Lectures on Spatio-temporal Point Processes

Peter J Diggle

Lancaster University and University of Liverpool

Outline

1. **Motivating examples**
   - The UK 2001 epidemic of foot-and-mouth disease (FMD)
   - Bovine tuberculosis in Cornwall, UK
   - Bramble canes
   - gastroenteric disease in Hampshire, UK

2. **Spatial point processes (briefly!)**
   - Regularity, randomness, aggregation
   - Standard model classes
   - First and second moment methods
   - Likelihood-based methods

3 Spatio-temporal point processes

- Discrete/continuous time/space
- Examples re-visited

4 Exploratory analysis

- Static plots
- Animations
- Second-moment summaries

5 Modelling strategies for temporally discrete processes

- Treat times as (ordered) categorical marks
- Build discrete-time transition models

6 Modelling strategies for temporally continuous processes

- Empirical modelling (Cox processes)
- Mechanistic modelling (conditional intensities)

7 Towards synthesis: de-compartmentalising spatial statistics

- Cressie's classification of spatial statistics
- Point processses and geostatistics
- Analyse problems, not data

8 Open question-and-answer session

The 2001 UK FMD epidemic

First confirmed case 20 February 2001

Approximately 140,000 at-risk farms in the UK
(cattle and/or sheep)

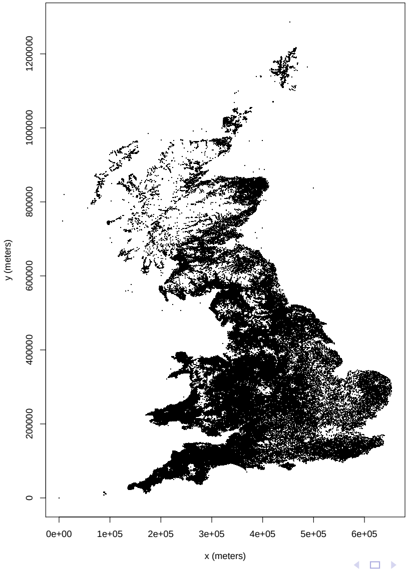
Outbreaks in 44 counties, epidemic particularly severe
in Cumbria and Devon

Last confirmed case 30 September 2001

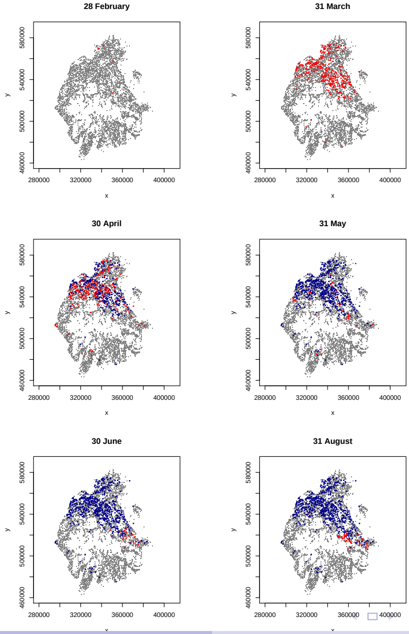
Consequences included:

- more than 6 million animals slaughtered (4 million for disease control, 2 million for "welfare reasons")
- estimated direct cost £8 billion

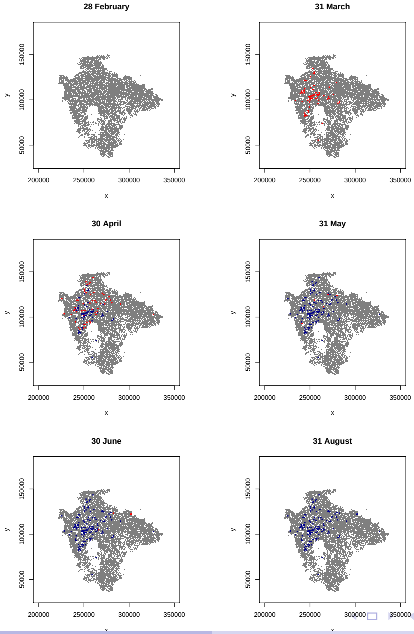
UK animal-holding farms in January 2001



Progress of the epidemic in Cumbria



Progress of the epidemic in Devon



Features of the epidemic

- **Predominantly short-range transmission between near-neighbouring farms**
- **Occasional long-range "jumps"**
- **Both reactive and pre-emptive culling strategies used to control spread of infection**

- **What factors affected the spread of the epidemic?**
- **How effective were control strategies in limiting the spread?**

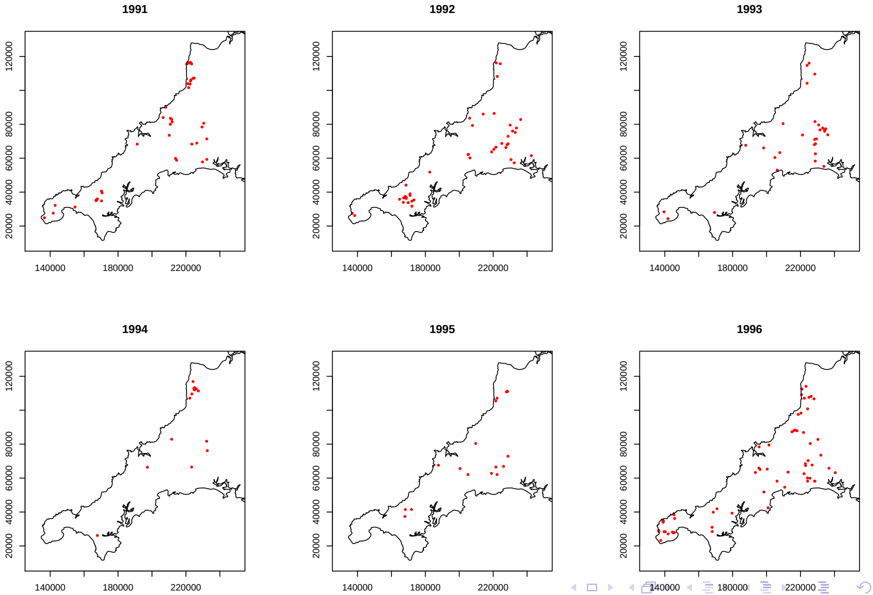
Animation at:

http://www.maths.lancs.ac.uk/staff/diggle/FMD

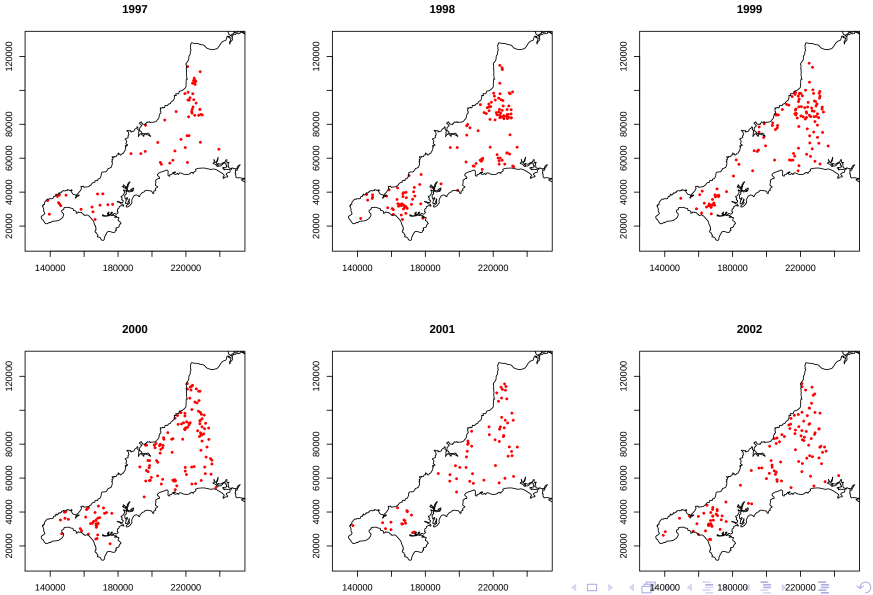
Bovine tuberculosis in Cornwall, UK

- **spatio-temporal distribution of disease in Cornwall, UK, over time-period 1991 to 2002**
- **cases identified from annual inspections**
- **several distinct genotypes**
- **prevalence has been increasing during the study-period**

Genotype 1


1991
1992
1993
1994
1995
1996
140000
180000
220000
20000
40000
60000
80000
120000

Genotype 2


1997
1998
1999
2000
2001
2002
20000
40000
60000
80000
120000
140000
180000
220000

Questions

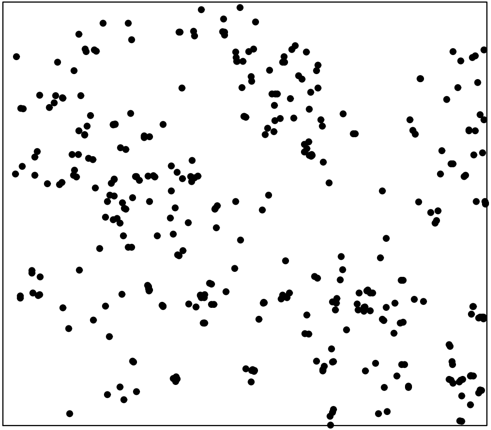
- **Is spatial variation in relative risk stable over time amongst all cases?**
- **Is spatial variation in relative risk stable over time within genotypes?**
- **Is spatial segregation between genotypes stable over time?**
- **Is primary mode of spread through direct contact between infected cows, or from local wildlife to cows?**

Bramble canes

- **positions of individual canes within a 9 metre by 9 metre plot**
- **newly emergent canes in year 0 generally survive to year 1**
- **a small proportion survive a second year**

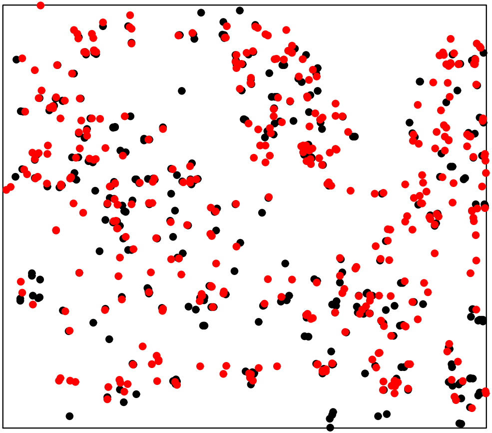
- **Is the spatial distribution in equilibrium?**
- **Is survival from year 1 to year 2 spatially neutral?**

Bramble canes, aged 0



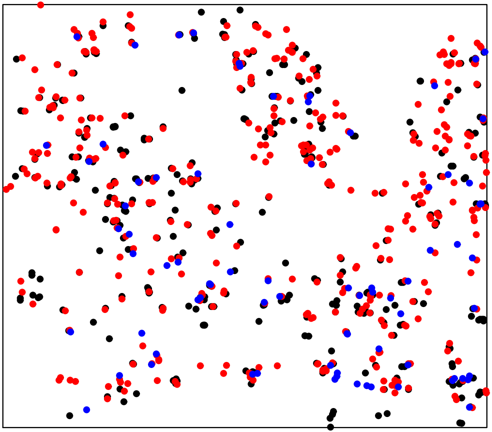
age 0 (n = 359)

