originate at a single (line) source ($s = 0$). Now suppose that the probability of spore deposition in a very small distance increment (Δs) is constant no matter how far the spore has moved, and that the probability is also the same for all spores in a population originating at the source. Call this probability $\zeta\Delta s$ (with ζ having units of 1/distance). Using standard methods for working with continuous distributions, it can be shown that the probability density function (*pdf*) for the random variable s (distance of spore movement until deposition) is given by:

$$D(s) = \zeta \exp(-\zeta s) \tag{7.8}$$

Note that the contact distribution is utilized again in Chapter 9 for certain types of spatial pattern analysis, and is a key component of all the modeling in Chapter 8. The integration of $D(s)$ from s to $s + \Delta s$ is the probability of a spore being deposited over this distance (when starting at distance 0); equivalently, $D(s)$ is the probability per unit distance of a spore being deposited at ("immediately around") s. Borrowing terminology from general theoretical epidemiology (Diekmann and Heesterbeek, 2000), $D(s)$ can be called the *contact distribution*, because, in a sense, it describes the probability of "contact" between a diseased individual at one location [i.e., the source in this case ($s = 0$)] and an individual at another location (the location identified by s). More precisely here, in terms of plant diseases, the contact distribution characterizes the probability of an infectious unit (e.g., spore) originating $s = 0$ coming in contact with a host individual at distance s (determined from the integration of $D(s)$ from s to $s + \Delta s$). Because $D(s)$ is a *pdf*, its integration from $s = 0$ to $s = \infty$ is 1, by definition. Biologically, this means for spore dispersal that a spore must end up somewhere. Additional refinements to the concept of a contact distribution will be found in the next chapter.

It should be pointed out that a spore being deposited at s is not quite the same thing as a spore contacting a host individual at s—there may not be a host individual exactly

at distance s (for instance). For practical purposes, both disease gradients and spore deposition gradients are utilized to determine the appropriate (or reasonable) contact distribution. When appropriate (see Chapter 8), one can multiply a contact distribution based on spore transport by a constant to reflect the fraction of deposited spores at s that actually contact a host.

Equation 7.8 is directly analogous to the exponential disease gradient model (equation 7.1), with $\zeta = b_E$. The difference is that equation 7.8 is scaled so that the integration over all distances is 1. Thus, the "source strength" parameter here is simply ζ. The expected value (mean) and the variance of a random variable with equation 7.8 as its *pdf* equal $1/\zeta$ and $1/\zeta^2$, respectively. The median s ($s_{50\%}$) equals $\ln(2)/\zeta$.

Now, suppose the probability of deposition per unit distance (ζ) is not constant, but is a random variable (with a possibly different value for each spore). This means that each spore has a different probability of being deposited, possibly because of its size or flight taken. The gamma distribution (Evans et al., 2000) is one useful distribution for ζ because it can take on any positive value. Using standard methods for mixing distributions, the following expression for flight distance is obtained:

$$D(s) = \Psi\lambda^{\Psi}(s+\lambda)^{-(\Psi+1)} \qquad (7.9)$$

Equation 7.9 is the *pdf* of a Pareto distribution of the second kind (Johnson et al., 1994), often simply called a Pareto distribution (Minogue, 1989). It is directly analogous to the modified power model for dispersal gradients (equation 7.5), with $b_P = \Psi + 1$. The only difference is that scaling is used here so that the integration of equation 7.9 over all distances equals 1. In contrast, there is no scaling of spore numbers (or infections) when typically using equation 7.4, and a_P gives the height of the Y curve. It is interesting to note that the mean and variance for the Pareto distribution are only defined for certain (large) values of Ψ. Minogue (1989) gives a very useful discussion of this property. If $\Psi > 1$ (i.e., $b_P > 2$), then the mean s is $\lambda/(\Psi - 1)$. The median s, which is defined for all Ψ, is $s_{50\%} = \lambda(2^{1/\Psi} - 1)$.

The differences between equations 7.8 and 7.9 are most apparent at very small and very large s values (near and far from the source), where the Pareto function gives larger values than the exponential. This is, essentially, the same concept as shown in the middle frame of Fig. 7.5 for the deterministic dispersal models. A plot of $\ln[D(s)]$ versus s has the same shape as a plot of $\ln(Y)$ versus s. When $\ln[D(s)]$ declines linearly with s, it is said that the distribution is *exponentially bounded*. Thus, the exponential distribution has this property. More details on the concept of exponential bounds for distributions are given in Chapter 8 and in Minogue (1989). Chapter 8 also describes generalizations of the distributions considered here.

With reference to the predicted values at large s, the Pareto distribution is not exponentially bounded, and therefore has a so-called "fatter tail" than the exponential. The implications of this difference can be seen in a simple simulation of spore flights. Using a random number generator, 2000 spore flight distances (from random positions along a line segment) were generated for an exponential and Pareto distribution. To make them equivalent in one sense, a median s ($s_{50\%}$) of 0.6 was chosen for both distributions. As shown in Fig. 7.6, points are clustered close to the source, and the density of points drops off fairly rapidly, for both models. With the exponential distribution, there is a very smooth drop off in density, with the largest distance around 8. However, with the Pareto model (i.e., power-model type dispersal), there is more irregularity in the drop off of point density, and there are points found at a very wide range of distances. In fact, 26 distances (out of 2000) were actually greater than the upper limit of $s = 1000$ used in the graph. It should be re-emphasized that half of the points in both graphs are for distances below 0.6.

One other feature of Fig. 7.6 is worth noting. For the Pareto distribution, not only are spores deposited far beyond the source, but some clumping or clustering of points at various distances are somewhat apparent. Thus, for pathogens with a Pareto-type distribution for spore dispersal, establishment of new foci of infection beyond are original source, and clustering of disease around the new infections, are standard features. These topics are dealt with in later chapters.

The exponential and Pareto contact distributions are emphasized in this sub-section because these correspond directly to the popular and useful exponential and power disease gradient models. Other distribution functions have been found to be useful, however. For instance, the normal, (half-)Cauchy, and other distributions have been used for $D(s)$ (Shaw, 1995; van den Bosch et al., 1988b, c; Pielaat and van den Bosch, 1998; Xu and Ridout, 1998). These all vary in their general shape, especially in terms of the "heaviness" or "fatness" of their tails at large distances (i.e., how fast $D(s)$ declines with s at large distances). The implication of tail fatness will become even more apparent later in this chapter and the next chapter when spatio-temporal disease development is discussed.

7.3.5 Some other dispersal models

As with temporal models of disease progress, there are other models than can be used to represent disease gradients away from an inoculum source. As mentioned above, spore dispersal is determined by a mixture of deposition and dilution (and escape) processes (Aylor, 1999), and one or the other may be more important under different conditions. It is possible that neither process will dominate, and that an observed gradient shows characteristics of both deposition and dilution. Lambert et al. (1980) made an initial step at expanding

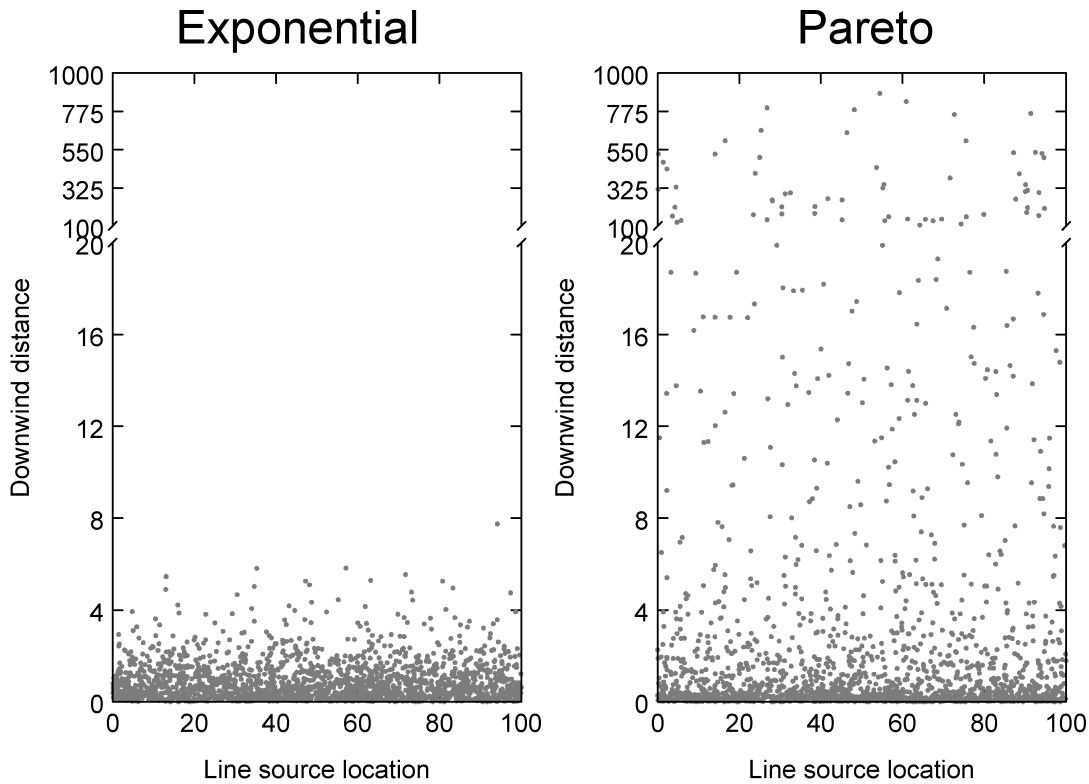

FIG. 7.6. Simulated dispersal of 2,000 spores (or other units of inoculum) from a line source (distance, s, of 0) according to the exponential (equation 7.8) contact distribution, with $\zeta = 1.15$, and Pareto distribution (equation 7.9), with $\lambda = 0.2$ and $\Psi = 0.5$. Both distributions have the same median, $s_{50\%} = 0.6$, meaning half of all points on the graphs are less than 0.6.

on the above models by developing a more general gradient model. Their function can be written as:

$$Y = a_1 \exp(-b_1 s^{\eta}) \tag{7.10}$$

where a_1, b_1, and η are parameters. When the unitless shape parameter, η, equals 1, equation 7.10 reduces to the exponential model (equation 7.1). When η approaches 0, equation 7.10 reduces in the limit to the power model (equation 7.4). Other values of η allow the model to take on other shapes.

In some ways, the generalization of equation 7.10 may not be needed. This is because the modified power model is, in fact, a generalized dispersal gradient model if λ can take on any value. In this situation, λ and b_P work together to indicate the shape of the gradient. This can be seen by the equation 7.7 for dY/ds. At large λ (for example, much larger than the largest distance considered), s in the denominator has very little influence on the rate. In other words, if λ is large, increasing s does not change the denominator to any extent, and the model can be written as $dY/ds \approx (b_P/\lambda)Y$, which is the formulation for the exponential model (with $b_E = b_P/\lambda$). In more mathematical terms, when λ becomes indefinitely large, with b_P/λ held to a finite value (meaning that b_P also becomes indefinitely large), the modified power model reduces to the exponential model.

A different generalization is possible by combining both exponential and power functions in the same model (Maffia and Berger, 1998). One formulation is:

$$Y = a_1 s^{-b_2} \exp(-b_1 s) \tag{7.11a}$$

in which a_1, b_1, and b_2 are parameters. This model is consistent with some physical models for that have been developed for spore dissemination (see equation 15 in Aylor, 1999). When $b_2 = 0$ (in principle, when dilution/escape of spores is of no or little importance relative to deposition processes), equation 7.11a reduces to equation 7.1. Likewise, when $b_1 = 0$ (in principle, when deposition is of little importance relative to dilution/escape processes), equation 7.11a reduces to equation 7.4. It is possible, however, for both b parameters to be non-zero. The equation can be linearized to:

$$\ln(Y) = \ln(a_1) - b_2 \ln(s)(-b_1 s) \tag{7.11b}$$

which facilitates in parameter estimation (see below). The model can be further generalized by adding a constant (λ) to s, as done with the power model (equation 7.5).

Another possible model is based on a deterministic version of the half-Cauchy statistical distribution for $D(s)$. One way that this model can be written is:

$$Y = \frac{a_1}{(1+(b_1 s)^2)} \tag{7.12}$$

Interestingly, Aylor and Ferrandino (1990) developed this same model for spore deposition near an inoculum source based on assumed mechanisms of spore trajectories. Like the power model (equation 7.4 or 7.5), this model involves raising s (actually, a function of s) to a power; thus, it may be appropriate as an alternative to the power model for gradients that are very steep near the source ($s \approx 0$) and flat at large distances. Unlike the power model (without modification; equation 7.4), however, equation 7.12 has a finite value at $s = 0$, namely $Y(0) = a_1$. As a contact distribution, the half-Cauchy distribution has been used heavily in simulation and theoretical studies (Shaw, 1995; Xu and Ridout, 1998), but has not gained much popularity (yet) for characterizing observed gradients. The interested reader can verify that with equation 7.12, $1/Y$ is a linear function of s^2.

Some other contact distributions are discussed in Chapter 8.

7.3.6 Some calculations

Readers who have learned how to perform calculations with the empirical disease progress models of Chapter 4 (see equation 4.4.6) will have no trouble performing similar operations with gradient models. It is sometimes convenient to consider that the two most common models (exponential and power) can be written generically as:

$$\ln(Y) = a^* - b^* s^* \tag{7.13}$$

where a^* represents either $\ln(a_E)$ or $\ln(a_P)$, b^* is either b_E (units of 1/distance) or b_P (unitless), and s^* is either just s (for exponential) or $\ln(s + \lambda)$ for (modified) power (with $\lambda = 0$ in many cases).

Just a couple of calculations are demonstrated here. Suppose that distance is measured in centimeters from a source, Y represents (mean) rust pustules per plant, and that $a_E = 500$ and $b_E = 0.02$/cm. *What is Y at 4 m (400 cm)?* Using equation 7.13, $\ln(Y) = \ln(500) - 0.02 \times (400) = 6.215 - 8 = -1.875$. Back-transformation gives a Y of 0.168. Such a simple calculation could be done directly with equation 7.1. Now, with the same parameters, *at what distance is Y = 1?* Rearranging equation 7.13, leads to:

$$s^* = \frac{-[\ln(1) - \ln(500)]}{0.02} = \frac{-[0 - 6.215]}{0.02}$$
$$= 310.8 \text{ cm } (\sim 3.1 \text{ m}).$$

The interested reader should be able to determine *how much a_E or b_E would have to be changed so that Y declines to 1 pustule/plant not at 3.1 m but at 1 m (100 cm)*. Answer: either new $a_E = 7.39$ (a reduction in source strength of ~496) or new $b_E = 0.062$/cm (an increase of gradient steepness of 0.042/cm).

When performing similar calculations with the power model, it is important to remember that one is dealing with a logarithmic transformation of distance. For instance, with $\lambda = 1$ cm, suppose that $b_P = 1.2$ and that $a_P = 500$ (i.e., that Y at the source is $500 \times 1^{-1.2} = 500$ pustules/plant). The distance at which $Y = 1$ can be determined from: $s^* = -[\ln(1) - \ln(500)]/1.2 = 5.18$. Given that $s^* = \ln(s + 1)$ here, s is calculated as: $\exp(5.18) - 1 = 176.7$ cm (~1.8 m).

7.4 Model Fitting

7.4.1 Choosing a model

As with disease progress curves (Chapter 4), a combination of interrelated graphic and statistical methods can (or should) be used to select a proper model for describing a disease gradient. As with all methods, the model selection process is constrained by the available data. That is, any conclusion regarding the 'best' model may be limited due to measurement error (see Chapter 2), sampling error (Chapter 10), and the range of distances over which Y is assessed. Nevertheless, choosing a reasonable model is essential for quantifying dispersal and disease spread (see later), and in comparing results for different treatments, locations, years, pathogens, etc.

The general methods presented in Chapter 4 can typically be applied to observed dispersal gradients to select a reasonable model for the data. We demonstrate the process with data on number of pustules per plant for rust of beans (Aylor and Ferrandino, 1989). A line source was used, and results for the first assessment time in 1986 in the south direction were read from the graph. As expected, there was a decline in Y with increasing s (Figs. 7.7 and 7.8), which is consistent with either exponential or power-type models for dispersal. The sharp decline in the first 3 m or so, and the very flat curve at large distances are suggestive of a power model, but this could be misleading if data are not recorded for very small and large s.

