spatial aspects of ecological dynamics
– an introduction to spatial point processes

Janine Illian
CREEM, University of St Andrews, Scotland

Edinburgh, June 2009
modelling the individuals’ perspective

modelling from the individuals’ perspective has become popular in (plant) ecology:

- spatial birth and death processes
- dynamic models
- mechanistic models...

How about statistical models? Statistical models taking on the plants’ eye view are **spatial point process models**:

- many motivating examples for spatial point processes have been derived from ecology
- **but**: spatial point processes have rarely been used to directly answer topical ecological questions
spatial point process modelling

**aim**

- *describe and model the spatial structure formed by the locations of individuals in space*

  *specifically:*

- *use existing spatial point process methods in the context of ecological research*

- *develop methodology that is suitable for this purpose*
a spatial pattern...

- Are the daisies randomly distributed in the lawn of our garden???

an exercise

- consider the square on the sheet of paper
- imagine it is a piece of the lawn in my garden
- try to draw a random pattern...
introduction
Assessing spatial randomness
Spatial point process models
muskoxen in Greenland

motivation

some examples

- similar exercise done with students last year
- if points are randomly distributed we would expect that the location of each point is independent of the locations of all the other points
- pattern is too regular; no points at small distances
- this pattern is unlikely to be "random"
- regularity, repulsion
Some examples

- If points are randomly distributed, we would expect them to occur anywhere in the plot with the same probability.
- No points at near the edges.
- This pattern is unlikely to be "random."
some examples

- much better...
- this pattern is more likely to be "random"
- how can we assess this?
- do we know for sure?
spatial point process modelling

aim

describe and model the spatial structure formed by the locations of individuals in space

- assess patterns as to whether they are likely to be "random"
- find a formal description of "randomness" – a suitable statistical model
- find a formal description of "non-randomness" – relative to the random case
spatial randomness

"What do you mean by "random pattern"?"

important distinction:

1) complete spatial randomness (CSR): the points are **independently** scattered in space

2) points are considered to be a realisation of a **random mechanism** that generates patterns with the same spatial properties; these properties can deviate from CSR (e.g. aggregation, regularity)
spatial point process

• a mathematical model that captures the characteristics of spatial point patterns in a finite number of parameters

• a specific model with a given set of parameters provides a mechanism by which spatial point patterns may be generated; all have the same spatial characteristics

• a spatial pattern generated from a spatial point process is termed a realisation of the process

• only consider spatial patterns in the two-dimensional space; can easily be generalised to patterns in more general spaces and higher dimensions

• the location of any object can be modelled – plants, animals, stars, cells, cities...
technical definition

A spatial point process is a random variable \( X \), with an observed pattern being a realisation \( x \) of this random variable. For each Borel set \( B \subset \mathbb{R}^2 \), let \( \phi_X(B) \) be the number of points of \( X \) in \( B \). We identify a **point configuration** with a counting measure \( \phi_X \) on Borel sets on \( \mathbb{R}^2 \).

Let \( N \) be the set of all such measures. On \( N \) define \( \mathcal{N} \) as the smallest \( \sigma \)-algebra generated by sets of the form

\[
\{ \phi \in N : \phi(B) = n \quad n \in \{0, 1, 2, \ldots\}, \quad B \text{ some bounded Borel set}\}.
\]
Technical definition

Let \((\Omega, \mathcal{A}, P)\) be some probability space. A **spatial point process** \(X\) may then be regarded as a measurable mapping from \((\Omega, \mathcal{A})\) into \((\mathbb{N}, \mathcal{N})\), i.e. as a random variable. A spatial point pattern \(x\) is then a realisation of this random variable.  

In applications, the process \(X\) lives in some subset \(W\) of \(\mathbb{R}^2\) and patterns are only observed in a bounded area \(S \subset W\).
marked point process

- in situations where additional data exist on the objects that form the spatial point pattern, these additional data are conventionally termed **marks**
- combining a spatial point process with marks yields a marked pointed process and hence marked point patterns
- the marks may be either quantitative variables, such as weight or height in the case of plants, or qualitative variables such as species, thus defining different **types** of points in the pattern
marked point processes

More formally, let $Z$ be a simple point process in $\mathbb{R}^2$. Attach a random mark $m_\xi \in \mathcal{M}$, where $\mathcal{M}$ is some mark space, to each point $\xi \in Z$. This yields a marked point process

$$X_m = \{(\xi, m_\xi) : \xi \in Z\}.$$ 

In most applications, the mark space $\mathcal{M}$ is a subset of $\mathbb{R}^d$ with $d \geq 1$, but more general mark spaces may be considered. Note that if $\mathcal{M} = \{1, \ldots, k\}$, $X$ is a multi-type point process with $k$ different types of points. Note further that a multi-type process can also be regarded as a $k$-tuple of different subprocesses $(X_1, \ldots, X_k)$. 