Bramble canes, aged 0, 1



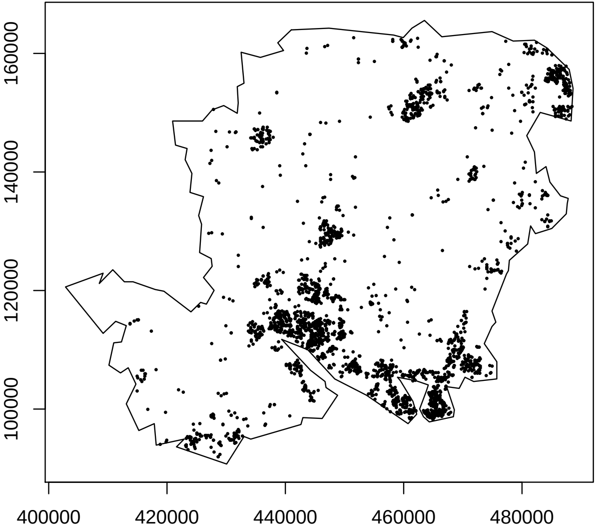
age 0 (n = 359) **age 1 (n = 385)**

Bramble canes, aged 0, 1, 2



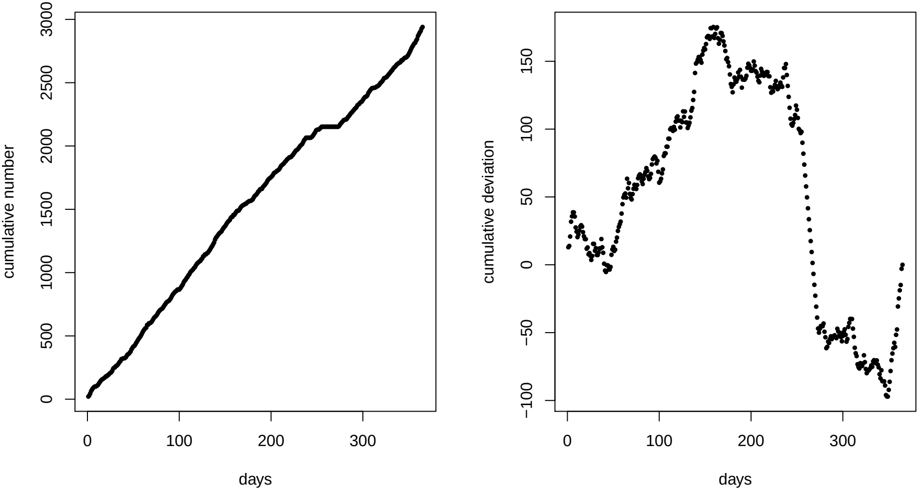
age 0 (n = 359) **age 1 (n = 385)** **age 2 (n = 79)**

Gastroenteric disease in Hampshire, UK: space



2940 incident cases, 1 January to 31 December 2001

Gastroenteric disease in Hampshire, UK: time



Gastroenteric disease in Hampshire, UK: questions

- **Can statistical modelling improve the timeliness of current UK health surveillance schemes?**
- **What is the normal spatio-temporal pattern of reported cases?**
- **Where and when do spatially and temporally localised anomalies in the incidence pattern occur?**
- **Can we achieve real-time surveillance?**

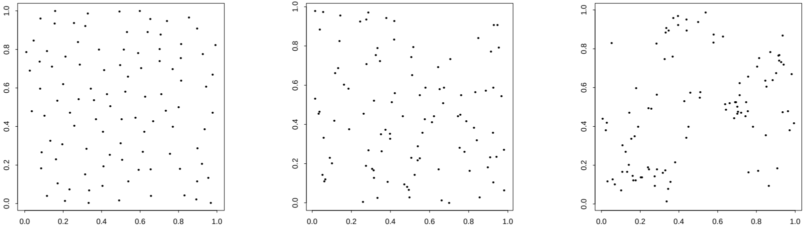
Animation at:

`http://www.maths.lancs.ac.uk/staff/diggle/aegiss`

2. Spatial point processes (briefly!)

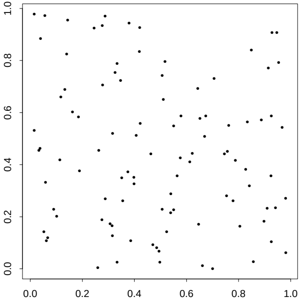
- Regularity, randomness, aggregation
- Standard model classes
- First and second moment methods
- Likelihood-based methods

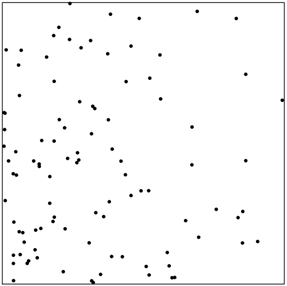
Regular → random → aggregated



Models: Poisson process

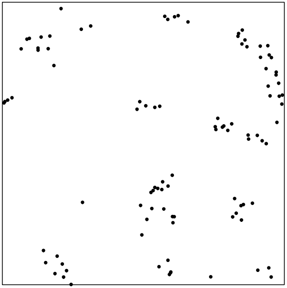
- **benchmark for complete spatial randomness**
- **events independently and uniformly distributed in space**
- **inhomogeneous process retains independence but allows non-uniform distribution in space (intensity function, $\lambda(x)$)**

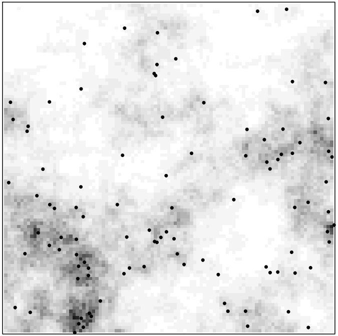




Models: Cox process

- **inhomogeneous Poisson process with stochastic intensity function**
- **more flexible than it sounds: often useful as model for aggregated patterns**
- **log-Gaussian Cox process a relatively tractable sub-class**

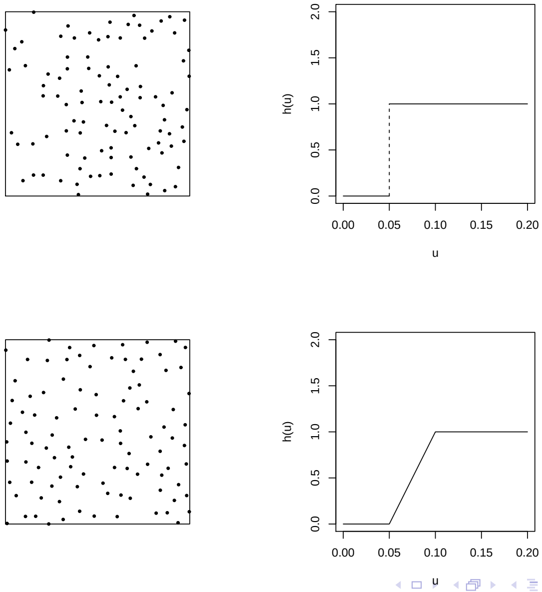




Models: Markov point process

- **conditional intensity at x depends on configuration of events in some neighbourhood of x**
- **usually defined by its likelihood ratio wrt a unit-intensity Poisson process**
- **properties generally intractable**
- **pairwise interaction process a useful sub-class for modelling regular patterns**

Pairwise interaction point processes



Point process intensities

First-order intensity function:

$$\lambda(x) = \lim_{|dx| \to 0} \left\{ \frac{E[N(dx)]}{|dx|} \right\}$$

Second-order intensity function:

$$\lambda_2(x, y) = \lim_{\substack{|dx| \to 0 \\ |dy| \to 0}} \left\{ \frac{E[N(dx)N(dy)]}{|dx||dy|} \right\}$$

Conditional intensity function:

$$\lambda_c(x|y) = \lambda_2(x, y)/\lambda(y)$$

Covariance density:

$$\gamma(x, y) = \lambda_2(x, y) - \lambda(x)\lambda(y).$$

Stationary, isotropic case

- $\lambda(x) \equiv \lambda = E[N(A)]/|A|$ (constant, for all A)
- $\lambda_2(x, y) \equiv \lambda_2(\|x - y\|)$ (depends only on distance)
- $\lambda_c(u|o) = \lambda_2(u)/\lambda$ (conditioning on event at the origin)
- $\gamma(u) = \lambda_2(u) - \lambda^2$

The K-function

$$K(s) = 2\pi\lambda^{-2}\int_0^s \lambda_2(r)rdr.$$

Ripley (1976, 1977)

Theorem. For a stationary, isotropic, orderly process,

$$K(s) = \lambda^{-1}E[\text{number of further events within distance } s \text{ of an arbitrary event}]$$

- gives a physical interpretation of $K(s)$,
- suggests a method of estimating $K(s)$ from data,
- explains why $K(s)$ is a useful descriptor of spatial pattern:
  - for clustered patterns, $K(s) > \pi s^2$
  - for regular patterns, $K(s) < \pi s^2$

Why is the K-function so useful?

Ease of interpretation

- $K(s)$ is a scaled expectation
- for a homogeneous, planar Poisson process, $K(s) = \pi s^2$
- $K(s)$ is invariant to random thinning.

Ease of estimation

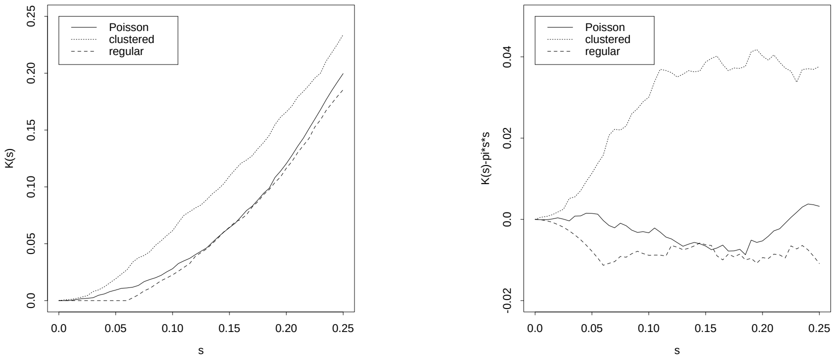
Data: $x_i \in A : i = 1, ..., n$

Recall that $K(s) = E(s)/\lambda$ where

$$E(s) = \text{E[number of further events within distance } s \text{ of an arbitrary event]}$$

Hence derive estimator $\hat{K}(s) = \hat{E}(s)/\hat{\lambda}$ by method of moments

Estimates of K(s) for three simulated patterns



Subtraction of πs^2 makes interpretation easier

Multivariate K-functions

λ_j = mean number of type j events per unit area.

$$\lambda_j K_{ij}(s) = \text{E[number of further type j events within distance s of an arbitrary type i event]}$$

Note: $K_{ij}(s) = K_{ji}(s)$

Benchmark results for multivariate K-functions

- type j events a homogeneous Poisson process $\Rightarrow K_{jj}(s) = \pi s^2$
- type i and type j processes independent $\Rightarrow K_{ij}(s) = \pi s^2$
- type i and type j events a random labelling
  $\Rightarrow K_{ii}(s) = K_{jj}(s) = K_{ij}(s) = K(s)$

Likelihood-based methods

- **Log-likelihood function for inhomogeneous Poisson process**

$$L(\lambda) = \sum_{i=1}^{n} \log \lambda(x_i) - \int_A \lambda(x)dx$$

- **For non-Poisson models, likelihood is generally intractable**
- **Monte Carlo methods used to enable classical or Bayesian inference**
  - **MCML**
  - **MCMC**

3 Spatio-temporal point processes

- Discrete/continuous time/space
- Examples re-visited

Classification of spatio-temporal point processes

Point process defined on $\mathbb{R}^2 \times \mathbb{R}^+$?

Point pattern a partial realisation on $A \times T \subset \mathbb{R}^2 \times \mathbb{R}^+$

Recall three earlier examples:

- **Gastroenteric disease in Hampshire, UK**
  Each reported case (x_i, t_i) could, in principle, have occurred at any place and time within $A \times T$.
- **The 2001 UK FMD epidemic**
  T is a continuous time-interval (1 January to 31 July 2001)
  A is a discrete set (locations of stock-holding farms)
- **Bovine tuberculosis in Cornwall, UK**
  A is continuous (Cornwall)
  T is discrete: incident cases only identified annually.