A plot of dY/ds versus s is of importance for theoretical distinctions between the two model processes (dY/ds goes rapidly to very large negative values for power, and less rapidly to finite negative values for exponential) (see Figs. 7.3 and 7.4). However, the plot of $\Delta Y/\Delta s$ typically is not of practical value in model selection. This is because it is generally unrealistic to have a sufficient number of sampling distances very close to the source to see if dY/ds is declining fast enough to be representative of the power model. Since estimated dY/ds is a ratio, there can be considerable imprecision in its estimate,

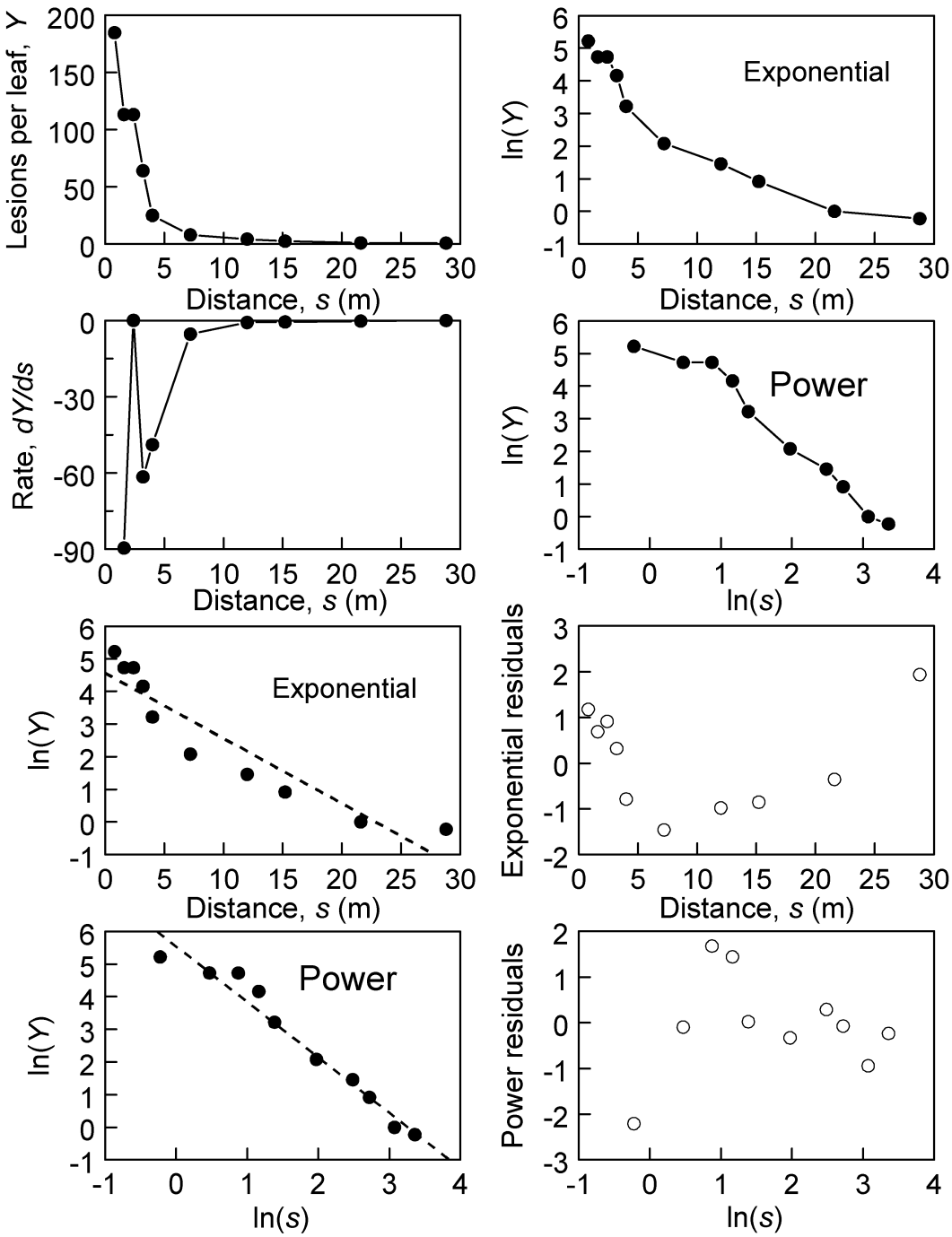

FIG. 7.7. Number of lesions (pustules) per leaf in relation to distance from a line source, for bean rust (from Aylor and Ferrandino, 1989). Upper four frames: data are shown on linear and transformed scales. Lower four frames: fit of exponential (equation 7.2) and power (equation 7.6) models to the transformed data, together with the residuals.

resulting in large scatter of points. This can be seen for the bean rust example: although the trend clearly is towards large negative numbers as s goes towards 0, there is also a wide range of rate estimates over the first three-to-four distances (Fig. 7.7).

Plots of $\ln(Y)$ versus s and $\ln(s)$ are especially useful for selecting a model. These graphs correspond to the linearized versions of the two dispersal models (equations 7.3 and 7.6)—a straight line results if a particular model is appropriate. For the bean rust example (Fig. 7.7), the plot of $\ln(Y)$ versus s shows considerable curvature, indicating that the exponential model would not be appropriate. The type of curvature is typical when power-model values are plotted on this scale (see Fig. 7.3). On the other hand, a plot of $\ln(Y)$ versus $\ln(s)$ (Fig. 7.7) was a fairly straight line, except very close to the source, where some curvature of the line was revealed. As discussed above, this curvature at small $\ln(s)$ occurs when a λ is greater than 0 in equation 7.6 (see Fig. 7.4). Using a value of λ equal to half the distance

between the source and the first sampling distance (0.4), reduced this curvature somewhat, but as we show later, the impact of this addition was minor on model fitting. Overall, the power model seems more reasonable than the exponential model for describing the example data set.

As with all evaluations of dispersal/disease gradients, the range of sampling distances can greatly influence the results (Ferrandino, 1996). Over a limited distance range, the model predictions can be very similar, even though there are fundamental differences in the two model types at very large and small distances from the source. This can be seen with the bottom frame of Fig. 7.5. The solid points represent disease values corresponding exactly to the power model (equation 7.4), but the data are plotted in the form of $\ln(Y)$ versus s—the standard scale for the exponential model (equation 7.1). The straight line represents one fit of the exponential model to these values. The fit is quite good, except close to the source, where the exponential model underpredicts the true Y values. If the first observation (at $s = 1$) was not present (e.g., when the first row of plants was 2 m from the source), a researcher would have little evidence that the exponential model was inappropriate. If data were available far beyond $s = 12$, the differences between the two models could become noticeable—the decrease in true $\ln(Y)$ with distance would continue to decline for this power-model relation (i.e., the curve would become flatter), whereas the exponential model would indicate a continuous linear decline in $\ln(Y)$.

The main point from this discussion is that the investigator should never over-interpret the results from a single data set. Given that different physical processes may dominate over different distance ranges (Aylor, 1999), and that observed gradients generally represent only a portion of an entire dispersal gradient, even a perfect fit of a simple model to a small data set does not indicate that this model would necessarily hold at all distances or for all dispersal events for the same pathogen/host combination. One *does* need to have reasonable model choices to quantify properly the spread of disease, and to compare treatment results. However, conclusions on mechanisms must always be based on cumulative evidence, and not on a limited dataset. Ferrandino (1996) further elaborates on the challenges in distinguishing between exponential-like and power-law-like gradients when disease is assessed over limited spatial ranges.

7.4.2 Estimating parameters—linear methods

A wide range of statistical methods can be used to fit disease-gradient models to data. It is assumed here that the reader has digested the presentation on model fitting for disease progress curves (section 4.6), so that considerably less detail is needed here. Least squares and maximum likelihood (as well as some other statistical methods) can be used for model fitting and parameter estimation. We focus in this chapter primarily on least squares methods.

As shown with equation 7.13, both the exponential and modified power models (with λ assumed known) can be expressed as linear models. Readers can see Chapter 3 for discussion of the meaning of models that are linear in the parameters. Statistical versions of these mathematical dispersal models can be written as:

$$\ln(Y)_j = \ln(a_E) - b_E s_j + \varepsilon_j \qquad (7.14a)$$

$$\ln(Y)_j = \ln(a_P) - b_P \ln(s + \lambda)_j + \varepsilon_j \qquad (7.14b)$$

in which the j subscript indicates the j-th data point or distance ($j = 1, \ldots, n$), and ε is the error (residual) term, a random variable (difference between observed transformed Y and that determined by the model). The intercept for the models is $\ln(a_E)$ or $\ln(a_P)$, and the slope is b_E or b_P. Equation 7.14b is not, of course, linear in terms of the parameter λ, because λ appears only as a logarithmic function of $s + \lambda$. One needs to pre-specify a value for λ to use linear statistical methods. The most common approach is to assume that $\lambda = 0$, which results in the simpler regular power dispersal-gradient model. However, one could specify a value roughly equal to the "size" (diameter, depth) of the source (Mundt, 1989).

To use linear least squares, the standard assumptions about the distribution of the error term (ε), and hence the response variable [$\ln(Y)$ in this situation], are made (see equation 4.6.2, and Neter et al., 1983). Normality usually is not overly important, except under extreme circumstances, but independence of the observations and constant variance are very important assumptions. With disease progress curves, the response (transformed disease intensity) is cumulative, in that Y at $t = 20$ equals Y at $t = 10$, plus any new disease that develops between times 10 and 20. Thus, it is very important to check for serial correlations with disease progress modeling. However, Y [or $\ln(Y)$] for disease gradients, as typically expressed, is not a cumulative value over space; for example, Y at $s = 20$ is *not* the accumulation of lesions (or infections) from distance values of 0 to 20. In this sense, one might not expect to find correlations of Y values for disease gradients. In fact, because Y *is* declining with distance (Fig. 7.1), there is (spatial) autocorrelation (lack of independence) because of this spatial trend; that is, Y is serially correlated over distance. In most cases when there is a strong gradient of disease from a (single) source, the correlation of Y values can be "removed" by fitting a gradient model to the data. For instance, $\varepsilon = Y - a_E \exp(-b_E s)$ may not be correlated (if the exponential model is reasonable) even if the Y values are correlated. Using time series descriptions, the trend is removed by fitting a particular model to gradient data. It is easy to check for independence by assessment of the residuals, as discussed in Chapter 4, just to be sure.

If Y equals number of deposited spores or number of lesions, it is unlikely that $var(Y)$ is constant. In general, $var(Y)$ can be expected to increase with mean Y (Schabenberger and Pierce, 2002). It turns out that for discrete data in the form of counts, the logarithmic transformation does a reasonable job of stabilizing variances. That is, the variable $\ln(Y)$ may have close to a constant variance for practical purposes, and weighting based on assumed (or calculated) variances is not necessary. Thus, we assume that $\ln(Y)$ does have a constant variance for model-fitting purposes here. Later in this chapter we deal with situations where the response variable (whether transformed or not) cannot be assumed to have a constant variance.

Fig. 7.8 shows the data points plotted in Fig. 7.7 in the form of a SAS program for least squares analysis. Fitting the exponential model (equation 7.14a) to the bean rust data resulted in the following equation:

$$\ln(\hat{Y}) = 4.58(0.352) - 0.20(0.027)s$$

in which the numbers in parentheses are estimated standard errors. The coefficient of determination (R^2) was 0.878, and the residual variance was estimated as 0.579. The estimate of a_E can be obtained by back-transformation of 4.58 as $\exp(4.58) = 97.5$, which is considerably below the observed Y at the first distance ($s = 0.8$). It should be noted that the estimate of $\hat{b}_E$ has units of 1/distance; that is, $b_E = 0.20$/m. Evaluation of the predicted line and the residuals in the lower frames of Fig. 7.7 reveal clearly that this model did not provide an appropriate fit over the range of observed data. In particular, the (standardized) residuals showed a very strong pattern. Of course, if the observations within 2 m of the source and beyond 25 m were not made, a good fit would have been obtained.

Fitting the modified power model (equation 7.14b, with $\lambda = 0$) to the data resulted in the following equation:

$$\ln(\hat{Y}) = 5.56(0.246) - 1.70(0.119)\ln(s)$$

```
/* Dispersal gradient analysis  */
data a;
            input s Y @@; /*pairs of (s,Y) strung across records */
            lnY=log(Y); lns=log(s); lns4=log(s+.4);
            datalines;
0.8  184.9   1.6 113.3   2.4 113.3    3.2  64.1
4.0   25.0   7.2 8.0     12.0  4.3    15.2  2.5
21.6 1.0     28.8  0.8
;
proc reg data=a;
            title 'Linear regression analysis';
            /* residual plot specified for each model */
            model lnY = s;  plot r.*s;  *exponential model;
            model lnY = lns;  plot r.*lns; *power model;
            model lnY = lns4; plot r.*lns4; *modified power model (c=0.4);
proc nlin data=a;
            title 'Power model (fixed c = 0)';
            parameters a=1000 b=2; /* Initial guesses of parameters */
            bounds a >= 0; bounds b >=0;
            c=0.0;
            model lnY = log( a*(s+c)**(-b) );
proc nlin data=a;
            title 'Modified power (with c unknown)';
            parameters a=1000 b=2 c=.4;
            bounds a >= 0; bounds b >= 0; bounds c >= 0;
            model lnY = log( a*(s+c)**(-b) );
proc nlin data = a;
            title 'Modified power (c = 0.4)';
            parameters a = 1000 b = 2;
            bounds a >= 0; bounds b >= 0;
            c=0.4;
            model lnY = log( a*(s+c)**(-b) );
run;
```

FIG. 7.8. Input program file of SAS to read in dispersal data and fit dispersal models, in linear and nonlinear forms. Note: in version 9.1 of SAS, the NLIN procedure no longer displays the total sum of squares, SST (see section 3.5.2), which is needed to calculate the pseudo-R^2. SST can be obtained by determining the variance of the response variable and multiplying by number of observations minus 1.

which had an R^2 of 0.962 and an estimated residual variance of 0.179. The slope here is unitless. The back-transformation of the intercept is given as exp(5.56) = 259.8, which is an estimate of Y at $s = 1$ m, *not* at the source. One could re-scale distance from meters to centimeters, as an example, by multiplying distance by 100, and then taking the log of s. Fitting the power model (with $\lambda = 0$) produces the same results, except for the intercept term which is now estimated as 13.38 (standard error = 0.766). Back-transformation gives exp(13.38) = 646,934, obviously a very large number, for the predicted Y at 1 cm. Re-scaling to millimeters would give an even larger number. This is because Y becomes indefinitely large as s goes towards 0 with this model (and a_P is the value of Y at $s = 1$). Using the results for the centimeter distance scale, predicted Y at 100 cm (=1 m) is given by:

$$13.38 - 1.70\ln(100) = 13.38 - 7.83 = 5.55$$

which is the same as the original intercept with distance in meters (except for rounding error).

Graphic evaluation of the predicted line and residuals showed that this model was clearly more appropriate than the exponential for the bean rust data set. The greatest errors were near the source, where the slight curvature on the log-log scale was previously identified. Improvement in fit could be accomplished by consideration of a value of λ greater than 0. Readers can evaluate a range of λ values, but we show results for just $\lambda = 0.4$ m. This is half the distance between the first sampling distance and the center of the source, which may be a reasonable choice if λ is interpreted physically as a measure of size of the source. Least squares resulted in the following equation:

$$\ln(\hat{Y}) = 6.10(0.228) - 1.88(0.108)\ln(s + 0.4)$$

which had an R^2 of 0.974, slightly larger than when $\lambda = 0$. The slope parameter was slightly larger (1.88 versus 1.70), and the intercept was also larger. Back-transformation of the intercept, exp(6.1) = 445.9, is an estimate of a_P, equaling the predicted Y when $s + 0.4 = 1$ m (or $s = 0.6$ m). One can determine the predicted Y at $s = 1$ m, with $Y(1) = a_P(1 + \lambda)^{-b_P} = 445.9(1.4)^{-1.88} = 236.9$. This estimate is close to the 259.8 value obtained for a_P with the simpler power model above ($\lambda = 0$), which is, by definition, estimated Y at $s = 1$ when $\lambda = 0$.

Other than possibly improving the fit to observed data, the advantage of the modified power model (with $\lambda > 0$) is that one can directly estimate Y at the source, $Y(0)$, as a measure of source strength. Using the formula given after equation 7.5, one obtains:

$$Y(0) = 445.9(0.4)^{-1.88} = 2497.$$

This value is considerably larger than the source-strength estimate based on the exponential model ($\hat{a}_E = 97.5$). This $Y(0)$ value, however, is considerably below the non-sensical estimate of Y at 1 cm [i.e., $Y(0.01)$], given above when the regular power model ($\lambda = 0$) was fitted to the data. The reader is reminded that the exponential model provided a rather poor fit to these data close to the inoculum source, and that the regular power model breaks down as s approaches 0.

In summary, for the example data set, the (modified) power model provided a better fit than the exponential model. This will not necessarily be true for other data sets. Using a value of λ greater than 0 resulted in a better fit than the use of $\lambda = 0$. However, it is not necessarily clear that one needs to use the more general model for these data. The improvements in R^2 and residual plot (not shown) were marginal when using $\lambda = 0.4$ m. The reader can determine that the fit can be even better with higher values for λ. If the goal of the research was to obtain a precise estimate of the "source strength" and the gradient of disease close to the source, then the use of a more complex model is certainly warranted. But if the goal was to provide a reasonably simple overall description of the gradient (starting around 0.7 m), to facilitate comparisons among groups (e.g., treatments, cultivars, years), then the more general model would not be needed. In fact, the use of the more general model where it is not fully needed may be disadvantageous because model fitting would be more complicated.

