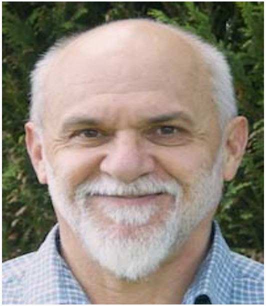
The Thinner Takes It All

Applications of thinned point processes in ecology

Andy Seaton

University of St Andrews

Acknowledgements



Dr Janine Illian,
University of St Andrews

Prof David Borchers,
University of St Andrews

Prof Finn Lindgren,
University of Edinburgh

Dr Fabian Bachl, University of Edinburgh

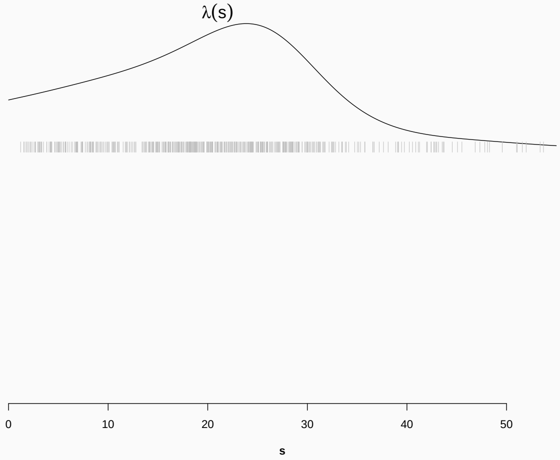
Rick Camp, University of St Andrews and US Geological Survey

