One-stage distance sampling using iterated integrated nested Laplace approximations

Andy Seaton

University of St Andrews

Acknowledgements

Prof Janine Illian, University of Glasgow

Prof David Borchers, University of St Andrews

Dr Richard Glennie, University of St Andrews

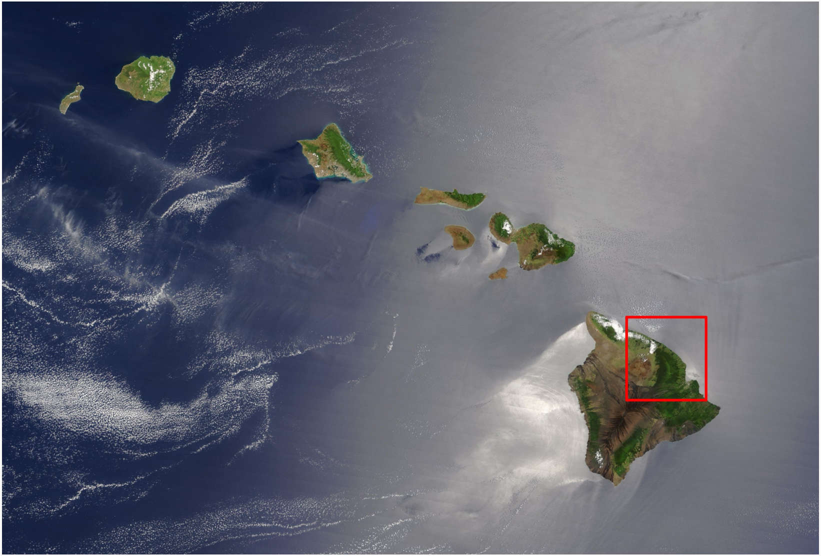
Prof Finn Lindgren, University of Edinburgh

Dr Fabian Bachl, University of Edinburgh

Dr Rick Camp, US Geological Survey

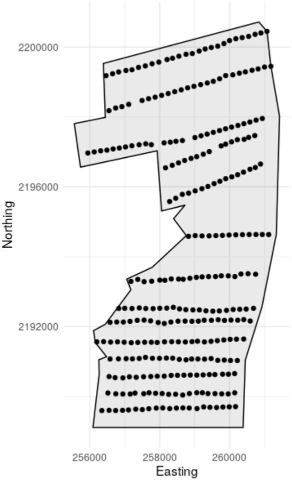
Dr David Miller, University of St Andrews

Akepa Case Study



Akepa Case Study

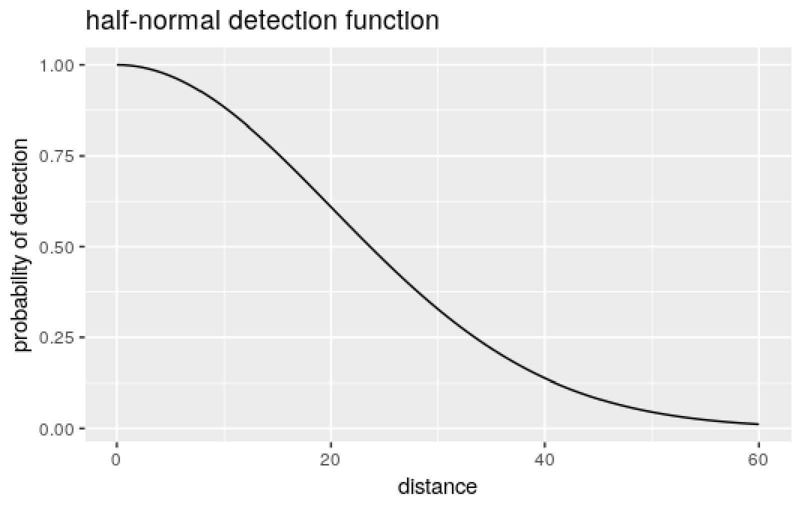
Hawaii Forest Bird Survey transect locations



Akepa Case Study

source: Jack Jeffrey, US Fish and Wildlife Service

What is distance sampling?



