Spatial analysis of epidemics:
Disease gradients and patterns

Space....The Frontier....Finally.

- Epidemics are dynamic population processes in time *and space*
- We have spent a few weeks dealing exclusively with temporal dynamics
- However, diseases usually spread in space as they increase in time
- Because of spread, disease intensity is not the same everywhere in a field or plot -- there is a *pattern* to disease
- Spread *may or may not* be relatively easy to measure and quantify.

"Given that the transmission of pathogens leading to disease requires the close juxtaposition of a susceptible [healthy, disease-free] individual with an infected conspecific, vector, or environmental source of pathogens, transmission dynamics are inherently spatial processes."

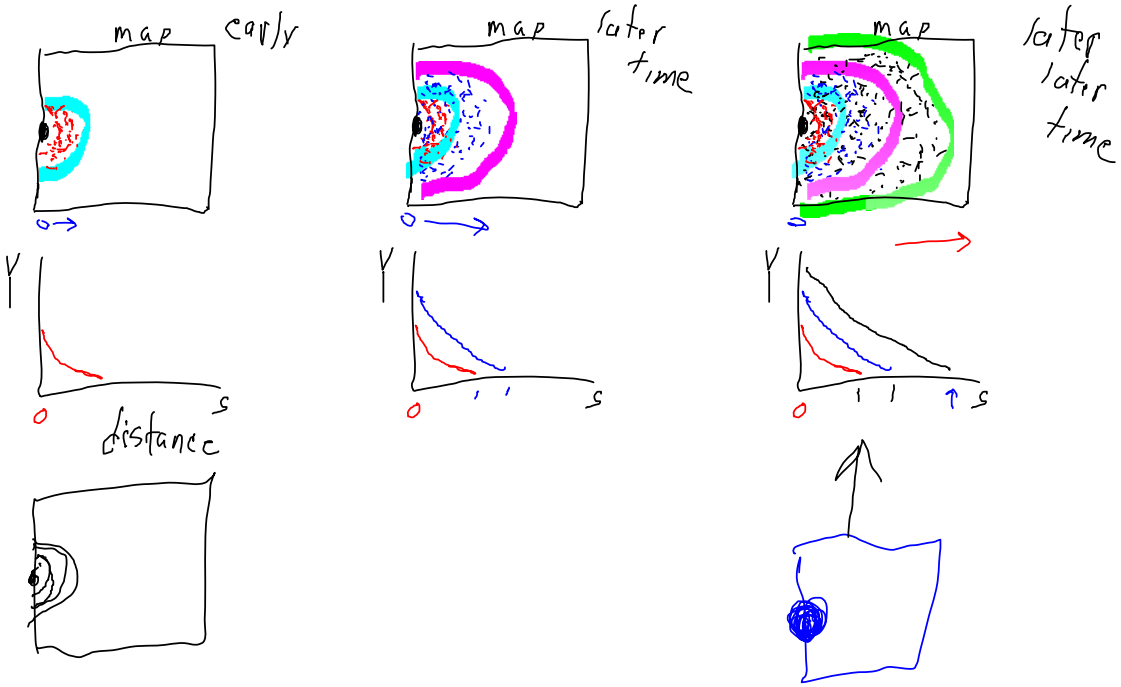
-- Ostfeld, Glass, and Keesing (2005). Spatial epidemiology: an emerging (or re-emerging) discipline. *Trends in Ecology and Evolution* 20: 328-336.

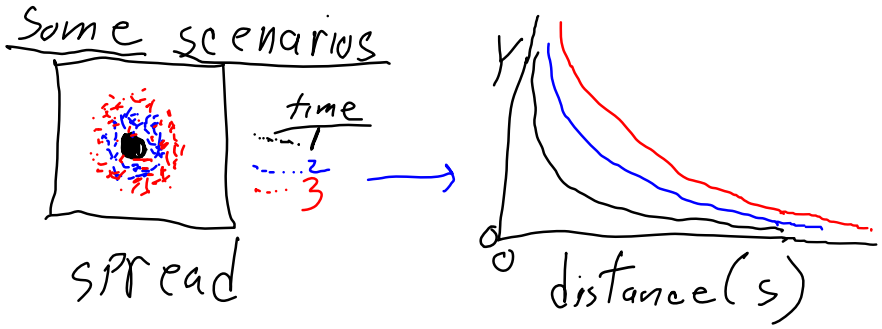
From HLIR model (coupled differential equations):

$$dH/dt = -\beta IH = -(\alpha\theta\psi)IH \quad \text{or} \quad dY/dt = +\beta IH = +(\alpha\theta\psi)IH$$