Dr David Miller, University of St Andrews

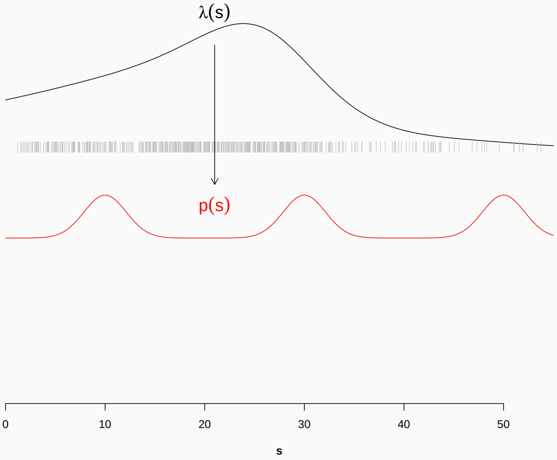
Overview

1. Distance Sampling and "Density Surface" Models
2. Spatial Capture-Recapture

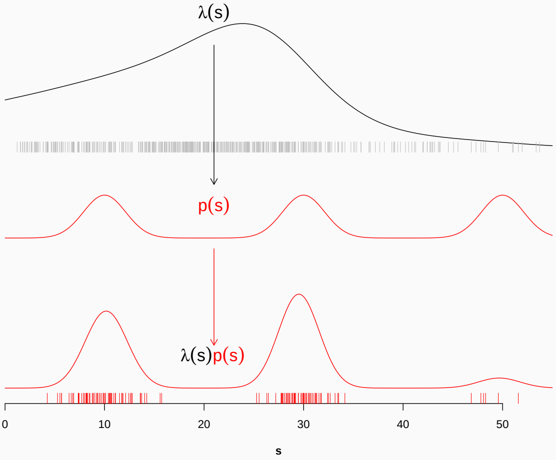
Thinned Poisson Processes



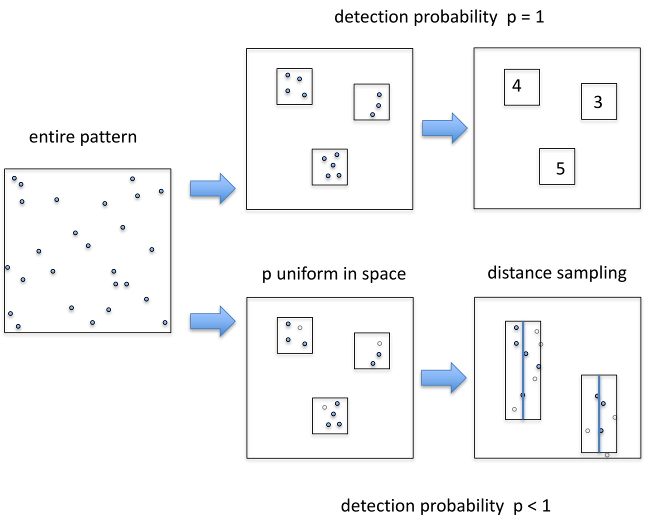
Thinned Poisson Processes



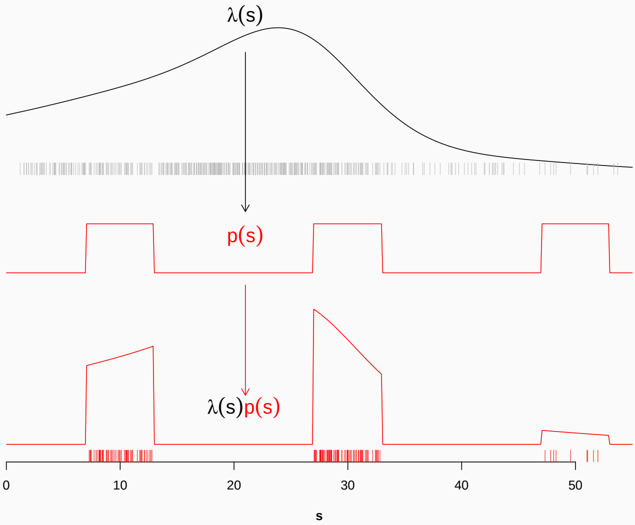
Thinned Poisson Processes



