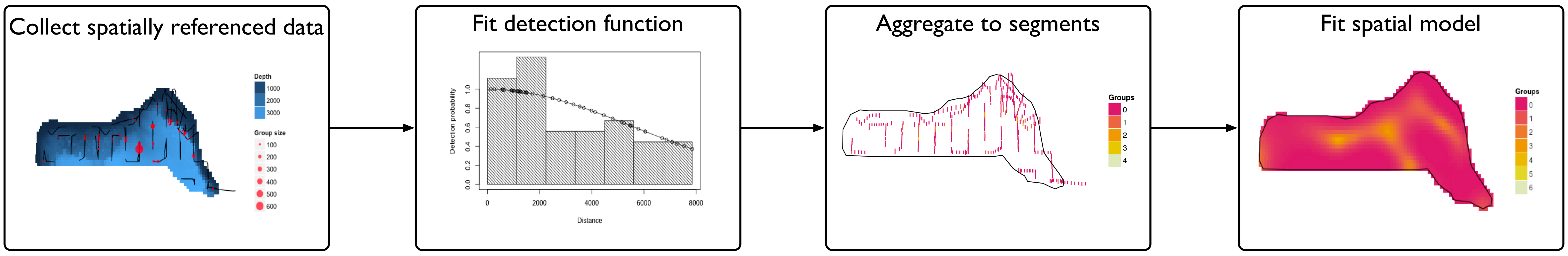
Spatial density surface estimation from distance sampling surveys

David L. Miller, Len Thomas, Laura Marshall
Centre for Research into Ecological and Environmental Modelling
University of St Andrews

Density surface modelling



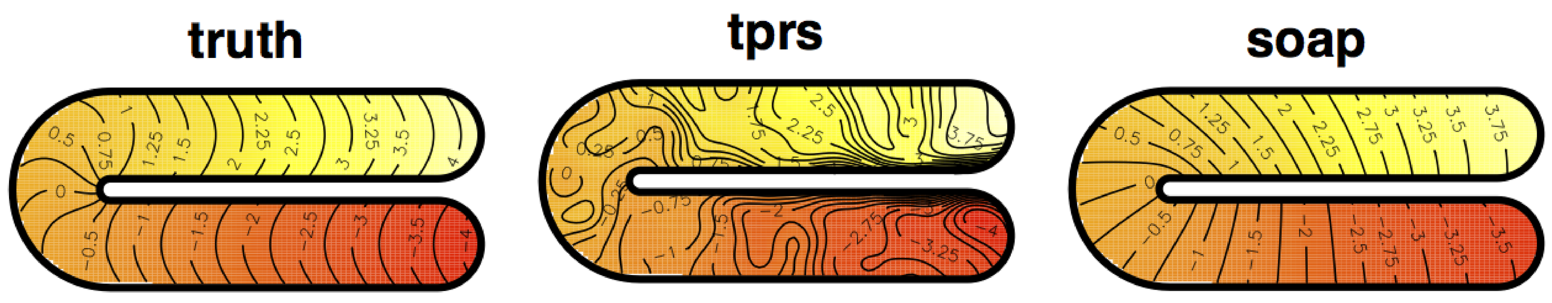
In Distance we follow the approach of Hedley and Buckland (2004). Having fit a detection function, we aggregate the effective strip widths (CDS) or estimated abundances to segments (MCDS). We then fit a spatially referenced model to the segment data. Models include:

$$\mathbb{E}(\mathbf{n_j}) = \exp\left[\log_e(2\hat{\mu}\mathbf{l_j}) + \sum_k \mathbf{f_k}(\mathbf{z_{jk}})\right] \text{ (CDS models)} \quad \textit{or} \quad \mathbb{E}(\hat{\mathbf{N}}_\mathbf{j}) = \exp\left[\log_e(2w\mathbf{l_j}) + \sum_k \mathbf{f_k}(\mathbf{z_{jk}})\right] \text{ (MCDS models)}$$

where $\mathbf{n_j}$ is count per segment, $\hat{\mu}$ is the effective strip width, $\mathbf{l_j}$ is the length of the segment, $\hat{\mathbf{N}}_\mathbf{j}$ is the (Horvitz-Thompson) estimated abundance in the segment and $\mathbf{w}$ is the truncation distance. $\mathbf{j = 1, \ldots, J}$ index the segments. The $\mathbf{f_k}$s are smooths of environmental covariates $\mathbf{z_{jk}}$.

Recent developments

Complex region smoothers



- Often the study region has an odd shape.
- This can lead to incorrect inference.
- Recent advances in spatial modelling allow us to work around this.
- We opt for the soap film smoother approach of Wood et al. (2008).