It must be pointed out that a direct comparison of the fit of the exponential and modified power models is, in a sense, not completely fair if λ is allowed to take on any positive value. In this situation, the modified power model contains three unknown parameters, whereas the exponential model has only two unknowns. Generally, a 3-parameter model will provide a better fit than a 2-parameter one, especially if the 2-parameter version is a special case of the 3-parameter model. This was seen by fitting a Richards model (with unknown η parameter) to disease progress data (Chapter 4). When λ is determined directly from the observed data, the modified power model becomes a flexible (general) gradient model that can take on multiple shapes (Minogue, 1989). One cannot directly estimate λ using linear least squares because the model (equation 7.24b) is not linear in terms of λ. Regressing $\ln(Y)$ on $\ln(s + \lambda)$ for a large number of possible λ values (say, $\lambda = 0$, 0.1, 0.2, …, 1), and selecting the λ with the highest R^2 is an indirect way of estimating this parameter. We show in the next sub-section how to estimate λ with nonlinear methods.

Specifying a value of λ greater than 0 in the power model *is* important if one is trying to quantify the gradient very near the source. The obvious question then arises from this discussion: how can one compare the fit of the exponential model and the modified power model (with non-zero λ) without giving unfair advantage to the latter? One answer to this is to choose a λ independent of the gradient. This could be done before collecting the data or before looking at a graph. If λ can be considered a measure of the size of the source, then this value can

certainly be approximated. The rule followed above was to use a λ equal to half of the distance between the center of the source and the first data point. Other possibilities exist. However, investigators should be aware that estimates of a_P and b_P are always conditional on the choice of λ.

7.4.3 Estimating parameters—nonlinear methods

The standard models for dispersal gradients are all nonlinear in the parameters. In common practice, as discussed above, the models are linearized so that linear least squares can be utilized to estimate parameters. Nonlinear estimation methods offer great flexibility, in that investigators are not forced to use models that can be linearized. There is a penalty associated with nonlinear methods, however, because the methods are more complicated, may fail or may produce biased estimates of parameters, and require a considerable number of data points compared with linear methods. Experienced users, however, can benefit greatly from nonlinear estimation methods. Interested readers should refer back to sections 3.5 and 3.6 of Chapter 3. We present some analyses here for the bean rust data set. Relevant code for fitting is given in Fig. 7.8.

A dispersal model can be expressed generally as: $Y_j = f(s_j; \Theta)$, where Θ represents a vector of parameters. Two ways of expressing the model in statistical form are:

$$Y_j = f(s_j;\Theta) + \varepsilon_j \qquad (7.15a)$$

and

$$Y_j = f(s_j;\Theta)\varepsilon_j \qquad (7.15b)$$

in which the j corresponds to the j-th observation (distance from the source). The comments made in Chapter 4 (section 4.6.3) regarding appropriateness of the different statistical forms of disease progress models are applicable here. In short, if equation 7.15a is correct, only nonlinear methods can be used, because this statistical model cannot be linearized (but see section 3.5), even though a transformation of $f(\bullet)$ is linear. On the other hand, if $f(\bullet)$ represents the exponential or power models (and some others), equation 7.15b can be linearized—by taking logs of both sides. It is not clear based on our understanding of the population biology of diseases whether the error term in a nonlinear statistical model should be additive or multiplicative.

To avoid excessive repetition of ideas, we only discuss one assumption about the ε, and hence Y, values. Because spore density or lesions per plant varies greatly with distance from a source (see Fig. 7.1), it is very unlikely that the variance of ε (or Y) will be constant. Rather, $var(Y)$ likely increases with increasing mean Y (or some monotonically increasing function of Y). One approach for incorporating this property of Y into the analysis is to use weighted least squares (or maximum likelihood). Reasonable choices for the weights are $1/Y$ or $1/Y^2$.

It can be shown that if $var(Y)$ is proportional to Y^2, then $\ln(Y)$ has an approximate constant variance (Neter et al., 1983). Given that the linearized forms of the exponential and (modified) power models involve the log of Y, unweighted linear least squares for model fitting can be justified. A similar approach is to use nonlinear methods, but with the model in log-transformed form. For instance, with equation 7.15b (but not 7.15a), one can write:

$$\ln(Y)_j = \ln\left[f(s_j;\Theta)\right] + \ln(\varepsilon)_j \qquad (7.15c)$$

in which $\ln(\varepsilon) = \varepsilon^*$ is a "new" error term, equivalent to the ε in equation 7.14. For the exponential and power models, algebraic manipulation of $\ln[f(\bullet)]$ results in a linear equation. However, equation 7.15c can be used directly with nonlinear parameter estimation methods with any model in which Y has a variance proportional to the square of Y.

In Fig. 7.8 we show how to fit a few versions of the modified power model to the rust data, with the models in the form of equation 7.15c. Many other statistical programs could be used. $\ln(Y)$ is created in the data step ($\ln Y$), and the right-hand side of equation 7.15c is given in the model statement. The error term is implied, and not written. The parameter statement identifies unknown parameters and gives initial guesses for their values. Bounds could have been placed on the parameters, and an initial search of parameter values for starting values in the optimization method could have been specified. Estimates of the parameters and coefficients of determination are given here, with standard errors in parentheses:

Power:	$\hat{a}_P = 260(64)$,	$\hat{b}_P = 1.70(0.12)$,
	$\lambda = 0$,	$R^2 = 0.96$
Modified Power:	$\hat{a}_P = 447(102)$,	$\hat{b}_P = 1.88(0.11)$,
	$\lambda = 0.4$,	$R^2 = 0.97$
Modified Power:	$\hat{a}_P = 1993(2462)$,	$\hat{b}_P = 2.36(0.38)$,
	$\hat{\lambda} = 1.7(1.1)$,	$R^2 = 0.983$.

Note that we use a hat over λ when it is estimated directly from the data (third case) and do not use a hat when a specific value of λ is simply inserted in the model (first two cases here).

The model with λ fixed at either 0 or 0.4 (first two cases) gave the same results as found with linear least squares (see above) because the log transformation results in a linear model. For nonlinear models in general, one

definitely cannot expect to find this correspondence between linear and nonlinear models. Note that $\ln(a_P)$ is estimated with linear models, and a_P is directly estimated here. By allowing λ to take on any value (third case), a considerable improvement in the coefficient of determination occurred, with nearly a perfect fit to the data. The estimate of λ ($\hat{\lambda} = 1.7$) was about equal to the distance from the source to the *second* sampling location. Thus, the concept of λ as a measure of source size probably does not hold in this case. Rather, the role of λ in influencing the shape of the dispersal gradient likely is dominant here. As expected, the estimates of a_P and b_P depended on the value of λ (either fixed or estimated).

The estimate of a_P for the general case (=1993) is so different from the other estimates, that one might question the validity of the estimate. One must recall, however, that a_P is the predicted Y at $s = 1 - \lambda$. With $\lambda = 1.7$, this means that a_P represents Y at a distance of -0.7 m, $Y(-0.7)$, which is *not* a physical location. It is more interesting to determine the predicted Y at $s = 0$ m. This is indeterminate when $\lambda = 0$. In the modified power model, $Y(0)$ is predicted as $\hat{a}_P\hat{\lambda}^{-\hat{b}_P}$. With λ fixed at 0.4, the source strength is determined as $447(0.4)^{-1.88} = 2503$; with λ estimated, the source strength is determined as $1993(1.7)^{-2.36} = 570$, a value smaller than found with the small (fixed) λ value. As discussed previously, the power model becomes more exponential like as λ gets very large, so it is not surprising that the $Y(0)$ predictions becomes smaller here as λ increases. In one sense, $s = 0$ is a hypothetical (theoretical) location for this data set in that the first data point occurs nearly a meter from the inoculum source. For the actual observed data points, all versions of the model fit well.

One interesting result when fitting the modified power model with unknown λ was the large estimated standard errors for estimated a_P and λ. This suggests a fair degree of imprecision in the estimates of these two parameters here. These two estimates also were highly and positively correlated, because of the nature of the model formulation. For instance, as shown in the formula for estimated $Y(0)$ in the previous paragraph, as $\hat{\lambda}$ increases, $\hat{\lambda}^{-\hat{b}_P}$ must decrease (for positive b_P). So, to achieve the same, $\hat{Y}(0)$, $\hat{a}_P$ must increase with increasing $\hat{\lambda}$. Ratkowsky (1990) indicates that in models of the modified power type, a_P and $\hat{\lambda}$ parameters have very nonlinear behavior (see Chapter 3 for some discussion). One way to improve the precision of the estimates, and reduce their bias, is to re-parameterize the model using the so-called "expected value" parameter in place of a_P. The issue is very important for those working on nonlinear modeling of dispersal gradients. The modeling presentation in Chapter 5 of Schabenberger and Pierce (2002) is very helpful in this regard.

Based on use of the nonlinear estimation procedure, other dispersal models described in section 7.3.5 can be fitted. For example, the estimate of the shape parameter in the Lambert et al. model (equation 7.10) was $\hat{\eta} = 0.24$ (s.e. = 0.12), close to, but slightly above the limiting value for a pure power model ($\eta = 0$). This suggests that the power model is more appropriate than the exponential, which agrees with the results already presented. Fitting the joint power-exponential model (equation 7.11a) to the data resulted in: $\hat{b}_1 = 0.05$ (s.e.= 0.034) for the exponential component, and $\hat{b}_2 = 1.31$ (s.e.= 0.271) for the power component. Theoretically, if the gradient was purely of the power type, b_1 would be 0. The estimated b_1 was not significantly different from 0, which indicates the appropriateness of the power model. This can also be tested based on the reduction in residual sum of squares between the fit of the power model and equation 7.11a. However, such a simple interpretation is hindered by the fact that s and $\ln(s)$ (see equation 7.11b) are correlated, and the correlation can be above 0.9 in many cases. Another way of looking at this is that equation 7.11a involves two functions of the *same* variable (s). Because of this, tests involving the parameters can be ambiguous. Ferrandino (1996) developed an ingenious approach to replace s and $\ln(s)$ with two uncorrelated functions, and test whether or not the power and exponential models equally fit the data. Readers should check the original article for details.

It is worth pointing out that equation 7.12 provided the best fit of all the models to the example data in terms of smallest residual sum of squares and largest coefficient of determination. This is noteworthy, because this model has only two parameters, and the next best models had three parameters (modified power with estimated λ). This suggests that more work is needed to assess the appropriateness of the Cauchy-like model for additional dispersal gradients.

7.5 Disease Gradients—Correcting for Maximum Intensity

As discussed at the beginning of the chapter, dispersal of propagules typically is assessed by quantifying deposition of spores (or other infectious units) at various distances from a source, or, more commonly, by measuring disease intensity (such as number of lesions or pustules) at each distance. The latter is relevant when number of spores is directly and linearly proportional to number of lesions, which is reasonable when density of lesions is not too high. But it may be difficult to specify with great precision what is meant by "too high". With increasing spore deposition, there is a decrease in the fraction of spores that cause countable lesions (Vanderplank, 1975). Spores may compete with each other, or spores may land on previously infected sites on plants. As shown with disease progress curves, one must account for the fact that disease intensity has an upper limit in an epidemic.

In this section, we address in detail how to make adjustments in models to account for an upper limit to disease intensity at a given location. The need for such

an adjustment has been known for decades (Gregory, 1973), yet surprisingly, papers are still published in plant pathology where no adjustments are made, leading to some questionable interpretations of results. We start with an *ad hoc* approach to altering the models. Then, we consider a more formal approach, similar to the protocol taken for disease progress, where the exponential model (equation 4.1) is expanded by the incorporation of a disease-free term. Because disease gradients are often determined for more than one time during an epidemic, we show how to quantify and compare gradients obtained at multiple times. This approach is general, and can be used to compare any gradients, not just those corresponding to different times. The developed models form the basis for additional theoretical characterization of disease in time and space.

7.5.1 Simple adjustment

Several decades ago, it was pointed out that a modification must be made to the common dispersal models if disease intensity is used as the dependent variable and intensity is not low (Gregory, 1973). So far in this chapter, we have operated on the assumption that Y in the exponential (equation 7.1), (modified) power (equation 7.4), and similar models is a random variable with 0 as a lower bound and *no* upper bound. However, Y really has M (number of plant individuals or host area that could be diseased) as a physical upper bound, which corresponds to an upper bound for y of 1 ($y = Y/M$). In many cases, the actual upper bound for y, K, may be even less than 1 ($K < 1$ or $KM < M$; see section 4.4.8), but we restrict our attention to the situation with $K = 1$. The simple no-upper-bound assumption, although reasonable if Y or y is not too high, is unrealistic when Y or y is moderate or high. The recommendation of Gregory is to define Z as the multiple infection transformation of disease intensity. If y represents disease on a 0–1 scale, then one replaces Y with $\ln[1/(1-y)]$; that is, one uses $Z = \ln[1/(1-y)] = \ln[M/(M-Y)]$ for Y in the above models. The rationale here is that there is close to unit increase in y or Y with unit increase spore deposition at low disease intensity, but the increase in y or Y with increase in spore deposition becomes less and less at higher disease levels (because spores land on lesions, or spores compete for infection sites). The linear relation between Z and spore deposition would be maintained at all disease densities, in principle. This transformation can be obtained directly from a consideration of the discrete statistical distributions that are discussed in Chapter 9.

This approach results in:

$$\ln\left(\ln\left(\frac{1}{1-y}\right)\right) = \ln(a_E) - b_E s \qquad (7.16a)$$

for the linearized version of the exponential model (compare with equation 7.2), and:

$$\ln\left(\ln\left(\frac{1}{1-y}\right)\right) = \ln(a_P) - b_P \ln(s+\lambda) \qquad (7.16b)$$

for the linearized version of the (modified) power model (compare with equation 7.6). Note that the left-hand side of the equations involves a double-log transformation, which is known as the complementary log-log transformation of y, CLL(y), in statistics. This transformation can also be written as: $\mathrm{CLL}(y) = \ln[-\ln(1-y)]$.

The multiple infection transformation commonly is used to represent disease development over time for monocyclic diseases (Chapter 4). Thus, it *might* appear that equations 7.16a and 7.16b are primarily of use for monocyclic diseases. This multiple-infection-transformation correction is, in fact, used for both monocyclic- and polycyclic-disease gradients. Under some circumstances, the CLL transformation *is* obtained from more mechanistic models of monocyclic disease development in space and time (see Minogue, 1986, and section 8.3.3). However, the reader is reminded that the pure monocyclic growth model results in $\ln[1/(1-y)]$ on the left hand side of the temporal equation (equation 4.11), not CLL(y).

At a minimum, investigators should use CLL(y) rather than $\ln(y)$ or $\ln(Y)$ in linear models whenever they are characterizing disease gradients and y is greater than ~0.2. We show in the next sub-section that possibly more appropriate transformations can be derived for describing gradients based on a more formal analysis of dy/ds.

7.5.2 Generalizations of the exponential and power models

For disease progress curves, one can model disease on an absolute scale, Y, or on a proportion scale, y ($=Y/M$). For dispersal studies, it is more common to work with Y than with y, but this is an arbitrary choice in many cases. In Chapter 4, it was shown that a direct approach to expanding the differential equation model for unlimited population growth was to multiply the level of disease intensity (y) by the level of disease-free intensity ($1-y$). This same idea can be used here to develop models for disease gradients. In the following sub-sections, we apply appropriately expanded models to an actual data set. Because we now focus directly on disease intensity, and not just dispersal processes, we explicitly call the following equations disease gradient models.

One can divide both sides of the dY/ds equations for the exponential and (modified) power models (equations 7.1 and 7.7) by M to obtain:

$$\frac{dy}{ds} = -b_E y \qquad (7.17a)$$

and

$$\frac{dy}{ds} = \frac{-b_P y}{s} \tag{7.17b}$$

for the rate of change in y with change in s. Just to simplify presentation, we assume that $\lambda = 0$ for the modified power model for the rest of the section. Based on the arguments made in Chapter 4, equations 7.17a and 7.17b can produce totally unrealistic values for dy/ds at large y, leading to possibly impossible predictions of y (e.g., $y > 1$) at small s. Multiplying the right-hand side of these rate equations by $1 - y$ can result in more realistic and interpretable representation of spread over the entire range of disease intensities. As in section 7.5.1, we assume that $K = 1$. The exponential model for spread becomes:

$$\frac{dy}{ds} = -b_L y(1-y) \tag{7.18}$$

after multiplication by the proportion of host that is disease free. This is, actually, a logistic model for a disease gradient. We use b_L for the gradient parameter (units of 1/distance) to be a reminder that this parameter is not quite the same as b_E (equation 7.17a); at low y, however, equations 7.17a and 7.18 are indistinguishable, in general (with $b_L \approx b_E$). Integration leads to:

$$y = \frac{1}{1 + A_L \exp(b_L s)} \tag{7.19}$$

in which A_L is the constant of integration, equaling $(1 - y_0)/y_0$, where y_0 is (theoretical) disease intensity at the inoculum source (i.e., at $s = 0$). Readers should note that the exponent in the denominator is positive, which is different from the temporal version of the logistic model (equation 4.15b). This is because of the negative sign in the differential equation (equation 7.18). An equivalent way to write equation 7.19 is:

$$y = \frac{1}{1 + \exp(-(\ln[y_0/(1-y_0)] - b_L s))}.$$

This can be compared with equation 4.15a for temporal disease progress, but now there is a negative sign before the spread parameter.