Janine Illian
CREEM, University of St Andrews, Scotland

spatial aspects of ecological dynamics – an introduction to spatial point processes
Poisson process

- **homogeneous Poisson process** provides a formal way of describing (or modelling) patterns that were termed "random" patterns above
- patterns are more correctly referred to as patterns exhibiting complete spatial randomness (CSR)
- other more complicated models can be viewed as a generalisation of the homogeneous Poisson process
- serves as a reference process when the first and second order spatial characteristics of a specific pattern are analysed
Poisson process

A homogeneous Poisson process $X$ with constant intensity $\lambda_0$ (number of points per unit area) has two properties:

1. The density of points is constant in the area under investigation.

2. The location of any point in the pattern is independent of the location of any of the other points in the pattern, i.e. there is no interaction between the points.
Poisson process – technically

Let $\mu$ be the Lebesgue measure on the Borel sets in $S$. Then the Poisson process on $S$ with intensity $\lambda_0$ has the properties:

1. For any Borel set $B \in S$ the cardinality of $B$, $X(B)$, follows a Poisson distribution with mean $\lambda_0 \cdot \mu(B)$.
2. For any disjoint Borel sets $B_1, \ldots, B_{n_0} \subseteq S$ with an arbitrary $n_0 \geq 2$, $X(B_1), \ldots, X(B_{n_0})$ are independent.
aim of this talk

aim

- *explain how randomness can be assessed formally with statistical methods*

summary characteristics

- *introduce statistical models that describe randomness and various types of non-randomness*

spatial point process models
summary characteristics (or statistics)

- when analysing a spatial point pattern, we seek to infer properties of the pattern
- properties will then be used to choose an appropriate point process model (see below); realisations of this model are patterns with properties similar to the original
- the initial analysis of a spatial point pattern consists of a descriptive analysis of the general properties of the pattern using summary statistics
- the information gained from the summary statistics will subsequently allow the choice of a specific point process model to fit to the data
summary characteristics (or statistics)

The summary statistics reflect the characteristics

i) "first order" characteristics, describing the density of points through space
   The density of points may be constant over the area considered, thus resulting in a homogeneous pattern, or instead exhibit spatial trend or other forms of non-constant density, resulting in an inhomogeneous pattern.

ii) "second order" characteristics, relating to the relative position or interaction among points
   Points may be randomly scattered in space, show regularity resulting from negative interaction (inhibition) among the points, or exhibit a clustered structure (aggregation) resulting from positive interaction (attraction) among points.
First order summary characteristics

**summary characteristics – first order**

- **first order summary statistics** are an analogue to the mean in standard, i.e. non-spatial, statistics in the sense that they describe the density or intensity of the spatial pattern in terms of the average number of points per unit area.
- The average number of points varies in space if the process is inhomogeneous.
- Unlike in standard statistics, the first order summary statistic is often not a single value but is a function of the spatial location, i.e. a function of $x$ and $y$, and can be plotted as a 3-dimensional surface.
summary characteristics – first order

More technically, for a point process $X$ the **intensity measure** is given by

$$\Lambda(B) = E[\phi(B)], \text{ for any Borel set } B,$$

where $\phi$ is a counting measure as defined above and $E[.]$ denotes the expected value. If $\Lambda$ is absolutely continuous with regard to the Lebesgue measure, a density function, the **intensity function** $\lambda : B \to \mathbb{R}^+$ exists, such that

$$\Lambda(B) = \int_B \lambda(x) \, dx.$$ 

If the intensity is constant, i.e. if $\lambda(x) = \lambda_0$ the point process $X$ is called a **homogeneous** or **(first order) stationary** process.
**summary characteristics – second order**