Variance propagation

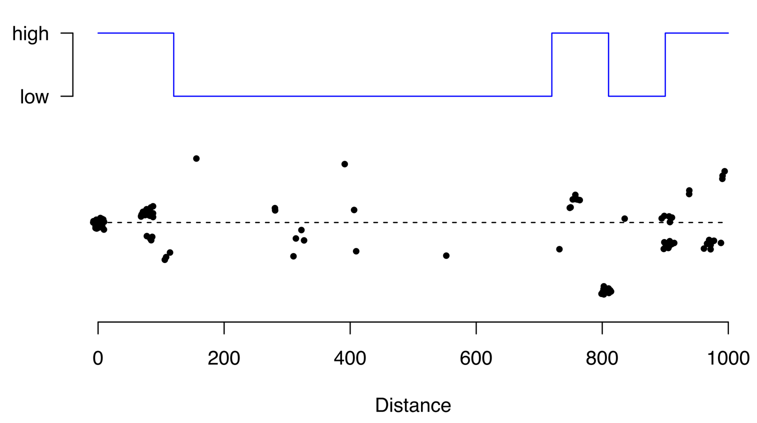
- Uncertainty in detection function estimation and the spatial model must be combined.
- Usually achieved using the *delta method* $\Rightarrow$ independence between detection process and the spatial process
- Clearly this is not the case!
- Williams et al. (2011) propose a method of *variance propagation*:
  - Fit a spatial model with

$$\mathbb{E}(\mathbf{n_j}) = \exp\left[\log_e(2\hat{\mu}\mathbf{l_j}) + \left[\frac{\partial \log_e \widehat{\mathbf{P_a}}(\theta; \mathbf{z_j})}{\partial \theta}\right]_{\theta=\hat{\theta}} \cdot \gamma + \sum_k \mathbf{f_k}(\mathbf{z_{jk}})\right]$$

  where $\gamma = \theta - \hat{\theta}$ ($\hat{\theta}$ is the MLE of θ).
  - Derivative term can then be thought of as a random effect with parameter $\gamma \sim \text{MVN}(\mathbf{0}, -\mathbf{H}_{\hat{\theta}}^{-1})$.
  - Resulting variance from the GAM *includes* detection function variability.
- Only works with detection functions with no covariates (CDS).

Markov modulated Poisson process – Skaug (2006)

- Often observe clustering
- 2-state (high/low) process
- Biologically motivated
- Can include GLM/GAM components



New R package: `dsm`

- Individual (MCDS) as well as environmental covariates
- Binned and continuous & group and individual data
- Faster bootstrap
- New bootstrap method incorporating detection function uncertainty
- Soap film smoothing for complex regions (see left)
- Variance propagation (see left)
- CV plotting
- Tutorial available at `http://github.com/dill/dsm/wiki/Examples`
- Talks to `mrds` and the new package `Distance`
- In Distance 7.0, on CRAN soon!

Other approaches

- DSpat - Johnson et al. (2010)
  - Directly model the point process
  - Setection function as thinning of the process
  - (Spatial) mixture of detection functions
  - Over-dispersion handled by post-hoc correction
- `unmarked`
  - Not full spatial modelling but can use transect-level spatial covariates
  - Hierarchical approach
  - Binned data only
- Bayesian point processes via (RJ)MCMC - Niemi and Fernández (2010)
  - Intensity function – product of a parametric function of the covariates
  - Mixture of Gaussian kernels as a spatial smooth (priors on knots select smoothing)
  - Single precision parameter $\Rightarrow$ cannot accommodate both small- and large-scale variation
  - "Known" detection function as a thinning of the process

Coming soon!

A review paper incorporating all this information and **more**: *Spatial models for distance sampling data: recent developments and future directions.* David L. Miller, Louise Burt, Eric Rexstad and Len Thomas.

References

Hedley, S. L. and S. T. Buckland (2004). Spatial models for line transect sampling. *Journal of Agricultural, Biological, and Environmental Statistics 9*(2), 181–199.
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*(1), 310–318.
Niemi, A. and C. Fernández (2010). Bayesian Spatial Point Process Modeling of Line Transect Data. *Journal of Agricultural, Biological, and Environmental Statistics 15*(3), 327–345.
Skaug, H. J. (2006, June). Markov Modulated Poisson Processes for Clustered Line Transect Data. *Environmental and Ecological Statistics 13*(2), 199–211.
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*(3), 526–535.
Wood, S. N., M. V. Bravington, and S. L. Hedley (2008). Soap film smoothing. *Journal of the Royal Statistical Society. Series B, Statistical Methodology 70*(5), 931–955.