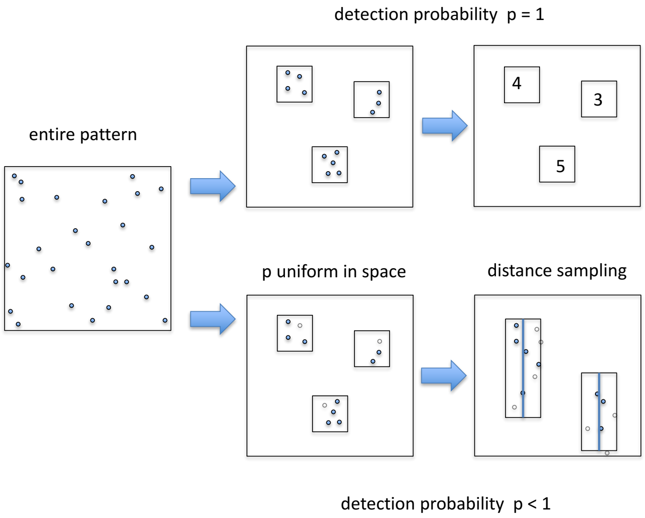
Distance Sampling



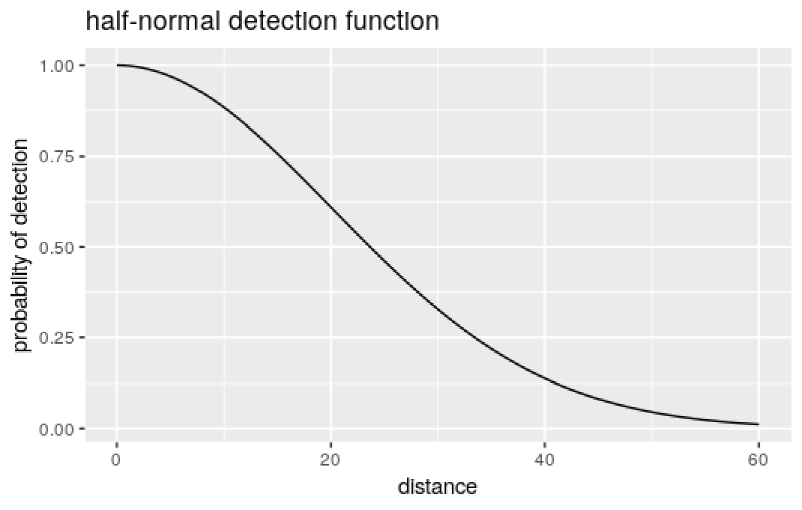
Distance Sampling



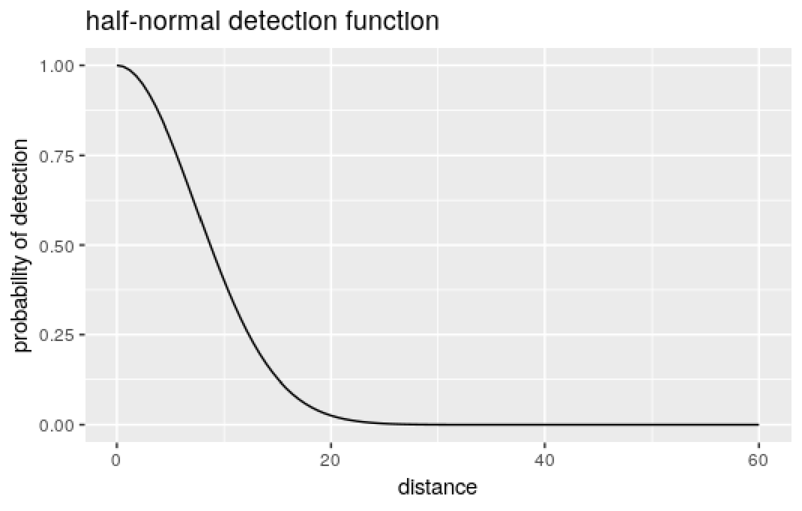
Distance Sampling



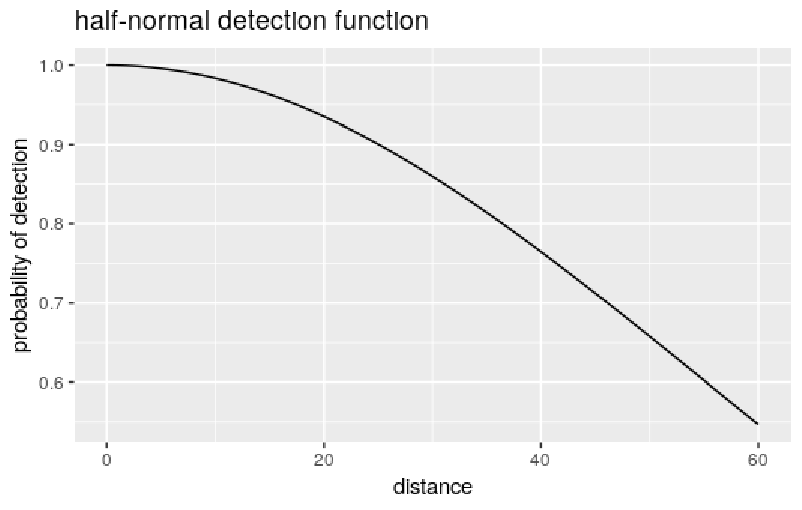
Distance Sampling - the detection function



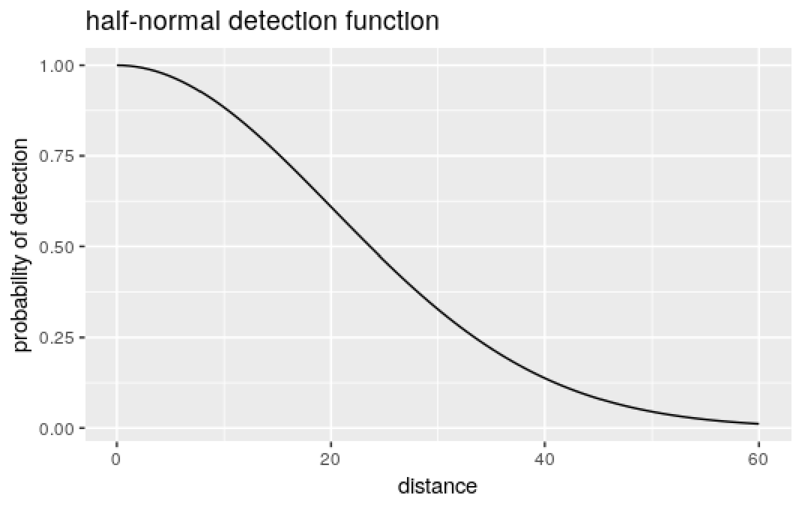
Distance Sampling - the detection function