- **second order summary statistics** are an analogue to measures of dispersion in non-spatial statistics, in the sense that they describe the location of an individual point relative to the other points.
  - considers the expected number of points at specific distances from each point,
  - does not yield a single measure but is a function of distance (and location)
summary characteristics – second order

- the pair correlation function $g$ considers the expected number of points at a distance $r$ from an arbitrary point.

$$g(r) = \frac{\text{intensity of points at dist. } r \text{ from arbitrary point}}{\lambda_0}$$

- the pair-correlation function with $g(r) > 1$ indicates clustering and $g(r) < 1$ regularity
For two Borel sets $B_1$ and $B_2$ we define the **second order factorial moment measure** $\mu^{(2)}$ as

$$\mu^{(2)}(B_1 \times B_2) = E \sum_{\xi \in X, \eta \in X, \xi \neq \eta} 1[\xi \in B_1, \eta \in B_2].$$

If $\mu^{(2)}$ has a density function $\rho^{(2)}$, i.e. if

$$\mu^{(2)}(B_1 \times B_2) = \int_{B_1} \int_{B_2} \rho^{(2)}(\xi, \eta) d\xi d\eta$$

then $\rho^{(2)}$ is called the second-order product density. If $\rho^{(2)}$ depends only on the distance $r$, the **homogeneous pair correlation function** is defined as:

$$g(r)_{\text{hom}} = \frac{\rho^2(r)}{\lambda^2}.$$
second order summary characteristics

summary characteristics – second order

- characteristics can be generalised to a multi-type setting
- in analogy to the univariate case, for two subprocesses $X_i$ and $X_j$, where $i \neq j$, $i, j \in M$ the cross summary statistics $g_{ij}$ consider the number of pairs of points from the two different subprocesses at a distance $r$ from an arbitrary point
- the interpretation of the multivariate second order summary statistics are similar to the interpretation of the univariate second order summary statistics
- summary statistics may be generalised to more than two types of points, but interpretation becomes increasingly difficult
• we can no detect and describe deviations from CSR
• next step: find a formal description (i.e. a statistical model) for patterns that deviate from the Poisson case
⇒ more general spatial point process models

generalisations of the Poisson case
• inhomogeneous Poisson process
• Cox processes
• Markov point processes
spatial point process models

- A specific model with a given set of parameters provides a mechanism by which **spatial point patterns** may be generated, all having the same spatial characteristics.
- A spatial pattern generated from a spatial point process is termed a **realisation** of the process.
- Modelling data; infer information on ecological processes from the spatial pattern formed by individuals.
- Model of the **spatial locations** of objects but NOT a model of properties of the patterns (size, growth) given the locations.
- Model of an entire pattern.
- In other words, one datum = one pattern.
statistical models

- overall properties of the patterns are represented by (a small number of) parameters in the model
- parameters reflect, e.g. density of the plants, the strength, direction and form of inter-and intra-species interaction
- allow for assessment of the goodness of fit of a model to a data set
- predictions can be made, and the uncertainty of these predictions can be quantified
inhomogeneous Poisson process

- a straightforward generalisation of the homogeneous Poisson process may be achieved through introducing inhomogeneity, but no interaction
- the two properties considered above are now:
  1. the intensity of the point pattern is not constant over the bounded region
  2. there is no interaction between the points
- the intensity of the pattern is now described by a non-constant intensity function $\lambda(.)$
model classes - Cox processes

- describe aggregation or clustering resulting from (unobserved) environmental variability (e.g. nutrient levels in the soil)
- variability is assumed to be a stochastic process in itself
- generalisation of Poisson process: intensity function is random
- often called "doubly-stochastic" processes
model classes - Cox processes

example:

log Gaussian Cox process: a Cox process with random intensity

\[ \Lambda(s) = \exp\{Z(s)\}, \]

where \( \{Z(s)\} \) is a (stationary and isotropic) Gaussian random field, \( s \in \mathbb{R}^2 \); properties:

- conditional on the (unknown) environmental conditions
  
  Poisson process; \( X|Z \sim \text{Poisson}(S, Z) \)

- \( Z(u) \) is a random variable for all \( u \in S \); regard \( \Lambda \) as a random intensity function

- choosing \( Z \) in the above definition to be deterministic yields a Poisson process