Algebraic manipulation of equation 7.19 results in the following linear equation:

$$\ln\left(\frac{y}{1-y}\right) = -\ln(A_L) - b_L s \tag{7.20}$$

Because $-\ln(A_L) = -\ln[(1-y_0)/y_0] = +\ln[(y_0/(1-y_0)]$, equation 7.20 can be written as:

$$\ln\left(\frac{y}{1-y}\right) = \ln\left(\frac{y_0}{1-y_0}\right) - b_L s.$$

which is an equation for a straight line with slope $-b_L$ and intercept given by the logit of disease intensity at the source (see equation 4.16 for comparison).

Example disease gradients corresponding to equation 7.19 are shown in Fig. 7.9 for selected values of y_0 (or A_L) and a single b_L value. These curves represent five different times in an epidemic; the choice of A_L values appears arbitrary here, but is explained in a later subsection. The curves are mirror images of logistic disease progress curves (because of the negative sign before b_L), with the highest y values at the smallest s values. Unlike the situation for no upper limit to the response variable, the curves level off at high y (or low s). For cases where y is not too high (e.g., <20%), the curves look the same as for the exponential dispersal model (equation 7.2).

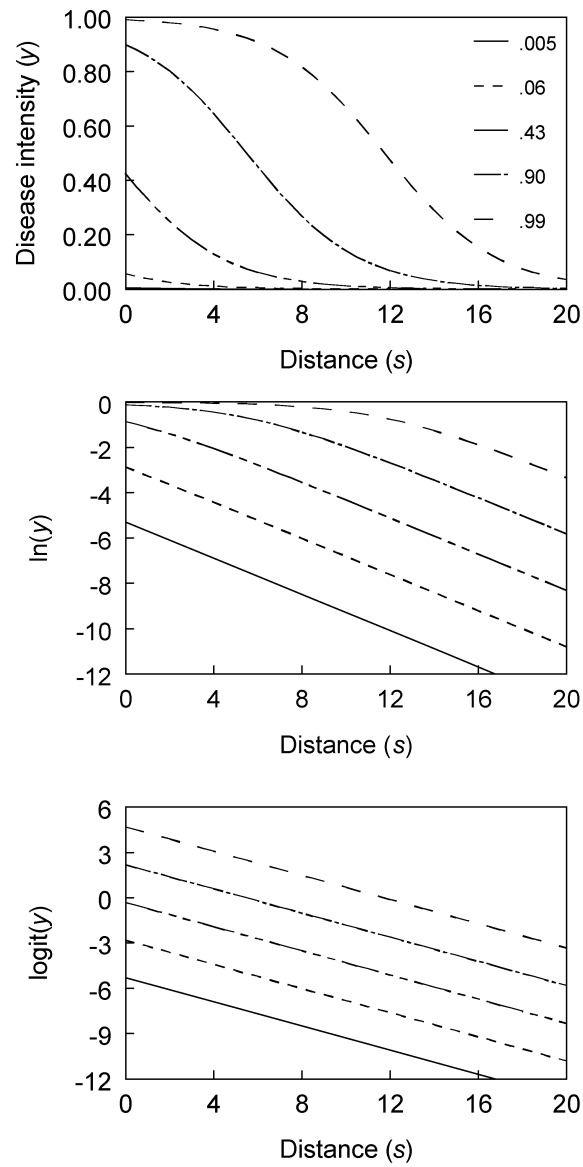

FIG. 7.9. Disease gradients according to the logistic model (equation 7.19), for five different levels of disease intensity at $s = 0$ (y_0; labeled in upper frame), all with the same spread parameter, b_L. Gradients shown on a linear and two transformed scales.

This is because the exponential is a good approximation of the logistic model at low y.

If one plotted the gradients on a log-scale (middle frame of Fig. 7.9), it appears that the gradients flatten over time, especially at distances relatively near the source ($s < 12$, here), even though a single spread parameter value was used. The idea of dispersal gradients flattening over time with polycyclic diseases has a long history in plant pathology. Gregory (1968) indicated that secondary spread flattens a primary gradient, which makes sense when untransformed disease intensity is used as the measure of infections (see top frame of Fig. 7.9). With an upper limit to disease intensity, the rate of change of y with change in s must slow down (with all other things being equal) as y increases, because there are fewer opportunities for infections when most of the plant host is already infected. However, Zadoks and Kampmeijer (1977), van den Bosch et al. (1988a–c), and Minogue and Fry (1983b), among others, posited that infection gradients do not necessarily flatten over time during an epidemic. We only point out a few issues here related to gradient flattening.

To determine whether or not there is flattening of the infection gradient, one must use the appropriate transformation of y (or a model that makes the proper adjustment for maximum y). If equation 7.18 is appropriate, then one must plot logits versus distance, as shown in the bottom frame of Fig. 7.9. Here, the correction for maximum disease is accounted for graphically, and one can see that the gradients are stable in the sense that there is no change in the slope for the different times. Because logit(y) is very similar to ln(y) at low y, one can also see the stable gradients at low y in the ln(y) : s graph in Fig. 7.9 if one just looks at the right-hand side of the graph. Of course, the steepness of the gradients does not change here precisely because equation 7.19 was used with a *fixed* (single) value of b_L at all values of y_0. If the gradients were becoming less steep over time, then b_L of equation 7.19 (or equation 7.20) would be getting smaller (closer to 0) with increasing time (if equation 7.19 was appropriate at *all* times). Tests of b_L, therefore, become a basis for evaluations of gradients over time.

The power model can also be easily generalized to accommodate a limitation to disease. The model can be written in differential-equation, integrated, and linearized forms as:

$$\frac{dy}{ds} = \frac{-b_{PL}y(1-y)}{s} \tag{7.21}$$

$$y = \frac{1}{1 + A_{PL}s^{b_{PL}}} \tag{7.22}$$

$$\ln\left(\frac{y}{1-y}\right) = -\ln(A_{PL}) - b_{PL}\ln(s) \tag{7.23}$$

in which A_{PL} equals $(1 - y_1)/y_1$, where y_1 is (theoretical) disease intensity at a distance of 1 [$=y(1)$]. The intercept term in equation 7.23 is equal to $+\ln[y_1/(1 - y_1)]$, the logit of disease at $s = 1$. Note that we label the gradient parameter as b_{PL} (for "power-logistic"; unitless) because the numerator in equation 7.21 is of the logistic form, but there is also a denominator of s, reflecting a power-type gradient. The model can also be referred to as the *log-logistic*. The link between this model and the logistic disease gradient model may be seen more clearly by an alternative expression for y of equation 7.22:

$$y = \frac{1}{1 + \exp(-(\ln[y_1/(1-y_1)] - b_{PL}\ln(s)))}$$

The log-logistic model, with a different parameterization, is used in section 3.5 of Chapter 3.

Disease gradients described by equation 7.22 are shown in Fig. 7.10, corresponding to five different values of y_1 (or A_{PL}), for a fixed (single) value of b_{PL}. Because of the nature of power-model-type gradients, in which dy/ds is a large negative number at small s, the decline in y with increasing s is relatively large near the source even when overall disease intensity is high (e.g., top curve in Fig. 7.10). Nevertheless, the curves show some appearance of becoming flatter near the source as y increases. This is expected since there is an upper limit to disease (=1) in this model. The flattening of the gradient is more pronounced when one plots ln(y) versus ln(s) (middle frame of Fig. 7.10), which is the proper scaling for the simpler power model (equation 7.6), but not the proper scaling when a correction for maximum disease is made. When one looks just at the right hand side of this graph, however, corresponding to relatively low y, one sees parallel lines (i.e., no flattening). When logit(y) is plotted against ln(s) (bottom frame of Fig. 7.10), which gives the proper scaling for the model introduced here (equation 7.23), lines with the same slope are found for each time, an indication of stable gradients. As with the logistic model above, the gradients *must* be stable in this theoretical example because they were generated with a single spread parameter in the equation. One can ask the question for observed gradients of disease: does steepness change (as measured by b_{PL}) over time? In other words, are gradients stable? As discussed in the next paragraph, the answer is more complicated than for an exponential-type gradient.

The discussion of the power model (equation 7.7) in terms of the concept of gradient steepness also applies to any model in which distance is in the denominator of the dy/ds equation. In short, there is always some ambiguity of the meaning of the equality (or inequality) of two gradients in which dy/ds (or dY/ds) involves both y and s. For the simpler situation of the exponential or logistic models of spread, dy/ds only involves y and a parameter. Then, if two gradients have the same b_L, steepness (as defined here

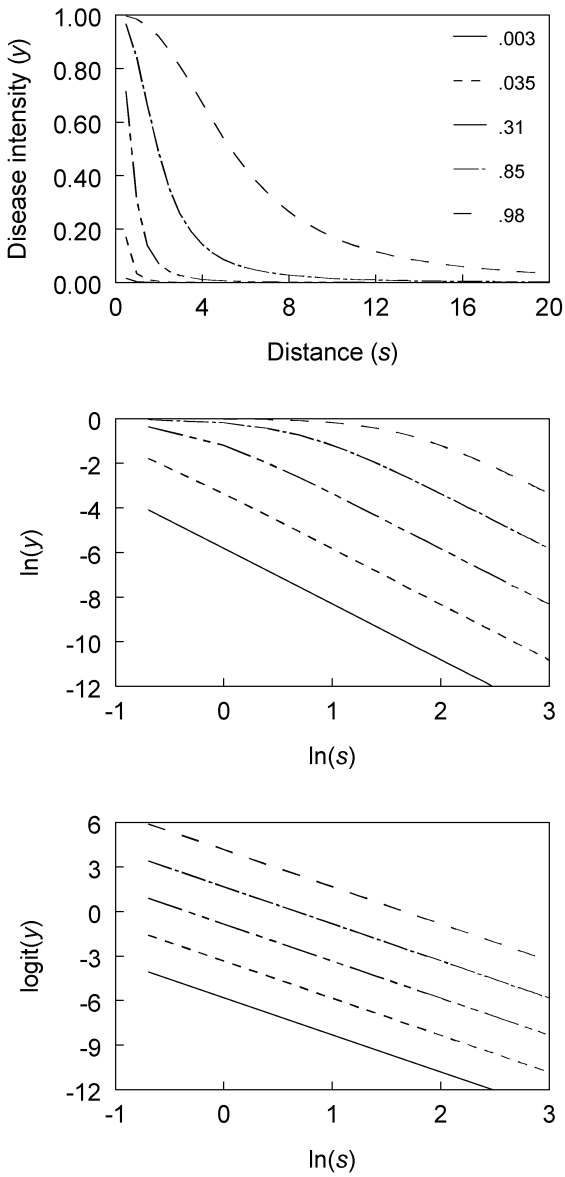

FIG. 7.10. Disease gradients according to the power-logistic model (equation 7.22), for five different levels of disease intensity at $s = 1$ (y_1; labeled in upper frame), all with the same spread parameter, b_{PL}. Gradients shown on a linear and two transformed scales.

by dy/ds) is the same at a given value of y. However, for equation 7.21, if two gradients have the same b_{PL}, steepness is not necessarily the same at a given y, because steepness depends also on distance from the source. Comparing estimates of b_{PL} is, nevertheless, common and useful, but one should be cautious in not over-interpreting the meaning of the parameter differences.

7.5.3 Other models

In addition to the models in the above sub-section, various other generalizations and expansions of the dispersal and disease gradient models can be utilized. Just as the logistic model can be modified to account for asymmetry in dy/dt over time (see section 4.4.4), one can use functions of $1 - y$ [e.g., $\ln(1) - \ln(y)$ for the Gompertz model (see equation 4.19)] rather than $1 - y$ directly in the equations for dy/ds. This approach is occasionally used in practice (e.g., Maffia and Berger, 1998). One could also expand the other gradient models in section 7.3.5 to adjust for maximum disease. We do not pursue these approaches here, although they can be valuable when one has the goal of a very accurate description of observed gradients.

One can also consider simplifications of the models presented for disease gradients, at least for comparison purposes. For instance, suppose that dy/ds is not affected by y, but there is still a maximum possible disease intensity equal to 1. If distance does not affect the gradients directly, then one obtains:

$$\frac{dy}{ds} = -b_M(1 - y) \tag{7.24}$$

in which b_M is now the spread parameter (units of 1/distance). This is analogous to the monomolecular model (equation 4.8) for temporal disease progress, and we use the M subscript for this reason. If distance does affect the gradient, then one obtains:

$$\frac{dy}{ds} = \frac{-b_{PM}(1 - y)}{s} \tag{7.25}$$

in which the spread parameter (b_{PM}) is unitless. These and other models were developed and considered by Jeger (1983). They have occasionally been evaluated (e.g., Jeger et al., 1983; Madden et al., 1990a). Solutions are easy to obtain, as either nonlinear of linearized equations.

7.5.4 Model fitting

7.5.4.1 General comments. As with the simpler exponential and power models, one can use various model fitting methods utilizing either the nonlinear or linear versions of the disease gradient models discussed here.

One issue that needs some attention, however, concerns the variance of the response (dependent) variable in the models. For a discrete count variable with no upper limit, $var(Y)$ can be expected to increase with increasing Y. We pointed out that the transformation used to linearize the nonlinear models for Y, $\ln(Y)$, also does a good job of stabilizing variances. With disease intensity on a proportion (0–1) scale, $var(y)$ is a more complex function of y. As discussed in detail later in Chapter 9, disease incidence most likely has a binomial or more general discrete distribution. If so, $var(y)$ is proportional to $y(1 - y)$. In fitting the nonlinear models (with y as the dependent variable) to data, one could then use weights equal to $1/[y(1 - y)]$.

When fitting the linearized versions of these models to data, the variance of the transformed y, y^*, is the relevant quantity. In section 4.6.2, we showed how to estimate the

variance of y^*, $var(y^*)$, when $var(y)$ was known (or assumed). For the logit transformation, $y^* = \ln[y/(1-y)]$, so $var(y^*) = [y(1-y)]^{-1}$. Thus, to fit either equation 7.20 or 7.23 to data (with an assumed additive error term, ε), one could use $y(1-y)$ as weights.

When dealing with disease severity data, it is not clear which statistical distribution is most appropriate. However, one can generally assume that $var(y)$ is still proportional to $y(1-y)$, and use the same weights as discussed for disease incidence.

For either disease incidence or severity, there may or may not be statistical evidence from the model fit that weights are required if there is one observation per distance. A residual plot can certainly provide strong evidence that variances are unequal, and hence the need for weights. However, with a limited number of data points, the residual plot may be ambiguous. Then the decision regarding weights is more subjective. One could argue that weights are always necessary for a variable that ranges between 0 and 1, especially if there are observed values close to either limit. On the other hand, if the empirical evidence from model-fit diagnostics (e.g., residual plot) does not give clear evidence of unequal variances, then there is not direct evidence to use any weights. In the latter situation, it could be argued that weighting is counter-productive, because one is obtaining a poorer fit (in terms of residual sum of squares) than without weights. In this chapter we use weights (or appropriate variance-stabilizing transformation in previous sections), but we do not take an inflexible view on this issue, and accept the merit of both arguments in cases where there is no clear evidence of unequal variances for the data being analyzed.

7.5.4.2 Example—graphical evaluation. We consider another bean rust example here, based on one of the data sets described in Maffia and Berger (1998). A line source of infected plants was used, and data originated in the spring-planting. Disease intensity was measured in various ways, but we only discuss results for leaf disease incidence, that is, the proportion of leaves with pustules. Results were obtained for six times, $t = 1$, 8, 15, 22, 29, and 36 days, in which $t = 1$ corresponded to the first day with observed pustules.

This example is typical of many in which gradients are observed at several times in the epidemic. For the first and second assessment times, there was a fairly steep decline in y with increasing s near the source (Fig. 7.11). At distances of 5 or more meters, incidence was very low. By later times, there was a considerable flattening of the gradients on the original scale (y), and y was close to 1 near the original inoculum source at the later times. Moreover, at these later times, y was still very high at the largest distance considered (8.5 m). Ideally, one would want to use large enough distances so that the "full" gradient could be observed at each time, but this is totally impractical in most field studies. In this case, one would need to use at least a four-fold increase in field size to observe low values of y at large distances from the source.

When one ignores the fact that there is an upper limit to y, and plots $\ln(y)$ versus s or $\ln(s)$, one sees a flattening of the disease gradients with time, especially at $s < 5$ m. Also note that incidence is so low for the first time that it is difficult to compare these results for large distances with results for the other times. We, thus, (mostly) focus on $t \geq 8$.

Considering that the upper limit of y (=1) is approached here, it is appropriate to utilize the logistic and power-logistic gradient models (equations 7.20 and 7.23) to represent these data. Plots of logit(y) versus s and $\ln(s)$ are shown on the right-hand side of Fig. 7.11. In general, a clear choice between the two models is not possible. There is slight curvature when logit(y) is plotted versus s, with decreasing slope at increasing distance. This is indicative of a power-type model. The plot of logit(y) versus $\ln(s)$ does produce a straight line, except for large distances (high s), where a slight downward bending of lines could be seen. Residual plots from linear regression analysis show some pattern with both model choices (not shown). Because of the straight lines obtained for logit(y) versus $\ln(s)$ at $s < 6$ m [$\ln(s) < 1.6$], we select the power-logistic as the primary model to quantify these gradients. This is consistent with other studies that found a power-law type relation for dispersal gradients of rust spores (Aylor and Ferrandino, 1989; Mundt, 1989). However, data are insufficient in this example to interpret the gradients in terms of dispersal mechanisms. This would require other data sets and larger fields (larger maximum s) in the studies. In particular, for the latter issue, it would be informative to know if that slight curvature of the logit line at $\ln(s) > 1.6$ (Fig. 7.11) would continue at much larger distances.