Shorthand terminology
spatially discrete/continuous, temporally discrete/continuous

4 Exploratory analysis

- Static plots
- Animations
- Second-moment summaries

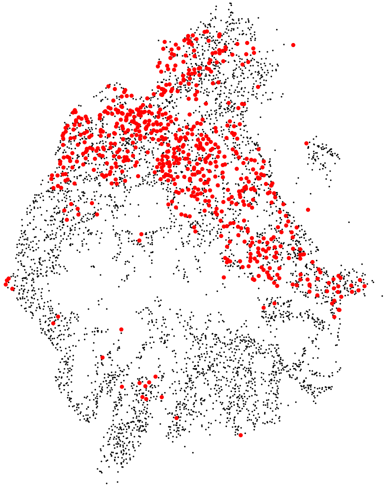
Static plotting

$$(x_i, t_i) : i = 1, ..., n$$

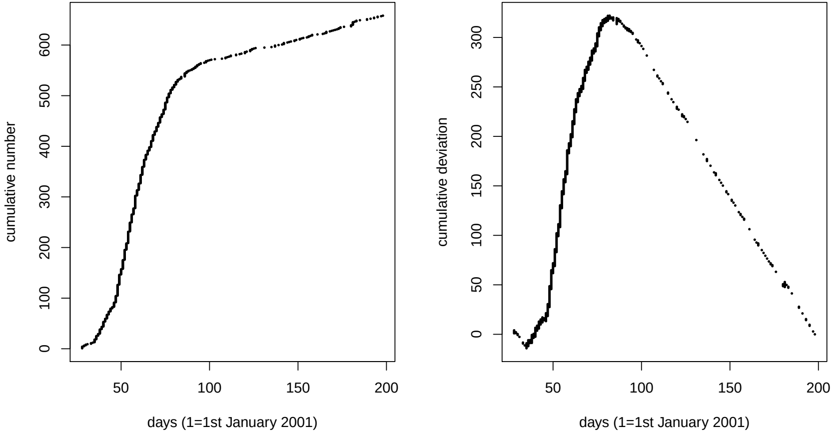
Static plots include:

- separate plots of $x_i : i = 1, ..., n$ and $t_i : i = 1, ..., n$
- plots of $x_i : i = 1, ...n_j$ within discrete time-intervals (t_{j-1}, t_j)
- plots of $x_i : i = 1, ..., n$ with $t_i : i = 1, ..., n$ codified through plotting symbol

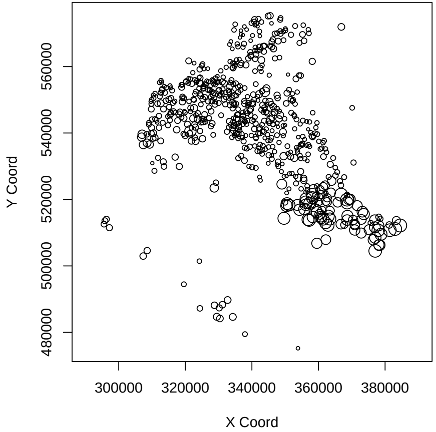
FMD2001: case-locations



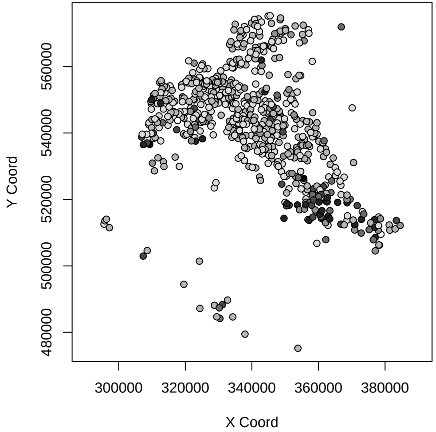
FMD2001: incident case-reports



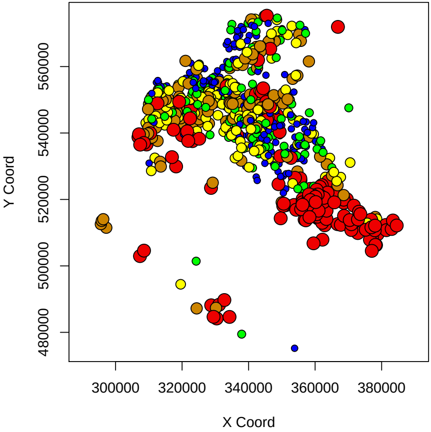
FMD2001: case-locations



FMD2001: case-locations

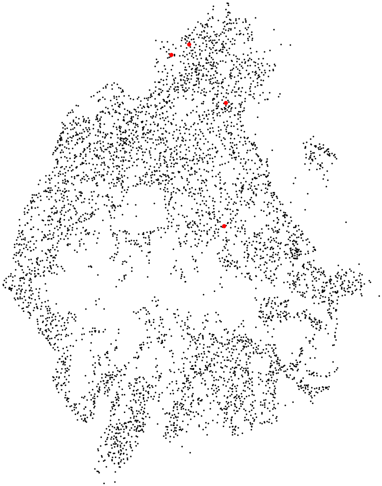


FMD2001: case-locations

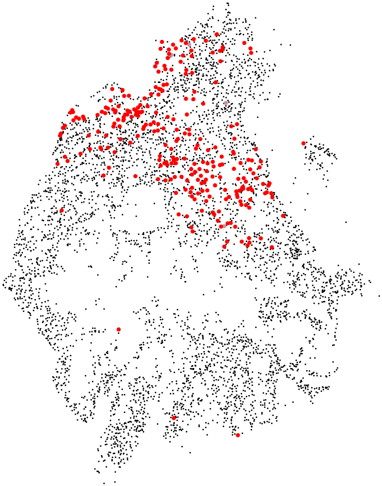


FMD2001: susceptible farms

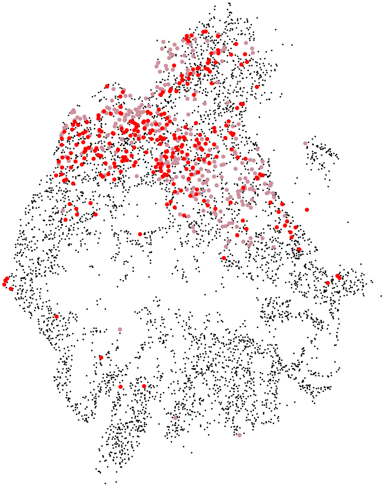
FMD2001: cases 0



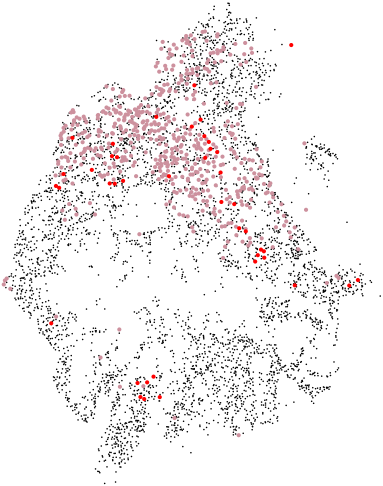
FMD2001: cases 0,1



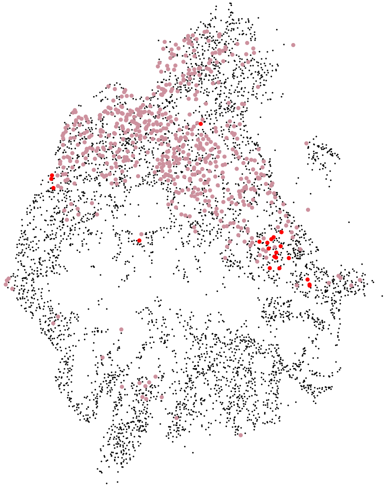
FMD2001: cases 0,1,2



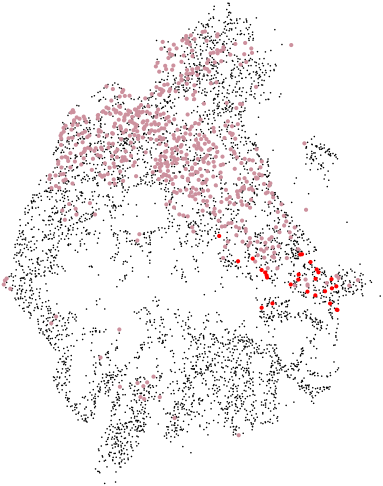
FMD2001: cases 0,1,2,3



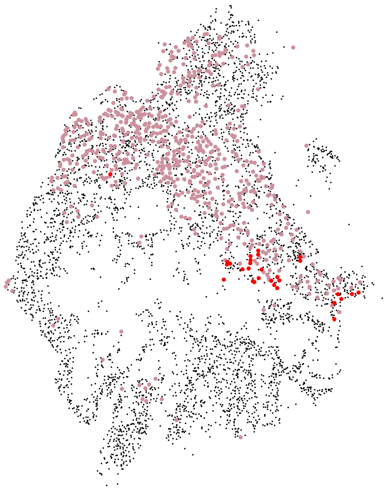
FMD2001: cases 0,1,2,3,4



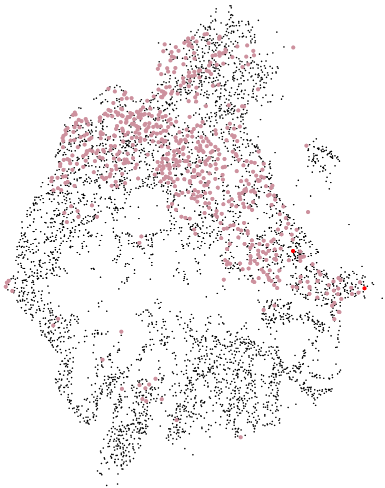
FMD2001: cases 0,1,2,3,4,5



FMD2001: cases 0,1,2,3,4,5,6



FMD2001: cases 0,1,2,3,4,5,6,7



Dynamic plotting

Dynamic plots:

- **animation**
- **link to GIS for more sophisticated displays**

Intensity-reweighted stationarity (spatial case)

Recall: intensity $\lambda(x)$, second-order intensity $\lambda_2(x, y)$

Define: pair correlation function

$$\rho(x, y) = \frac{\lambda_2(x, y)}{\lambda(x)\lambda(y)}$$

Process is intensity-reweighted stationary if $\rho(x, y) = \rho(||x - y||)$

Point process analogue of real-valued process with stationary variation about a non-constant mean

Continuous spatio-temporal point process intensities

intensity function

$$\lambda(x,t) = \lim_{|\,dx\,|,|dt| \to 0} \left\{ \frac{E[N(dx,dt)]}{|dx||dt|} \right\}$$

second-order intensity function

$$\lambda_2(x,y,s,t) = \lim_{|dx|,|dy|,|ds|,|dt| \to 0} \left\{ \frac{E[N(dx \times ds)N(dy \times dt)]}{|dx||dy||ds||dt|} \right\}$$

second-order conditional intensity

$$\lambda_c(x,s|y,t) = \lambda_2(x,y,s,t)/\lambda(y,t)$$

No explicit account taken of time's directional nature

The K-function(s)

Intensity-reweighted stationary case

$$\lambda_2(x,s,y,t)/\lambda(x,s)\lambda(y,t) = \rho(u,v)$$

Two-definitions: non-directional or directional in time

$$K(u,v) = 2\pi \int_{-v}^{v} \int_{0}^{u} \rho(x,t) x dx dt$$

$$K(u.v) = 2\pi \int_{0}^{v} \int_{0}^{u} \rho(x,t) x dx dt$$

- estimation by method of moments, analogous to spatial case
- non-directional or directional versions affect edge-correction
- joint non-parametric estimation of $\lambda(x,t)$ and $K(u,v)$ from a single realisation is fundamentally an ill-posed problem