Spatial modelling with distance sampling data

The traditional two-stage approach:

Spatial modelling with distance sampling data

The traditional two-stage approach:

1. Estimate the detection function

Spatial modelling with distance sampling data

The traditional two-stage approach:

1. Estimate the detection function
2. Bin data into counts

Spatial modelling with distance sampling data

The traditional two-stage approach:

1. Estimate the detection function
2. Bin data into counts
3. For each bin, calculate the 'effective sampled area' (probability of detection x size of bin)

Spatial modelling with distance sampling data

The traditional two-stage approach:

1. Estimate the detection function
2. Bin data into counts
3. For each bin, calculate the 'effective sampled area' (probability of detection x size of bin)
4. Fit a count GLM/GAM with an offset term

Spatial modelling with distance sampling data

The traditional two-stage approach:

1. Estimate the detection function
2. Bin data into counts
3. For each bin, calculate the 'effective sampled area' (probability of detection x size of bin)
4. Fit a count GLM/GAM with an offset term

Pick your favourite GLM/GAM software (`mgcv` is a popular choice)

`R-INLA`

`R-INLA` allows us to fit GLMs/GAMs

`R-INLA`

`R-INLA` allows us to fit GLMs/GAMs

Why bother?

`R-INLA` allows us to fit GLMs/GAMs

Why bother?

1. You want to be **Bayesian** but you are too impatient for MCMC (INLA is fast approximate Bayesian method)

R-INLA

`R-INLA` allows us to fit GLMs/GAMs

Why bother?

1. You want to be **Bayesian** but you are too impatient for MCMC (INLA is fast approximate Bayesian method)
2. **Incorporate prior knowledge** of the species (who does this though?)

`R-INLA` allows us to fit GLMs/GAMs

Why bother?

1. You want to be **Bayesian** but you are too impatient for MCMC (INLA is fast approximate Bayesian method)
2. **Incorporate prior knowledge** of the species (who does this though?)
3. **One-stage model fit** - simultaneously estimate detection and spatial models

`R-INLA` allows us to fit GLMs/GAMs

Why bother?

1. You want to be **Bayesian** but you are too impatient for MCMC (INLA is fast approximate Bayesian method)
2. **Incorporate prior knowledge** of the species (who does this though?)
3. **One-stage model fit** - simultaneously estimate detection and spatial models
4. **Sample from joint and marginal posterior distributions** - uncertainty in detectability is accounted for naturally

R-INLA

`R-INLA` allows us to fit GLMs/GAMs

Why bother?

1. You want to be **Bayesian** but you are too impatient for MCMC (INLA is fast approximate Bayesian method)
2. **Incorporate prior knowledge** of the species (who does this though?)
3. **One-stage model fit** - simultaneously estimate detection and spatial models
4. **Sample from joint and marginal posterior distributions** - uncertainty in detectability is accounted for naturally
5. Use popular `R-INLA` features (**point process likelihood**, **scalable sparse random effects**, spatio-temporal models, barrier model, extreme value distributions, penalised complexity priors)

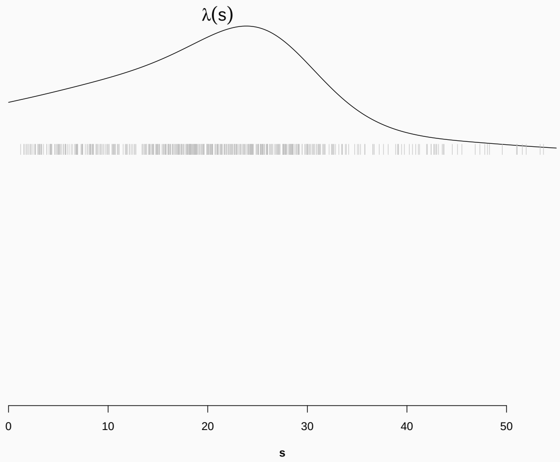
R-INLA

`R-INLA` allows us to fit GLMs/GAMs

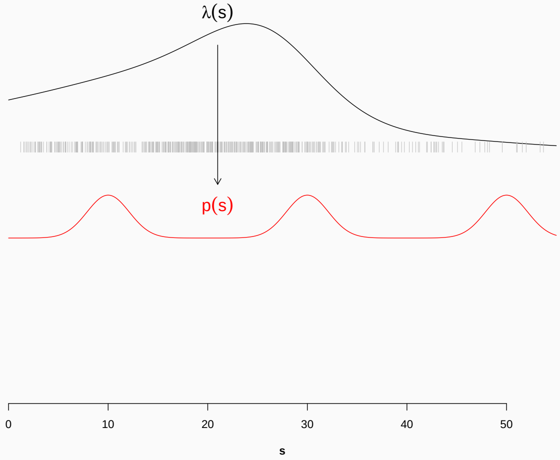
Why bother?