θ: Ultimately depends on distance between *I* and *H*

Results of spatial analysis allows, ultimately, for the expansion of temporal models to spatio-temporal models of epidemics (later)



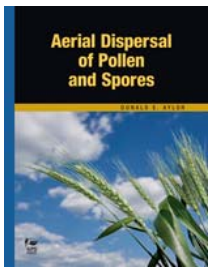


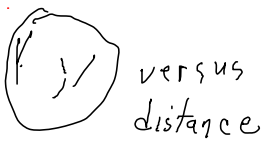
see figure 7.1, 7.17

Concepts

- Disease spread *implies* movement, but diseased plants (leaves, etc.) do not move (typically), just the inoculum
- We consider ***spread to be the result of dispersal***
- **Dispersal**:
  - Movement of 'inoculum' (such as spores, infectious units, propagules, viruliferous vectors) from one place to another
  - Movement of units of inoculum from the place they were formed to other locations
    - Involves **liberation, transport, deposition**
    - Physical processes (not dealt with here)
      - Pioneer: **P. H. Gregory** (see his pioneering book)
      - See articles by **Don Aylor**, A. McCartney, L. Huber, S. Isard, others
- **Because of dispersal, the intensity of disease varies systematically with distance from the source of inoculum**
  - *There is more disease in some places than others*
- We primarily deal with disease intensity and not the 'inoculum'
  - There is a vast literature on the physics of inoculum dispersal (**a very worthy topic**), but we do not discuss here (an entire course can be devoted to this topic!)

Don Aylor (2017).APS Press.

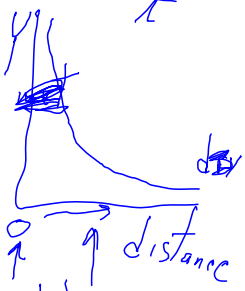




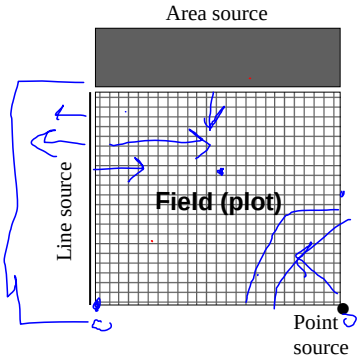
versus distance

- For temporal dynamics, we always considered the **disease progress curve** (*y* versus *t*; or *Y* versus *t*; or *L, I, R,* and *L+I+R* versus *t*)
- For spatial dynamics, we consider the **disease gradient**:
  The change (generally, *decline*) in disease intensity over distance from an inoculum source
  Intensity: incidence, severity (including counts)

- **Disease gradients** can have multiple causes (which serves as one convenient way of classifying them)
  - **Environmental gradient**
    Disease gradient due to physical or other external factors not related to dispersal (or the biology of the disease)
  - **Dispersal gradient**
    Disease gradient caused by variation in deposition of 'inoculum' in relation to distance from an inoculum source
    The type of gradient of interest to us
  - **Deposition gradient**
    - (not a disease gradient at all). 'Inoculum' deposited per unit area (or length) in relation to distance from inoculum source
      - Deposition gradients *lead* to dispersal gradients