5 Modelling strategies for temporally discrete processes

- Treat times as (ordered) categorical marks
- Build discrete-time transition models

Time as a categorical mark

- **formally analyse as a multivariate spatial point process**
- **but respect time-ordering when interpreting results**

Example: BTB data

Analysing the BTB data

Case-control analogy

- estimate intensity surfaces $\lambda(x, t)$ using kernel smoothing
- assess year-on-year changes using case-control methods:

$$\hat{\rho}(x) = \hat{\lambda}(x, t + 1)/\hat{\lambda}(x, t)$$

BTB data are also multivariate in another sense

- each case is genotyped
- main scientific interest is to quantify spatio-temporal patterns of segregation amongst different genotypes

Diggle, Zheng and Durr (2004)

Frequency distribution of BTB cases

Year	genotype						All
	1	2	3	4	5	6	
1989	4	10	11	2	3	2	30
1990	6	7	11	2	1	0	27
1991	23	7	7	4	0	0	41
1992	19	13	7	2	1	0	42
1993	19	1	7	2	0	0	29
1994	12	0	2	0	1	1	15
1995	9	3	5	0	0	0	17
1996	33	5	5	15	3	0	61
1997	42	5	8	6	3	1	64
1998	81	17	8	19	7	5	132
1999	79	14	17	6	14	5	130
2000	74	9	35	18	11	4	147
2001	37	5	15	9	1	1	67
2002	56	13	28	19	1	1	117
Total	494	109	166	104	26	20	919

Poisson process model

- genotyped case are independent Poisson processes with respective intensity functions $\lambda_k(x) : k = 1, .., m$
- factorisation of intensity functions: $\lambda_k(x) = \lambda_0(x)\rho_k(x)$
- type-specific probabilities,

$$p_k(x) = \frac{\lambda_k(x)}{\sum_{j=1}^{m} \lambda_j(x)} = \frac{\rho_k(x)}{\sum_{j=1}^{m} \rho_j(x)}$$

- relative risk surfaces,

$$\rho_{jk}(x) = \rho_j(x)/\rho_k(x) = \lambda_j(x)/\lambda_k(x) = p_j(x)/p_k(x)$$

Hence estimate relative risks or type-specific probabilities without modelling $\lambda_0(x)$

Spatial segregation

Multivariate process is completely unsegregated if:

- $\rho_k(x) = \alpha_k \rho(x)$
- $\rho_{jk}(x) = \rho_{jk}$
- $p_k(x) = p_k$

Multivariate process is completely segregated if $p_k(x) = 0$ or 1 for every location x

Kernel regression estimation

Approach is to treat spatial segregation as a non-parametric multinomial regression problem.

- **Kernel function, eg $w_0(x) = \exp(-||x||^2/2)$**
- **Bandwidths h_j, scaled kernels $w_j(x) = w_0(x/h_j)/h_j^2$**
- **Kernel regression weights, $w_{ij}(x) = w_j(x - x_i)/\sum_{k=1}^{n} w_j(x - x_k)$**

$$\hat{p}_j(x) = \sum_{i=1}^{n} w_{ij}(x) I(Y_i = j)$$

Choosing the bandwidth

- log-likelihood function

$$L(p_1, ..., p_m) = \sum_{i=1}^{n} \sum_{j=1}^{m} I(Y_i = j) \log p_j(x_i)$$

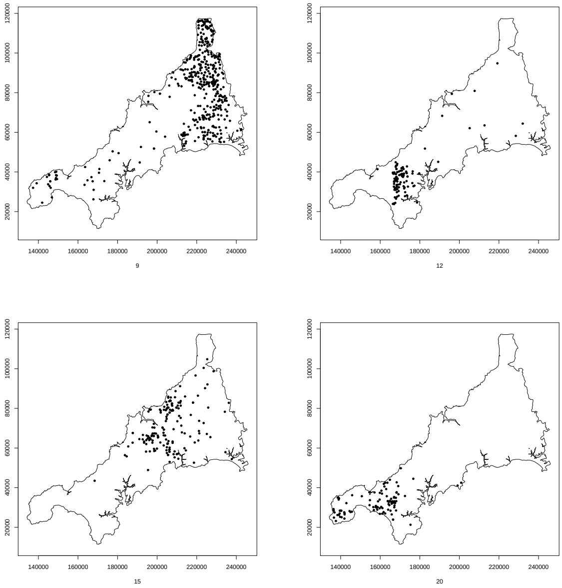
- cross-validated form of log-likelihood avoids $\hat{h} = 0$

$$L_c(h) = \sum_{i=1}^{n} \sum_{j=1}^{m} I(Y_i = j) \log \hat{p}_j^{(i)}(x_i)$$

$\hat{p}_j^{(i)}(x_i)$ is kernel estimator based on all data except (x_i, Y_i).

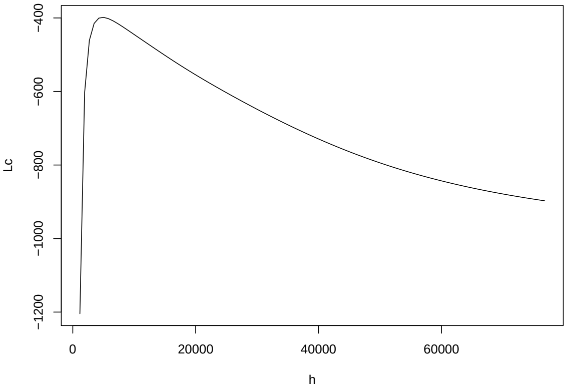
- common band-width for all m components gives $\sum_{j=1}^{m} \hat{p}_j(x) = 1$, for every location x.
- common bandwidth across genotypes and years eases interpretation

Spatial distribution of common genotypes, 1989 to 2002



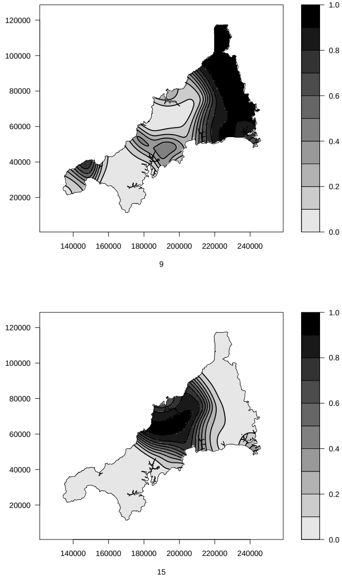
Spatial segregation of common genotypes

Cross-validated log-likelihood for four most common genotypes

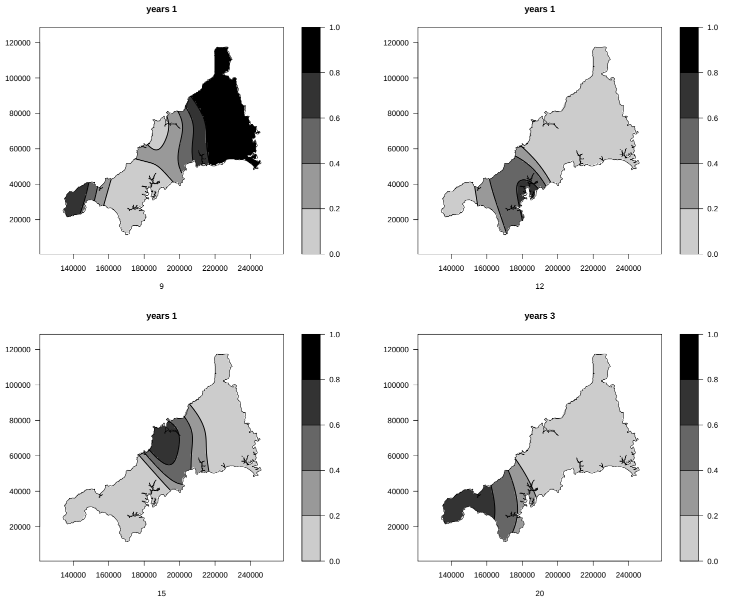


Suggests well-defined optimum band-width $\hat{h}$

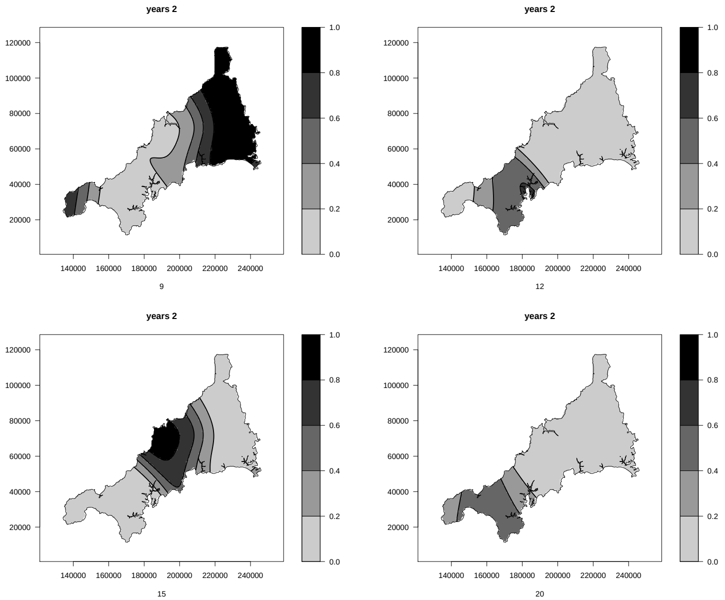
Estimated type-specific probability surfaces



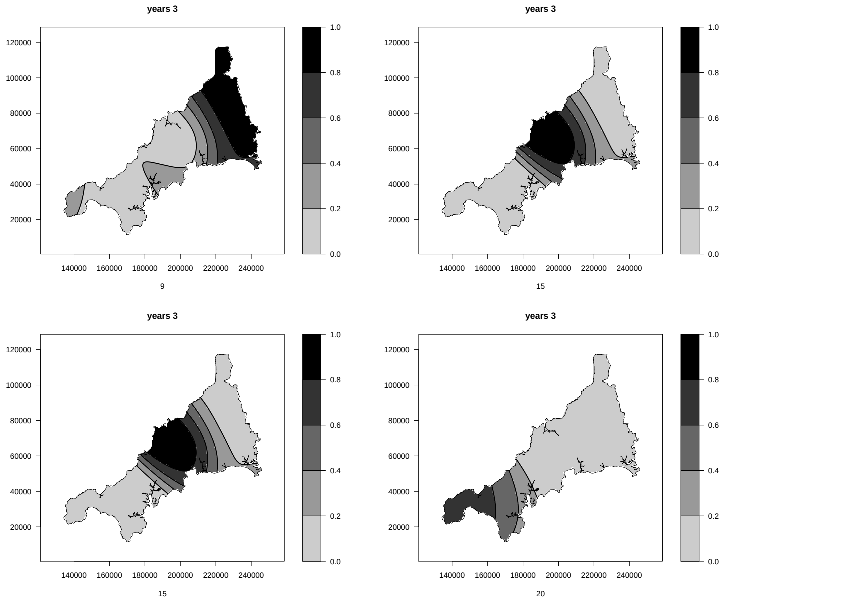
Does degree of segregation change over time? 1997/98



