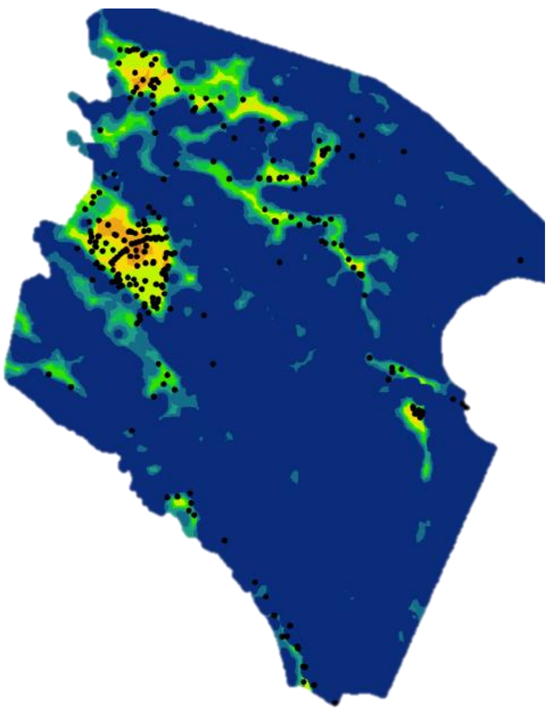
Modelling habitat selection across multiple spatial scales using varying coefficient regression

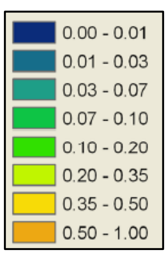
Thomas Cornulier - University of Aberdeen, UK

Alex Villers - CNRS-CEB Chize, France

Species-habitat relationship

Abundance = $f_1(\text{habitat}_1) + f_2(\text{habitat}_2) + \ldots + \varepsilon$

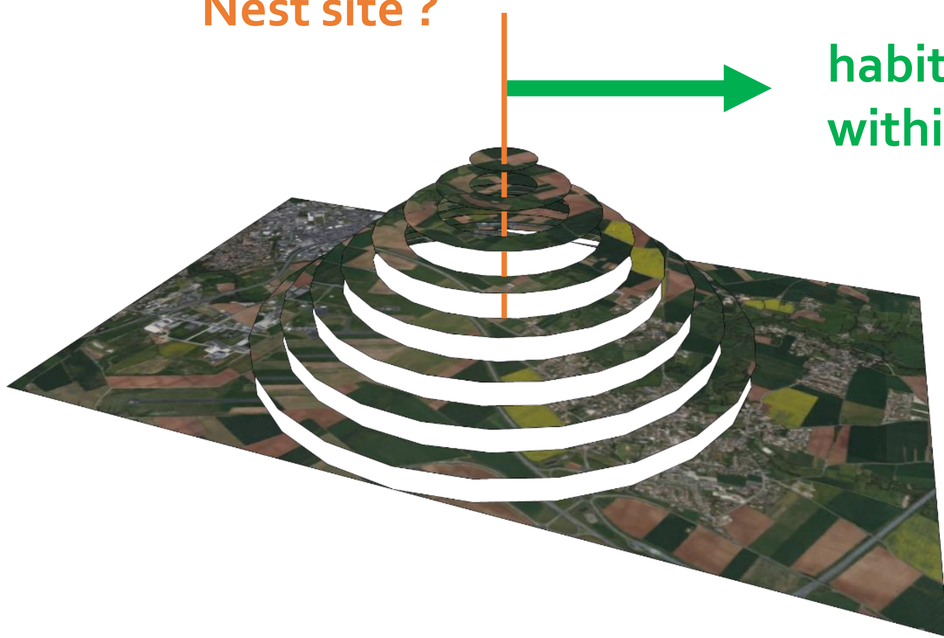




Predictor variables	Coefficient	SE
(Intercept)	−11.05	0.77
AgriPast	−0.71	0.97
AgriPast^2	0.43	0.30
WaterSurf	0.55	0.18
Log(AgriPerm)	−1.19	0.48
Aquatic	−0.28	0.07
Urban	−0.29	0.07
AgriIntens	−0.60	0.09
Stream	0.15	0.05
DistRoad	0.26	0.08
AgriHetero	−0.33	0.09

(Gadenne et al. 2014)

Scale-dependence



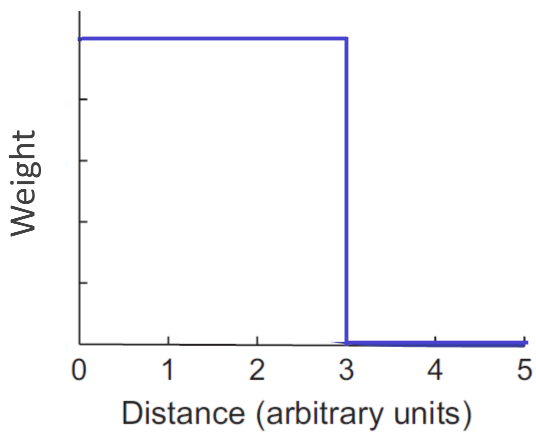
Statistical considerations (1)

