Some other approaches (and software) for spatial modelling of distance sampling data

Primary reference:

- Miller, D.L. et al. 2013. Spatial models for distance sampling data: recent developments and future directions. Methods in Ecology and Evolution. 4:1001-1010

More references at end…

General formulation[1]

- Point process model
  - Location of animals are a realization of an underlying process with intensity $D(x,y)$
  - Then, number of animals in an area A is Poisson(λ) where

$$\lambda = \int_A D(x,y)dxdy$$

  i.e. an inhomogeneous Poisson process (IPP)

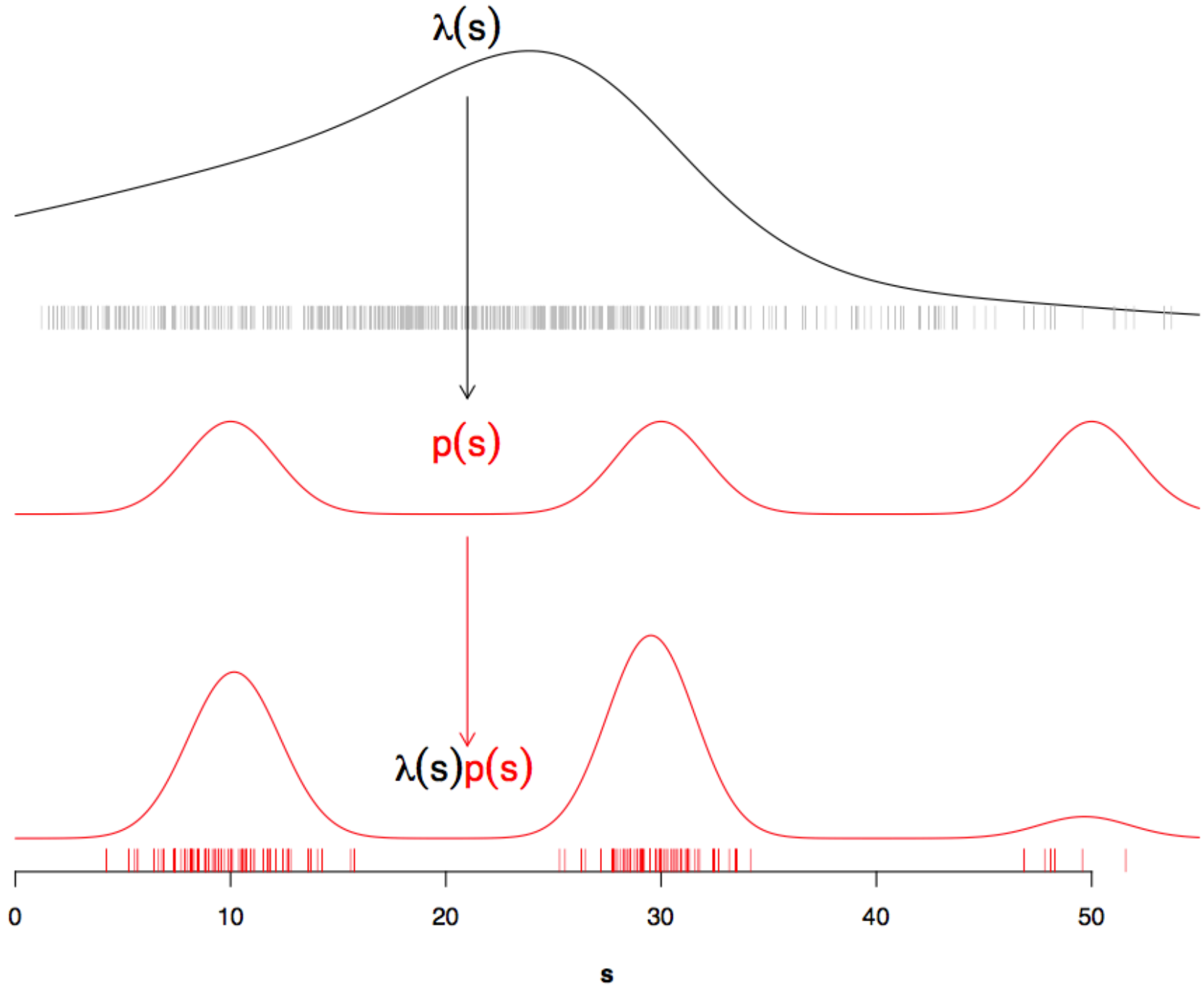
[1]see Hedley (2000) and Johnson et al. (2010)