- **Disease gradients due to dispersal (i.e., dispersal gradients) are of importance here**
- More concepts: **Inoculum source**
  - For polycyclic diseases, any infected (specifically, infectious [*I*]) individual is an inoculum source for other disease-free (*H*) individuals
  - For monocyclic diseases, inoculum source can be infected individuals at another location ('external' to the current epidemic) or in another year; or spores in the soil, or spores originating elsewhere (another epidemic)
  - But, to *characterize (i.e., **study**)* dispersal gradients, we *need to be more restrictive* in what we consider an **inoculum source (an *operational* definition):**
    - **A spatially-restricted concentration of spores, infectious units, other units of 'inoculum', or diseased individuals that can produce 'inoculum' (disease focus)**
      - Generally, the source is much smaller in area than the area of interest (field) (in our operational definition)
      - Typically consider the source as '***separate***' from the area of interest (for convenience of experimentation only)
  - Classify inoculum sources based on size and geometry



Kinds of inoculum sources

(see Chapter 7 in MHV for more detailed definitions)

You must know definitions of these terms .

Point source: width (size) of source is <1% of the distance over which spread is measured

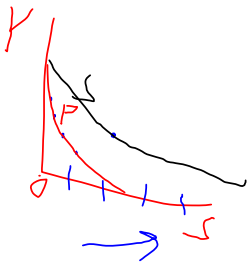
Line source: point source stretched over the entire width (length) of the field or plot, perpendicular to the direction of spread

Area source: A "line source with depth"; width is considerably more than 1% of the distance of the field.

← Partial definitions

More classifications of dispersal gradients (or disease gradients resulting from dispersal gradients):

- **Primary gradient:**
  - **All diseased individuals (infections) are due to spores (or other infectious units) originating at the original inoculum source**
    - Essentially: dealing with **primary infections** (whether the epidemic is monocyclic or polycyclic)
    - May continue over short or long times
- **Secondary gradient:**
  - **Diseased individuals (infections) are due to inoculum produced outside of the original inoculum source (and in the host population of interest)**
    - Essentially: dealing with **secondary infections** and polycyclic epidemics
- Some gradients may be a mixture of both



Models of disease (dispersal) gradients

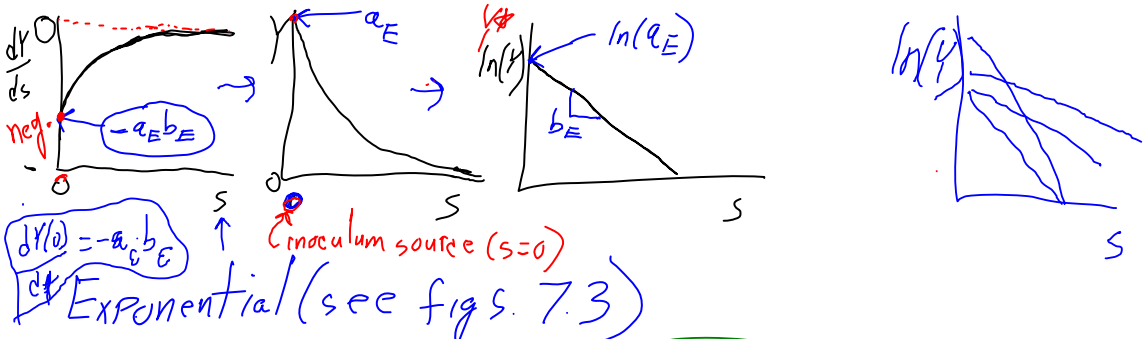
- As with disease progress curves, epidemiologists routinely use models (both simple and very complex) to characterize disease gradients
- Most common (and simple) models with acceptable biological realism are deterministic, either in the form of differential equations, integrated nonlinear models, or linearized equations
- **Notation:**
  - ***s***: distance (as in space)
  - ***Y***: disease intensity in absolute units (lesions, infections), or even spore density
    - Since many gradients are concerned with numbers of lesions (or even spores), we start with *Y*
    - ***Y(s)***: ***Y*** at distance ***s***; also Y_s
  - ***y***: disease intensity as a proportion
  - ***a*** and ***b*** (with subscripts): parameters
  -
  - ***dY/ds***: absolute rate of change in ***Y*** with change in distance (***s***); steepness of the gradient
- Model forms will be (**should be!**) quite familiar to you (by now)
  - These do **not** necessarily align with monocyclic and polycyclic diseases (so, *be careful* in interpretation!)