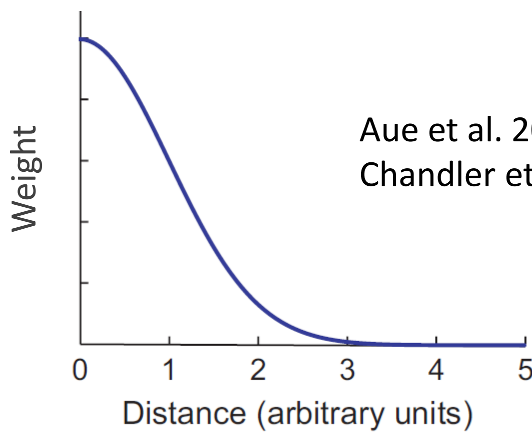
- Selecting optimal scale(s)
  - Model selection: likelihood profiling of fixed optimal scale?
  - Optimization / multiple testing
  - Control for effective # parameters / Overfitting?

Statistical considerations (1)

- Selecting optimal scale(s)
  - Model selection: likelihood profiling of fixed optimal scale?
  - Optimization / multiple testing
  - Control for effective # parameters / Overfitting?

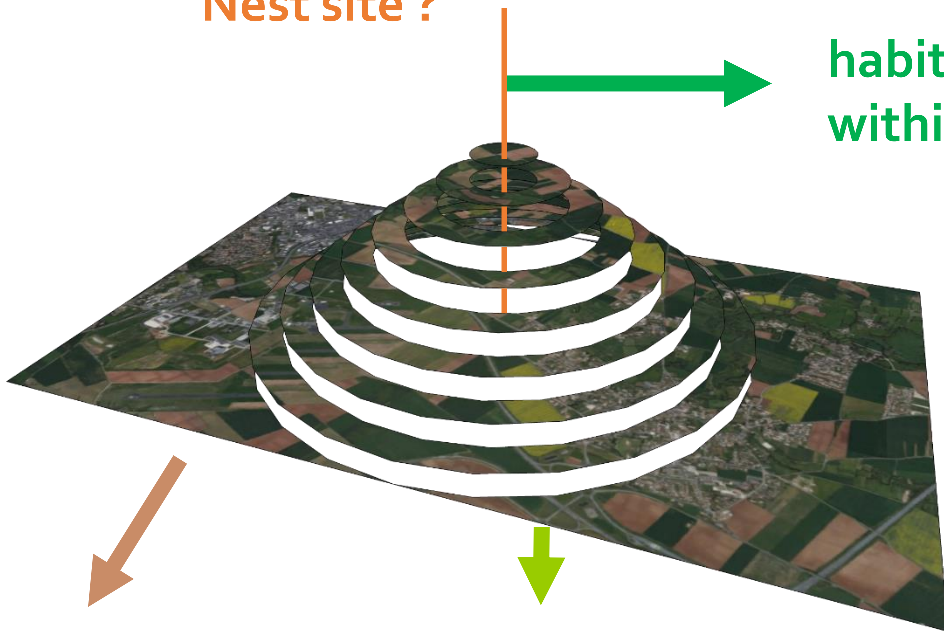
- Selecting optimal shape: step function, gradual decay?





Aue et al. 2012 *MEE*,
Chandler et al. 2016 *Landsc. Ecol.*

Scale-dependence



Habitat A

spatial location

$$\begin{bmatrix} a_{1,1} & \cdots & a_{1,P} \\ \vdots & \ddots & \vdots \\ a_{N,1} & \cdots & a_{N,P} \end{bmatrix}$$

scale

Habitat B

$$\begin{bmatrix} b_{1,1} & \cdots & b_{1,P} \\ \vdots & \ddots & \vdots \\ b_{N,1} & \cdots & b_{N,P} \end{bmatrix}$$

scale

...

