Step-Selection Functions

John Fieberg, Associate Professor

Department of Fisheries, Wildlife and Conservation Biology

Outline

Step-selection functions

- What are they and why were they developed?
- Data development and model fitting
- Parameter interpretation

Telemetry Data and Independence

Historically, biologists would:

- Sample less frequently or subsample data until location data are "independent"
- Justify treating data as independent if an individual could have moved anywhere in its home range between sampling times ("biological independence").

In the later case, it would make sense to use something like the outer 95% contour of the estimated home range to determine availability.

RSF to SSF

Use-availability likelihood:

$$f^u(x) = \frac{\exp(x\beta)f^a(x)}{\int_{s\in A}\exp(x\beta)f^a(x)}$$

How should we model availability with GPS data collected on a fine temporal scale?

Is it OK to assume locations are independent?

Step-Selection Functions

Determine availability using:

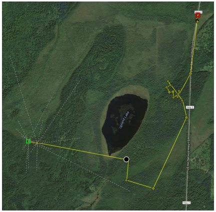
- Location at previous time point = $u(x', t)$
- $\phi(x, x')$ = resource-independent movement kernel (describes how the animal would move in homogeneous habitat)

$$u(x, t+\tau) = \frac{w(x\beta)\phi(x,x')u(x',t)}{\int_{x \in A} w(x\beta)\phi(x,x')u(x',t)dx'}$$

- $w(x\beta) = \exp(x\beta)$ spatial preference function

Available Points

Determine available points by simulating movements from previous locations



Results in time-dependent availability distributions

Conditional Logistic Regression

If we knew $\phi(x, x')$, we could take a random sample of n_a points from this distribution and evaluate:

$$\prod_{i=1}^{n} \prod_{t=1}^{T} \frac{exp(x_{it}^{u}\beta)}{exp(x_{it}^{u}\beta)+\sum_{j=1}^{n_a} exp(x_{jt}^{a}\beta)}$$

- x_{it}^{u} = the t^{th} used point taken on the i^{th} animal.
- x_{jt}^{a} is the j^{th} available point associated with the t^{th} used point taken on the i^{th} animal.

This is the same as the likelihood of a conditional logistic regression model

We can fit using `clogit` function in R `survival` library:

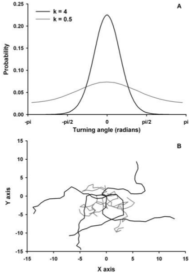
```
clogit(y~ x+ strata(strataID), data=)
```

Problem: We don't know $\phi(x, x')$!

Options:

- Resample observed step lengths and turn angles
- Use step-lengths and turn angles to parameterize statistical distributions
  - Step lengths: gamma distribution
  - Turn angles: von Mises distribution (or uniform on $-\pi$ to π)

von Mises distribution



Another Potential Issue

$$u(x, t+\tau) = \frac{w(x\beta)\phi(x,x')u(x',t)}{\int_{x \in A} w(x\beta)\phi(x,x')u(x',t)dx'}$$

Problem:

- $\phi(x, x')$ is meant to reflect movement in the absence of resource selection.
- We observe movements that reflect the combination of $\phi(x, x')$ AND $w(x, \beta)$

Can give biased estimates of movement parameters

Forester, J.D., Im, H.K. & Rathouz, P.J. (2009). Accounting for animal movement in estimation of resource selection functions: Sampling and data analysis. Ecology, 90, 3554–3565.

Estimating movement parameters

Solution:

- Use parametric distributions for step length and turn angles
- Include step length, ln(step length), and cos(turn angles) to modify (re-estimate) movement parameters

Opportunity:

- Can include interactions with predictors to model how movement is influenced by habitat
- Equivalent to fitting a biased correlated random walk model

Avgar, T., Potts, J.R., Lewis, M.A. & Boyce, M.S. (2016). Integrated step selection analysis: Bridging the gap between resource selection and animal movement. Methods Ecol. Evol., 7, 619–630.

Duchesne, T., Fortin, D. and Rivest, L.P. (2015) Equivalence between step selection functions and biased correlated random walks for statistical inference on animal movement. PloS one, 10, e0122947.

amt

The `random_points()` function in the `amt` package

- Uses maximum likelihood to fit gamma and von Mises distributions to step length and turn angles
- Generates random steps, turn angles using the fitted distributions
- Uses these to form random steps, and thus choose new available locations

Can use clogit() function to fit the model (or fit_issf() in amt package)

See Signer, J., Fieberg, J. and Avgar, T., 2019. Animal movement tools (amt): R package for managing tracking data and conducting habitat selection analyses. Ecology and Evolution, 9(2), pp.880-890.

Interpretation of Parameters

Slope parameters characterize *relative risk* of using a point as a function of its spatial characteristics (*given equal availability and holding everything else constant!*).

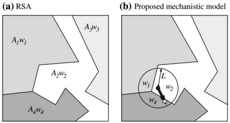
	coef	exp(coef)	se(coef)	z	Pr(>|z|)
Elevation	0.000	1.000	0.000	-0.450	0.652
PopDens	0.000	1.000	0.000	-0.323	0.747
forest	0.073	1.075	0.029	2.516	0.012
sl_	0.000	1.000	0.000	-0.288	0.774
log(sl_)	0.003	1.003	0.007	0.461	0.645

Consider forest: holding everything else constant, an animal would more likely to be in forest than not...

Interpretation

Although coefficients have similar interpretation:

- they reflect decisions on a more local (time and space) scale
- require heterogeneity in spatial predictors at this scale



Barnett, A. H., and P. R. Moorcroft. 2008. Analytic steady-state space-use patterns and rapid computations in mechanistic home-range analysis. Journal of Mathematical Biology 57:139–159.

Coefficients and Scaling

Estimates of habitat selection will depend on the modeled time scale!

- $\hat{\beta}_{RSF} \neq \hat{\beta}_{SSF}$
- $\beta_{RSF} > \beta_{SSF}$
- $\hat{\beta}_{SSF}$ should increase as we sample less frequently

Barnett, A. H., and P. R. Moorcroft. 2008. Analytic steady-state space-use patterns and rapid computations in mechanistic home-range analysis. Journal of Mathematical Biology 57:139–159.

Signer, J., J. Fieberg, and T. Avgar. 2017. Estimating utilization distributions from fitted step-selection functions. Ecosphere 8(4):e01771.

Estimating Utilization Distributions

We can estimate utilization distributions by simulating our parameterized movement model:

$$u(x, t+\tau) = \frac{w(x\beta)\phi(x,x')u(x',t)}{\int_{x\in A} w(x\beta)\phi(x,x')u(x',t)dx'}$$

Steady-state distribution: $u* = \lim_{t\to\infty} u(\cdot, t)$ (analogous to range distribution)

We can also estimate transient distributions to answer:

- Where might it go next week?
- How might it change its movement if we alter the environment?

`amt` has some functionality for simulating movements from fitted models, but still a work in progress...

Step-Selection Functions

- Availability is determined by considering movement
  - Note: need to have data collected at a regular sampling interval
- Allow one to relax the independence assumption (to independent “steps” rather than independent locations)
- Similar parameter interpretation, but coefficients depend on the modeled scale
- Active area research, so stay tuned for new developments