Exponential model (negative exponential)

- First used by Frampton (1942) and Gregory and Read (1949) for gradients
  - But, many call it the Kiyosawa & Shiyomi model (after their 1972 paper)
- ***dY/ds = - b_EY***
-
  - Absolute rate is *negative*, because *Y* declines with distance
  - Rate is proportional to *Y*, meaning: the higher the *Y*, the steeper the decline in *Y* ***(the greater the density of disease, the greater the decline over distance)***
  - ***b_E*: dispersal or spread parameter**; units of *1/distance*
    - Function of dispersal process (physics) and properties of host and pathogen
    - Higher susceptibility, or higher aggressiveness (higher infection efficiency), or less dense crop, for example, may result in lower b_E (closer to 0)
    - Large b_E means steep gradient

Exponential model

- ***dY/ds = - b_EY*** ⟶ ***Y = $a_E$exp[- b_Es]***
  - ***a_E:***
    - parameter; constant of integration; value of *Y* at *s*=0 , *Y(0) or* Y_0 (because exp[-b_E0] = exp[0] = 1)
    - Indicator of height of the curve
    - Measure of **'source strength'** -- *estimated* (predicted) amount of inoculum **or disease** at the source
    - Remember: in experiments, one often does not directly observe *Y* at the source
- ***ln(Y) = ln(a_E) - b_Es***
- Linear model in terms of ***ln(a_E)*** and ***b_E*** : ***Y* = a_* - b_*s***
  - Note: in regression analysis: it is assumed that there are pluses (+) between terms. Thus, one will obtain a negative slope, even though b_* is really positive (the minus is part of the model)



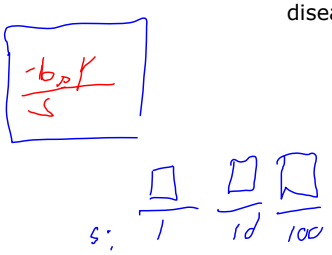
$$\frac{dY}{ds} = -b_\varepsilon Y \rightarrow Y = a_\varepsilon \cdot e^{-b_\varepsilon s} \rightarrow$$

$$\ln(Y) = \ln(a_\varepsilon) - b_\varepsilon \cdot s$$

$$Y^* = a_\varepsilon^* - b_\varepsilon \cdot s$$

Power model (power-law or inverse power model)

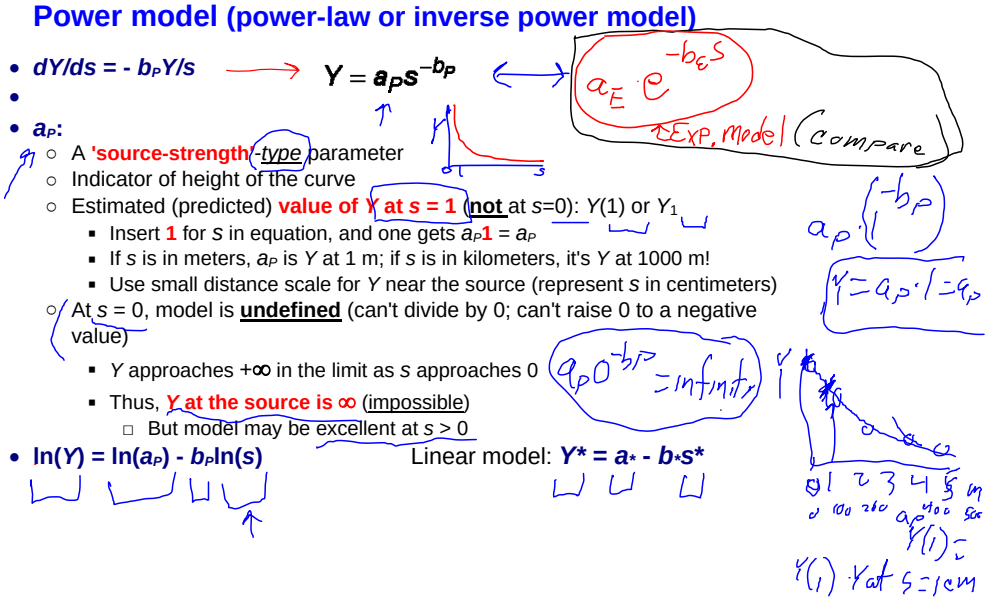
Compare with (negative) exponential model for disease gradient,