Distance Sampling - the detection function



Distance Sampling - the detection function

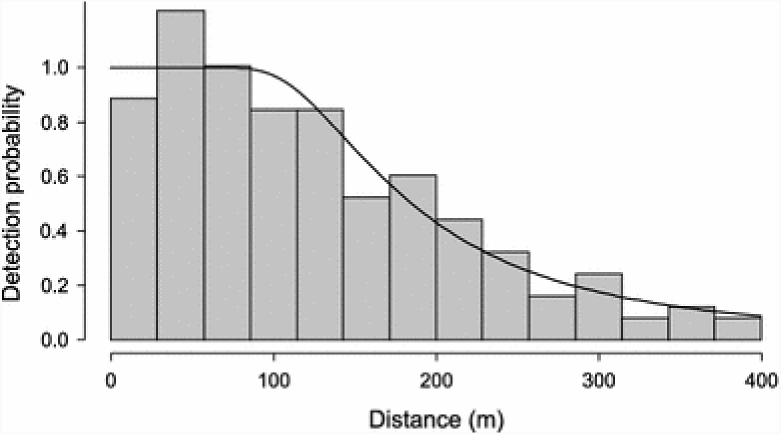


Note the intercept assumed equal to 1

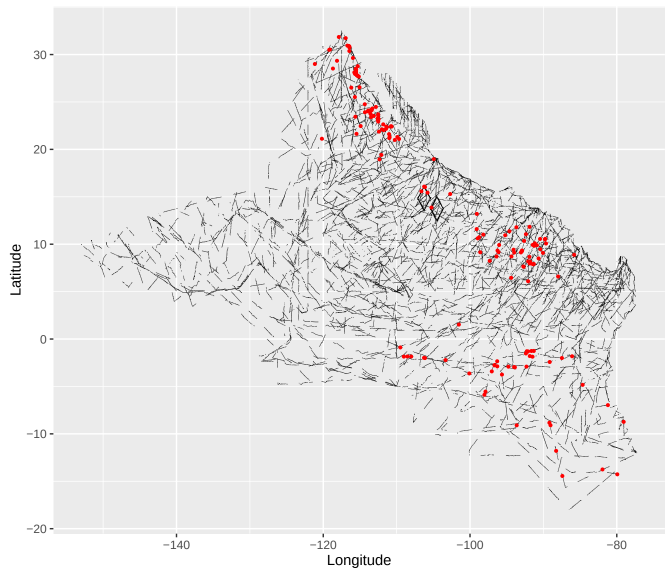
Distance Sampling - the detection function

source: the birdist.com

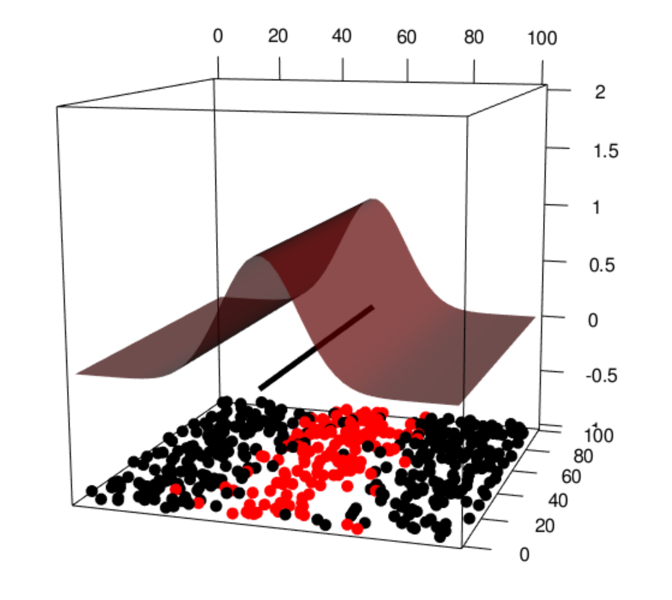
Distance Sampling - the detection function