Does degree of segregation change over time? 1999/00



Does degree of segregation change over time? 2001/02



Conclusions for BTB data

- very strong spatial segregation between genotypes
- spatial segregation broadly consistent over time
- points towards local reservoirs of infection during development of epidemic
- alternative analysis strategy is likelihood-based predictive inference within multivariate log-Gaussian Cox process

$$\Lambda_k(x) = \exp\{\alpha_k + S_k(x)\}$$

- may also be possible to include covariate effects:

$$\alpha_k = \beta_{0k} + \beta_{1k} z(x)$$

Transition models between successive times

$$[\mathcal{P}] = [\mathcal{P}_1][\mathcal{P}_2|\mathcal{P}_1]...[\mathcal{P}_t|\mathcal{P}_{t-1},...\mathcal{P}_1]$$

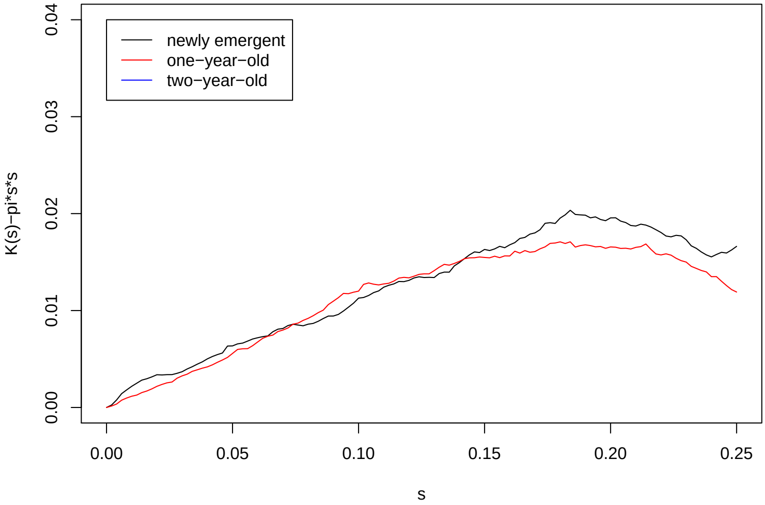
- **Markov in time a convenient working assumption**
- **may have some mechanistic justification when times represent successive generations**

Example: Brambles data

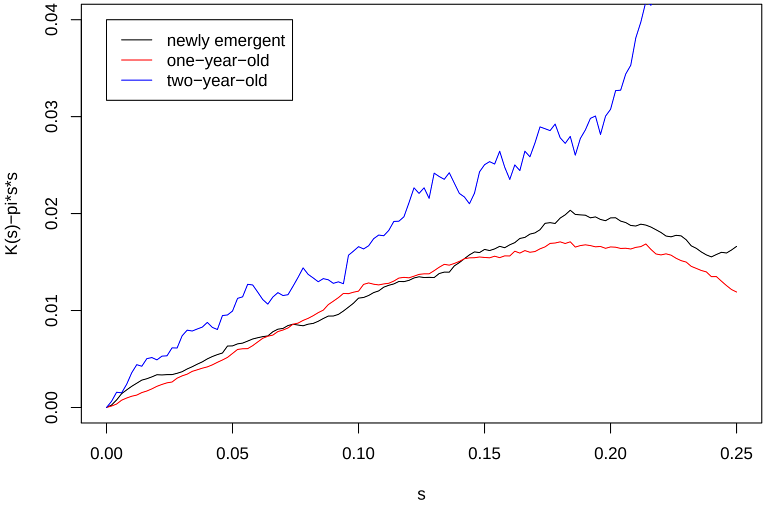
Exploratory analysis of brambles data

- **need to sample in successive years to capture spatio-temporal behaviour**
- **available data are a cross-sectional snap-shot**
- **can use K-functions to give partial answers to questions of scientific interest**

K-functions for brambles data, years 0, 1

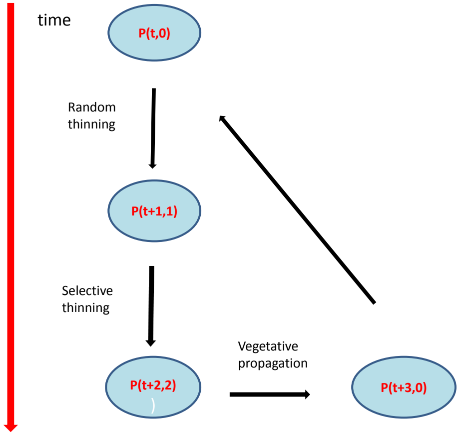


K-functions for brambles data, years 0, 1, 2



Analysis strategy for (hypothetical) longitudinal data

- **data show spatial aggregation on two distinct spatial scales**
- **most canes survive for two years**



6 Modelling strategies for temporally continuous processes

- Empirical modelling (Cox processes)
- Mechanistic modelling (conditional intensities)

Empirical

trans-Gaussian Cox process models – Poisson process with stochastic intensity

$$\Lambda(x, t) = \mathcal{F}\{S(x, t)\}$$

Mechanistic

conditional intensity function $\lambda(x, t|\mathcal{H}_t)$.

- $\mathcal{H}_t$ = complete history (locations and times of events) up to $t-$
- $\lambda(x, t|\mathcal{H}_t)$ = conditional intensity (hazard) for new event at location x, time t, given history $\mathcal{H}_t$

Empirical modelling of continuous processes

• **trans-Gaussian Cox process model relatively tractable, especially so if log-Gaussian: closed-form expressions for moment structure**

• **also flexible: able to generate a wide range of aggregated patterns**

- **scientifically natural if major determinant of pattern is environmental variation**
- **otherwise, often still a sensible empirical model**

Real-time disease surveillance

Problems with current surveillance system in UK include:

- **under-reporting by general practitioners (GP's)**
- **inconsistencies in reporting rates between GP's**
- **delays between onset and confirmation of cases**

The AEGISS project: can spatio-temporal statistical modelling improve the early detection of anomalies in the observed pattern of incident cases?

Possible data-sources

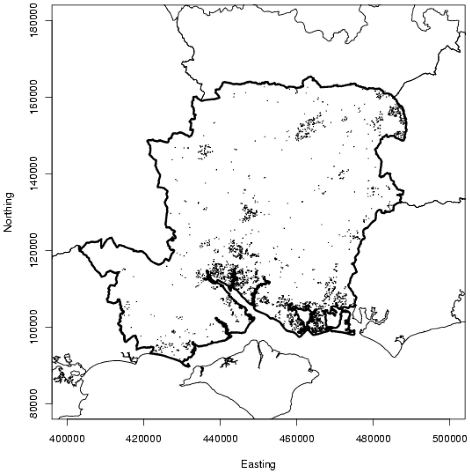
GP-reported data

- GP locations form a discrete spatial network
- known number of patients registered with each GP
- possibility of individual follow-up
- but not all GPs take part in the project
- and those who do take part do not report consistently

NHS Direct data

- 24hr phone-in advisory service introduced ca2000
- date and location (post-code) recorded electronically for each inquiry
- but spatial and temporal variation in pattern of usage is unknown

Gastroenteric disease in Hampshire, UK



3374 incident cases, 1 August 2000 to 26 August 2001

Statistical formulation

Notation

$$\begin{aligned} \lambda_0(x,t) &= \text{normal intensity of incident cases} \\ \lambda(x,t) &= \text{actual intensity of incident cases} \\ R(x,t) &= \text{spatio-temporal variation from normal pattern} \end{aligned}$$

$$\lambda(x,t) = \lambda_0(x,t)R(x,t)$$

Scientific objective

- Use incident data up to time t to construct predictive distribution for current "risk" surface, $R(x,t)$,
- hence identify potential anomalies, to be investigated further.

Diggle et al (2003), Diggle, Rowlingson and Su (2004)

Spatio-temporal model formulation

$$\lambda(x, t) = \lambda_0(x, t)R(x, t)$$

- $\lambda_0(x, t) = \lambda_0(x)\mu_0(t)$
- $R(x, t) = \exp\{S(x, t)\}$
- $S(x, t)$ = spatio-temporal Gaussian process:

  $\mathrm{E}[S(x, t)] = -0.5\sigma^2 \quad \mathrm{Var}\{S(x, t)\} = \sigma^2 \Rightarrow \mathrm{E}[R(x, t)] = 1$

  $\mathrm{Corr}\{S(x, t), S(x - u, t - v)\} = \rho(u, v)$
- conditional on $R(x, t)$, incident cases form an inhomogeneous Poisson process with intensity $\lambda(x, t)$

Estimating $\lambda_0(x)$: adaptive kernel smoothing

Data

$x_i : i = 1, ..., n$: locations of cases

Adaptive bandwidth

- $\tilde{\lambda}_0(x_i)$ = pilot estimator
- $\tilde{g}$ = geometric mean of $\tilde{\lambda}_0(x_i) : i = 1, ..., n$
- $h_i = h_0 \{\tilde{\lambda}_0(x_i)/\tilde{g}\}^{-0.5}$

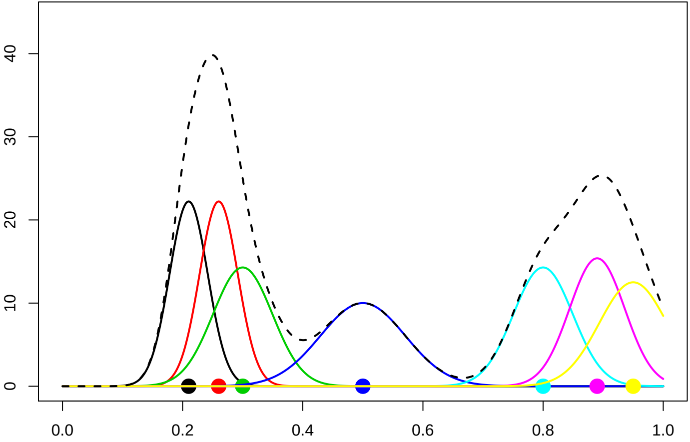
Kernel function

$\phi(u) = (2\pi)^{-1} \exp\{-0.5||x||^2\}$

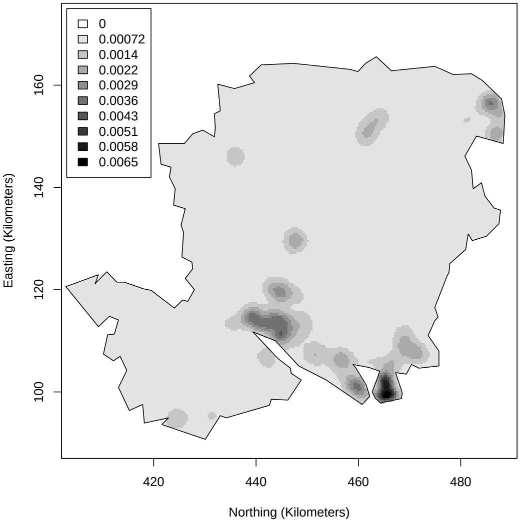
Estimator

$\hat{\lambda}_0(x) = n^{-1} \sum h_i^{-2} \phi\{(x - x_i)/h_i\}$

Locally adaptive kernel estimation



Estimate $\hat{\lambda}_0(\mathbf{x})$



Estimating $\mu_0(t)$: Poisson log-linear regression

- **Day-of-week effects**

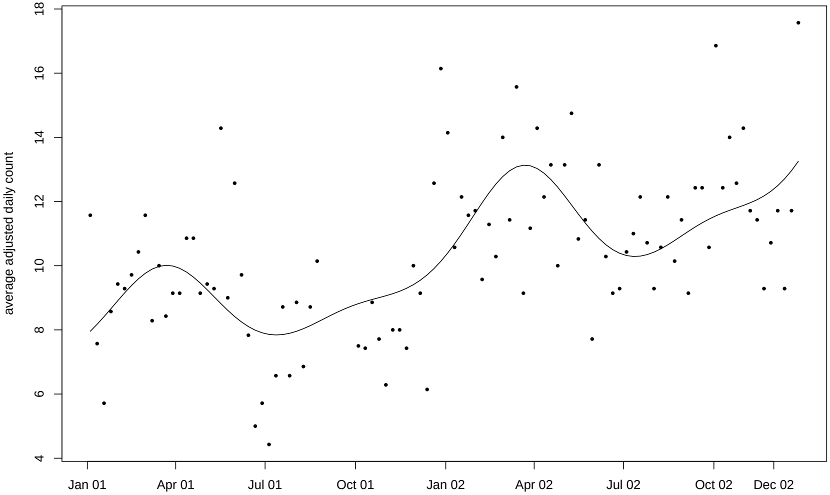
  7-level factor

- **Time-of-year effects**

  sine-cosine wave at frequency $\omega = 2\pi/365$, plus first harmonic

$$
\begin{aligned}
\log \mu_0(t) &= \delta_{d(t)} + \alpha_1 \cos(\omega t) + \beta_1 \sin(\omega t) \\
&+ \alpha_2 \cos(2\omega t) + \beta_2 \sin(2\omega t) + \gamma t
\end{aligned}
$$