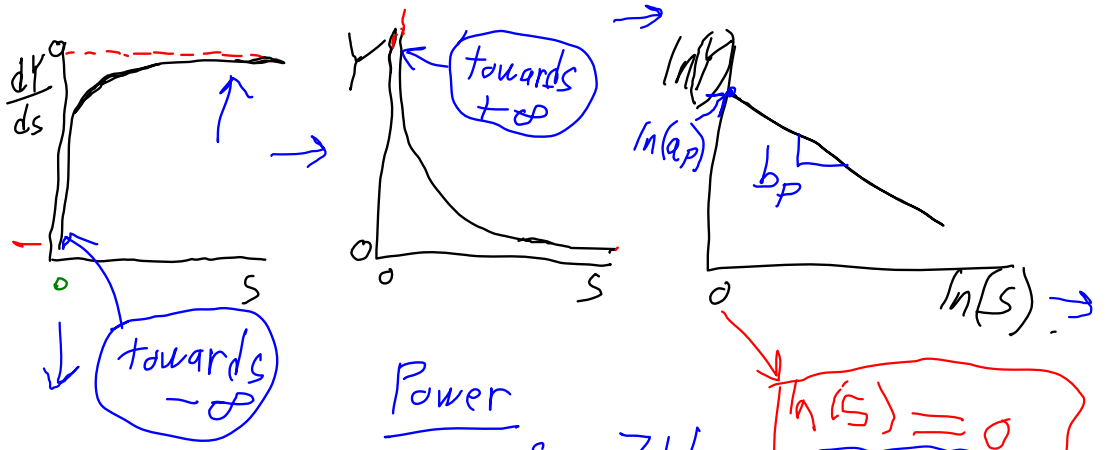


- Sometimes called the **Gregory** model because he was an early user of the equation (reviewed in 1968)
- ***dY/ds = - b_PY/s***
-
  - Absolute rate is negative, because *Y* declines with distance
  - Rate is proportional to *Y*, meaning: the higher the *Y*, the steeper the decline
  - Rate is *also* inversely proportional to *s*, meaning: rate gets smaller (closer to 0) at increasing distance **(see the 'dilution' argument later)**
    - (but *Y* also gets smaller; thus more complicated)
  - ***b_P***: dispersal or spread parameter; ***unitless*** (1/distance is explicitly in the model)
  - Function of dispersal process and properties of host and pathogen
  - Direct interpretation is trickier (as will be shown later)

Power model (power-law or inverse power model)