7.5.4.3 Example—linear regression. For the interested reader, the actual data points are given in Fig. 7.12 in the form of a SAS program for regression analyses based on equations 7.20 and 7.23. Before showing results for the logistic power model, we present some results that are obtained if one (incorrectly) ignored the fact that y is limited from above by 1. That is, we fitted the regular power model (equation 7.6, with y, rather than Y, used for disease intensity) to the data and compared slopes (estimates of b_L) among times. This analysis is not specified in the SAS program. Estimated slopes were 0.69, 0.41, 0.20, 0.12, 0.10, and 0.07 for times 1 through 36, based on weighted linear regression of $\ln(y)$ on $\ln(s)$, with weight equal to $y/(1-y)$ [see section 4.6.2 for variance of $\ln(y)$]. Results indicate a strong tendency for decreasing slope over time. In fact, the slope for day 8 was about six times larger than the slope for day 36 (0.41/0.07).

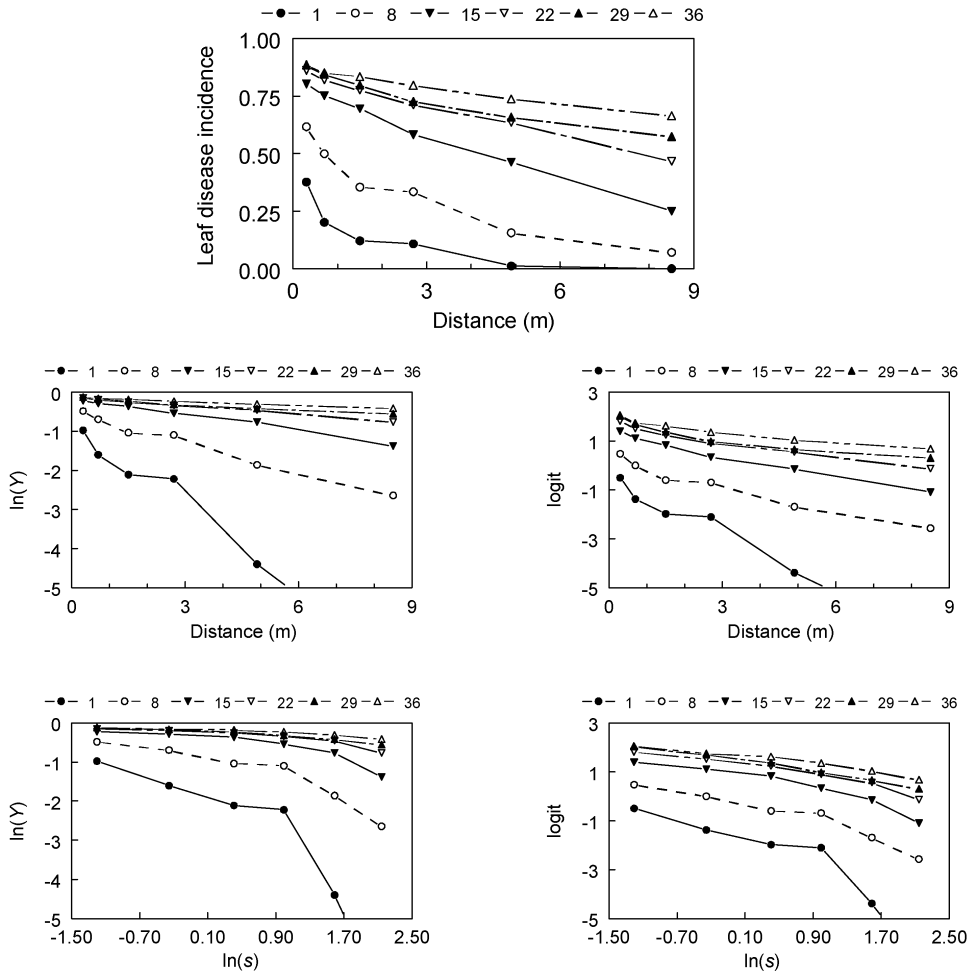

FIG. 7.11. Disease gradients at six assessment times for bean rust (Maffia and Berger, 1998). Observations of leaf disease incidence (proportion of leaves with pustules) in relation to distance from a line source were taken for the Fall planting of the study. Data shown on a linear and various transformed scales.

For later purposes, it is helpful to indicate explicitly that the statistical form of power-logistic model (equation 7.23) can be written as:

$$\ln\left(\frac{y}{1-y}\right)_j = -\ln(A_{\mathrm{PL}}) - b_{\mathrm{PL}}\ln(s)_j + \varepsilon_j \qquad (7.26a)$$

or

$$y_j^* = A_{\mathrm{PL}}^* - b_{\mathrm{PL}}\ln(s)_j + \varepsilon_j \qquad (7.26b)$$

in which the j subscript refers to the j-th distance (e.g., the first distance is $s_1 = 0.3$ m), y^* is the symbol for the transformation that produces a linear model (logit here), and $A_{\mathrm{PL}}^* = -\ln(A_{\mathrm{PL}}) = \ln[y_1/(1-y_1)]$ for the power-logistic model. Fitting equation 7.26b to the data for each time, using weighted least squares [with $y(1-y)$ as weight], produces the results in Table 7.1.

The error degrees of freedom (df) equals 4 for each time. As shown in the top frame of Fig. 7.13, the fitted lines are very reasonable summaries of the gradients on a logit:ln(s) scale. The $\hat{A}_{\mathrm{PL}}^*$ values increased with time, as expected, since y increased over time. Back-transformation of estimated A_{PL}^* can be used to estimate y at 1 m. For instance, at $t = 29$ days, $\hat{y}_1 = 1/\left(1+\exp(-A_{\mathrm{PL}}^*)\right) = 0.81$. The estimates of b_{PL} do become smaller over time, but the trend is much weaker than when the simpler power model (with no correction for maximum disease) was used. Although the R^2 values were high, there was some pattern to the residual plot, indicating that the power-logistic model was not ideal. For comparison, some results from fitting the statistical version of the logistic gradient model (equation 7.20) are also shown in Table 7.1. The estimates of b_{L} declined faster with time than the estimates of b_{PL}. This and other models resulted in undesirable residual plots, although the R^2 was very high for some assessment times with the logistic model. As discussed above, we concluded that the power-logistic model was a reasonable (but not ideal) choice for analysis.

7.5.4.4 Example—comparing parameter estimates. When one wants to compare two gradients, one can test the equality of the spread parameters or the intercepts (transformation of the constant of integration) using

```
data rust;
                    input t s y @@; * several grous of t-s-y per record;
                    logit=log(y/(1-y)); lns=log(s);
                    wt = y*(1-y);  * weight for linear regression of logits;
                    datalines;
1 0.3 0.377   8 0.3 0.616   15 0.3 0.802   22 0.3 0.860   29 0.3 0.883   36 0.3 0.886
1 0.7 0.202   8 0.7 0.500   15 0.7 0.751   22 0.7 0.820   29 0.7 0.845   36 0.7 0.851
1 1.5 0.122   8 1.5 0.355   15 1.5 0.695   22 1.5 0.775   29 1.5 0.798   36 1.5 0.835
1 2.7 0.109   8 2.7 0.334   15 2.7 0.582   22 2.7 0.711   29 2.7 0.726   36 2.7 0.796
1 4.9 0.012   8 4.9 0.156   15 4.9 0.463   22 4.9 0.635   29 4.9 0.658   36 4.9 0.738
1 8.5 0.0009  8 8.5 0.071   15 8.5 0.251   22 8.5 0.467   29 8.5 0.574   36 8.5 0.664
;
proc sort data=rust out=rust; by t;
proc reg data=rust;
                    title 'Bean rust, leaf disease incidence; logit; WEIGHT; separate times';
                    by t;  /*separate weighted least squares regression analysis for each time*/
                    weight wt;
                    model logit = s; plot r.*s;  *plot of residual (r.) versus distance (s);
                    model logit = lns; plot r.*lns;
proc mixed data=rust;
                    title 'Bean rust, leaf disease incidence; logit; WEIGHT; COVARIANCE';
                    weight wt;
                    class t; /*covariance analysis, with separate intercepts & slopes*/
                    model logit = lns t lns*t / solution ddfm=satterth;
proc reg data=rust;
                    title 'Bean rust, leaf disease incidence; logit; WEIGHT; SPATIO-TEMPORAL';
                    weight wt;
                    model logit = lns t; plot r.*p.; *residuals vs. predicted;
run;
```

FIG. 7.12. Input program of SAS to read in data shown in Fig. 7.11 and fit power-logistic (equation 7.23) models to the data using linear (weighted) least squares. Program also performs a covariance analysis to determine the effect of time on gradients, and fits the power-logistic spatio-temporal model (equation 7.36) to the data.

TABLE 7.1. Fit of the logistic (equation 7.20) and power-logistic (equation 7.23) models to the bean rust data shown in Fig. 7.12, together with the estimated distance ($\hat{s}'$) where disease intensity equals 0.6, for the different disease assessment times.

	Power logistic				*Logistic*			
t	$\hat{A}^*_{PL}$	$\hat{b}_{PL}$	R^2	$\hat{s}'$	$\hat{A}^*_{L}$	$\hat{b}_{L}$	R^2	$\hat{s}'$
1	−1.6 (0.15)	0.88 (0.17)	0.88	–	−0.5 (0.23)	0.72 (0.15)	0.85	–
8	−0.3 (0.12)	0.77 (0.12)	0.91	0.4	0.3 (0.15)	0.37 (0.05)	0.92	–
15	0.8 (0.15)	0.71 (0.12)	0.90	1.7	1.3 (0.09)	0.29 (0.02)	0.98	3.1
22	1.3 (0.11)	0.58 (0.07)	0.93	4.7	1.6 (0.09)	0.21 (0.02)	0.97	5.7
29	1.5 (0.04)	0.53 (0.03)	0.99	7.9	1.7 (0.16)	0.18 (0.03)	0.89	7.2
36	1.7 (0.06)	0.41 (0.12)	0.96	23.5	1.9 (0.09)	0.15 (0.02)	0.94	10.0

various methods. The gradients may be from disease assessments made at different times in a single epidemic (as here), or correspond to contemporaneous assessments of disease for different treatments (e.g., cultivars, fungicide applications). The approaches discussed in Chapter 4 are applicable here, and we directly address only a limited number of approaches.

The simplest approach is to use a *t*-test to compare the parameter estimates. For instance, if one had two gradients, each characterized by the power-logistic model (equation 7.26a), one can test the null hypothesis:

$$H_0: b_{PL1} = b_{PL2}, \text{ or } H_0: b_{PL1} - b_{PL2} = 0$$

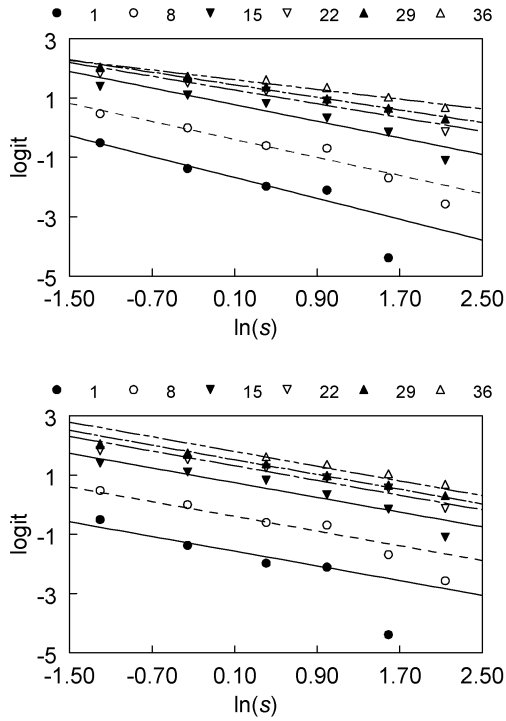

FIG. 7.13. Fit of the power-logistic model (equation 7.23) to the bean rust gradient data (see Fig. 7.11) for each of six assessment times. Upper frame: a separate intercept and slope for each time. Lower frame: fixed slope for all times.

versus the alternative hypothesis that the parameters are not equal. The *t*-statistic can be written as:

$$t = \frac{\hat{b}_{\mathrm{PL1}} - \hat{b}_{\mathrm{PL2}} - 0}{s(\hat{b}_{\mathrm{PL1}} - \hat{b}_{\mathrm{PL2}})}$$

in which the denominator is the standard error of the difference of the two estimated slopes. This can be estimated (for uncorrelated slopes) as:

$$s(\hat{b}_{\mathrm{PL1}} - \hat{b}_{\mathrm{PL2}}) = \left(s^2(\hat{b}_{\mathrm{PL1}}) + (s^2(\hat{b}_{\mathrm{PL2}})\right)^{1/2}$$

in which the two terms on the right hand side are the squares of the standard errors of the estimated slopes. For the comparisons of estimated slopes for days 8 and 15 of the bean rust example, the *t*-statistic is given by:

$$t = \frac{0.77 - 0.71}{(0.12^2 + 0.12^2)^{1/2}} = \frac{0.06}{0.17} = 0.35.$$

The degrees of freedom for the *t*-test is equal to the sum of the error degrees of freedom for the two disease progress curves ($df = 4 + 4 = 8$). This achieved *t*-statistic is much below the critical value for 8 *df* (2.31) to be significant at $\alpha = 0.05$. Thus, the slopes are not significantly different (at $\alpha = 0.05$) for these two gradients.

As an aside, readers should note that *s* is used for both the distance variable and for an estimated standard error or standard deviation, and that *t* is used for the time variable as well as the *t*-statistic. Other than using unusual symbols for time and space variables (or for standard errors and *t*-statistics), this double interpretation cannot be avoided. The meaning of the symbols should be apparent based on context of the material.

The two intercepts can be compared in the same way as the comparisons of the slopes. Then, other pairs of gradients can be compared. The reader should be able to determine that the intercepts clearly are different between many times (as expected), and that some slopes are also different from each other. This approach is very useful when there are only a small number of gradients, or the investigator only has information on parameter estimates and standard errors (say, from a published paper). With multiple gradients, it may be difficult to gain a general understanding of disease spread based on pairwise comparisons. Thus, it is generally more efficient to use a covariance analysis to test whether the heights of the lines (intercepts) or their slopes vary among groups (Milliken and Johnson, 2002). This approach also allows the investigator to incorporate features of the experimental design, such as blocking, into the analysis.

Assuming that the power-logistic model is appropriate, we can expand equation 7.26b to:

$$y^*_{ij} = A^*_{\mathrm{PL}i} - b_{\mathrm{PL}i} \ln(s)_{ij} + \varepsilon_{ij} \qquad (7.27a)$$

in which the *i* subscript represents the *i*-th group (time in the example). There is an *i* subscript on ln(*s*) (in addition to the *j* subscript) because it is possible that the distances are not the same for each group (although this is not the case here). Thus, the potential for each group to have a different intercept (A^*) and slope (b_{PL}) is incorporated into this statistical model. An alternative way of writing this equation is:

$$y^*_{ij} = \gamma + A^*_{\mathrm{PL}i} - \iota \ln(s)_{ij} - b_{\mathrm{PL}i} \ln(s)_{ij} + \varepsilon_{ij} \qquad (7.27b)$$

in which γ and ι represent overall intercepts and slopes, respectively. In this formulation, $\gamma + A^*_i$ is the intercept for the *i*-th group, and $\iota + b_{\mathrm{PL}i}$ is the slope for the *i*-th group. Readers should compare this with equation 4.38b, and reread the discussion about use of this type of equation. The formulation in equation 7.27b is most useful in a covariance analysis for testing whether or not there is a significant effect of group on the intercept and slope. This model is fitted to the bean rust data, with time as the group, using the MIXED procedure of SAS in Fig. 7.12. Note that time (*t*) is given as a class (category or factor) variable in the program. An overall intercept (γ) is implied by the model statement and is not identified; *t* specifies the effect of time on the intercept [a "main" effect term ($A^*_{\mathrm{PL}i}$)]. The overall slope and the group effect on the slope are specified with lns (ι) and lns**t* ($b_{\mathrm{PL}i}$), respectively. That is,

in the latter case, the meaning of a group-dependent slope is that the slopes are not the same (interaction of group and the continuous variable).