Estimate $\hat{\mu}_0(t)$



Estimating σ^2 and $\rho(\mathbf{u}, \mathbf{v})$: moment-matching or MCML

Separability

$$\rho(\mathbf{u}, \mathbf{v}) = \rho_x(\mathbf{u})\rho_t(\mathbf{v})$$

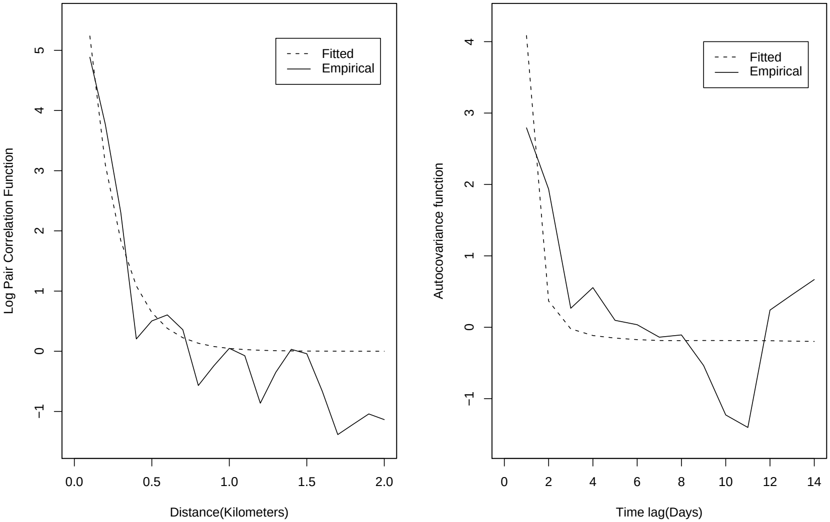
Spatial correlation

- arbitrary, but exponential gives reasonable fit
- $\rho_x(\mathbf{u}) = \exp(-||\mathbf{u}||/\theta)$

Temporal correlation (Markovian)

$$\rho_t(\mathbf{v}) = \exp(-|\mathbf{v}|/\phi)$$

Estimates $\hat{\rho}_x(\mathbf{u})$ and $\hat{\rho}_t(\mathbf{v})$



Spatial prediction

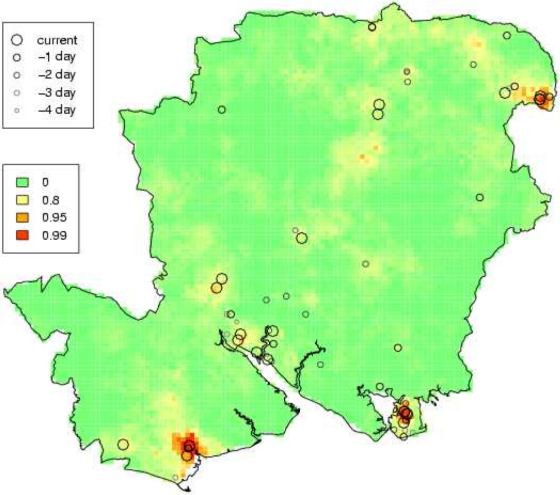
- **plug-in for estimated model parameters**
- **MCMC to generate samples from conditional distribution of $S(x, t)$ given data up to time t**
- **choose critical threshold value $c > 1$**
- **map empirical exceedance probabilities,**

$$p_t(x) = \mathrm{P}\left(\exp\{S(x,t)\} > c | \text{data}\right)$$

- **web-reporting with daily updates**

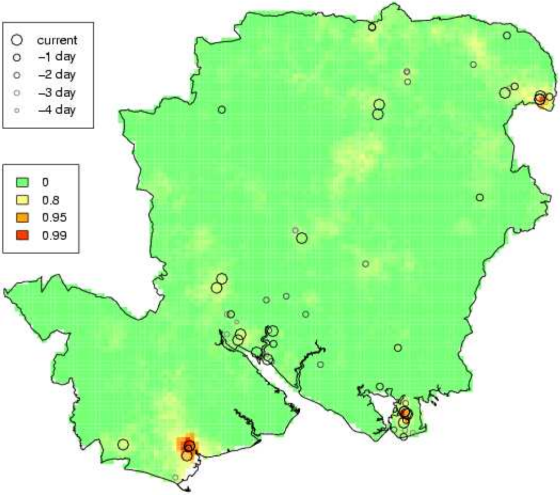
Do we need to take account of parameter uncertainty?

Spatial prediction: results for 6 March 2003



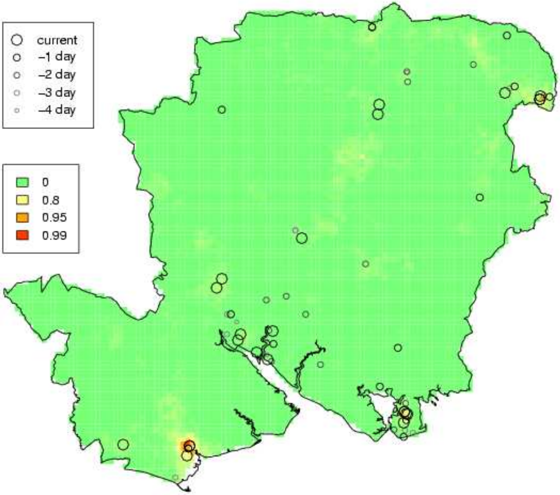
$\mathbf{c = 2}$

Spatial prediction: results for 6 March 2003



$c = 4$

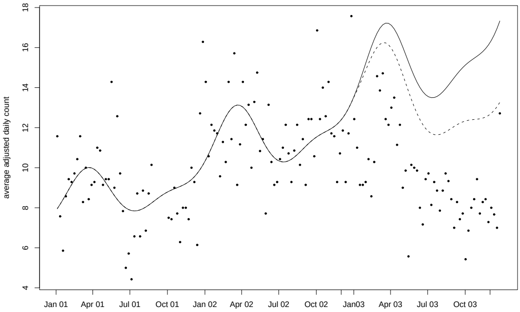
Spatial prediction: results for 6 March 2003



c = 8

Refining the methodology: AEGISS2

- **Monte Carlo maximum likelihood for parameter estimation**
- **tuning to balance sensitivity vs specificity**
- **stochastic model to replace $\mu_0(t)$**



Mechanistic modelling of continuous-time processes

Model-specification through the conditional intensity function exploits the directionality of time

- $\mathcal{H}_t$ = complete history (locations and times of events) up to $t-$
- $\lambda(x, t|\mathcal{H}_t)$ = conditional intensity (hazard) for new event at location x, time t, given history $\mathcal{H}_t$

Likelihood analysis

Log-likelihood for data $(x_i, t_i) \in A \times [0, T] : i = 1, ..., n$, with $t_1 < t_2 < ... < t_n$, is

$$L(\theta) = \sum_{i=1}^{n} \log \lambda(x_i, t_i | \mathcal{H}_{t_i}) - \int_0^T \int_A \lambda(x, t | \mathcal{H}_t) dxdt$$

- dynamic transformation to a unit-intensity Poisson process
- but usually analytically tractable, and numerically challenging

Partial likelihood analysis: spatially discrete process

Data $(x_i, t_i) \in A \times [0, T] : i = 1, ..., n,\ t_1 < t_2 < ... < t_n$

Potential locations, $\mathcal{X} = \{x_j^* : j = 1, ..., N\} \quad (x_i = x_j^*)$

Condition on locations x_i and times t_i, derive log-likelihood for observed ordering $1, 2, ..., n$

- $\mathcal{R}_i \subseteq \{1, , ..., N\}$ = risk-set at time t_i
- $$p_i = \frac{\lambda(x_i, t_i | \mathcal{H}_{t_i})}{\sum_{j \in \mathcal{R}_i} \lambda(x_j^*, t_j | \mathcal{H}_{t_j})}$$
- partial log-likelihood is

$$L_p(\theta) = \sum_{i=1}^{n} \log p_i$$

Partial likelihood analysis: spatially continuous process

Data $(x_i, t_i) \in A \times [0, T] : i = 1, ..., n,\ t_1 < t_2 < ... < t_n$

Potential locations, $\mathcal{X} = \subseteq \mathbb{R}^2$

Condition on locations x_i and times t_i, derive log-likelihood for observed ordering $1, 2, ..., n$

- $$p_i = \frac{\lambda(x_i, t_i | \mathcal{H}_{t_i})}{\int_{\mathcal{X}} \lambda(x, t_j | \mathcal{H}_{t_j}) dx}$$

- partial log-likelihood is

$$L_p(\theta) = \sum_{i=1}^{n} \log p_i$$

A model for the FMD epidemic (after Keeling et al, 2001)

Transmission of infection

$\lambda_{ij}(t)$ = conditional rate of transmission from farm i to farm j

Farm-specific covariates

$$\begin{aligned} n_{1i} &= \text{number of cows} \\ n_{2i} &= \text{number of sheep} \end{aligned}$$

Reporting delay

Reporting date is infection date plus τ
(latent period of disease plus reporting delay if any)

Statistical model

$$\lambda_{jk}(t) = \lambda_0(t) A_j B_k f(||x_j - x_k||) I_{jk}(t)$$

- **Baseline hazard:** $\lambda_0(t)$ (arbitrary)
- **Infectivity and susceptibility:**

$$\begin{aligned} A_j &= \alpha n_{1j} + n_{2j} \\ B_k &= \beta n_{1k} + n_{2k} \end{aligned}$$

- **Transmission kernel:** $f(u) = \exp\{-(u/\phi)^{0.5}\} + \rho$
- **At-risk indicator for transmission:** $I_{jk}(t)$

$I_{jk}(t) =$ 1 if farm k not infected and not slaughtered by time t and farm j infected and not slaughtered by time t; 0 otherwise

Fitting the model

- Rate of infection for farm k at time t is

$$\lambda_k(t) = \sum_j \lambda_{jk}(t)$$

- Partial likelihood contribution from ith case is

$$p_i = \lambda_i(t_i) / \sum_k \lambda_k(t_i)$$

- $\lambda_0(t)$ indeterminate, estimate identifiable parameters by maximising partial likelihood

FMD results

Common parameter values in Cumbria and Devon?