Statistical considerations (2)

- Combining effects at multiple scales
  - Multicollinearity
  - Multiple sets of coefficients equally plausible
  - -> ill-defined/unstable/overfitting models

Statistical considerations (2)

- Combining effects at multiple scales
  - Multicollinearity
  - Multiple sets of coefficients equally plausible
  - -> ill-defined/unstable/overfitting models

- Regularization:
  - Penalizing excessive variance in the coefficients by imposing a smoothness constraint

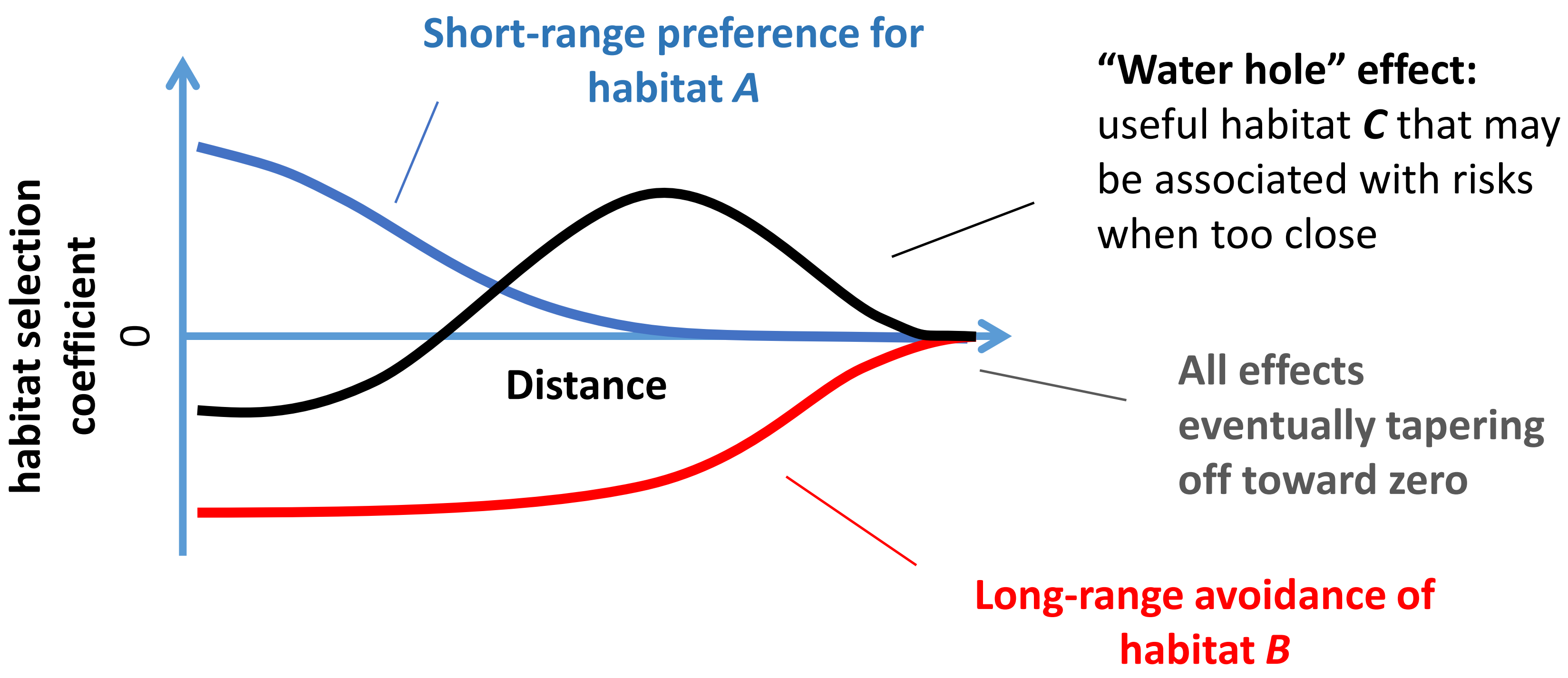
Insights from ecology

Scale of habitat effect
should depend on function

- Habitat for breeding
- Habitat for hiding
- Habitat for feeding
- …

Insights from ecology

- Habitat selection strength varies with spatial scale.
- Shape and scale of selection habitat-specific.



We need

a practical approach to gain ecological insight about the strength, shape and scale of habitat selection, and able to deal with multicollinearity

Varying coefficient regression

See e.g. Hastie and Tibshirani (1993)

- Coefficients α_p for habitat A_p varying smoothly with distance p.