1. You want to be **Bayesian** but you are too impatient for MCMC (INLA is fast approximate Bayesian method)
2. **Incorporate prior knowledge** of the species (who does this though?)
3. **One-stage model fit** - simultaneously estimate detection and spatial models
4. **Sample from joint and marginal posterior distributions** - uncertainty in detectability is accounted for naturally
5. Use popular `R-INLA` features (**point process likelihood**, **scalable sparse random effects**, spatio-temporal models, barrier model, extreme value distributions, penalised complexity priors)
6. Support for multiple likelihoods (e.g. multiple data sources)

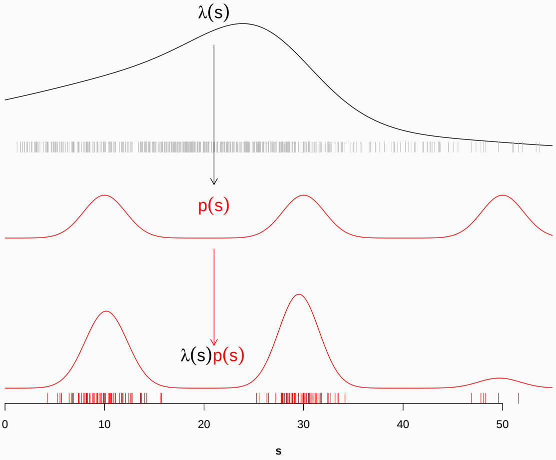
A Point Process Perspective



A Point Process Perspective



A Point Process Perspective



A Point Process Perspective

The intensity for *detected* points is $\tilde{\lambda}(s) = \lambda(s)p(s)$

A Point Process Perspective

The intensity for *detected* points is $\tilde{\lambda}(s) = \lambda(s)p(s)$

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

A Point Process Perspective

The intensity for *detected* points is $\tilde{\lambda}(s) = \lambda(s)p(s)$

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

$\log \lambda(s)$ is the **spatial process model**

A Point Process Perspective

The intensity for *detected* points is $\tilde{\lambda}(s) = \lambda(s)p(s)$

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

$\log \lambda(s)$ is the **spatial process model**

$\log p(s)$ is the **observation process model**

A Point Process Perspective

The intensity for *detected* points is $\tilde{\lambda}(s) = \lambda(s)p(s)$

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

$\log \lambda(s)$ is the **spatial process model**

$\log p(s)$ is the **observation process model**

This model is implemented in `inlabru` as a `"cp"` (Cox process) likelihood

Iterated INLA

$p(r) = \exp\left(-r^2/2\sigma^2\right)$ where $\sigma^2 > 0$

Iterated INLA

$p(r) = \exp\left(-r^2/2\sigma^2\right)$ where $\sigma^2 > 0$

$\log p(r)$ is non-linear in σ^2

Iterated INLA

$p(r) = \exp\left(-r^2/2\sigma^2\right)$ where $\sigma^2 > 0$

$\log p(r)$ is non-linear in σ^2

If we knew $p(r)$ exactly then this would act like an offset term

Iterated INLA

$p(r) = \exp\left(-r^2/2\sigma^2\right)$ where $\sigma^2 > 0$

$\log p(r)$ is non-linear in σ^2

If we knew $p(r)$ exactly then this would act like an offset term

1. Guess $p(r)$
2. Fit a model with $p(r)$ as offset
3. **Search** for a better guess
4. Iterate until **convergence**

Iterated INLA

$p(r) = \exp\left(-r^2/2\sigma^2\right)$ where $\sigma^2 > 0$

$\log p(r)$ is non-linear in σ^2

If we knew $p(r)$ exactly then this would act like an offset term

1. Guess $p(r)$
2. Fit a model with $p(r)$ as offset
3. **Search** for a better guess
4. Iterate until **convergence**

Search is Newton-Raphson + some tweaks

Iterated INLA

$p(r) = \exp\left(-r^2/2\sigma^2\right)$ where $\sigma^2 > 0$

$\log p(r)$ is non-linear in σ^2

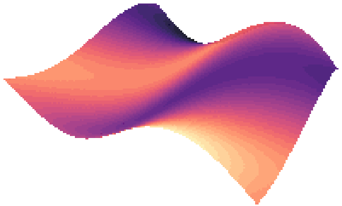
If we knew $p(r)$ exactly then this would act like an offset term