Partial likelihood ratio test: $\chi^2_4 = 2.98$

Parameter estimates

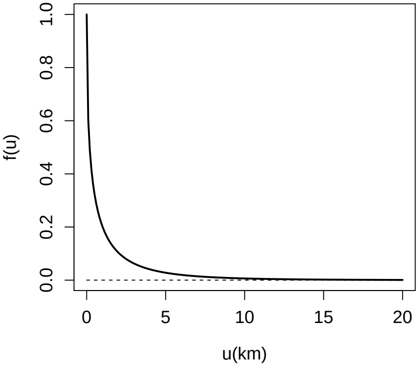
$$(\hat{\alpha}, \hat{\beta}, \hat{\phi}, \hat{\rho}) = (4.92, 30.68, 0.39, 9.9 \times 10^{-5})$$

(but partial likelihood ratio test rejects $\rho = 0$)

Standard errors

Available via usual asymptotic argument, but numerical estimates of information matrix are unreliable

Fitted transmission kernel

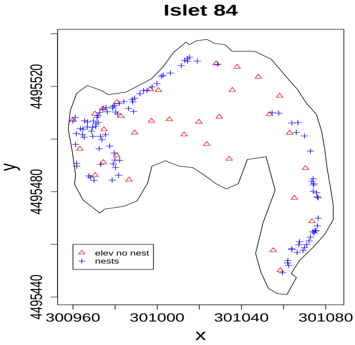


Qualitatively similar to estimate given in Keeling et al (2001)

Nesting colonies of common terns

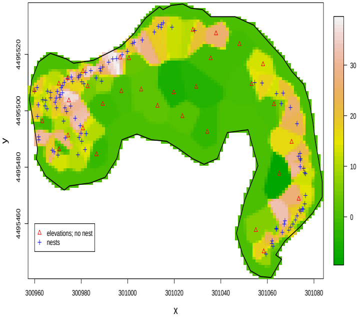
- **do birds exhibit a more complex form of interaction than simple inhibition?**

Islets 23 and 84



Coast boundaries (—), spatial locations of the nests (+), and other locations for which elevation is recorded (△) for islets 84 (left panel) and 23 (right panel)

Approximation of the elevation surface



Piece-wise constant approximation based on Voronoi tessellation of measurement locations

Conditional intensity formulation

$$\lambda(x, t|\mathcal{H}_t) = \lambda_0(t) \exp\{\beta z(x)\} g(x, t_i|\mathcal{H}_t)$$

- $g(x, t|\mathcal{H}_t)$ models dependence on locations of earlier nests
- $\beta z(x)$ models log-linear effect of elevation

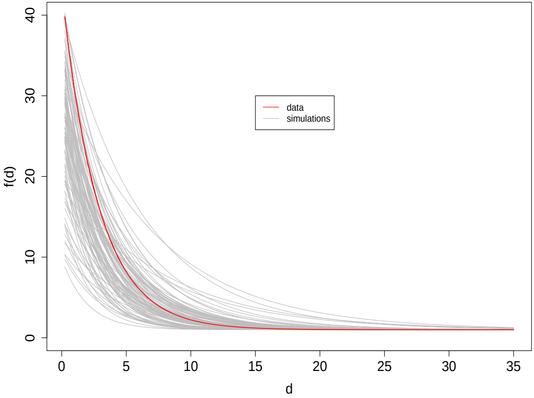
Models for $\mathbf{g}(\cdot)$

$$h(u) = \begin{cases} 0, & u \leq d_0 \\ 1 + \theta \exp\left\{-\frac{(u-d_0)^c}{\phi}\right\}, & u > d_0 \end{cases}$$

M1	$g(x, t \mid \mathcal{H}_t)$	$=$	$h\left(\min_{j:t_j<t}(\lvert\lvert x_j - x \rvert\rvert)\right)$	nearest neighbour
M2	$g(x, t \mid \mathcal{H}_t)$	$=$	$\prod_{j:t_j<t} h(\lvert\lvert x - x_j \rvert\rvert)$	multiplicative

Partial likelihood criterion favours M1 (nearest neighbour) with $c = 1$ (exponential)

Precision of estimation of $\mathbf{h}(\cdot)$

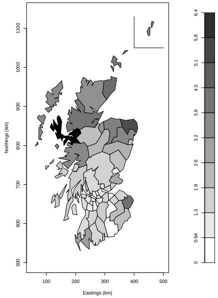


Data (—) and simulations of model (—)

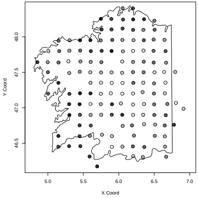
7 Towards synthesis: de-compartmentalising spatial statistics

- Cressie's classification of spatial statistics
- Point processses and geostatistics
- Analyse problems, not data

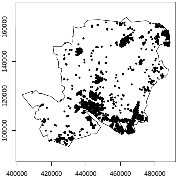
Spatial statistics according to Cressie



Lattice data

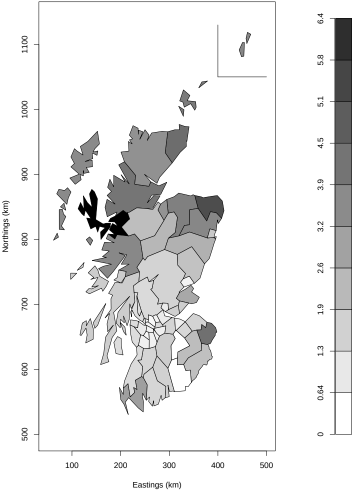


Geostatistics



Point processes

Lattice data

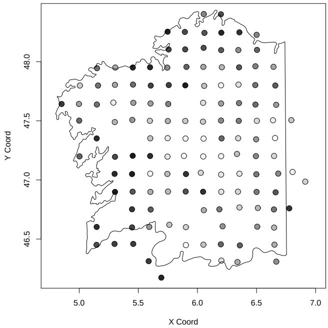


Data: outcomes $Y_i : i = 1, ..., n$

Model: Markov random field:
$[Y_i | \{Y_j : j \neq i\}] : i = 1, ..., n$

"...or indeed, for any multivariate distribution at all"
Hawkes, in discussion of Besag (1974)

Geostatistics

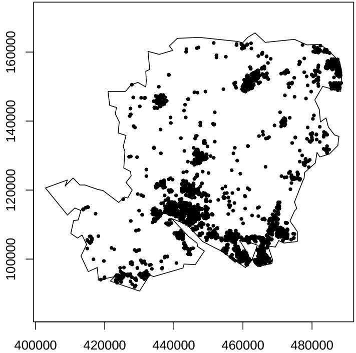


Data: outcome and locations
$(Y_i, x_i) : i = 1, ..., n$ (variable n)

Model: spatially continuous stochastic process, $Y(x) : x \in \mathbb{R}^2$

Presumed scientific interest in $Y(x)$ at non-sampled locations

Point process



Data: **outcomes** $\mathbf{x_i \in A : i = 1, ..., n}$ $\mathbf{(A \subset \mathbb{R}^2)}$

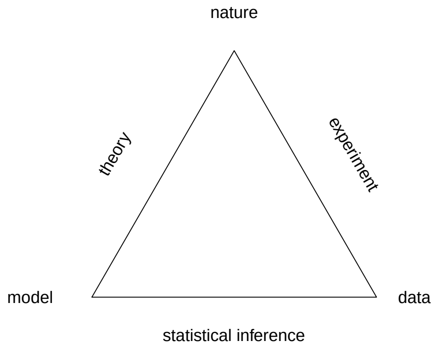
Model: **countably infinite set of points,** $\mathbf{x_i \in \mathbb{R}^2 : i = 1, 2, ...}$

Locations are of scientific interest in themselves.

A classification of

- **processes?**
- **models?**
- **methods?**
- **data-formats?**

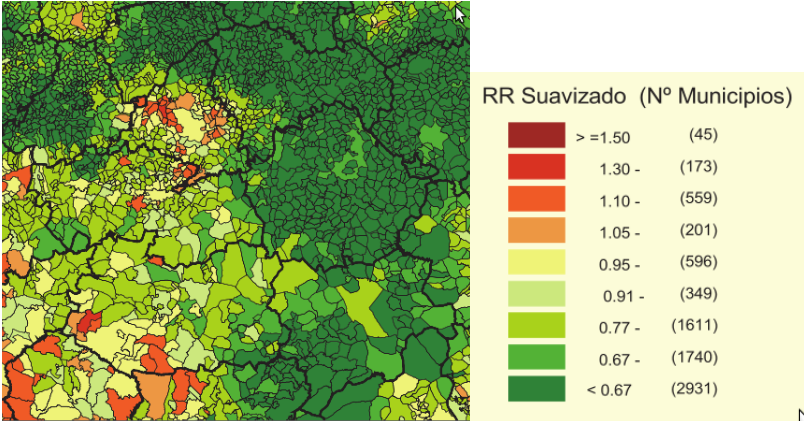
Analyse problems, not data



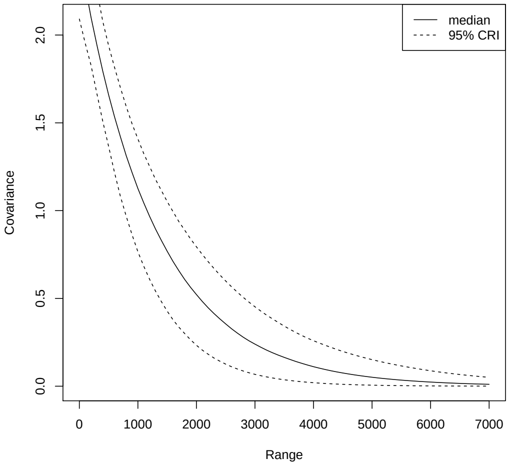
A statistical model is:

- **a device to answer a question**
- **a bridge between scientific theory and empirical evidence**
- **a framework to enable principled inference in the presence of uncertainty**

Lung Cancer mortality in the Castile-La Mancha, Spain



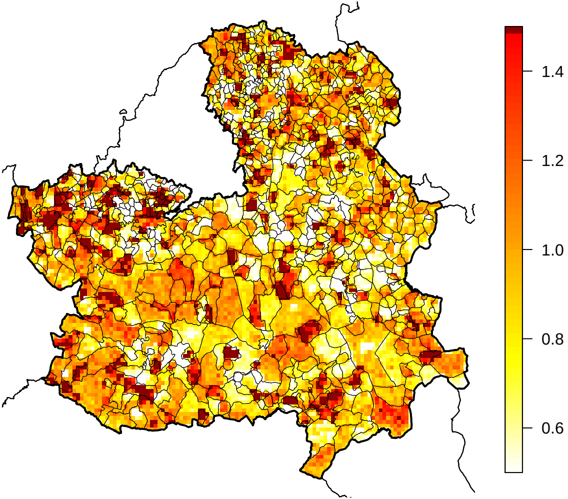
Lung Cancer mortality: fitted spatial correlation function



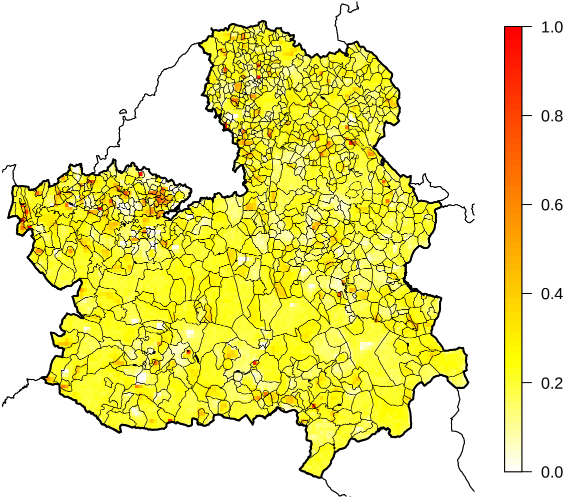
Lung Cancer mortality: standardised covariate effects

	Quantile		
Parameter	**0.50**	**0.025**	**0.975**
Percentage illiterate	**1.13**	**1.03**	**1.24**
Percentage unemployed	**0.92**	**0.8**	**1.03**
Percentage farmers	**0.88**	**0.76**	**1.00**
Percentage of people over 65 years old	**1.2**	**0.96**	**1.51**
Income index	**1.19**	**1.03**	**1.39**
Average number of people per home	**0.98**	**0.75**	**1.26**