Fitting this model to the data showed the following: a highly significant effect of distance ($F = 246.6$; $df = 1, 24$; $P < 0.001$), meaning that ι was different from 0; a highly significant effect of time ($F = 103.0$; $df = 5, 24$; $P < 0.001$), indicating that (A^*_{PLi}) varied with time; and a marginal effect of the interaction of time and distance ($F = 2.6$; $df = 5, 24$; $P = 0.05$), meaning some evidence that b_{PLi} was not the same for all times. In this model formulation, the actual estimates of the intercepts and slopes for each time can be obtained by the estimates of γ, ι, and the group-specific parameter estimates (using the solution option). For instance, the slope for group three (third time) would be $\hat{\iota} + \hat{b}_{PL3}$ and the intercept would be $\hat{\gamma} + \hat{A}^*_{PL3}$. A more direct way of obtaining these estimates would be to use the following model statement:

model logit = *t t* *ln *s*/noint solution;

where noint option eliminates γ from the model ("no intercept"). Because there is also no individual ln*s* term specified (thus, no ι), one is essentially fitting equation 7.27a to the data. The *F* tests displayed for group and the interaction of group and the continuous variable, however, are no longer applicable for testing group and interaction effects when these changes to the computer code are made (even though the estimates of the slopes and intercepts, and their standard errors are appropriate). Using this approach, as an example, one obtains the following parameter estimates for 15 days:

$$t = 15: \quad \hat{A}^*_{PL3} = 0.85(0.11), \quad \hat{b}_{PL3} = 0.71(0.08)$$

which agrees with the individual regression results shown above. The agreement does not have to be this good, since the covariance analysis, as done here, is assuming a single error variance for all the data, not a separate one for each gradient (each assessment time, in this case).

Since there is only marginal effects of group on the slope (marginal interaction), it is natural to consider fitting a covariance model with a single (group independent) slope. Equation 7.27a can be simplified to:

$$y^*_{ij} = \gamma + A^*_{PLi} - \iota \ln(s)_{ij} + \varepsilon_{ij} \qquad (7.27c)$$

Note this formulation indicates that $b_{PLi} = 0$ for all i (all groups). In other words, ι now represents a single slope (i.e., $\iota \equiv b_{PL}$, with no i subscript). This model can be fitted using the following code:

model logit = *t* ln *s*/solution;

With this analysis, the (single) estimated slope was: $\hat{b}_{PL} = 0.62$ (s.e.= 0.043), intermediate among the individual slopes for the different times. Estimates of the different intercepts can be obtained from $\hat{\gamma} + \hat{A}^*_{PLi}$. By giving the noint option as well, one directly obtains $\hat{A}^*_{PLi}$ (because $\gamma = 0$). The estimated intercepts for the six times were: −1.5, −0.3, 0.8, 1.4, 1.6, and 1.8, similar to those obtained when analyzing data for each time separately. The residual variance for this model was 0.0125, which is not a lot larger than the residual variance found for the fit of equation 7.27b (0.0098), indicating that a reasonable fit of this multi-time data set is obtainable with a single spread parameter. As shown in the bottom frame of Fig. 7.13, a single slope is reasonable for describing these data, although the fit is slightly better with time-dependent slopes (top frame).

To conclude this section, analysis of the bean rust gradients shows that logistic and power logistic models could be used, but latter was chosen based on common use of a power-type model for rust diseases. On the original y scale, there was clear flattening of the gradients with increasing time, especially near the inoculum source. There also was some evidence that the gradients flattened over time on a logit(y) : ln(s) scale, but a constant gradient steepness (b_{PL}) over time was a reasonable approximation for the data. Through the use of the logit transformation (in this case), the limitation to disease intensity (i.e., $y < 1$) was accounted for, so that the relationship between (transformed) disease intensity and (transformed) distance could be characterized.

7.6 Spatio-Temporal Dynamics of Disease Spread

7.6.1 General comments

By characterizing disease gradients at several times in an epidemic in the previous section, we have been casually considering the spatio-temporal dynamics of disease development in host populations. That is, we considered how y changed with s, in the form of disease gradients, and how y changed with t, through the change in gradient height and (potentially) gradient steepness with increasing time. Here we start to formalize the relationship between y and both t and s.

As mentioned in section 7.1, in previous chapters we mostly considered a plant disease epidemic as a temporal process, characterized by dy/dt, even though we recognized that space was important. With monocyclic diseases, for instance, we indicated that $dy/dt = r_M(1 - y)$ (equation 4.8), where r_M equals the product of inoculum density or abundance (x'), and the probability per unit time that a unit of inoculum causes an infection (ϕ = the probability per unit time that a unit of inoculum contacts a host individual multiplied by the probability that the unit of inoculum in contact with a disease-free host actually infects the host). If x' is clustered at a single

location (line, point, etc.), then ϕ is unlikely to be constant, but rather would be function of distance from x', $\phi(s)$. This function, $\phi(s)$, is actually analogous to a contact distribution (see section 7.3.4). For polycyclic diseases, the situation is (considerably) more complicated because any infected individual can then serve as an inoculum source, at least during the diseased individual's infectious period (see Chapter 5). Instead of a constant probability of any produced spore coming into contact with a host individual (θ; see sections 5.2.2 and 5.2.3), a spore produced by an infected individual at location ξ has a much greater chance of contacting a host individual at location s if the locations are close together ($s - \xi$ is small) than if the locations are far apart ($s - \xi$ is large). In other words, θ is actually a function of $s - \xi$, $\theta(s - \xi)$. Thus, the logistic model (equation 4.14) and the more complex models of Chapter 5 for y (e.g., equation 5.30) are inadequate for representing epidemics in space and time when the epidemics start from a well-defined inoculum source (or sources). Although the temporal models of the previous chapters were quite useful for describing and understanding epidemics, the next level of quantification requires a formal consideration of space and time in disease development. The approach presented in this section is the first step towards developing a better understanding of the spatio-temporal processes within an epidemic. Through the models developed, the concept of disease spread hopefully becomes clearer, and empirical approaches for characterizing spread are obtained that are natural consequences of the developed models. However, the specific models shown here should be considered a stepping stone towards the fuller characterization of spread that is presented in Chapter 8.

7.6.2 Two spatio-temporal models

Previously, we sometimes found it useful to use the notation $y(t)$ instead of just y to represent disease intensity at time t. With the simultaneous consideration of space and time, we can use the notation $y(t, s)$ to refer to disease intensity at location s and time t. Representation in two spatial dimensions can be done by replacing s with s_1 and s_2 (for instance) to represent row and column locations, $y(t, s_1, s_2)$. The latter approach is not considered here. A general formulation to represent $y(t, s)$ can be given as:

$$y(t, s) = f(t, s; \Theta). \tag{7.28a}$$

where Θ is a vector of parameters. It is quite possible that $f(\bullet)$ is not a simple (or even complicated) single equation, but can only be written as an integral

$$f(t, s; \Theta) = \iint \Xi \, d\hat{s} d\hat{t} \tag{7.28b}$$

in which Ξ is an arbitrary general mathematical expression, $\hat{s}$ and $\hat{t}$ are specific locations (between 0 and s) and times (between 0 and t), respectively, and the integrations are over space and time. In this situation, numerical solutions would be required, although approximate analytical solutions may exist for certain circumstances (see Chapter 8). Here we consider situations where a solution to the integration exists, so that $f(\bullet)$ is a single equation.

Disease intensity is now a function of two continuous independent variables, s and t. Mathematically, partial derivatives ($\partial\bullet$) are used instead of ordinary derivatives to represent rates of change in models with more than one independent variable. One can take the partial derivatives of equation 7.28a to determine rates of change of y with respect to time and space. In particular, we write $\partial y(t, s)/\partial t$ for the absolute rate of change in y with time at location s, and $\partial y(t, s)/\partial s$ for the absolute rate of change in y with space (distance) at time t. For some presentation circumstances, these can be simplified to $\partial y/\partial t$ and $\partial y/\partial s$, where the dependence on s and t, respectively, is assumed. For this approach to work, there is a mathematical requirement that $f(\bullet)$ is a continuous function.

For polycyclic diseases, a logistic model is often an appropriate description of disease progress over time (Chapters 4 and 5). Thus, it is natural to consider a form of equation 7.28 in which the partial derivative with respect to time is given by:

$$\frac{\partial y}{\partial t} = r_L y(1 - y) \tag{7.29}$$

Other functions could be considered here, such as the Gompertz, for temporal disease progress. Likewise, a H-L-I-R-type model (Chapter 5), with the first y term on the right hand side of equation 7.29 replaced by infectious disease, could be used for temporal progress. For monocyclic diseases, it is natural to consider a form of equation 7.28 in which the partial derivative for disease with time is given by:

$$\frac{\partial y}{\partial t} = r_M (1 - y) \tag{7.30}$$

We have shown in previous sections that the rate of change in disease with distance can be reasonably represented by the logistic or power-logistic models. Thus, we consider forms of equation 7.28 in which partial derivatives with respect to s can be written as:

$$\frac{\partial y}{\partial s} = b_{\mathrm{L}} y(1 - y) \tag{7.31}$$

and

$$\frac{\partial y}{\partial s} = \frac{b_{\mathrm{PL}} y(1 - y)}{s}. \tag{7.32}$$

Of course, other models could also be considered. This is a good place to point out that a logistic (or power-logistic) gradient model does not necessarily mean that the disease is polycyclic (compound interest) in nature.

The next step is to consider *pairs* of partial differential equations, one for space and one for time. For instance, consider logistic equations for temporal (equation 7.29) and spatial (equation 7.31) dynamics of disease. Solution leads to the following equation:

$$y = \frac{1}{1 + A_{LL}\exp(b_L s - r_L t)} \tag{7.33}$$

in which A_{LL} is the constant of integration, equal to $(1-y_0)/y_0$, where y_0 is the theoretical y at $t = 0$ and $s = 0$. That is, $f(\bullet)$ of equation 7.28a is given by the right hand side of equation 7.33. Another way of looking at this is that the partial derivatives of equation 7.33 equal equations 7.29 and 7.31. A little algebra leads to the following linear equation:

$$\ln\left(\frac{y}{1-y}\right) = -\ln\left(A_{LL}\right) - b_L s + r_L t \tag{7.34}$$

The intercept term can also be written as $+\ln[y_0/(1 - y_0)]$. Thus, equations 7.33 or 7.34 can be used to describe disease in space and time.

Fig. 7.14 shows one example epidemic that follows equation 7.33, with $r_L = 0.25$/day and $b_L = 0.4$/m. The 3D graph in the top frame visually shows disease development in the two dimensions. Note, for visual clarity, the time axis runs into the page (at a slight angle), with the smallest times in the foreground. At any selected distance (e.g., $s = 4$ m), there is a logistic increase in disease over time, with low y at $t = 0$, and y near 1.0 at large time (when s is small). The temporal curves are sigmoid in shape, although high disease intensity is not reached at large distances (within the time frame considered here). At $s = 0$, one can see the logistic increase in y at the inoculum source.

Reading the 3D graph from left to right, one can see the disease gradients over distance. These are reverse-sigmoid in shape as well, because equation 7.31 was specified for the gradient. The middle frame of Fig. 7.14 shows disease gradients for selected (labeled) times, although the curve for $t = 0$ days barely rises above the 0 axis. At early times, the gradients are exponential in shape, as expected, but they flatten close to the source because of the limitations to disease. The bottom frame of the figure shows that logit-transformed disease intensity declines linearly with distance (see equation 7.34). Moreover, because the times selected were equally spaced (every 10 days), the vertical difference of the logits is constant over all distances ($10 \cdot r_L$). It should be noted that the theoretical values plotted in Fig. 7.14 are

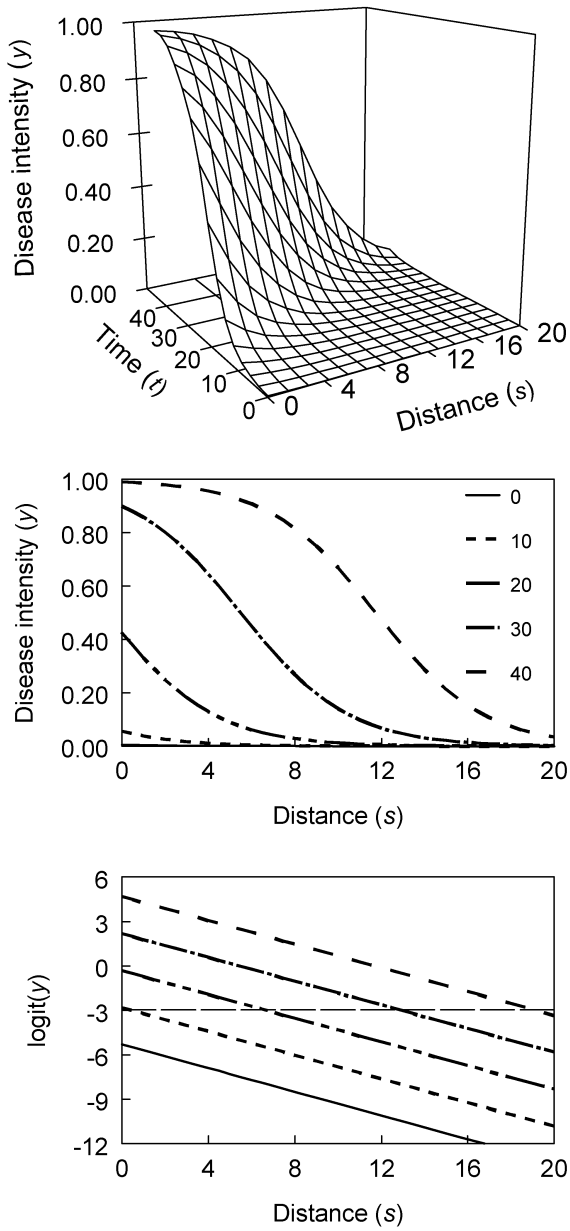

FIG. 7.14. Spatio-temporal development of disease intensity based on logistic equation 7.33. Disease gradients for six labeled times are shown on linear and transformed scales in the middle and lower frames. Horizontal line is isopath of $y' = 0.05$.

the same as those in Fig. 7.9. The values of y_0 chosen in the earlier figure were simply theoretical disease intensities at $s = 0$ for 10-day increments with the model parameters given here.

Another natural pairing of partial differential equations is the logistic one for temporal dynamics (equation 7.29) and the power-logistic one (equation 7.32) for spatial dynamics. This corresponds to the following equation for y:

$$y = \frac{1}{1 + A_{LPL} s^{b_{PL}} e^{-r_L t}} \tag{7.35}$$

where A_{LPL} is the constant of integration, $(1 - y_1)/y_1$, where y_1 is y at $t = 0$ and $s = 1$. In other words, the right hand side of equation 7.35 equals $f(\bullet)$ of equation 7.28a. Equation 7.35 can also be written as:

$$y = \frac{1}{1 + A_{LPL}\exp(b_{PL}\ln(s) - r_L t)}$$

Algebra leads to the linear equation:

$$\ln\left(\frac{y}{1-y}\right) = -\ln(A_{LPL}) - b_L \ln(s) + r_L t \qquad (7.36)$$

which indicates straight-line relations between logits and both t and $\ln(s)$. For equations 7.34 and 7.36, we can write A^*_{LPL} for $-\ln(A_{LPL})$.

An example corresponding to equation 7.35, with $r_L = 0.25$/day and $b_{PL} = 2.5$, is shown in Fig. 7.15. There is a sigmoid-type increase in y at all distances beyond 0, because of the logistic nature of $\partial y/\partial t$, with y approaching 1 near the source (over the time span considered here). Compared with the example in Fig. 7.14, there was a sharper decline in y as s increases, because of the s term in the denominator of $\partial y/\partial s$ (see middle frame of Fig. 7.15). Moreover, logits decreased in a linear manner with increasing $\ln(s)$, with the same slope for all times, as required by equation 7.36 (bottom frame of Fig. 7.15) in this example. In particular, the vertical difference between the lines is constant at all distances because the difference in times for the lines was 10 days ($10 \cdot r_L$). As pointed out for the previous figure, the values in Fig. 7.15 are the same as those in the previous sections of the chapter and shown in Fig. 7.10. What looked like arbitrary choices for y_1 (or A) in Fig. 7.10 were actually theoretical values for y at $s = 1$ at the selected time values.

Many other models for $y(t, s)$ can be specified by considering pairs of partial differential equations for space and time dynamics. Before discussing the usefulness of some of these possibilities, we must first present an additional concept of great importance when characterizing spatio-temporal dynamics.

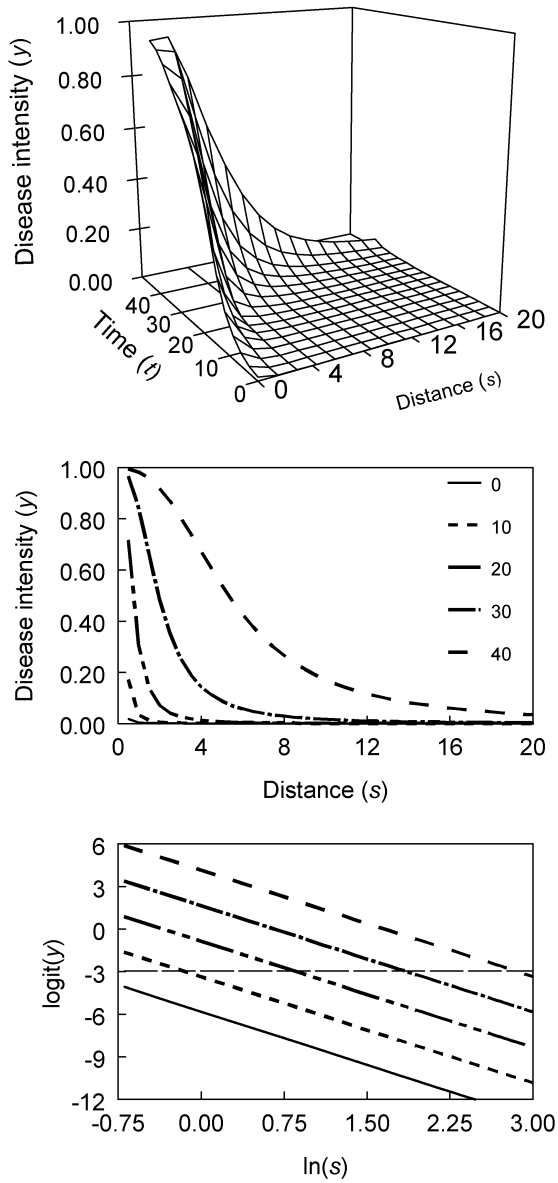

FIG. 7.15. Spatio-temporal development of disease intensity based on power-logistic equation 7.35. Disease gradients for six labeled times are shown on linear and transformed scales in the middle and lower frames. Horizontal line is isopath of $y' = 0.05$.