1. Guess $p(r)$
2. Fit a model with $p(r)$ as offset
3. **Search** for a better guess
4. Iterate until **convergence**

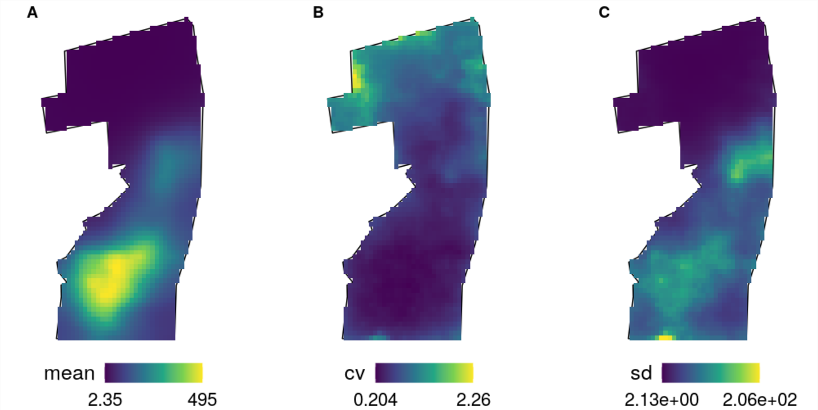
Search is Newton-Raphson + some tweaks

Convergence defined as 'the modes of the marginal posteriors don't change'

Spatial Model

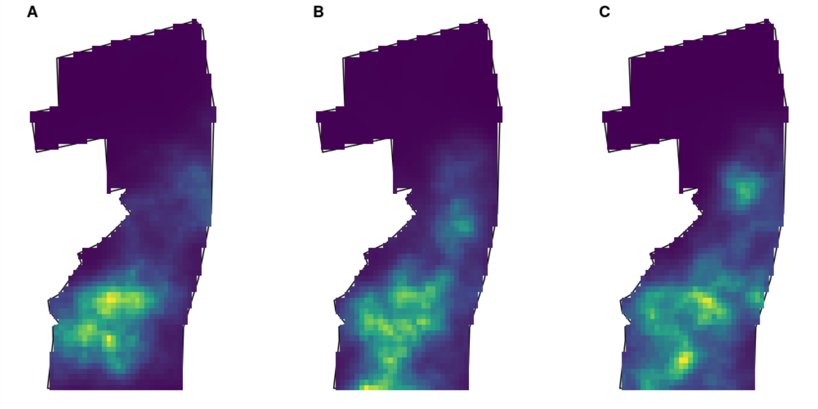


Akepa case study



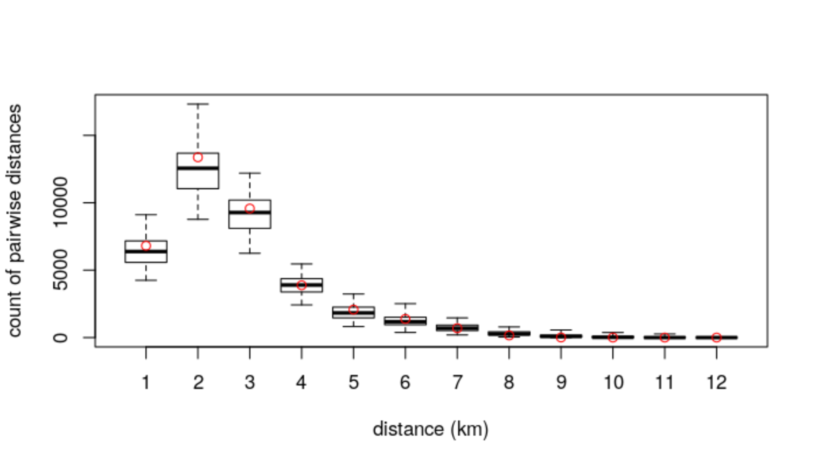
Summary statistics of the posterior intensity field