$$Y_i \sim Poisson(\lambda_i)$$

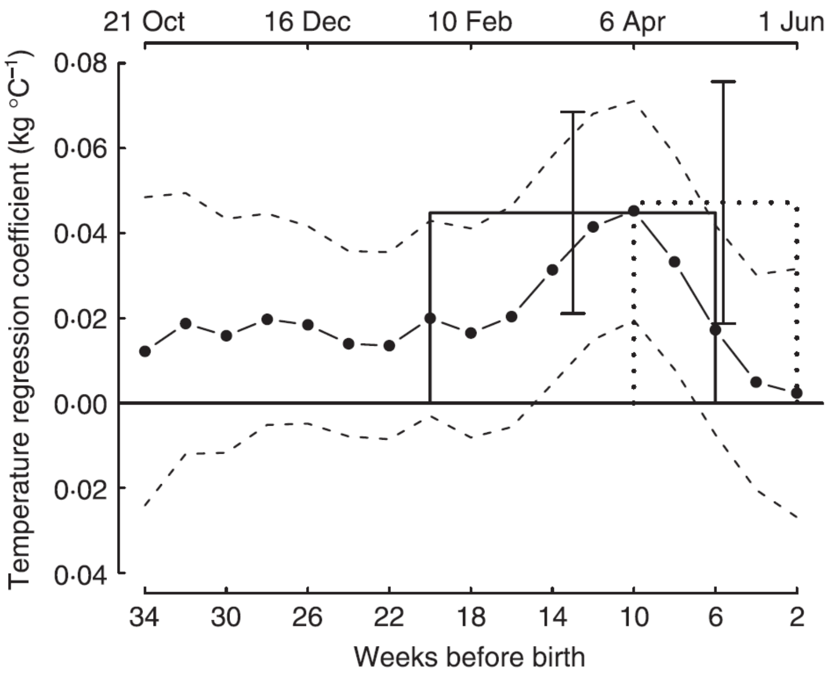
$$log\begin{pmatrix}\lambda_1\\ \vdots \\ \lambda_N\end{pmatrix} = \begin{bmatrix}1\\ \vdots \\ 1\end{bmatrix}\gamma + \begin{bmatrix}a_{1,1} & \cdots & a_{1,P}\\ \vdots & \ddots & \vdots \\ a_{N,1} & \cdots & a_{N,P}\end{bmatrix}\begin{bmatrix}\alpha_1\\ \vdots \\ \alpha_P\end{bmatrix}$$

$$\begin{bmatrix}\alpha_1\\ \vdots \\ \alpha_P\end{bmatrix} = f\left(\begin{bmatrix}1\\ \vdots \\ P\end{bmatrix}\right)$$

- f is some smooth function

Sims, Elston et al. (2007) Journal of Animal Ecology

- Red deer weight at birth ~ temperature prior to birth
- Coefficients for temperature varying smoothly with time lag.



- First-order random walk

"Difference Penalty Regression"

- Coded in R (about 60 lines)

Varying coefficient regression

- Coefficients α_p for habitat A_p varying smoothly with distance p.

$Y_i \sim Poisson(\lambda_i)$

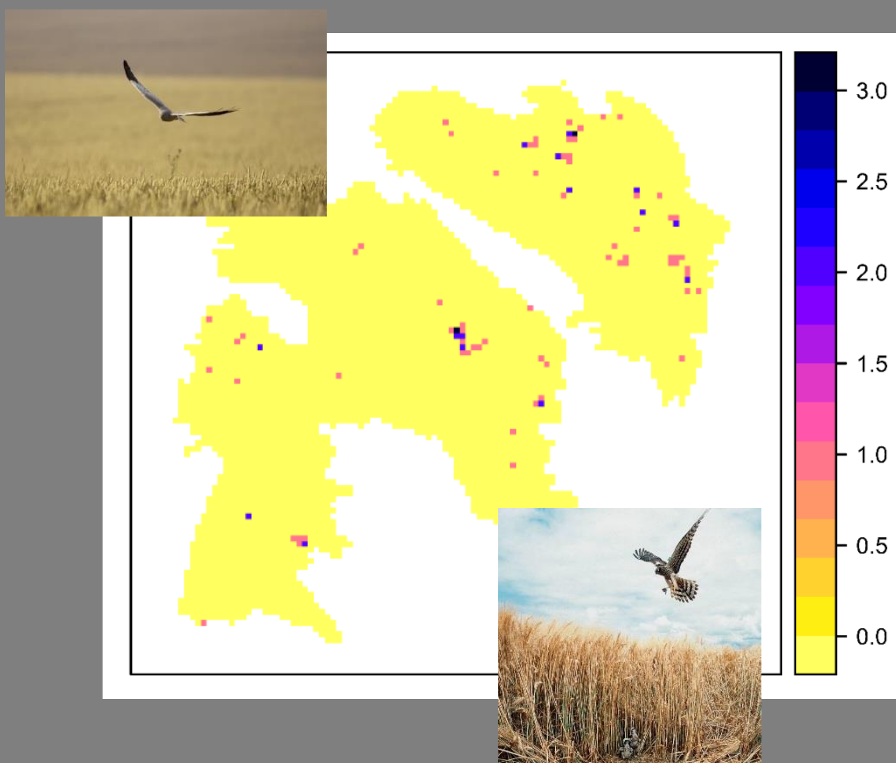
$$log\begin{pmatrix}\lambda_1\\ \vdots \\ \lambda_N\end{pmatrix} = \begin{bmatrix}1\\ \vdots\\ 1\end{bmatrix}\gamma + \begin{bmatrix}a_{1,1} & \cdots & a_{1,P}\\ \vdots & \ddots & \vdots\\ a_{N,1} & \cdots & a_{N,P}\end{bmatrix}\begin{bmatrix}\alpha_1\\ \vdots\\ \alpha_P\end{bmatrix}$$