Adding uncertain detection

- Let $g(x,y)$ be the probability of detection of an animal given it is at location x,y
- Detection of animals represents a “thinning” of the IPP, which also yields an IPP, with

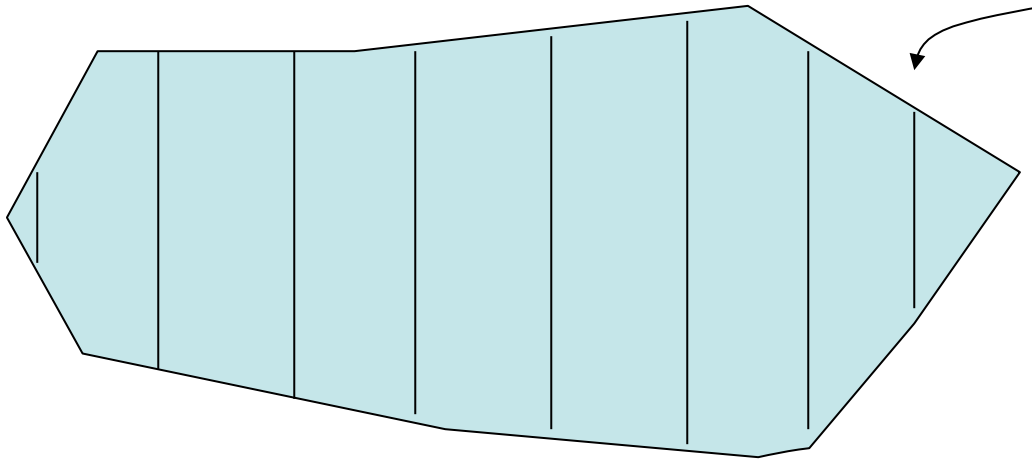
$$\lambda = \int_A D(x,y)g(x,y)dxdy$$

(Can use marked point process framework to deal with animal-level covariates in $g()$)


λ(s)
p(s)
λ(s)p(s)
0
10
20
30
40
50
s

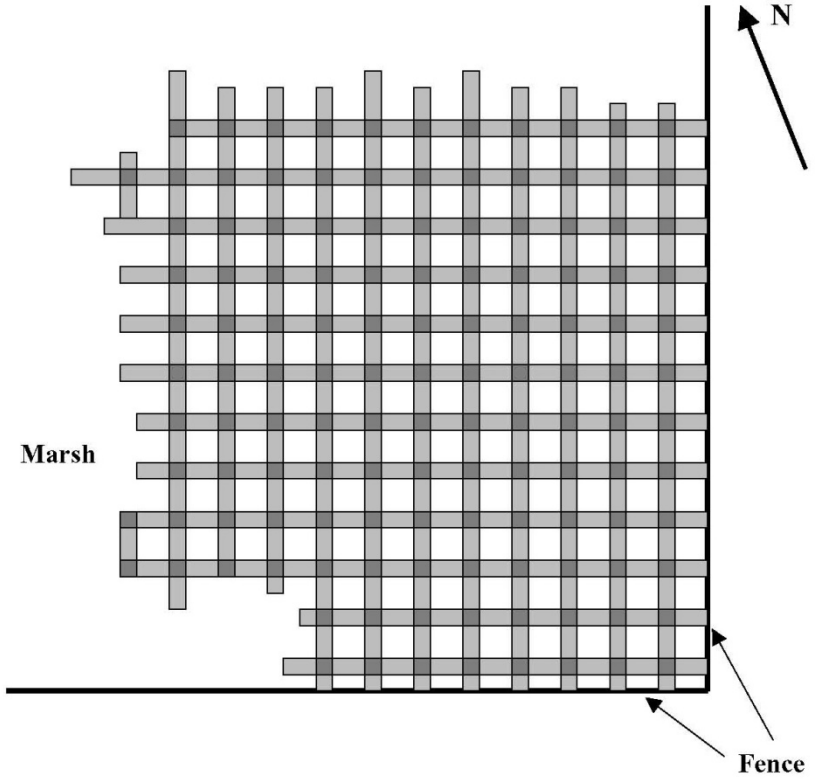
Full likelihood approach

- Simultaneously estimate parameters of $D()$ and $g()$
- Is it worth it? Often little information about $D()$ contained in the distances (which are used to fit the $g()$)