Akepa case study



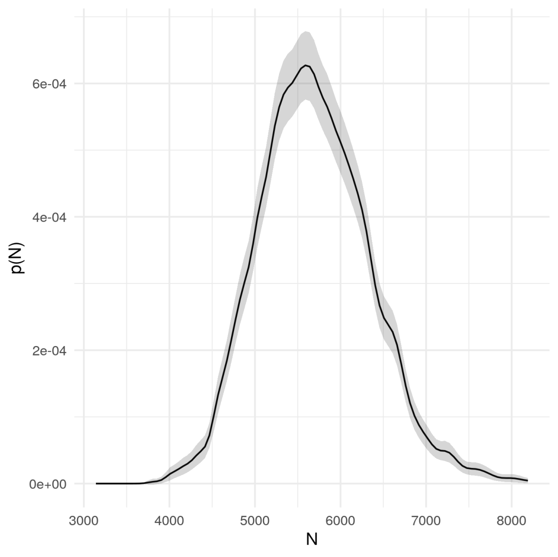
Three realisations of the posterior intensity field

Akepa case study



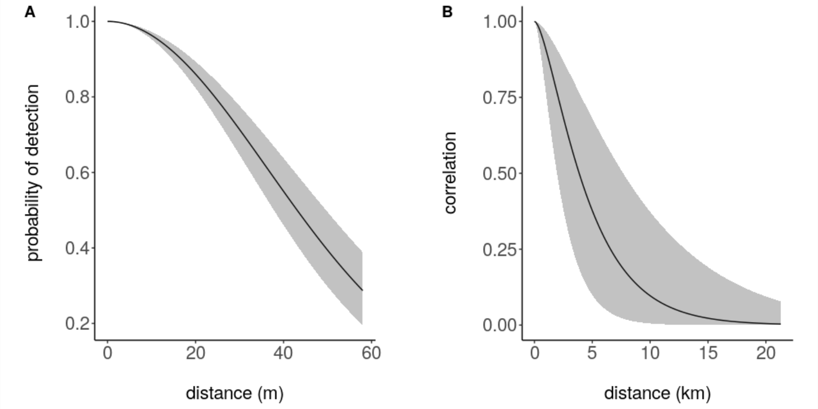
Analagous to Ripley's K-function

Akepa case study



Posterior of realised abundance $\pm$ 2 Monte Carlo standard errors

Akepa case study



Posterior detection function and Matérn covariance function

References

F. E. Bachl, F. Lindgren, D. L. Borchers, and J. B. Illian.
Inlabru: An R package for Bayesian spatial modelling from ecological survey data.
Methods in Ecology and Evolution, 2019.

M. V. Bravington, D. L. Miller, and S. L. Hedley.
Reliable variance propagation for spatial density surface models.
arXiv:1807.07996 [stat], July 2018.

R. J. Camp, D. L. Miller, L. Thomas, S. T. Buckland, and S. J. Kendall.
Using density surface models to estimate spatio-temporal changes in population densities and trend.
Ecography, 2020.

D. L. Miller, R. Glennie, and A. E. Seaton.
Understanding the Stochastic Partial Differential Equation Approach to Smoothing.
Journal of Agricultural, Biological and Environmental Statistics, Sept. 2019.

Y. Yuan, F. E. Bachl, F. Lindgren, D. L. Borchers, J. B. Illian, S. T. Buckland, H. Rue, and T. Gerrodette.
Point process models for spatio-temporal distance sampling data from a large-scale survey of blue whales.
The Annals of Applied Statistics, 2017.

Extras!


The sea of non-linearity
The land of latent gaussian models
Here be Dragons
Here we worry Identifiability? Multi-modality?
Here we trust Laplace approximations

Non-linear model components in spatial ecology

Functional responses:

Non-linear model components in spatial ecology

Functional responses:

- Diminishing returns (saturating responses)

Non-linear model components in spatial ecology

Functional responses:

- Diminishing returns (saturating responses)
- Trade-offs between resources (prey-choice literature)

Non-linear model components in spatial ecology

Functional responses:

- Diminishing returns (saturating responses)
- Trade-offs between resources (prey-choice literature)
- Selectivity curves for fisheries data

Non-linear model components in spatial ecology

Functional responses:

- Diminishing returns (saturating responses)
- Trade-offs between resources (prey-choice literature)
- Selectivity curves for fisheries data
- Force positive / negative / quadratic relationships

Non-linear model components in spatial ecology

Functional responses:

- Diminishing returns (saturating responses)
- Trade-offs between resources (prey-choice literature)
- Selectivity curves for fisheries data
- Force positive / negative / quadratic relationships

Replace random effects with parametric dependence structures:

Non-linear model components in spatial ecology

Functional responses:

- Diminishing returns (saturating responses)
- Trade-offs between resources (prey-choice literature)
- Selectivity curves for fisheries data
- Force positive / negative / quadratic relationships

Replace random effects with parametric dependence structures:

- Logistic growth

Non-linear model components in spatial ecology

Functional responses:

- Diminishing returns (saturating responses)
- Trade-offs between resources (prey-choice literature)
- Selectivity curves for fisheries data
- Force positive / negative / quadratic relationships

Replace random effects with parametric dependence structures:

- Logistic growth
- Ricker growth

Non-linear model components in spatial ecology

Functional responses:

- Diminishing returns (saturating responses)
- Trade-offs between resources (prey-choice literature)
- Selectivity curves for fisheries data
- Force positive / negative / quadratic relationships

Replace random effects with parametric dependence structures:

- Logistic growth
- Ricker growth
- The population ecology literature is vast

Non-linear model components in spatial ecology

Functional responses:

- Diminishing returns (saturating responses)
- Trade-offs between resources (prey-choice literature)
- Selectivity curves for fisheries data
- Force positive / negative / quadratic relationships

Replace random effects with parametric dependence structures:

- Logistic growth
- Ricker growth
- The population ecology literature is vast

Other things:

Non-linear model components in spatial ecology

Functional responses:

- Diminishing returns (saturating responses)
- Trade-offs between resources (prey-choice literature)
- Selectivity curves for fisheries data
- Force positive / negative / quadratic relationships

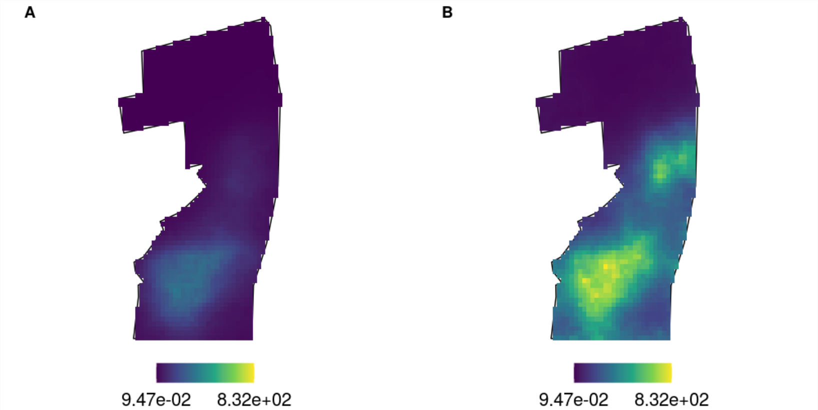
Replace random effects with parametric dependence structures:

- Logistic growth
- Ricker growth
- The population ecology literature is vast

Other things:

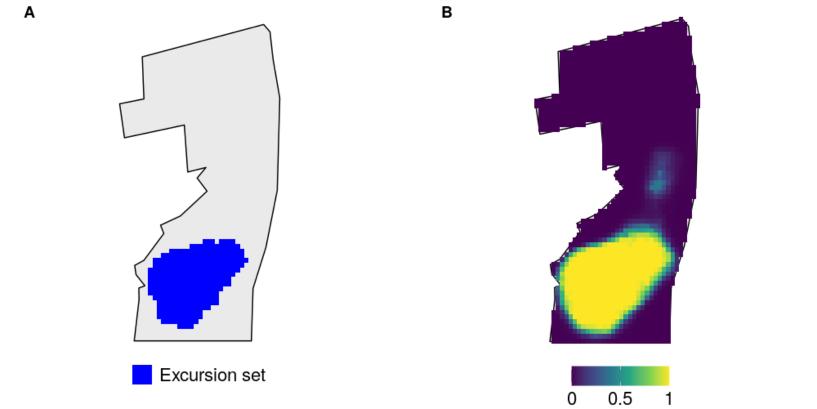
- Level set Cox process - allow a mixture of random fields

Extras!



0.025 and 0.975 pointwise prediction quantiles for the posterior intensity field

Extras!



Excursion set for 1 bird per hectare and 95% probability level and corresponding excursion function