- summary characteristics are known analytically
model classes – probability density

- Often convenient to express the general point process model in terms of its probability density, most notably for the class of Markov processes, see below.
- Define these densities with regard to the unit rate homogeneous Poisson process.
- The density describes the probability for a point pattern relative to the homogeneous Poisson process.
- Not to be confused with the entirely different concept of an intensity (or point density), where the expected number of points in a given location is being considered.
model classes – probability density

- for a general Poisson process

\[ f(x) = \prod_{i=1}^{n(x)} \lambda(x_i) \exp \left(- \int_S \left[ \lambda(\mu) - 1 \right] d\mu \right) \]

is the density function.

- for the **homogeneous Poisson process** on \( S \) with intensity \( \lambda_0 > 0 \) this simplifies to

\[ f(x) = \exp\left(- (\lambda_0 - 1)|S|\right) \cdot \lambda_0^{n(x)}, \]

with \( n(x) \) the number of points in \( x \) and \( |S| \) the volume of \( S \).

- the density of the inhomogeneous Poisson process is known in its full form but for more complicated models, the density is known only up to a normalising constant.
model classes – Markov processes

- the class of spatial Markov point processes models patterns exhibiting aggregation (or inhibition) due to interaction between points
- termed *Markov* due to a spatial Markov property referring to interaction with spatial neighbours only
- implies a generalisation of the properties (1) and (2) of the Poisson process in as far as that the points are now no longer independent
- often also called Gibbs processes
- NOT to be confused with Markov chains
model classes – Markov processes

• simplest example of a Markov process is the **Strauss process**
• interaction (repulsion) is constant within a fixed interaction radius $R$ around each point
• strength of the interaction can range from no interaction to complete inhibition within the radius $R$ around each point
• in the case of no interaction the process is equivalent to a homogeneous Poisson process.
model classes – Markov processes

- the density of the Strauss process is given by

\[ f(x) = \alpha \beta^{n(x)} \gamma^{s_R(x)}, \]

where \( n(x) \) is the number of points in \( X \), \( s_R(x) \) the number of distinct pairs of points in \( X \) with a distance \( r < R \) and \( \alpha \) the normalising constant.

- the parameter \( \beta > 0 \) reflects the intensity of the process and the parameter \( \gamma \) the strength of the interactions between points with a distance \( r \leq R \).

- for \( \gamma = 1 \), \( X \sim \text{Poisson}(S, \beta) \).
model classes – Markov processes

- more general class of Markov point processes are the **pairwise interaction processes**
- the interaction around each point changes with the distance from the point according to an interaction function

Its density is of the following form:

\[
f(x) = \alpha \prod_{i=1}^{n(x)} \beta(\xi) \prod_{\xi,\eta \in X, \xi \neq \eta} h(\|\xi - \eta\|),
\]

where \(\alpha\) is the normalising constant, \(\beta(.)\) is a function describing the intensity of the process and \(h(.)\) is a non-negative interaction function.

- typically the strength of interaction decreases with distance from each point within a specific interaction radius
Fitting a model

1. describe the first and second order characteristics of a pattern
2. select appropriate spatial point process model for pattern, informed by the results of step 1
3. fit the model to pattern by parameter estimation
4. assess the goodness of fit of the model run simulations and compare the simulated patterns with the original pattern, e.g. based on first and second order characteristics
in reality...

- spatial point process models are mathematically elegant models of spatial organisation of individuals
- however spatial point process application in ecology have largely been restricted to descriptive analysis of forest plots
- development of spatial point process methodology has mainly been theory-driven rather than problem-driven
- theoreticians have rarely considered:
  - species-rich data sets
  - environmental heterogeneity
  - properties of individuals (i.e. marks)
  - spatio-temporal processes
- these are common features in ecological data sets
my work...

issues:

• standard parameter estimation methods (maximum likelihood) fail in this context (likelihoods not explicitly known)
• alternative methods are available, such as Markov Chain Monte Carlo methods (MCMC)
• however: realistic but complex models impossible to fit due to immensely long running times

my aim: develop application-driven spatial point process methodology suitable for modelling ecological data sets, including

• models for large data sets with many different species
• models for data sets with (large numbers of) covariates and marks
• patterns observed over time
muskoxen data – replicates over time