- width of transects small compared with width of study area,
- little information about change in density in the x direction
- contained in the observed distances from the transect

Exceptions



- Covered area is a large proportion of study area

Example Applications

- Hedley (2000)
- Royle et al. (2004) (nicely illustrates lack of information about $D()$ from the $g()$) – `distsamp in R` (package `unmarked`)
  - see also `gdistsamp` (same package)
- Royle and Dorazio (2008, section 5.5.5) [1] - `WinBUGS`
- Johnson et al. (2010) – `dspat in R (DSpat library)`
- `inlabru R` package; Yuan et al. (2017) software and training materials at `https://sites.google.com/inlabru.org/inlabru/download`

[1] See also Kéry and Royle 2015. Applied Hierarchical Modelling in Ecology Vol 1.

Two-stage methods

- First fit $g()$ using the distances
- Then fit $D()$ using the locations of observations, and the fitted $g()$
- Example: method we've been using on this workshop ("Count method" of Hedley (2000), Hedley and Buckland (2004), Hedley et al. (2004))

Advantages

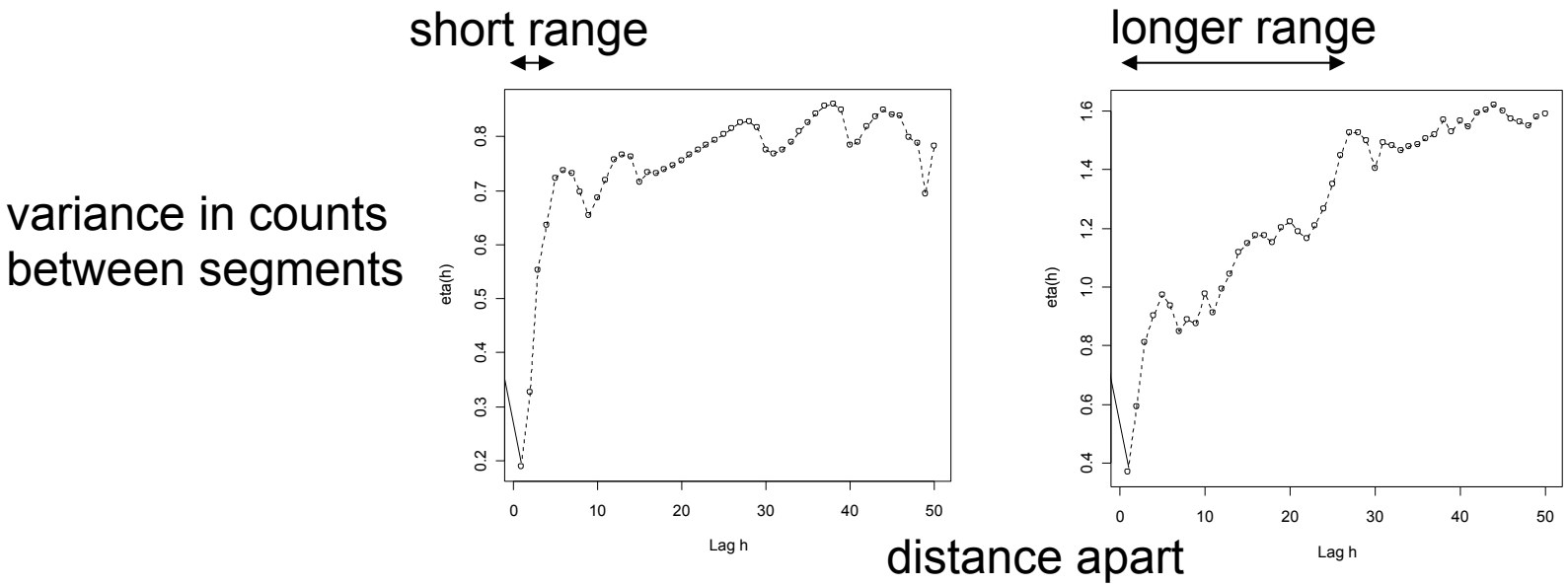
- 2 stage methods in general:
  - Divide and conquer
- Count method:
  - Uses standard, well accepted tools for fitting
  - GLM/GAM is reliable, robust
  - Software available
  - Conceptually simple