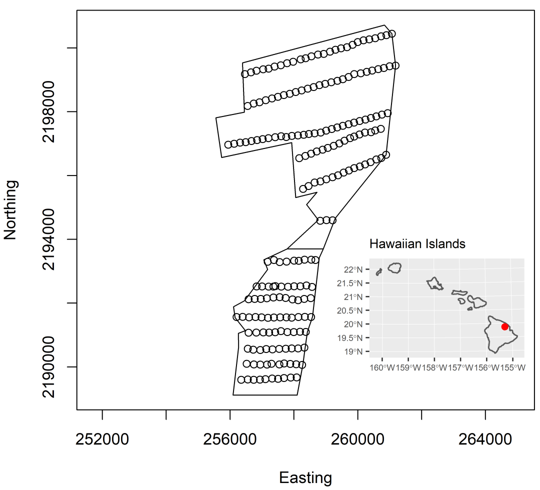
Line transect example - whale survey



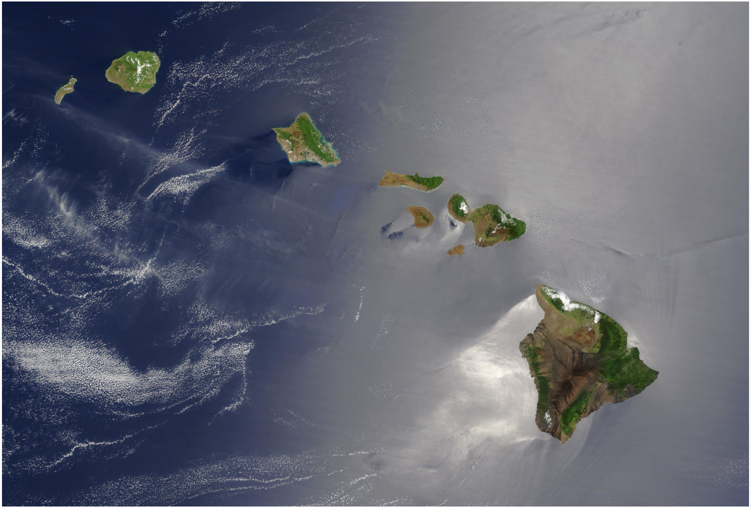
Line transect example - whale survey



Point transect example - Hawaiian Akepa survey



Point transect example - Hawaiian Akepa survey

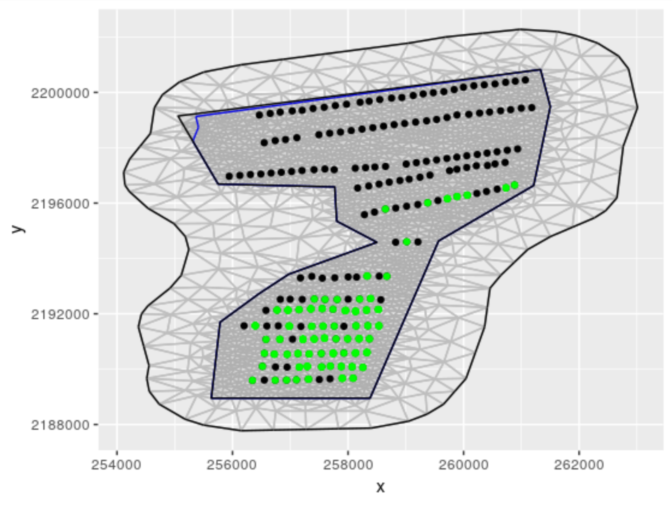


source: wikicommons

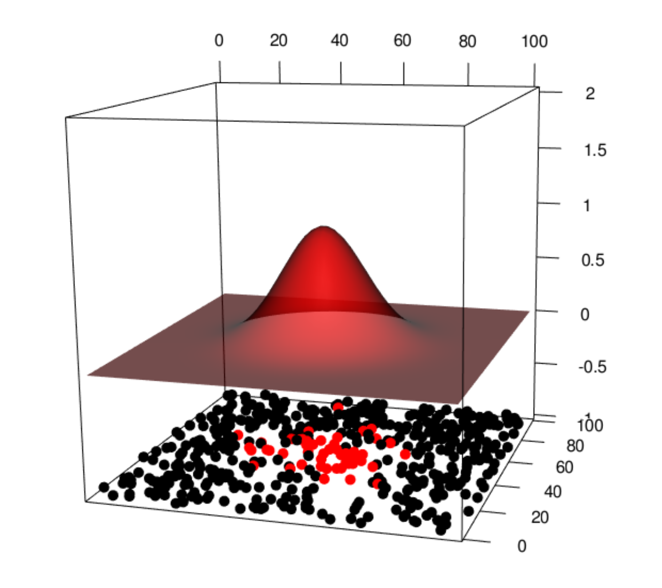
Point transect example - Hawaiian Akepa survey

source: Jack Jeffrey, US Fish and Wildlife Service

Point transect example - Hawaiian Akepa survey



Point transect example - Hawaiian Akepa survey



Point transect example

Recall that the intensity for detected points is $\tilde{\lambda}(s) = \lambda(s)p(s)$

Therefore,

$$\log \tilde{\lambda}(s) = \log \lambda(s) + \log p(s)$$

But $\log p(s)$ is typically not linear in it's parameters! (e.g. half-normal requires strictly positive variance parameter)