Zackenberg research station, east Greenland

- spatial locations of muskoxen herds recorded at different time points throughout different years
- habitat choice: does the spatial distribution depend on habitat quality?
- behavioural ecology: do male and female herds interact differently with other herds?
muskoxen data – replicates over time

Zackenberg research station, east Greenland

- spatial locations of muskoxen herds recorded at different time points throughout different years
- habitat choice: does the spatial distribution depend on habitat quality?
- behavioural ecology: do male and female herds interact differently with other herds?
an example – replicated spatial point patterns

- animals move in and out of the observation area
- habitat changes over time (e.g. snow melts)
- group interactions change over time (e.g. in mating season)
- patterns at any one time too small for analysis
- fit one "big" model to all time points together

⇒ treat "time" as random factor and habitat and group type as fixed factors in a spatial point process model
Assessing spatial randomness

Spatial point process models

Muskoxen in Greenland

An example – replicated spatial point patterns

- Animals move in and out of the observation area
- Habitat changes over time (e.g., snow melts)
- Group interactions change over time (e.g., in mating season)
- Patterns at any one time too small for analysis

⇒ Treat "time" as random factor and habitat and group type as fixed factors in a spatial point process model
mixed models for spatial point patterns

At time $t$ consider a pairwise interaction process for $x_t$ with conditional intensity

$$\lambda_{\theta, \phi_t}(u; x_t) = b_{\theta, \phi_t}(u) \prod_{i=1, x_i \neq u}^{n(x_t)} h_{\theta, \phi_t}(u, x_{ti}),$$

where $n(x_t)$ is the number of points in point pattern $x_t$.

Consider the loglinear case where

$$b_{\theta, \phi_t}(u) = \exp(\theta^T B_1(u) + \phi_t^T B_2(u))$$
$$h_{\theta, \phi_t}(u, v) = \exp(\theta^T H_1(u, v) + \phi_t^T H_2(u, v)).$$

Here $\theta$ is a parameter vector of fixed effects and $\phi_t$ is a parameter vector of random effects, realisations from $\phi \sim \mathcal{N}(0, \sigma^2 I)$.
mixed models for spatial point patterns

density with respect to the homogeneous Poisson process is then

\[ f(x_t; \theta, \phi_t) = \alpha(\theta, \phi_t) \prod_{i=1}^{n(x_t)} b_{\theta,\phi_t}(x_{ti}) \prod_{i<j} h_{\theta,\phi_t}(x_{ti}, x_{tj}), \]

where \( n(x_t) \) is the number of points in point pattern \( x_t \) and \( \alpha(\theta, \phi_t) \) is a (typically intractable) normalising constant. Assuming independence the density for a model of the replicated patterns \( x_1, \ldots, x_T \) is

\[ f(x_1, \ldots, x_T; \theta, \phi) = \tilde{\alpha}(\theta, \phi) \prod_{t=1}^{T} \prod_{i=1}^{n(x_t)} b_{\theta,\phi_t}(x_{ti}) \prod_{i<j} h_{\theta,\phi_t}(x_{ti}, x_{tj}), \]

where \( \tilde{\alpha}(\theta, \phi) \) is a different but still intractable normalising constant.
mixed models for spatial point patterns

pseudolikelihood for replicated patterns $\mathbf{x}_1, \ldots, \mathbf{x}_T$

$$PL_A(\theta, \phi; \mathbf{x}_1, \ldots, \mathbf{x}_T) = \prod_{t=1}^{T} \left( \prod_{i=1}^{n(x_t)} b_{\theta, \phi_t}(x_{ti}) \prod_{i \neq j}^{n(x_t)} h_{\theta, \phi_t}(x_{ti}, x_{tj}) \right) \times \exp \left\{ - \int_A b_{\theta, \phi_t}(u) \prod_{i=1}^{n(x_t)} h_{\theta, \phi_t}(u, x_{ti}) du \right\}. $$
**parameter estimation**

- apply Berman-Turner device; approximate the integral in the pseudolikelihood by a finite sum

$$\Rightarrow$$ log pseudolikelihood is formally equivalent to the log likelihood of independent weighted Poisson variables

- since the conditional intensity is expressed as the sum of fixed and random factors, employ a **generalised linear mixed model** with log link and Poisson outcome

- estimates based on software for generalised linear mixed models using penalised quasi-likelihood
example – Strauss process

\[ b_{\theta,\phi_t}(u) = \exp(\theta_1 + \phi_1 t) = \beta \times \tilde{\beta}_t, \]

where \( \theta_1 \) is a constant that describes the intensity common to all time points and \( \phi_1 t \) is a sample from a random variable \( \phi_1 \sim N(0, \sigma_\phi^2) \) reflecting variation in intensities among different time points, \( \beta + \tilde{\beta}_t > 0 \).

