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."


From HLIR model (coupled differential equations):

\[ \frac{dH}{dt} = -\beta I H = - (\alpha \theta \psi) I H \quad \text{or} \quad \frac{dY}{dt} = + \beta I H = + (\alpha \theta \psi) I H \]

\( \theta \): 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)
**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).
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.

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 \( Y \)
- \( y \): disease intensity as a proportion
- \( a \) and \( b \) (with subscripts): parameters
- \( \frac{dY}{ds} \): 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)
- \( \frac{dY}{ds} = -bEY \)
  - 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)
  - \( bE \): 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 \( bE \) (closer to 0)
    - Large \( bE \) means steep gradient
Exponential model

\[ \frac{dY}{ds} = -bEY \quad \Rightarrow \quad Y = aE \exp[-bEs] \]

- **\( aE \):**
  - Parameter; constant of integration; value of \( Y \) at \( s=0 \), \( Y(0) \) or \( Y_0 \) (because \( \exp[-bE0] = \exp[0] = 1 \))
  - 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

- \( \ln(Y) = \ln(aE) - bEs \)

- Linear model in terms of \( \ln(aE) \) and \( bE \):
  \[ Y^* = a^* - b^*s \]

- 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)
Power model (power-law or inverse power model)

- Sometimes called the Gregory model because he was an early user of the equation (reviewed in 1968)

\[ \frac{dY}{ds} = -bPY/s \]

- 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)

\[ \ln(Y) = \ln(a_P) - b_P \ln(s) \]

Linear model: \( Y^* = a - b\cdot s^* \)

Compare with (negative) exponential model for disease gradient,
**Modified power model (add a constant, \( \lambda \))**

- \( dY/ds = -bPY/(s+\lambda) \)
- \( \ln(Y) = \ln(aP) - bP\ln(s+\lambda) \)

- \( aP \): Indicator of height of the curve
- \( \lambda \): either arbitrary constant or approximate size of the source
- At \( s = 0 \), all relevant calculations are possible...

\[ Y = aP(\lambda)^{-bP} \]

is the 'true' source strength (comparable to \( aE \) of exponential)

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

- \( \frac{dY}{ds} = -b_P Y/(s+\lambda) \)
- \( \ln(Y) = \ln(a_P) - b_P \ln(s+\lambda) \)

\( a_P \): Indicator of height of the curve
\( \lambda \): 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 \( \lambda \): 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 \( \lambda \) is inside the log function)

\( \lambda = 20 \implies a_P = 0.99999 \)

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

\[
\begin{align*}
\ln(s+10) & \leftarrow \text{choose } \lambda : \\
\ln(s+1) & \leftarrow \text{highest } R^2 \\
\ln(s+2) & \leftarrow \text{lowest } MSE
\end{align*}
\]
Thus, there are two *fundamental* model forms for disease gradients:

**Exponential:**
\[
\frac{dY}{ds} = -bEY \\
Y = a_E e^{-b_E s} \\
\ln(Y) = \ln(a_E) - b_E s
\]

**Modified power \((\lambda)\), or power:**
\[
\frac{dY}{ds} = -b_P Y/(s+\lambda) \\
Y = a_P (s + \lambda)^{b_P} \\
\ln(Y) = \ln(a_P) - b_P \ln(s+\lambda)
\]

- Spore deposition
- Spore dilution & escape from canopy
  - Low air turbulence
  - High air turbulence

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.

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 $\ln(Y)$ versus $s$ and $\ln(Y)$ versus $\ln(s)$

- Unlike the case with temporal analysis, $Y^*$ is the same for the two models, but $s^*$ is not
  - Either $s$ or $\ln(s + \lambda)$

- Model fitting
  - Ordinary linear least squares
    - $\ln(Y)$ versus $s$ or $\ln(s)$
      - Use $\lambda=0$ (standard)
      - Or, try several $\lambda$ 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 $\lambda$
    - Evaluation of residuals

\[ \frac{dY}{ds} = \frac{\Delta Y}{\Delta s} \]
\[ \ln(Y) = a_0 - b_e - d \]
\[ \ln(Y) = a_p - b_p \ln(\lambda) \]