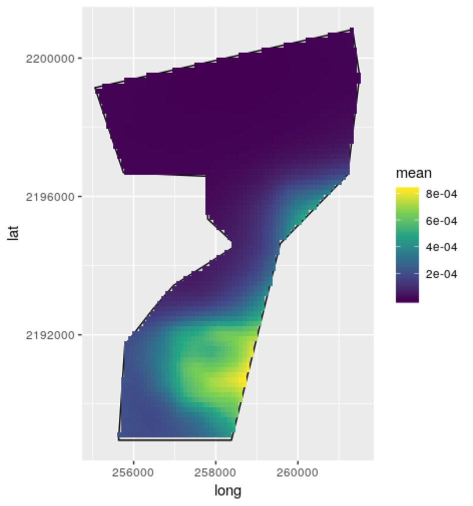
Solution: iterated INLA

Point transect example - iterated INLA

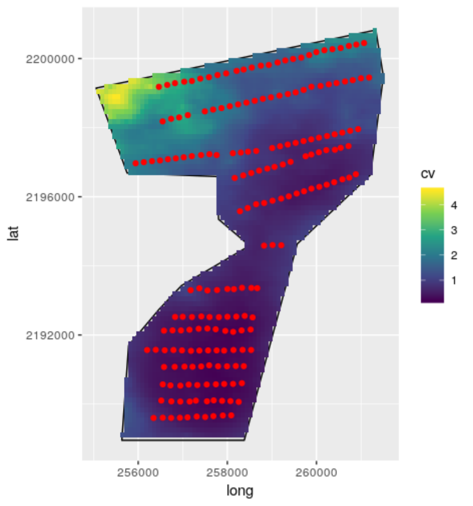
inlabru syntax example:

```
thinning = function(distance, log_sigma){
  exp(-distance^2 / 2 exp(log_sigma)^2
}

formula = coordinates + distance ~
            Intercept + ... + log(thinning(distance, log_sigma))

components = ~ Intercept + ... + log_sigma

lgcp(components,   # tell lgcp what are parameters to be estimated
     data,         # data as sp object
     formula,      # tell lgcp how to construct the linear predictor
     samplers)     # sampling locations as sp object
```

Point transect example - Hawaiian Akepa survey



Point transect example - Hawaiian Akepa survey



Point transect example - Hawaiian Akepa survey

A slight problem: we did not know the exact location of the point, only the distance from the observer.

Solution: derive the appropriate intensity for this partial data

Point transect example - Hawaiian Akepa survey

For a single point transect at location $\boldsymbol{s}_0$, letting $\boldsymbol{s}(r,\theta) = \boldsymbol{s}_0 + r[\cos\theta, \sin\theta]^T$, the intensity for points at distance r from $\boldsymbol{s}_0$ is:

$$
\begin{aligned}
\tilde{\lambda}(r) &= \int_{c(r)} \lambda(\boldsymbol{s}(r,\theta))p(r)\mathrm{d}\boldsymbol{s} \\
&= \int_0^{2\pi} r\lambda(\boldsymbol{s}(r,\theta)p(r)\mathrm{d}\theta \\
&= 2\pi r\lambda(\boldsymbol{s_0})p(r)
\end{aligned}
$$

Add a $\log(2\pi)$ offset for not knowing θ and a $\log r$ offset to account for the fact that we consider a larger area with increasing distance.

Point transect example - take home messages

- Conceptually nice one-stage model - avoids binning points into counts and uncertainty propagation between two stages

Point transect example - take home messages

- Conceptually nice one-stage model - avoids binning points into counts and uncertainty propagation between two stages
- Intensities can be derived for data even where you cannot draw a point on a map (more on this next)

Point transect example - take home messages

- Conceptually nice one-stage model - avoids binning points into counts and uncertainty propagation between two stages
- Intensities can be derived for data even where you cannot draw a point on a map (more on this next)
- Iterated INLA a general tool for more than just fitting a thinning probability function

Spatial Capture-Recapture

- capture-recapture methods have a long history of being used to estimated the size of a population

Spatial Capture-Recapture

- capture-recapture methods have a long history of being used to estimated the size of a population
- **spatial** capture-recapture uses the location information of captures and recaptures

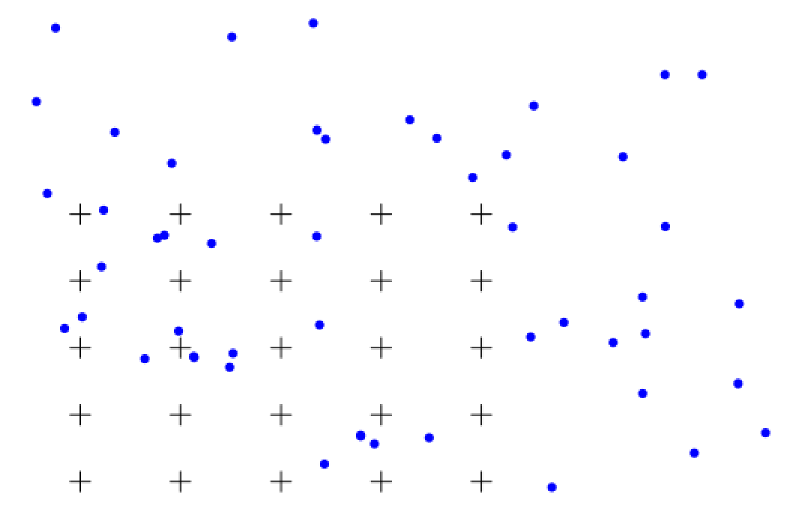
Spatial Capture-Recapture

- capture-recapture methods have a long history of being used to estimated the size of a population
- **spatial** capture-recapture uses the location information of captures and recaptures
- a natural way to join capture-recapture data and spatial modelling