7.6.3 $\partial s/\partial t$

7.6.3.1 Isopaths. Suppose a pathogen is introduced into a previously disease-free (perennial) host population at a single location (point, line) and the disease becomes established (see Chapter 5). With invasive diseases, it is natural to ask: How fast is disease spreading beyond the point of introduction? For this thought exercise, suppose further that one has absolute knowledge of all diseased individuals at the disease assessment times (no missed individuals due to sampling errors and no misdiagnoses of individuals for disease status). Given that the time of introduction is defined as $t = 0$, suppose that the farthest distance (s) from the initial inoculum source with a diseased individual at various times (weeks) was found to be:

$$t = 2 \text{ wk}, \quad s = 14 \text{ m};$$
$$t = 4 \text{ wk}, \quad s = 28 \text{ m};$$
$$t = 6 \text{ wk}, \quad s = 42 \text{ m}.$$

The distances can be considered to represent the front of the epidemic at the listed times. As in Chapter 4, we use Δ here to indicate a calculated difference between successive values. The calculated rate of spatial spread of disease with time ($\Delta s/\Delta t$) can be determined by dividing the difference of distances by differences in times.

Between weeks 2 and 4, the disease spread 14 m; that is, disease was 14 m farther away from the source at week 4 compared with week 2. Thus, $\Delta s/\Delta t = 14/2 = 7$ m/wk, or 1 m/day. With the hypothetical example, this rate is the same for all pairs of times. If this same rate continued indefinitely, one could calculate when disease would be, for example, 1 km from the original source. The answer is 1000/7 = 142.9 wk, or ~2.7 years. Likewise, one could ask: how far will the disease spread in the next 5 years? Assuming a continuous crop in all directions and no seasonal effects, and given that there are 260 wk in 5 years, the answer is $7 \times 260 = 1{,}820$ m or ~1.8 km.

This rate of spatial spread of disease over time tells us nothing about the intensity of disease at any distance from the original source, unless other assumptions are made or additional information is available (see below). Moreover, it is often impractical (because of costs and time) to assess disease at great enough distances from a source to find the farthest diseased individual. For instance, early in the bean rust example (Fig. 7.11), assessing leaves for pustules over 9 m from the source was satisfactory, but by 3 wk into the epidemic, about 50% of the leaves were infected at 9 m. Thus, it is quite possible that the front of the epidemic is unknown.

The practical approach to determining $\Delta s/\Delta t$ is to choose a (generally low) fixed value of y, y', and determine the distance where y' is found at selected times. It may be more practical to work with transformed disease intensity, and select a specific y^*, $y^{*\prime}$, for calculations. Consider, again, Fig. 7.14 for logistic-type disease gradients over time. In the bottom frame, a thin horizontal line is shown, corresponding to $y^{*\prime} = -2.944$, or $y' = 0.05$. A line (or curve) of fixed levels of disease intensity is called an *isopath*. The horizontal line is not shown in the middle frame because it is hard to see (being so close to 0). The points where the horizontal line intersects the gradient lines (or curves) are the distances where 5% disease intensity are found for the different times on the graph. At $t = 0$, there is no distance where 5% intensity is found. For the other shown times, the following distances are found (assuming that time is measured in days and distance in meters):

$$t = 10\,\text{days}, \quad s' = 0.38\,\text{m};$$
$$t = 20\,\text{days}, \quad s' = 6.63\,\text{m};$$
$$t = 30\,\text{days}, \quad s' = 12.88\,\text{m};$$
$$t = 40\,\text{days}, \quad s' = 19.13\,\text{m}.$$

Based on these specific results, one calculates $\Delta s/\Delta t$ as 0.625 m/day. For instance, between days 30 and 40, one obtains: $(19.13 - 12.88)/10 = 0.625$. The distances here are shown to two decimal places, although with observed gradients (e.g., Fig. 7.13), one will not be able have such precision (later, we show how we obtained such precise values for distance here). Plus, given that

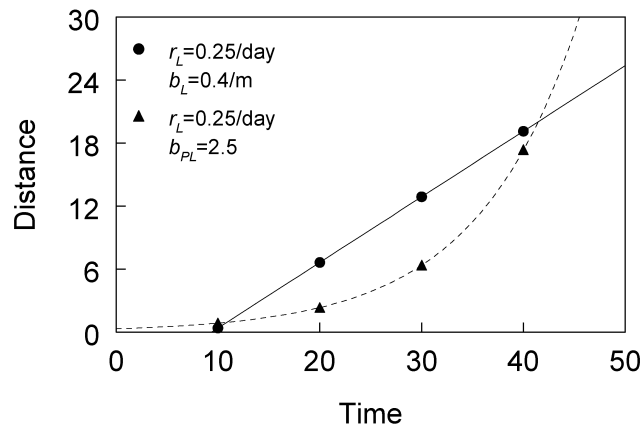

FIG. 7.16. Distance where $y' = 0.05$ is found in Figs. 7.14 (logistic spatio-temporal model) and 7.15 (power-logistic spatio-temporal model). Lines are solutions to equations 7.38 (solid) and 7.39 (dashed).

observations of disease are made at selected (fixed) values of distance and time, it is very possible that exactly 5% intensity is not found at any distance. This requires interpolation or use of model descriptions.

Before discussing model results, we reconsider Fig. 7.15 for the power-logistic type of gradients. The horizontal line is, as before, $y^{*\prime} = -2.944$ (for $y' = 0.05$). The values of log of distance, $\ln(s')$, can be read from the intersection points, and s' determined from $\exp(\ln(s'))$:

$$t = 10\text{ days}, \quad \ln(s') = -0.145, \quad s' = 0.86\text{ m};$$
$$t = 20\text{ days}, \quad \ln(s') = 0.855, \quad s' = 2.35\text{ m};$$
$$t = 30\text{ days}, \quad \ln(s') = 1.850, \quad s' = 6.39\text{ m};$$
$$t = 40\text{ days}, \quad \ln(s') = 2.855, \quad s' = 17.38\text{ m}.$$

Although the difference in $\ln(s')$ between times is constant here, the difference in s' increased with each successive time period. That is, the farther the distance from the source, or the longer the time since the start of the epidemic, the greater the $\Delta s/\Delta t$. For instance, between 30 and 40 days, $\Delta s/\Delta t$ equaled 1.1 m/day, but between 10 and 20 days, $\Delta s/\Delta t$ was only 0.15 m/day. The distance values for the two cases (Figs. 7.14 and 7.15) are shown in Fig. 7.16, together with lines that will become apparent shortly.

Since we are now dealing with changes in distance where a *specific* non-zero y (y') is found, it would be more exact to represent the rate of change as $\Delta s'/\Delta t$. This is because the result could depend on the specific disease intensity chosen. However, with a few exceptions, such additional notation is not needed.

7.6.3.2 Two models. Using the models for temporal ($\partial y/\partial t$) and spatial ($\partial y/\partial s$) disease development in section 7.6.2, it is straightforward to characterize the rate of spatial spread of disease over time. As discussed in section 7.6.2 above, partial derivatives are used with the

deterministic models for rates of change of a variable with two "independent" variables (t and s). It can be shown (Jeger, 1983) that the rate of isopath movement, $\partial s/\partial t$, is given by:

$$\frac{\partial s}{\partial t} = -\frac{\partial y/\partial t}{\partial y/\partial s} \tag{7.37}$$

for a general spatio-temporal process such as that represented by equation 7.28a. It is useful to think of $\partial s/\partial t$ as a *velocity* since it has units of distance per time. Specifically, $\partial s/\partial t$ is the velocity (or speed) of isopath movement, how fast a given level of disease moves from one location to another, or the rate at which the disease focus expands. The $\Delta s/\Delta t$ value calculated above for specific times and distances is replaced by the partial derivative, $\partial s/\partial t$. Because we are assuming that $y(t, s)$ is represented by a continuous function, one can think of a $\Delta s/\Delta t$ calculation as an estimate of $\partial s/\partial t$. With the adoption of a particular model form for equation 7.28a—which is equivalent here to selection a pair of functions, $\partial y/\partial t$ and $\partial y/\partial s$—one can directly define spatial movement or spread using equation 7.37.

Using the logistic equation 7.29 for $\partial y/\partial t$ and logistic equation 7.31 for $\partial y/\partial s$ (with solution given by equation 7.33), one obtains the very simple expression for the velocity based on equation 7.37:

$$\frac{\partial s}{\partial t} = \frac{r_L}{b_L} \tag{7.38}$$

Since both the numerator and denominator are constants, equation 7.37 indicates that $\partial s/\partial t$ is a constant ($C = r_L/b_L$), with units of distance per time. This means that any given fixed level of disease (y') or its transformation ($y^{*\prime}$) moves a fixed distance over a fixed time. For the example in Fig. 7.14, the horizontal difference between successive lines or curves (which are for 10-day increments) is the same, and equal to C multiplied by the time difference: $(0.25/0.4) \times 10 = 6.25$ m. Although the example was worked through for $y' = 0.05$, other isopaths could have been chosen. The actual distances will be different (e.g., $y' = 0.5$ will be closer to the source than $y' = 0.01$), but the change in distance over time will be the same (for this model).

Integration of equation 7.38 leads to the straight-line expression $s_0 + (r_L/b_L)t$, where s_0 is a constant of integration, corresponding to the distance where y' is found at $t = 0$. For the example, y at $t = 0$ is less than 0.05 at all distances, even at the source. Thus, s_0 does not necessarily have a simple physical interpretation. The straight line (with a slope of r_L/b_L) is shown on Fig. 7.16 with the previously obtained specific distances. For a selected y' less than 0.05, the points and the line will be higher (because the low disease intensity is found farther from the source); for a selected y' greater than 0.05, the points and the line will be lower (because the high disease intensity will be found closer to the source).

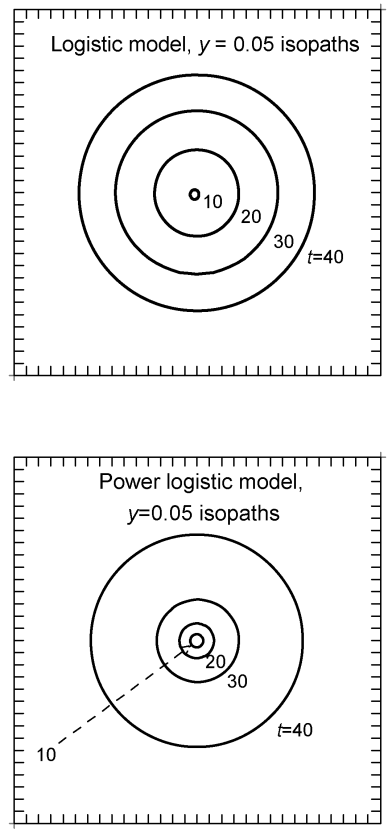

FIG. 7.17. Theoretical radial expansion of the $y' = 0.05$ isopath for a logistic (equation 7.33) and power-logistic (equation 7.35) spatio-temporal model at labeled time points.

When $\partial s/\partial t$ is constant, it is said that disease spreads in a *wave-like* manner (Diekmann, 1978; Thieme, 1977), and that epidemic develops as a traveling wave with velocity C. This can be seen by the top frame of Fig. 7.17, which corresponds to the same values in Fig. 7.16, but specifically for a point source (at the center of a field). The isopaths of $y' = 0.05$ increase from the center in a uniform manner at rate $C = 0.625$ m/day.

Now consider the situation where the power-logistic model describes the disease gradients (equation 7.32) and logistic model for temporal development (equation 7.29). In other words, we consider equation 7.35 as the solution for $y(t, s)$ in equation 7.28a. Utilization of equation 7.37 produces:

$$\frac{\partial s}{\partial t} = \frac{r_L s}{b_{PL}} = (r_L/b_{PL})s = \upsilon s \tag{7.39}$$

for the rate of isopath movement. Note that because b_{PL} is unitless, υ $(=r_L/b_{PL})$ has units of 1/time; thus, $\upsilon \cdot s$ is the velocity, and not just υ. Equation 7.39 is similar to the expression found in equation 7.38, except that the velocity is directly proportional to distance. That is, all other

things being equal, $\partial s/\partial t$ (velocity) increases as s increases. This is exactly the situation found for the situation in Fig. 7.15. Solution to this equation is given by:

$$s_0 \exp[(r_L/b_{PL})t],$$

where the constant of integration (s_0) is the theoretical distance at $t = 0$ where y' occurs. Taking the derivative of this equation with respect to time gives another way of expressing the disease expansion velocity:

$$C(t) = (r_L/b_{PL})s_0 \exp[(r_L/b_{PL})t]$$

Note that we explicitly labeled the velocity as $C(t)$ to emphasize that the implication of equation 7.39 is that velocity increases with time since the start of the epidemic. Because of the dependence of the velocity on distance, the isopaths increase exponentially with *time*, as can be seen by the broken curve in Fig. 7.16. For an idealized situation of a single point source in Fig. 7.17, one can see that the differences of the circle radii increase with time.

Epidemics characterized by equation 7.39 are said to spread in a non-wave-like manner because of the acceleration of isopath movement (Minogue, 1989). Ferrandino (1993) refers to this as *dispersive wave* spread.

It is important to note for both equations 7.38 and 7.39, $\partial s/\partial t$ is directly proportional to r_L. This means that the velocity of spatial spread increases with the temporal increase in disease (where the disease is present). For instance, if infection efficiency is higher, latent period shorter, or sporulation rate per lesion higher for disease 1 compared with disease 2, then $\partial s/\partial t$ will also be higher for disease 1 than for 2. On the other hand, $\partial s/\partial t$ is inversely proportional to b_* (b_L or b_{PL}) for both models. Since b_L is a measure of the gradient steepness, it is logical that isopath velocity increases as b_L decreases. For instance, if spores travel farther (on average) for pathogen 1 than for 2, then $1/b_L$ will be larger for 1 than for 2, producing a larger $\partial s/\partial t$ for 1 than for 2. For the unitless b_{PL} parameter, interpretation is a little more difficult, as discussed for the b_P in the dispersal gradient models and contact distribution models. However, since b_P characterizes the steepness of the gradient on a log distance scale, one can (in this situation) give roughly the same interpretation to b_P as to b_L, in terms of velocity of spread (in logs).

7.6.4 Other models

The spatio-temporal models of equations 7.33 and 7.35 were chosen for expository purposes in the previous section, because they built on common models for temporal and spatial components of epidemics and they capture some of the essential features of velocity of disease movement. Of course, other pairings of partial differential equations can (or should) be considered as the components of equation 7.28a. For instance, consider the situation where disease progress over time is of the monomolecular form (equation 7.30) and that the disease gradients are represented by $\partial y/\partial s = b_{PM}(1-y)/s$ (a power-monomolecular type). Utilizing equation 7.37, one obtains:

$$\frac{\partial s}{\partial t} = \frac{r_M s}{b_{PM}} = \left(\frac{r_M}{b_{PM}}\right)s$$

which is identical to equation 7.39, except that different symbols are used for the parameters. Likewise, pairing the monomolecular temporal model with a pure monomolecular model for the gradients [$b_M(1-y)$] results in a constant velocity (r_M/b_M).

However, what if temporal disease progress was of the monomolecular form (equation 7.30) and that the gradients were of the logistic type (equation 7.31)? This results in:

$$\frac{\partial s}{\partial t} = \frac{r_M}{b_L y} = \frac{(r_M/b_L)}{y} \qquad (7.40)$$

This velocity equation has some essential characteristics in common with the others we considered: $\partial s/\partial t$ increases with increasing temporal disease increase (r_M) and with decreasing steepness of the gradient ($1/b_L$). However, the velocity is also related to the reciprocal of disease intensity. That is, $\partial s/\partial t$ is a function of y, with greater $\partial s/\partial t$ at low intensity than at high intensity. For this situation, one would obtain a different disease expansion rate for every specific disease value (y'). For this circumstance, one could be more explicit and refer to the velocity as $\partial s'/\partial t$.

To understand disease-dependent velocities, consider a graph similar to the middle frame of Fig. 7.14 corresponding to the solution of equation 7.40. If one drew horizontal lines at different y' values, and calculated $\Delta s'/\Delta t$ for each, one would find greater $\Delta s'/\Delta t$ at low y' than at high y'.