Similarly, for \( \|u - v\| \leq r \)

\[ h_{\theta,\phi_t}(u, v) = \exp(\theta_2 + \phi_2 t) = \gamma \times \tilde{\gamma}_t, \]

where \( \theta_2 \) describes the interaction strength common to all time points and \( \phi_2 t \) is a sample from a random variable \( \phi_2 \sim N(0, \sigma_\phi^2) \) reflecting variation in interaction strength among different time points, \( 0 \leq \gamma + \tilde{\gamma}_t \leq 1 \). For \( \|u - v\| > r \), \( h_{\theta,\phi_t}(u, v) = 0 \).
example – Strauss process

At time $t$, the conditional intensity is then

$$\lambda_{\beta, \gamma}(u; x_t) = (\beta \times \tilde{\beta}_t)(\gamma \times \tilde{\gamma}_t)^{s(u, x_t)},$$

where

$$s(u, x_t) = \#\{x_{ti} \in x_t : 0 < \|x_{ti} - u\| \leq r\}.$$

We then estimate the parameters by fitting the model

$$\log \lambda_{tj} := \theta_1 + \phi_{1t} + \theta_2 v_{tj} + \phi_{2t} v_{tj},$$

where $v_{tj} = s(u_{tj}, x_t)$ and $\phi_{1t}$ and $\phi_{2t}$ are realisations from a random variable $\phi_1 \sim N(0, \sigma_{\phi_1^2})$ and $\phi_2 \sim N(0, \sigma_{\phi_2^2})$ respectively. The model can then be fitted using standard software for generalised linear models.
simulation study

Replicated realisations of Strauss processes with \( r = 0.1 \) in the unit square were generated (10 replicates, 100 simulation runs) with

- \( \beta \times \tilde{\beta}_t \), where \( \beta = 50 \) and \( \tilde{\beta}_t \sim N(0, 20) \) and \( \gamma \times \tilde{\gamma}_t \), where \( \gamma = 0.5 \) and \( \tilde{\gamma}_t \sim N(0, 0.2) \)

- comparison between analysis of individual patterns and analysis of all patterns with random effect for replicate

main result:
variance of estimators much smaller for mixed model than when each replicate is analysed separately
more suitable model

the simple Strauss process is too simplistic for the muskoxen data

- assumption of constant interaction within a fixed radius is not realistic; use function with decreasing interaction strength
- pattern highly inhomogeneous; use altitude and vegetation index as covariate, i.e. as a fixed effect

Interaction function:

\[ h_\theta(r) = \begin{cases} 
(1 - (r/R)^2)^2 & \text{if } 0 < r \leq R \\
0 & \text{else}
\end{cases} \]

for \( r \geq 0, R \geq 0 \)
Assessing spatial randomness

Spatial point process models

muskoxen in Greenland

results for the year 2005

\texttt{glmmPQL(y \sim v + altitude + vegetationindex, random = 1|time/v, family=poisson, weights=w)}.

- results indicate that intensity and interaction vary strongly with time
- both altitude and vegetation index (and their interaction) highly significant
discussion and outlook

- spatial point process models may be applied in ecology to infer information on ecological processes
- to do this suitable modelling approaches have to be developed to incorporate
  - marked patterns
  - inhomogeneity and large numbers of covariates
  - spatio-temporal data
  - interaction dependent on location...
discussion and outlook

- recently, **Bayesian approaches** to modelling have proved to be rather flexible with complex models
- parameter estimation via MCMC often computationally prohibitive
- currently, approximate yet well-behaved methods ("INLA") are being considered for spatial point process models
- likely that much more complicated models can be fitted within reasonable computational time in the near future
in the next episode...

truth about the daisies revealed!
references