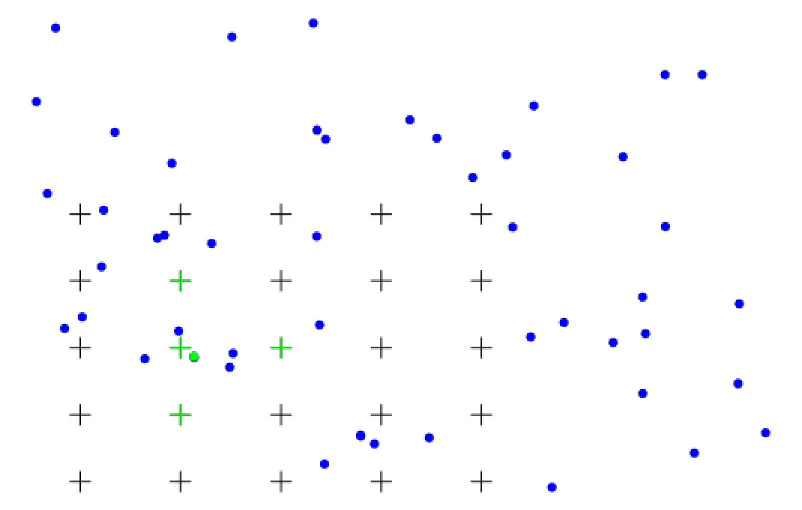
Spatial Capture-Recapture

source: snow leopard conservancy trust

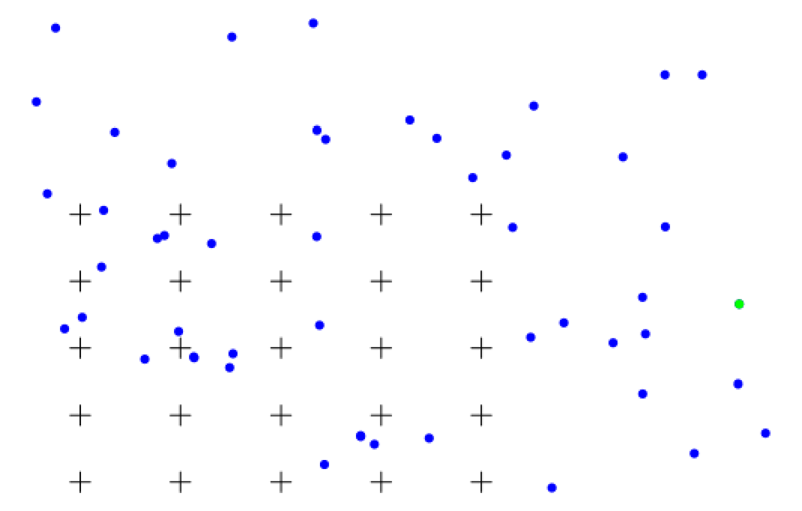
Spatial Capture-Recapture



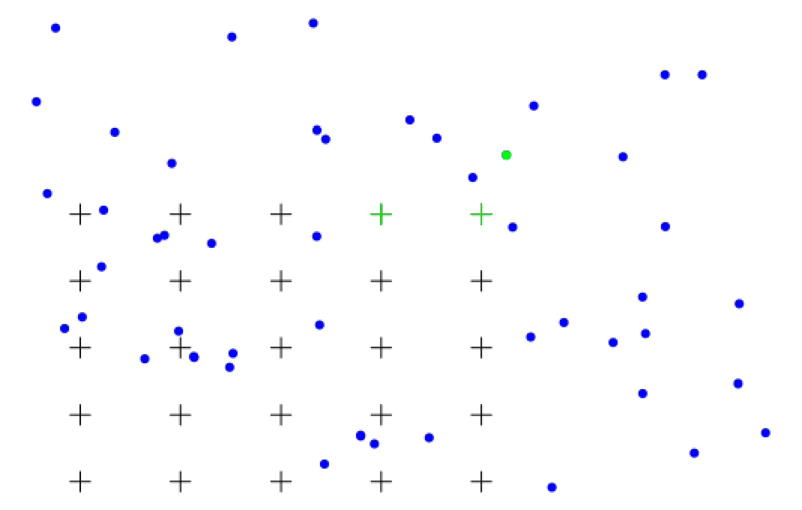
Spatial Capture-Recapture



Spatial Capture-Recapture



Spatial Capture-Recapture



Spatial Capture-Recapture - the likelihood

$$\mathcal{L}(\Omega) \propto \exp\left(-\int \lambda(\boldsymbol{s})p(\boldsymbol{s}|\phi)\mathrm{d}\boldsymbol{s}\right)\prod_{i=1}^{n}\int \pi(\Omega_i|\phi, \boldsymbol{s}_i)\lambda(\boldsymbol{s}_i)\mathrm{d}\boldsymbol{s}_i$$

Spatial Capture-Recapture - the likelihood

$$\mathcal{L}(\Omega) \propto \exp\left(-\int \lambda(\boldsymbol{s}) p(\boldsymbol{s}|\phi) \mathrm{d}\boldsymbol{s}\right) \prod_{i=1}^{n} \int \pi(\Omega_i|\phi, \boldsymbol{s}_i) \lambda(\boldsymbol{s}_i) \mathrm{d}\boldsymbol{s}_i$$

- The thinning and the estimation of the activity centre location share parameters ϕ

Spatial Capture-Recapture - the likelihood

$$\mathcal{L}(\Omega) \propto \exp\left(-\int \lambda(\boldsymbol{s})p(\boldsymbol{s}|\phi)\mathrm{d}\boldsymbol{s}\right)\prod_{i=1}^{n}\int \pi(\Omega_i|\phi, \boldsymbol{s}_i)\lambda(\boldsymbol{s}_i)\mathrm{d}\boldsymbol{s}_i$$

- The thinning and the estimation of the activity centre location share parameters ϕ
- Inference usually maximum-likelihood or Bayesian approach in MCMC

Spatial Capture-Recapture - the likelihood

$$\mathcal{L}(\Omega) \propto \exp\left(-\int \lambda(\boldsymbol{s}) p(\boldsymbol{s}|\phi) \mathrm{d}\boldsymbol{s}\right) \prod_{i=1}^{n} \int \pi(\Omega_i|\phi, \boldsymbol{s}_i) \lambda(\boldsymbol{s}_i) \mathrm{d}\boldsymbol{s}_i$$

- The thinning and the estimation of the activity centre location share parameters ϕ
- Inference usually maximum-likelihood or Bayesian approach in MCMC
- Typically $\lambda(\boldsymbol{s})$ is assumed constant or linear combination of fixed effect covariates

Spatial Capture-Recapture - the likelihood