Issues

- Residual small-scale spatial clustering
- Extrapolation problems
- Animals in clusters (groups, pods, etc)

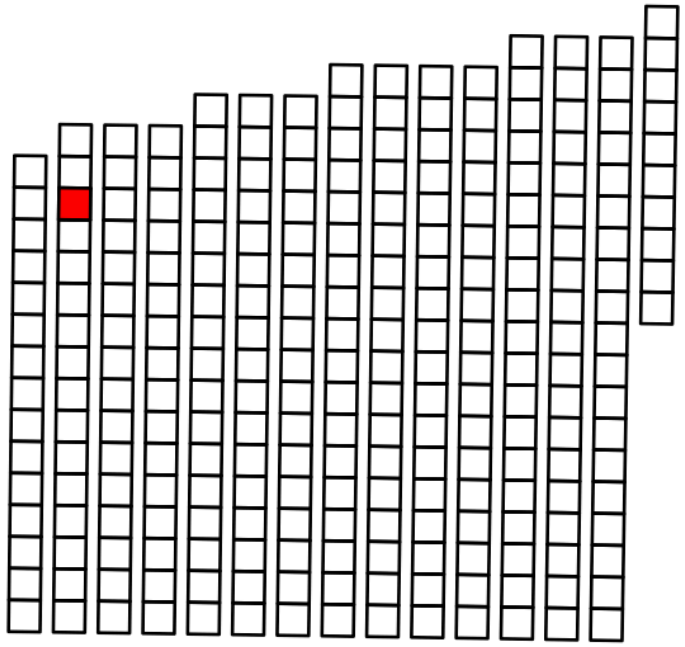
Residual small-scale clustering

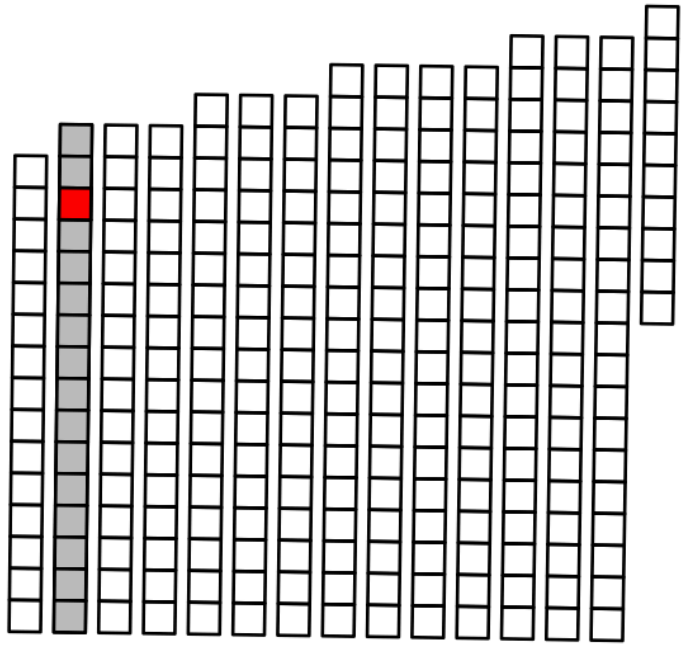
- Our models assume large-scale smooth pattern of $D()$ but there is often residual unexplained variation at a smaller scale



plots courtesy of Louise Burt

“Natural” structure





Solutions

- Better/different models
  - Within GLM/GAM framework (see also next slide):
    - mixed models (GLMM/GAMM) with spatial correlation in the error structure (similar to extensions of kriging)
    - auto-regressive models
    - estimating equations (GEE) with spatial correlation in error structure
  - Extensions/alterations to IPP
    - spatial shot-noise Cox process (Waagepeterson and Schweder 2006)
    - Markov-modulated Poisson process (Skaug 2006)
    - Bayesian tesselations. (Niemi and Fernandez 2010; Kaimi and Fernandez 2008)