Jeger (1983) showed some other interesting pairings of equations for the temporal and spatial components of epidemics. In general, however, the pairings of partial differential equations may not even lead to an analytical solution for $y(t, s)$ (equations 7.28a, b). The situation can be even more complicated. For instance, in this chapter we assumed that the population terms for temporal and spatial rates of change (r_* and b_*) are constants. But this is not necessarily so. For instance, even if y increased logistically everywhere in the population, r_L might not be fixed but depend on distance from (original) inoculum source; in particular, r_L could either increase or decrease with s. One would write this as $r_L(s)$ to show that r_L is a function of s. This could be directly assessed by analyzing the disease progress curves at *each* sampling location

(e.g., the vertical slices of y in Fig. 7.11) in the field to see if estimates of r_L changed systematically with distance. To add to the complexity, r_L could also vary with time, such that r_L values for the early and late parts of the epidemic were different. This could be written as $r_L(t, s)$. It should be noted that if r_L did vary systematically with time, the disease progress curve likely would appear to be non-logistic in shape.

Additionally, disease gradients might be logistic (or power-logistic) in nature at all times, but with a b_L (or b_{PL}) term that is a function of time, $b_L(t)$ [or $b_{PL}(t)$]. We assessed this directly for the power-logistic example in Fig. 7.11 and saw some evidence that the slope did decrease with time. Moreover, b_L (or b_{PL}) could vary with distance, so that the decline in y near the source has a different rate than far from the source. This could be written as $b_L(s)$ [or $b_{PL}(t, s)$]. This latter situation would probably appear as a non-logistic (or non-power-logistic) type of disease gradient at individual times.

The main points here are that the spatio-temporal dynamics of disease can be very complicated, and it may be very challenging to find a relatively simple model (equation 7.33 or 7.35) for representing the dynamics of disease in space and time that can be justified based on population biology. Even when the purely temporal dynamics and disease gradients are well described by standard models (e.g., logistic), characterizing the resulting spatio-temporal dynamics is a difficult endeavor. Ultimately, diffusion and/or integral modeling is necessary, as presented in the next chapter. Nevertheless, the models presented here do provide a reasonable starting point for understanding epidemics and for describing and comparing their development in space and time.

7.6.5 Analysis

As demonstrated above, one can (empirically) determine the velocity of isopath movement for a given epidemic, $\Delta s/\Delta t$, by choosing a specific value of y, y' (or $y^{*\prime}$), and determine the distance where this value occurs at the different assessment times. The change in distance divided by the change in time is a measure of the velocity. Generally, one should use a small value for y', because one is usually most interested in the movement of the front of the epidemic. A reasonable choice is $y' = 0.05$ or $y' = 0.10$. Much smaller values can lead to some misleading results because of the low precision of the estimated y (see Chapter 10). To determine whether or not $\Delta s'/t$ is affected by disease intensity, one could perform the calculations for other y' values (e.g., 0.25, 0.5).

In attempting to apply this protocol to the example data set in Fig. 7.11, one can see some immediate difficulties. Using $y' = 0.05$ (or $\text{logit}(y') = -2.94$), one can only determine $\Delta s/\Delta t$ between times 1 and 8 (with interpolation), since all y values are above 0.05 at later times. If one chose $y' = 0.6$, then the isopath movement rate could be determined between several times. However, if the rate depends on y, then it may be less interesting to know the movement of the 0.6 isopath than the 0.05 isopath (i.e., nearer the front).

One can use the model fitted to the gradient data to determine the distance where y' is obtained. For instance, assuming that the power-logistic gradient model is appropriate (equation 7.23), and using the parameter estimates shown above for the data in Fig. 7.13, one can determine that the distance corresponding to $y^{*\prime} = -2.94$ at $t = 8$ days as:

$$\hat{s}' = \exp\left(\frac{y^{*\prime} - \hat{A}^{*}_{PL}}{-\hat{b}_{PL}}\right) = \exp\left(\frac{-2.94 - (-0.3)}{-0.77}\right)$$
$$= \exp(3.43) = 31\text{m}$$

Note that the exponential function is used here because simple rearrangement of equation 7.23 produces an expression for $\ln(s')$. This value can also be obtained by rearranging nonlinear equation 7.22 [where one uses $\hat{A}_{PL} = \exp(-\hat{A}^{*}_{PL}) = 1.35$]. This actually is the type of calculation used to obtain the s' values to two decimal places in section 7.6.3.1 for the idealized disease gradients of Figs. 7.14 and 7.15. If this calculation is done for the logistic spatio-temporal model (equation 7.34), one would not perform the exponential operation.

The calculated distance here is considerably beyond the largest distance used in the study (Fig. 7.13), and given the uncertainty about the appropriateness of the model at large distances, may not be reliable. The reason the estimated s' is so large is because of the very shallow gradient (small dy/ds) at large distances for power-type model processes. For the data in Fig. 7.13, the researcher who wishes to estimate disease expansion over time will be much better off using disease intensity values above 0.25 for y'. This is because the calculated distances will be (mostly) within the range of distances used in the study.

Using calculated s' values for each time, based on the chosen disease-gradient model parameter estimates, one can then calculate $\Delta s/\Delta t$ values. In other words, the calculated s' values from the equation rearrangements are treated as observed distances. If we use $y' = 0.6$ (so that $y^{*\prime} = 0.405$), and the parameter estimates obtained when there was a different slope for each time, one obtains the s' estimates shown in Table 7.1. The differences in s' between times were: 1.3, 3.0, 3.2, and 15.6 m over the last five assessment times. One divides these by 7 days (the difference in time between the disease assessments on which the gradients are based) to obtain the following velocity values: 0.19, 0.41, 0.46, and 2.22 m/day. Clearly, there is an increasing velocity for $y' = 0.6$. This is partly because of the power-type relation (s in the numerator for $\partial y/\partial s$) and also because b_{PL} was decreasing somewhat over time.

Using the models developed in section 7.6.2, one does not need to base isopath calculations on observed gradients or estimated gradients for individual times. In particular, if one can assume a specific spatio-temporal

model, or find empirical evidence that such a model is reasonable, one can directly use model parameters for determining (or estimating) $\partial s/\partial t$. For the bean rust example (Fig. 7.15), we fitted the power-logistic model (equation 7.36) to the data using least squares. The statistical model can be written as:

$$y^*_{ij} = A^*_{\mathrm{LPL}} + r_{\mathrm{L}} t_i - b_{\mathrm{PL}} \ln(s)_j + \varepsilon_{ij} \qquad (7.41)$$

where y^*_{ij} is the logit of disease intensity for the i-th time and j-th distance, and ε_{ij} is the error. The SAS code for fitting this model is shown at the bottom part of Fig. 7.12. This model should be compared with equations 7.27a,b where a separate intercept and slope (on a transformed scale) were specified for *each* group, in which "group" corresponded to time (a class/factor variable). Here, there is one intercept term, both time and space are continuous variables, and there is no interaction of time and ln(s).

Fitting the model resulted in:

$$\hat{y}^* = -0.97(0.150) + 0.088(0.007)\ t - 0.564(0.071)\ \ln(s)$$

where the numbers in parentheses are standard errors of the parameter estimates. The R^2 value was 0.84. The estimated residual variance was 0.034. Using the parameter estimates, estimated velocity of isopath movement ($\partial s/\partial t$) was (0.088/0.564) $\cdot\ s = 0.156s$ m/day. For instance, the estimated velocity at 1, 3, 9, and 18 m equals 0.156, 0.47, 1.4, and 2.8 m/day. Distance of a fixed y at time t is estimated by:

$$\hat{s}' = \hat{s}_0 \exp\left(\frac{\hat{r}_{\mathrm{L}}}{\hat{b}_{\mathrm{PL}}} t\right)$$

where estimated s_0 (distance at $t = 0$ where y' is found) can be estimated by:

$$\exp\left(\frac{\hat{A}^*_{\mathrm{LPL}} - y^{*\prime}}{\hat{b}_{\mathrm{PL}}}\right)$$

in which $y^{*\prime}$ is the logit of transformed y'. This expression for predicted s_0 can be found by algebraic rearrangement of equation 7.36 (with $t = 0$). Using the parameter estimates and $y' = 0.6$ ($y^{*\prime} = 0.405$), one obtains:

$$\hat{s}' = 0.087 \exp(0.156t)$$

For days 8, 15, and 22, 29, and 36, one obtains estimated distances of: 0.3, 0.9, 2.7, 8.0, and 24.0 m. Based on an assessment of the data in Figs. 7.11 and 7.13, it would appear that these values are little too small for the early times. The problem lies with the fit of equation 7.36 to these data. In particular, at each distance, disease progress is not logistic, at least not logistic with $K = 1$. This can be seen by the plot of logit(y) versus ln(s). Since disease was assessed every 7 days, a logistic increase would correspond to *equal* vertical separation of the transformed gradients. For example, at $s = 0.7$ m, a purely logistic increase would mean equal changes in logits between times 1 and 8, 8 and 15, and so on. If r_{L} and b_{PL} were fixed for the entire epidemic, the vertical differences ($= r_{\mathrm{L}} \cdot \Delta t$) would be the same at each distance on the logit scale. The interested reader can construct disease progress curves for each distance by taking the vertical slices of logits and plotting these versus time. These data exhibit a slowing down of the epidemic as time increased, which corresponds to decreasing vertical differences of the logits at each time. A residual plot of the multiple regression confirmed the logistic model for temporal progress was not ideal here.

One could also fit the logistic-logistic model (equation 7.34) to the data to describe the epidemics. Then, one would use:

$$\hat{s}' = \hat{s}_0 + (\hat{r}_{\mathrm{L}} / \hat{b}_{\mathrm{L}})t$$

to represent isopath values. Here, s_0 is estimated as $[\hat{A}^*_{\mathrm{LL}} - y^{*\prime}]/\hat{b}_{\mathrm{L}}$. However, this model also would not do any better than the power-lotgistic because of the declining temporal increase. As an alternative, one could use a different temporal model (e.g., monomolecular, Gompertz), but it is difficult to combine either of these with a power-logistic (or logistic) disease gradient model. This is, in part, because $\partial s/\partial t$ is then a function of y. As an alternative to using a different temporal model, one could specify a temporal model where $\partial y/\partial t$ is inversely proportional to time, which results in logit(y) being a linear function of ln(t). This is the same as specifying a logistic model but with r_{L} a function of time, $r_{\mathrm{L}}(t)$, rather than a constant. Jeger (1983), Campbell and Madden (1990), and Madden et al. (1990b) show how to use this type of temporal model.

The above discussion underlines the challenges faced with spatio-temporal modeling and analysis of epidemics. Analytical solutions to partial differential equations for temporal and spatial development (equation 7.28a) may only be obtainable by making strong assumptions about the process being modeled (e.g., pairs of partial differential equations), which may or may not be reasonable. When the assumptions are reasonably met, the solutions make it quite convenient to describe, compare, and understand epidemics. That is, models such as equations 7.33 and 7.35, and their linearized forms, offer an elegant way of characterizing disease progress in space and time. As will become apparent in the next chapter, these models, especially for polycyclic diseases, can only be considered an intermediate step in developing and understanding of epidemics in time and space.

7.7 Disease Management

Knowledge of dispersal gradients can be of great value in developing and testing disease controls (McCartney and Fitt, 1998). For instance, one can decide how far a host crop must be separated from an inoculum source (e.g., an alternative host) in order to maintain disease intensity below a critical value. Or, one can determine the effects of resistant cultivars, cultivar mixtures, and other control treatments on the level of disease at, say, s' meters from an inoculum source. By having knowledge of disease development in time and space, one can determine how long it will take for disease to reach a certain critical value (say, y') under a given set of crop and environmental conditions. Then, through experiments, one can determine which treatments will delay the occurrence of the critical disease value.

The evaluation of controls on dispersal and disease spread is done through the appraisal of (estimated) model parameters. Calculations done using example parameters throughout this chapter are, thus, of value in evaluation of controls, and we do not repeat them here. With primary gradients, the effect of many controls will be to reduce a_* but not b_* (Minogue, 1986). That is, if a control (e.g., a resistant cultivar) reduces disease intensity by a constant fraction throughout a field (because, for instance, the probability of infection by a deposited spore is reduced by 50% compared with another cultivar), then the only difference in gradients is in the height of the curves (a_*). Thus, the spread parameter has no relation to control in this case. However, b_* may or may not be affected by some agricultural practices that can have an effect on disease intensity at any time, for instance, planting density (see, e.g., Madden and Boudreau, 1997). Certainly, the increased number of plants or leaves filters out spores due to barrier effects; however, spores would only have to travel short distances to reach, and possibly infect, other plants. Moreover, the crop environment may change with change in plant population density, leading to a change in the probability of infection by deposited spores.

For secondary spread, it is more difficult to make generalizations about the height of the disease gradient or the steepness of the gradient in reference to the original inoculum source. This is because the number of infections at a given location could be due to inoculum produced at many locations in a field, not just at the original source. If the control affects temporal disease development (r_*), then the control does affect spread of disease ($\partial s/\partial t$; see equations 7.38 and 7.39), whether or not b_* is affected.

Despite the documented evidence that dispersal is a key factor in epidemic development, Jeger (1999) has observed that knowledge of dispersal has been inadequately used in developing and implementing actual integrated pest management programs. Thus, there are considerable opportunities to improve disease control strategies and tactics by the utilization of pathogen dispersal and disease spread information obtained from epidemiological research.

7.8 Concluding Comments and a Prelude to the Following Chapters

In this chapter, we have considered the dispersal of spores and other infectious units (including inoculative vectors) from one location to another. More than for any other topic in botanical epidemiology, the study of dispersal involves a combination of principles and methods from physics and biology (Aylor, 1999; Madden, 1992). As is often done in epidemiology, in this chapter we condensed the complex physical models for spore transport to (relatively) simple ones for spore deposition (per unit area) or infection (e.g., pustule number per leaf) as a function of distance from an inoculum source—the dispersal gradient. These models are also scaled versions of the contact distribution, which specifies the probability of a spore (or any infectious unit) originating at location 0 arriving (being deposited) at location s.

Dispersal gradients are often of an exponential or (modified) power-law type (equations 7.1 and 7.4). Both model types can produce very similar predictions over a narrow-to-moderate distance range, and it can be difficult to choose between the models with many data sets (Ferrandino, 1996; Fitt et al., 1987). At distances very close to the inoculum source and far away, gradients corresponding to these two model types can deviate substantially. In particular, a power model is characterized by steeper declines in Y near the source and shallower declines in Y far from the source, relative to an exponential model (Fig. 7.5). Both linear and nonlinear statistical methods can be used to fit these models to data. Through the estimated parameters, one can compare gradients and evaluate the effects of imposed treatments or uncontrolled factors (e.g., differences between years).

When one is characterizing dispersal gradients through measurements of disease intensity (number of infections), one must be careful to adjust for the maximum possible disease whenever disease intensity is moderate or high. That is, if infections are represented on a proportion scale (y: $0 - 1$), one should not use the simple gradient models without modification. The modification typically involves multiplying the expression for dy/ds by $1 - y$ (or a function of $1 - y$). The resulting disease gradient models can be used to determine, among other things, how disease gradients change over time in an epidemic, and how far disease spreads over time. It is especially useful to determine the rate of movement (velocity: C or $C(t)$) of a fixed amount of disease from a source, which depends on the form of the dispersal gradient or contact distribution. One can characterize the simultaneous spatio-temporal development of disease by linking partial differential equations for disease progress in time and disease gradient over distance. For some situations,

analytical solutions that simplify epidemic descriptions are possible.

Ultimately, more elaborate modeling is required to represent the spatio-temporal epidemic process, especially for polycyclic diseases. For instance, in the derived models for disease intensity (equations 7.33 and 7.35), y at a given time is related only to distance from the *original* source. Once the primary spread of disease has occurred, however, every infected individual in the field can serve as an inoculum source—once the latent period has been passed and before the infectious period has ended. Thus, there is no unique (individual) distance of relevance, but many distances to consider. Infected individuals near location s are more likely to cause infections at s than infected individuals very far from s. This complicates the matter dramatically, as is shown in the next chapter.

Adding to the complexity of the quantification problem is spread of disease from multiple foci at the start of the epidemic rather from one inoculum source. Other than through simulation (as discussed partly in Chapter 8), the more common approach is to take a different perspective to the analysis, and to quantify the pattern of disease at one or more assessment times during an epidemic. This is dealt with in detail in Chapter 9.

7.9 Selected Readings

Aylor, D. E. 1999. Biophysical scaling and the passive dispersal of fungus spores: relationship to integrated pest management strategies. *Agric. For. Meteorol.* 97: 275–292.

Jeger, M. J. 1999. Improved understanding of dispersal in crop pest and disease management: current status and future directions. *Agric. For. Meteorol.* 97: 331–349.

McCartney, H. A., and Fitt, B. D. L. 1998. Dispersal of foliar fungal plant pathogens: mechanisms, gradients, and spatial patterns. In: *Plant Disease Epidemiology* (D. G. Jones, editor). Kluwer, London, pp. 138–160.

Pielaat, A., Madden, L. V., and Gort, G. 1998. Spores splashing under different environmental conditions: a modeling approach. *Phytopathology* 88: 1131–1140.