$$\mathcal{L}(\Omega) \propto \exp\left(-\int \lambda(\boldsymbol{s}) p(\boldsymbol{s}|\phi)\mathrm{d}\boldsymbol{s}\right) \prod_{i=1}^{n} \int \pi(\Omega_i|\phi, \boldsymbol{s}_i)\lambda(\boldsymbol{s}_i)\mathrm{d}\boldsymbol{s}_i$$

- The thinning and the estimation of the activity centre location share parameters ϕ
- Inference usually maximum-likelihood or Bayesian approach in MCMC
- Typically $\lambda(\boldsymbol{s})$ is assumed constant or linear combination of fixed effect covariates
- Watch this space for $\lambda(\boldsymbol{s})$ a realisation of log-Gaussian Cox process...

Summary

- Thinning functions are a natural way to account for how ecologists observe point patterns

Summary

- Thinning functions are a natural way to account for how ecologists observe point patterns
- Complex observation processes can be represented as a thinning - conditioning on auxiliary data (such as distances and other covariates) or being derived from a more complicated observation model (as in SCR)

Summary

- Thinning functions are a natural way to account for how ecologists observe point patterns
- Complex observation processes can be represented as a thinning - conditioning on auxiliary data (such as distances and other covariates) or being derived from a more complicated observation model (as in SCR)
- General software for specifying thinning functions has the potential to be widely used

Summary

- Thinning functions are a natural way to account for how ecologists observe point patterns
- Complex observation processes can be represented as a thinning - conditioning on auxiliary data (such as distances and other covariates) or being derived from a more complicated observation model (as in SCR)
- General software for specifying thinning functions has the potential to be widely used
- Potential for thinning to share information between for multiple observation processes e.g. citizen science, combining multiple data sources etc

Summary

Thanks for listening!