Animals in clusters/groups/pods

- When object detected is a group of animals, can:
  - Model individuals directly, rather than clusters (can use this to indirectly get a density surface of clusters)
  - Model density surface of clusters, and then scale up by mean cluster size
  - Separately model density surface of clusters and density surface of cluster size, then put together (therefore 3 stage modelling – e.g., BWH approach)

The Future: DenMod

- CREEM + Duke + NOAA + others (US Navy funded)
- Next generation of spatial modelling in Distance
  - New methodology, R packages, practical advice
  - Multiple survey analysis
  - Segment size issues (MMPP etc)
  - Extrapolation
  - Group size uncertainty & more!

Selected additional references

- Buckland, S.T., D.L. Borchers, A. Johnston, P.A. Henrys and T.A. Marques. 2007. Line transect methods for plant surveys. Biometrics 63: 989-998.
- Ferguson, M.C., J. Barlow, P. Fiedler, S.B. Reilly and T. Gerrodette. 2006. Spatial models of delphinid (family Delphinidae) encounter rate and group size in the eastern tropical Pacific Ocean. Ecological Modelling 193: 645-662.
- Kaimi, I. and C. Fernandez. 2008. Bayesian spatial modelling of line transect data. PhD chapter (Lancaster University).
- Hedley, S.L. 2000. Modelling heterogeneity in cetacean surveys. PhD thesis, University of St Andrews. http://www.creem.st-and.ac.uk/sharon/public_html/thesis.pdf
- Hedley, S.L. and S.T Buckland. 2004. Spatial models for line transect sampling. Journal of Agricultural, Biological and Environmental Statistics 9: 181-199.
- Hedley, S.L., S.T. Buckland and D.L. Borchers. 2004. Spatial distance sampling models. Chapter 4 of Advanced Distance Sampling.
- Johnson, D.S., J.L. Laake and J.M. Ver Hoef. 2010. A model-based approach for making ecological inference from distance sampling data. Biometrics 66: 310-318
- Kastdalen, L, H.C, Pedersen, G. Fjone, and H.P. Andreassen. 2003. Combining resource selection functions and distance sampling: an example with willow ptarmigan. Proceeding Symp. Remote sensing and Resource Selection. (builds on book by B.F.J. Manly et al. 2002. Resource selection by animals.)
- Niemi A. and C. Fernandez. 2010. Bayesian spatial point process modelling of line transect survey data. Journal of Agricultural, Biological and Environmental Statistics 15: 327.345.
- Royle, J.A., D.K. Dawson and S. Bates. 2004. Modelling abundance effects in distance sampling. Ecology 85: 1591-1597.
- Royle, J.A. and R.M. Dorazio. 2008. Hierarchical Inference and Modelling in Ecology. Academic Press.
- Skaug, H.J. 2006. Markov modulated Poisson processes for clustered line transect data. Environmental and Ecological Statistics. 13: 199-211.
- Waagepetersen, R. and T. Schweder. 2006. Likelihood-based inference for clustered line transect data. Journal of Agricultural, Biological and Environmental Statistics 11: 264-279.
- Williams, R., S.L. Hedley, T.A. Branch, M.V. Bravington, A.N. Zerbini and K.P. Findlay. 2011. Chilean Blue Whales as a case study to illustrate methods to estimate abundance and evaluate conservation status of rare species. Conservation Biology 25: 526-535.
- Yuan, Y. D.E. Bachl, D.L. Borchers, F. Lindgren, J.B. Illian, S.T. Buckland, H. Rue and T. Gerrodette. Submitted. Point process models for spatio-temporal distance sampling data. http://goo.gl/QUVhoK

Note: A more complete (but slightly out-of-date) list is contained in Miller et al. (2013).