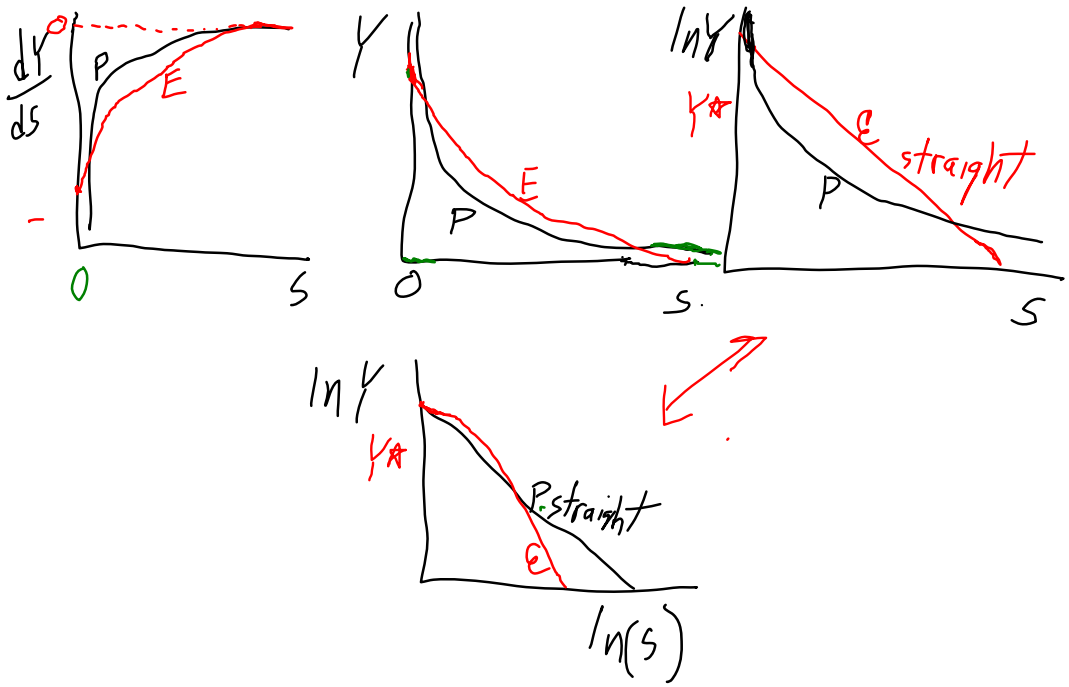
- ***dY/ds = - b_PY/s*** → $Y = a_P s^{-b_P}$
-
- ***a_P*:**
  - A **'source-strength'-*type*** parameter
  - Indicator of height of the curve
  - Estimated (predicted) **value of *Y* at *s* = 1** (**not** at *s*=0): *Y*(1) or Y_1
    - Insert **1** for *s* in equation, and one gets $a_P 1 = a_P$
    - If *s* is in meters, a_P is *Y* at 1 m; if *s* is in kilometers, it's *Y* at 1000 m!
    - Use small distance scale for *Y* near the source (represent *s* in centimeters)
  - At *s* = 0, model is **undefined** (can't divide by 0; can't raise 0 to a negative value)
    - *Y* approaches +∞ in the limit as *s* approaches 0
    - Thus, ***Y* at the source is ∞** (impossible)
      - But model may be excellent at *s* > 0
- **ln(*Y*) = ln(a_P) - $b_P$ln(*s*)** Linear model: ***Y** = a_* - b_*s****





$\frac{dY}{dS}$ $S = 1$

dS 0.1
0.01
0.001
0.0001
0.00001



Modified power model (add a constant, λ)

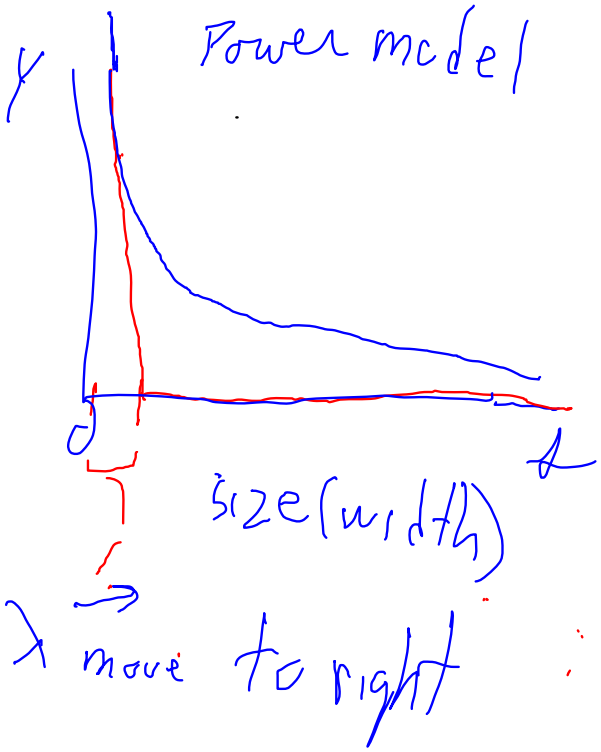
- $dY/ds = - b_P Y/(s+\lambda)$ → $Y = a_P(s+\lambda)^{b_P}$
- $\ln(Y) = \ln(a_P) - b_P \ln(s+\lambda)$
- a_P: Indicator of height of the curve
- λ: either arbitrary constant or approximate size of the source
- At s = 0, all relevant calculations are possible…

$$Y_0 = a_P(\lambda)^{-b_P}$$

is the '*true*' **source strength** (comparable to a_E of exponential)