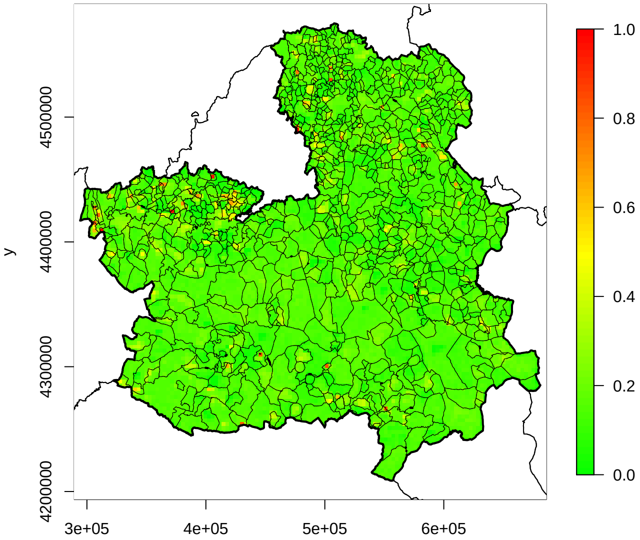
Lung Cancer mortality: covariate-adjusted relative risk

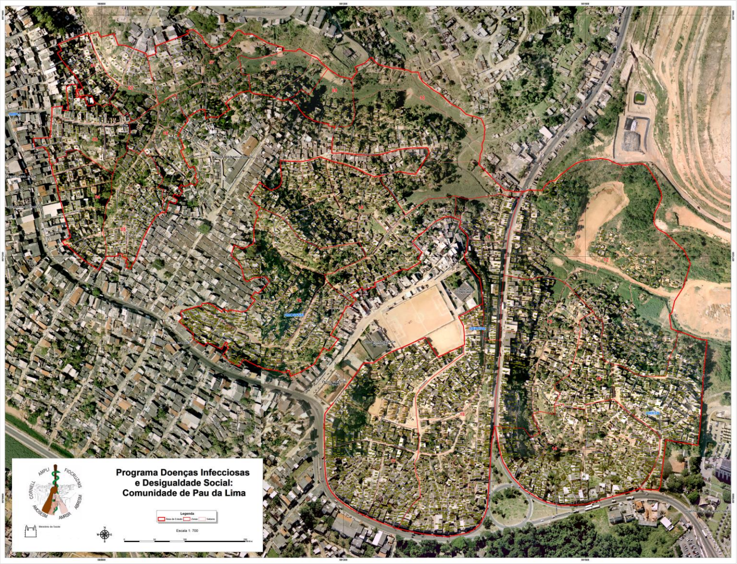


Lung Cancer mortality: $\mathrm{P}(\mathbf{RR} > \mathbf{1.1})$



Lung Cancer mortality: $\mathrm{P}(\mathbf{RR} > \mathbf{1.5})$




Programa Doenças Infecciosas
e Desigualdade Social:
Comunidade de Pau da Lima
AMPLI
FIOCRUZ/MS
CORNELL
ABRESM
AMOSESM
AMRSR
Legenda
Área de Estudo
Zonas
Setores
Escala 1: 700
N
S
W
E
0
50
100
200

Leptospirosis cohort study: Pau da Lima

- subjects i at locations x_i, blood-samples taken at times $t_{ij} \approx 0, 6, 12, 18, 24$ months
- sero-conversion defined as change from zero to positive, or at least four-fold increase in concentration
- data consist of:
  - $Y_{ij} = 0/1 : j = 1, 2, 3, 4$ (seroconversion no/yes)
  - $r_i(t)$ known and hypothesised risk-factors

Longitudinal data, binary outcome $\Rightarrow$ GEE? GLMM?

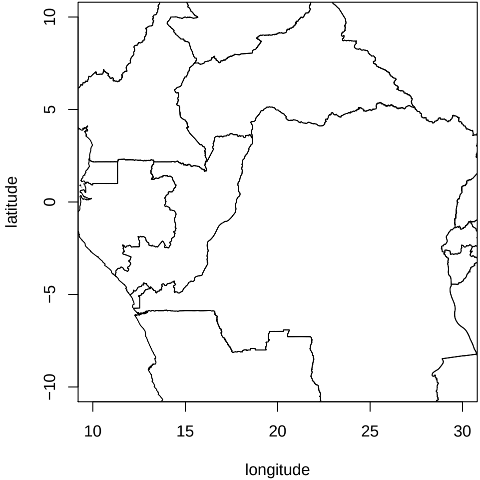
Leptospirosis cohort study: model formulation

- $\mathbf{Y_{it} = 1 \Leftrightarrow}$ **at least one infection event**
- **model infection events as person-specific, inhomogeneous Cox processes,**

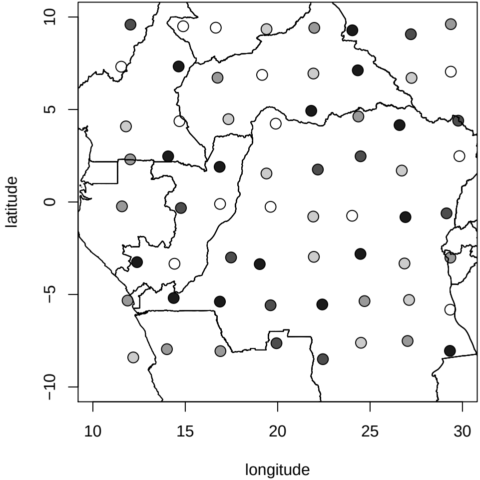
$$\Lambda_i(t) = \exp\{r_i(t)'\beta + U_i + S(x_i)\}$$

- $\mathrm{P}(Y_{it} = 1 | \Lambda_i(\cdot)\} = 1 - \exp\left\{- \int_{t_{i,j-1}}^{t_{ij}} \Lambda_i(u) du\right\}$

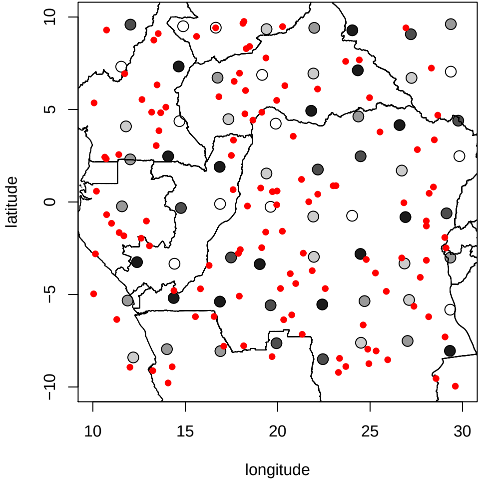
Data-synthesis



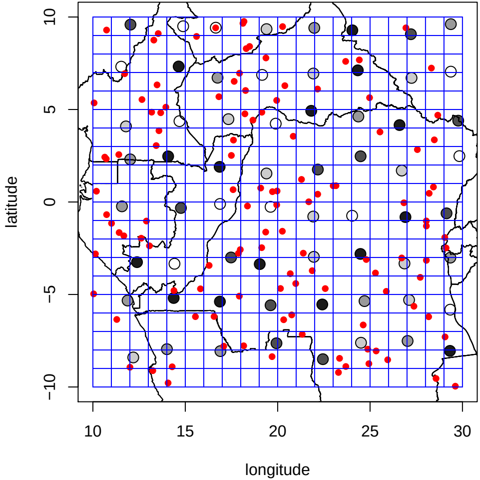
Data-synthesis



Data-synthesis



Data-synthesis



Data-synthesis

$$\begin{aligned} S &= \text{state of nature} \\ Y &= \text{all relevant data} \\ T &= \mathcal{F}(S) = \text{target for prediction} \end{aligned}$$

Model: $[S, Y] = [S][Y|S]$

Predictive inference: $[S, Y] \Rightarrow [T|Y]$

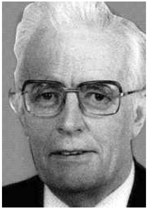
Statistics and scientific method

"Far better an approximate answer to the right question, which is often vague, than an exact answer to the wrong question, which can always be made precise."

John Tukey (1915–2000)

"...the importance of making contact with the best research workers in other subjects and aiming over a period to establish genuine involvement and collaboration in their activities."

Sir David Cox (b 1924)



References

Besag, J. (1974). Spatial interaction and the statistical analysis of lattice systems (with Discussion). *Journal of the Royal Statistical Society* B **36**, 192–225.

Cressie, N.A.C. (1991). *Statistics for Spatial Data*. New York: Wiley.

Diggle, P.J. (2013). *Statistical Methods for Spatial and Spatio-temporal Point Processes*. Boca ration: Chapman and Hall/CRC

Diggle, P.J., Eglen, S.J. and Troy, J.B. (2006). Modelling the bivariate spatial distribution of amacrine cells. In *Case Studies in Spatial point Processes*, ed A. Baddeley, J. Mateu, D. Stoyan, 215–233. New York: Springer.

Diggle, P.J., Knorr-Held, L., Rowlingson, B., Su, T., Hawtin, P. and Bryant, T. (2003). Towards On-line Spatial Surveillance. In *Monitoring the Health of Populations: Statistical Methods for Public Health Surveillance.*, ed R. Brookmeyer and D. Stroup. Oxford: Oxford University Press.

Diggle, P.J., Moraga, P., Rowlingson, B. and Taylor, B. (2013). Spatial and spatio-temporal log-Gaussian Cox processes: extending the geostatistical paradigm. *Statistical Science* (to appear)

Diggle, P.J., Rowlingson, B. and Su, T-L. (2005). Point process methodology for on-line spatio-temporal disease surveillance. *Environmetrics*, **16**, 423–434.

Diggle, P.J., Zheng, P. and Durr, P. (2005). Non-parametric estimation of spatial segregation in a multivariate point process. *Applied Statistics*, **54**, 645–658.

Geyer, C.(1999). Likelihood inference for spatial point processes. In *Stochastic Geometry: likelihood and computation*, ed O.E. Barndorff-Nielsen, W.S. Kendall and M.N.M. van Lieshout, 79–140.

Keeling, M.J., Woolhouse, M.E.J., Shaw, D.J., Matthews, L., Chase-Topping, M., Haydon, D.T., Cornell, S.J., Kappey, J., Wilesmith, J. and Grenfell, B.T. (2001). Dynamics of the 2001 UK foot and mouth epidemic: stochastic dispersal in a heterogeneous landscape. *Science*, **294**, 813–817.

López-Abente, G., Ramis, R., Pollán, M, Aragonés, N., Pérez-Gómez, B., Gómez-Barroso, D., Carrasco, J.M., Lope, V., García-Pérez, J., Boldo, E. and García-Mendizábal, M.J. (2006). ATLAS municipal de mortalidad por cáncer en España, 1989-1998. Madrid: Instituto de Salud Carlos III.

Møller, J. and Waagepetersen, R.P. (2004). *Statistical Inference and Simulation for Spatial Point Processes.* London: Chapman and Hall.

Ripley, B.D. (1976). The second-order analysis of stationary point processes. *Journal of Applied Probability*, **13**, 255–266.

Ripley, B.D. (1977). Modelling spatial patterns (with discussion). *Journal of the Royal Statistical Society* B **39**, 172–212.

Closing remarks

- **application-specific models with interpretable parameters, or generic classes of model?**
- **discrete-time or continuous time?**
- **exploiting the natural ordering of the time-dimension**
- **equilibrium behaviour of models is often of limited relevance?**
- **model-specification via conditional intensity as a route to relatively straightforward likelihood-based inference**

Over to you