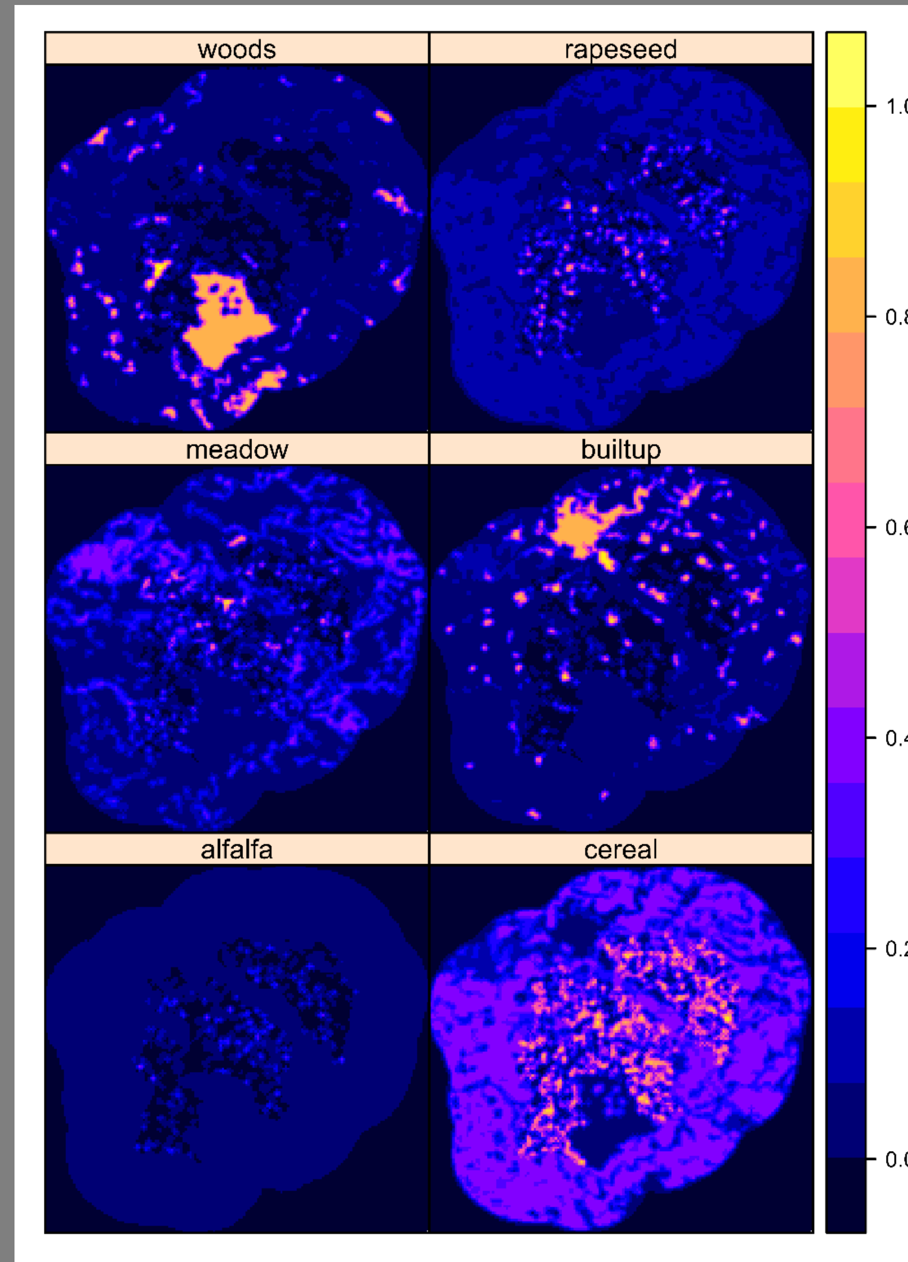
$$\begin{bmatrix}\alpha_1\\ \vdots\\ \alpha_P\end{bmatrix} = f\left(\begin{bmatrix}1\\ \vdots\\ P\end{bmatrix}\right)$$

- f is some smooth function (e.g. thin-plate spline).
- Estimated with mgcv::gam (Wood 2006) in R using GCV (1 line of R code)