- If one inserts $s = 1-\lambda$ into model, one gets a_P
  - Thus, a_P is predicted Y at $1-\lambda$ units of distance from center of source (shifting disease gradients λ to the left ($-\lambda$).
    - Think, then, of λ as '*diameter*' of inoculum source

$$a_P(0+\lambda)^{-b_P} = a_P(\lambda)^{-b_P}$$



Modified power model (add a constant, λ)

- $dY/ds = - b_P Y/(s+\lambda)$ $\quad Y = a_P(s+\lambda)^{b_P}$
- $\ln(Y) = \ln(a_P) - b_P \ln(s+\lambda)$ nonlinear
- a_P: Indicator of height of the curve
- λ: either arbitrary constant or approximate size of the source
- **Plots are similar to standard power model (except *very* close to source)**
- **Standard power model is just a special case (with $\lambda = 0$)**

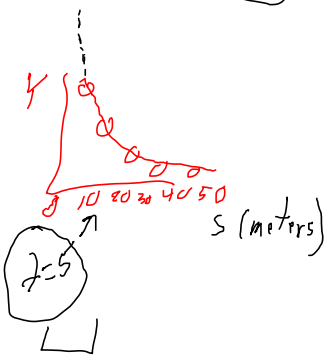
- Ad hoc method when one needs a finite source-strength estimate:
- For λ: use one-half the distance between 0 and the first distance with observed *Y*

Or, fit model with non-linear regression **(model is NOT linear because λ is inside the log function)**

$\lambda => 0. \rightarrow 9.999999$

Power law: $Z = \alpha X^{\beta}$

$\ln(Y) = \ln(a_p) - b_p \ln(s+5)$



$\ln(Y) = \ln(a_p) - b \ln(s+.01)$

$\ln(s+.10)$ ← choose λ:
$\ln(s+1)$ ←
$\ln(s+2)$ highest R^2
↓ lowest MSE (SSE)
$\ln(s+9)$ ↓
$\ln(s+9.9)$

Thus, there are two **fundamental** model forms for disease gradients

Exponential:

$dY/ds = - b_E Y$

$Y = a_E e^{-b_E s}$

$ln(Y) = ln(a_E) - b_E s$

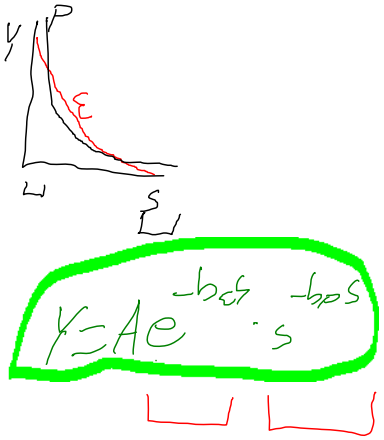
Modified power (λ), or power:

$dY/ds = - b_P Y/(s+\lambda)$

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

$\ln(Y) = \ln(a_P) - b_P \ln(s+\lambda)$

$\lambda = 0$ regular power



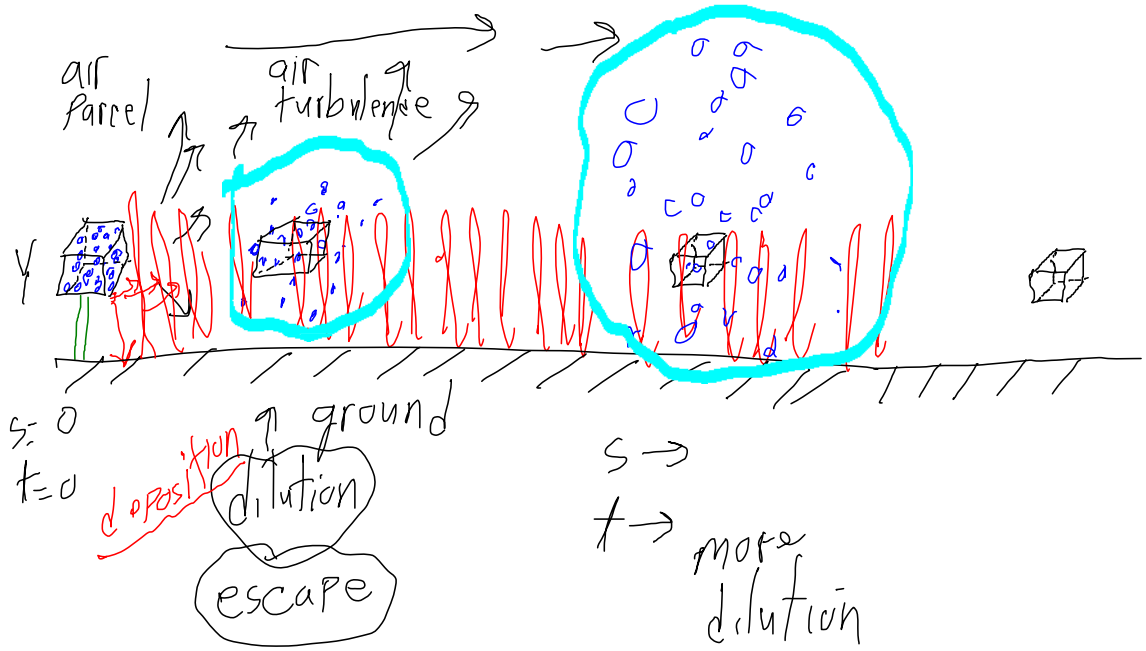
Spore **deposition** (low air turbulence)

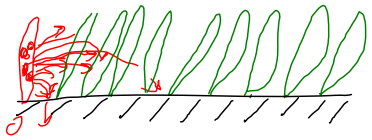
Spore **dilution & escape** from canopy (high air turbulence)

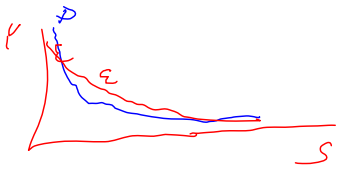
- high plant density
- low plant density
- Other models have been proposed, including combination of exponential and power functions (for mixture of deposition and dilution/escape). These two are sufficient for many studies

See: Aylor (1999) -- citation in book;
Also: **Reynolds (Dec. 2011). *Phytopathology* 101: 1465-70.**

splash dispersal

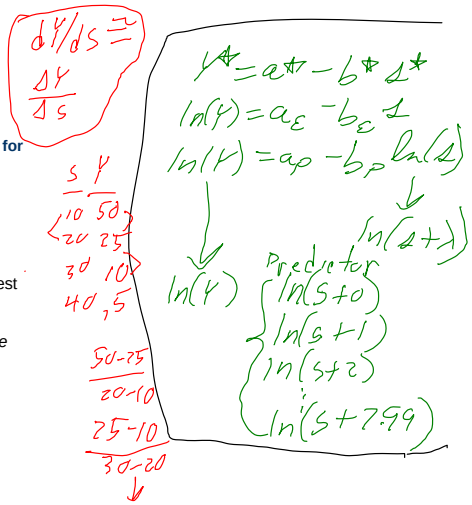






Use of models

- **Model selection**
  - **Graphical and statistical approaches** (as with disease progress curves)
  - **Plots** of *dY/ds* (estimated) versus *s*, *Y* versus *s*;
    and ***ln(Y):s*** **and** ***ln(Y):ln(s+λ)*** **λ=0 <---best way**
    - **Unlike the case with ~~temporal~~ analysis, *Y** is the same for the two models, but *s** is not**
      - Either *s* or ln(*s*+**λ**)
  - **Model fitting**
    - Ordinary linear least squares
      - ln(*Y*) versus *s* or ln(*Y*) versus ln(*s*+λ)
        - Use **λ**=0 (standard)
        - Or, ~~try~~ several **λ** values, from 0 up to the shortest observed distance
        - Analogous to trying different models
        - My preference: *1/2 between 0 and first distance*
    - Nonlinear least squares (or maximum likelihood or Bayesian)
      - Can directly estimate **λ**
    - Evaluation of **residuals**



Reading assignment:

Chapter 7
Pages 173-190