Empirical Modelling of Spatio-temporal Variation in Meningitis Incidence

Niamey, November 2009

Peter J Diggle, Barry Rowlingson and Michelle Stanton

School of Health and Medicine, Lancaster, UK

Supported by:

Ethiopian Ministry of Health,
Madeleine Thomson and colleagues, IRI Columbia NY,
Mohammed Yassin, Liverpool School of Tropical Medicine
Where did 2009 go?

- **Key objective:** short-term forecasting
  - spatial scale: district-level
  - time-scale: weekly
  - forecast lead-time: 1, 2, ... weeks

- **Funding applications:**
  - Menigitis Research Foundation: unsuccessful
  - Medical Research Council: pending
  - ... ?

- **PhD recruitment:** Michelle Stanton, Lydiane Agier
Outline

• **Ethiopian data**: weekly time-series at woreda-level, July 2002 – June 2008:
  – number of incident cases
  – land-surface temperature
  – rainfall

• **Exploratory analysis**:
  – **temporal**: country-wide incidence
  – **spatial**: woreda-level incidence
  – **spatio-temporal**: animation
Proposed modelling framework:

- multiplicative decomposition of incidence into temporal and spatio-temporal components
- regression adjustments for time-lagged environmental variables
- stochastic model to mimic short-term spread of epidemic amongst neighbouring woredas
Ethiopian data: temporal structure

Time series at weekly intervals in each of 567 woredas:

- incident counts, July 20002 to June 2008
- average land surface temperature (LST)
- total rainfall (TR)
- LST and TR mapped onto woredas, using the IRI Data Library
Ethiopian data: weekly national incidence
Incidence on log-scale
Land surface temperature

![Graph showing land surface temperature changes over weeks. The graph has a y-axis labeled LST (Land Surface Temperature) and an x-axis labeled week. The data shows variability over time with some peaks and troughs.]
Total rainfall

![Graph showing total rainfall over weeks. The x-axis represents weeks from 0 to 300, and the y-axis represents rainfall from 0.00 to 0.35.]
Cross-correlation: incidence vs LST
Cross-correlation: incidence vs TR
Regression model for country-wide incidence

\[ \text{INCIDENCE} = \text{SEASONAL} \times \text{RESIDUAL} \]
Regression model for country-wide incidence

INCI DENCE = SEASONAL $\times$ RESIDUAL
Regression model for country-wide incidence

\[ \text{INCIDENCE} = \text{SEASONAL} \times \text{RAINFALL} \times \text{RESIDUAL} \]
Correlated residuals
Correlated residuals
Correlated residuals

![Graph showing autocorrelation function (ACF) of residuals with lags on the x-axis and ACF values on the y-axis. The ACF plot shows significant correlations at certain lags.]
A dynamic seasonal model

\[ Y_t = \log \text{incidence in week } t = A_t + B_t \cos(2\pi t/52) + C_t \sin(2\pi t/52) \text{ residual} \]

Regression coefficients modelled stochastically:

\[
\begin{align*}
A_t &= A_{t-1} + \epsilon_t^A \\
B_t &= B_{t-1} + \epsilon_t^B \\
C_t &= C_{t-1} + \epsilon_t^C
\end{align*}
\]

Can treat environmental variables similarly, but beware of over-elaboration
Seasonal variation in black smoke

Static model:
Seasonal variation in black smoke

Dynamic model:

![Graph showing log average black smoke level over weeks]
Ethiopian data: spatial structure
Ethiopian data: spatio-temporal structure

Animation of incident counts
Ethiopian data: spatio-temporal structure

Animation of incident counts shows familiar epidemic structure:

- long, quiescent periods punctuated by localised outbreaks
- local spread over short time-periods
Dynamic spatio-temporal seasonal model

\[ R_t(x) = \text{risk in week } t \]
\[ \log R_t(x) = A_t(x) + B_t(x) \cos(2\pi t/52) + C_t(x) \sin(2\pi t/52) \]

Regression coefficients modelled as random fields, for example:

\[ A_0 = A_0(x) \quad A_t(x) = \int A_{t-1}(x-u)w(u)du + \epsilon_t^A(x) \]

Conditionally independent Poisson counts:

\[ Y_t(x) = \text{incidence in week } t \]
\[ N_t(x) = \text{population in woreda } x \]

\[ Y_t(x) \mid R_t(x) \sim \text{Poisson}\{N_t(x)R_t(x)\} \]
Take-home messages

1. Spatial scale:
   - analyse at finest available spatial resolution
   - interpret at policy-relevant scale

2. Information synthesis:
   - environmental covariates:
     - spatially sparse ground-truth (eg met-stations)
     - and spatially dense surrogates (eg satellite data)
     - and physically based models
   - GIS layers (eg transport routes)
   - social context (eg major population movements)
Take-home messages

3. Correlation is your friend:
   - what is happening here and now
   - can help to predict what will happen somewhere else

4. An honest answer to any prediction problem is a probability distribution
   - sensitivity (true positive)
   - specificity (false positive)
   - timeliness (forecast horizon)

5. Open-source software implementation for access and portability

Examples: www.lancs.ac.uk/staff/diggle