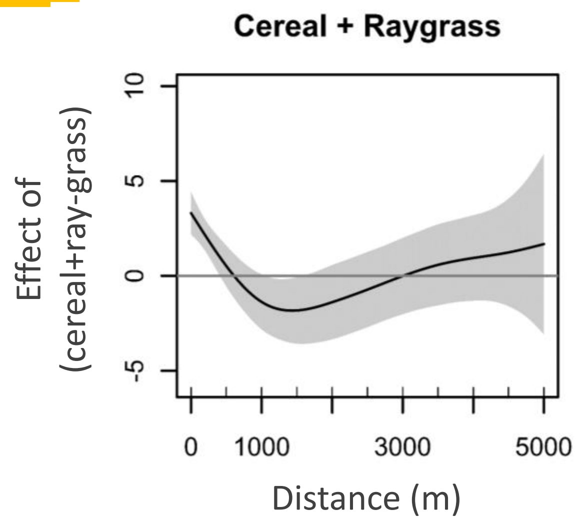
Illustration

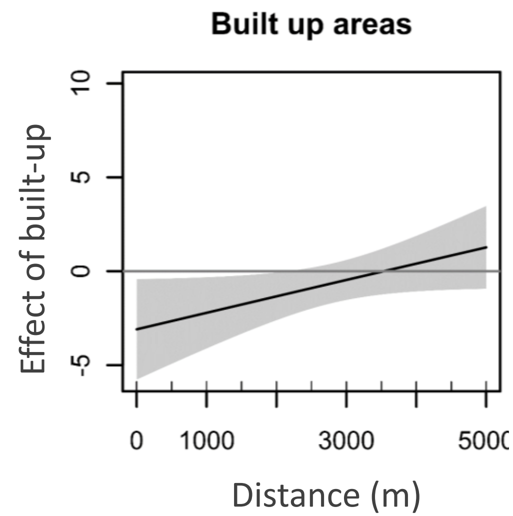
Nest site selection in Montagu's harrier (Semi-colonial raptor)

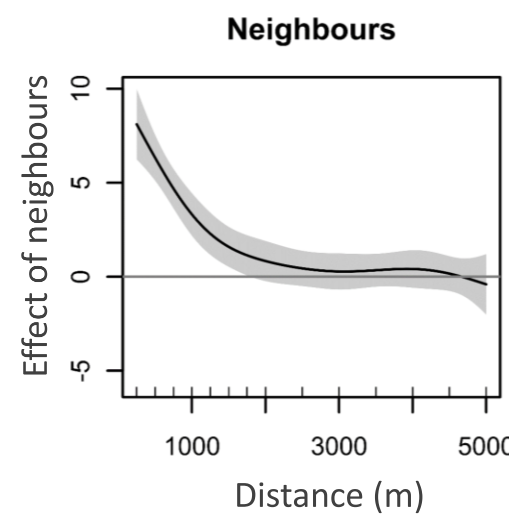




Illustration



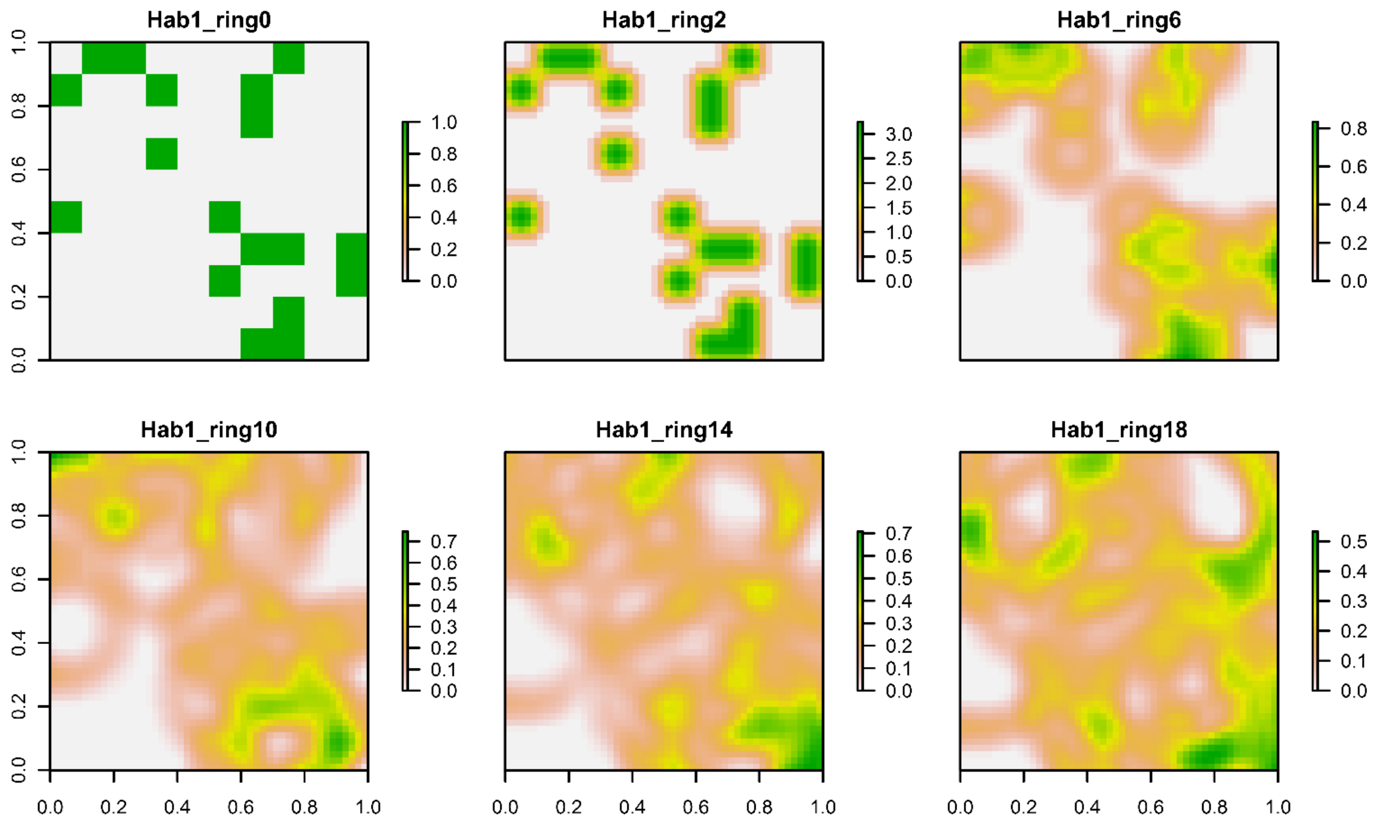




- Nesting and foraging habitats selected at short range (<1 km) with respect to foraging range (~5 Km).
- Built up areas avoided up to 2-3 Km.
- Seeks proximity to conspecifics within 1-1.5 Km.

Simulations

(1) Random landscapes: compute habitat availability within concentric rings (radius $x+\delta x$)



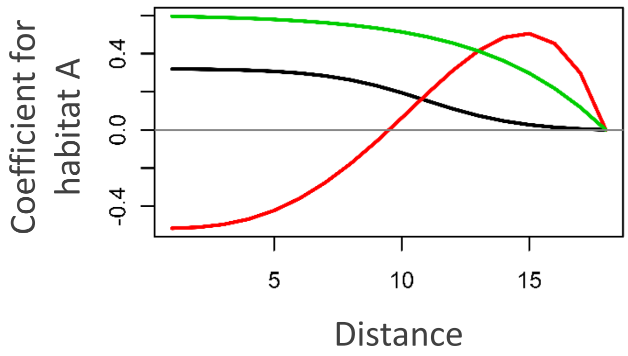
Simulations

(2) Random habitat selection functions $f(p)$

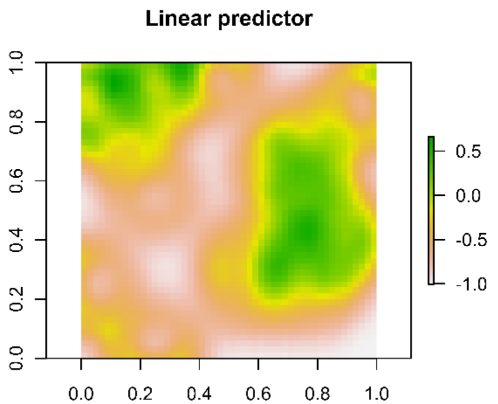
Either **Logistic** or **Beta**(U(0.5,10), U(0.5,10))

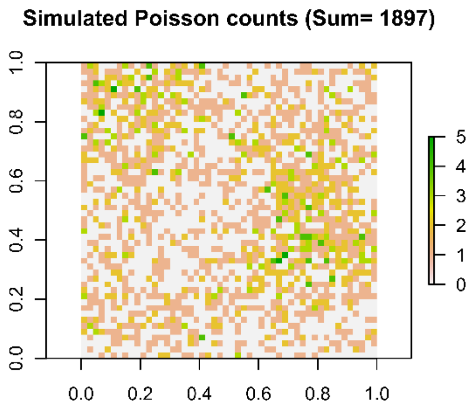
Last coefficient always zero.

Direction ~ binomial(0.5)



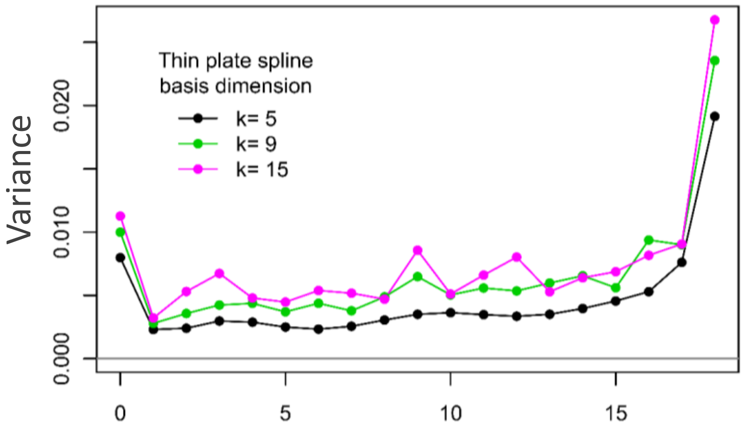
(3) Simulated species distributions

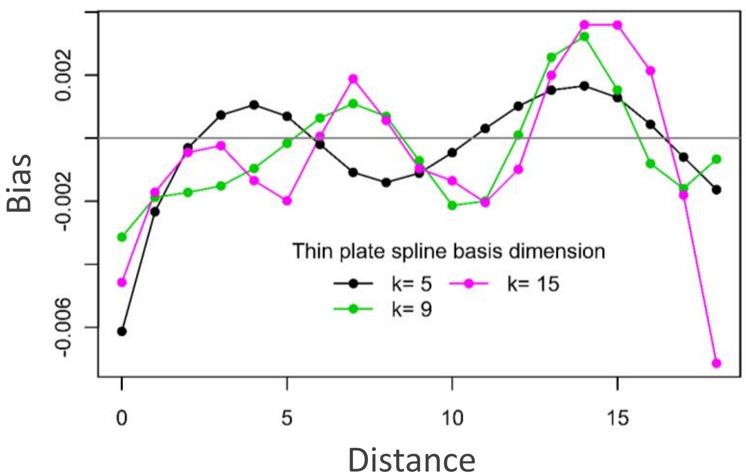




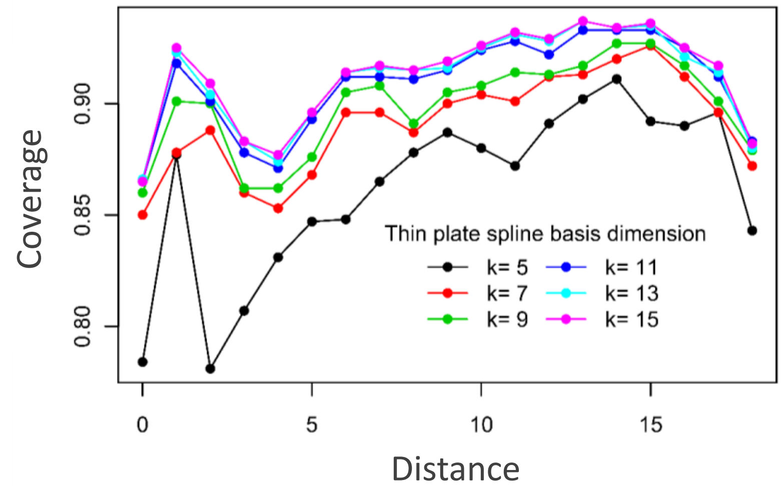
Model performance

Point-wise variance and bias of coefficients α

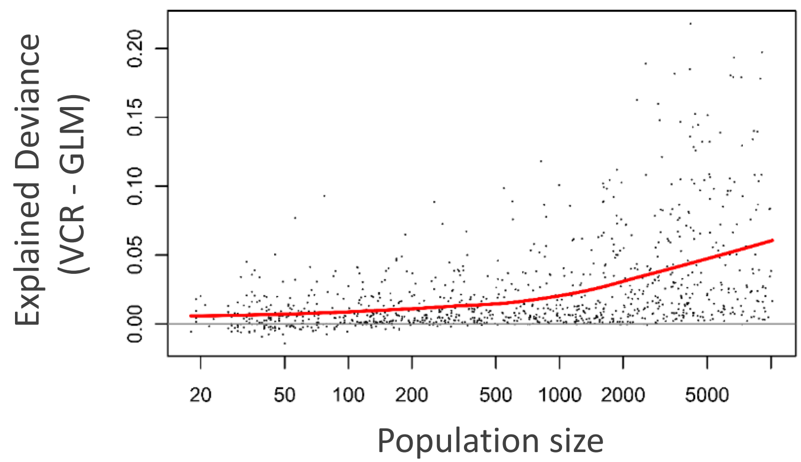




Point-wise coverage for α



Explained Deviance



Conclusions

- **Multicollinearity** across spatial scales **taken care of** by smoothing function *f*.
- Great **simplification of model selection:** 1-stop estimation of **strength, shape and scale** of habitat selection.
- **Gains in stability and explanatory power** thanks to borrowing information across spatial scales.
- **Simple implementation** with standard software.
- **Ecologically meaningful**.

Future directions

- Investigate performance of different smoothing functions, e.g. difference penalty (Sims *et al*.2007) or parametric smoothers?
- Control for scale-dependent collinearity between different resources in a multi-habitat regression setting.
- Impose explicit zero-effect constraint at large distances?