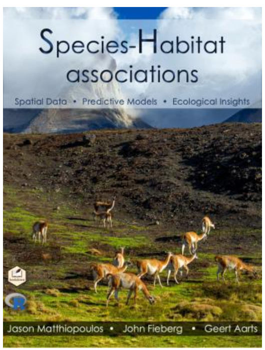
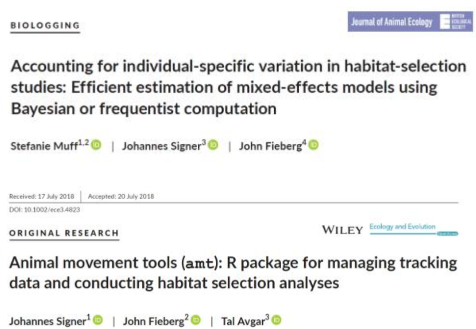
An Introduction to Resource and Step-Selection Functions

John Fieberg, Professor,

Department of Fisheries, Wildlife, and Conservation Biology





BIOLOGGING

Journal of Animal Ecology

Accounting for individual-specific variation in habitat-selection studies: Efficient estimation of mixed-effects models using Bayesian or frequentist computation

Stefanie Muff[1,2] | Johannes Signer[3] | John Fieberg[4]

Received: 17 July 2018 | Accepted: 20 July 2018

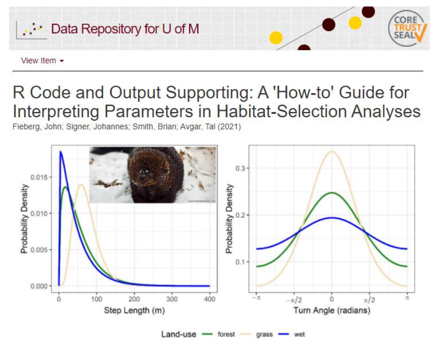
DOI: 10.1002/ece3.4823

ORIGINAL RESEARCH

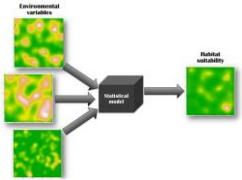
WILEY Ecology and Evolution

Animal movement tools (amt): R package for managing tracking data and conducting habitat selection analyses

Johannes Signer[1] | John Fieberg[2] | Tal Avgar[3]



Outline



Species-Habitat-Association Models

Species Distribution Models

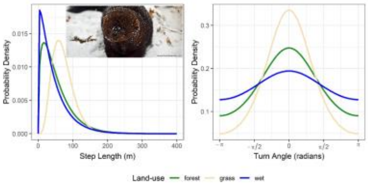
Habitat-Selection Models

$$f_u(x) = \frac{w(x,\beta) f_a(x)}{\int_{z \in E} w(z,\beta) f_a(z) dz}$$

Frameworks for Interpreting Parameters

Inhomogeneous Spatial Point Process

Weighted Distribution Theory



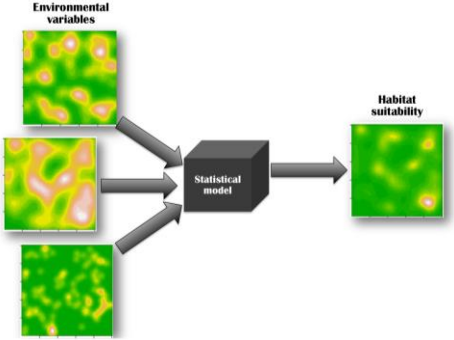
Modeling Animal Movement

Integrated Step-Selection Functions

Species-Habitat-Association Models

Link Species to their Environment

- Where do species occur?
- Why are there and where else they might be?

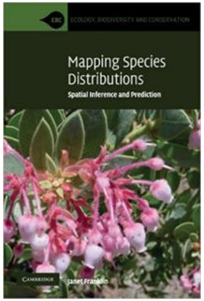


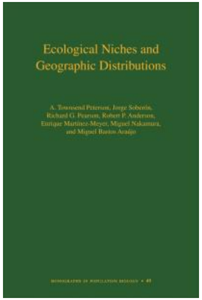
Applications

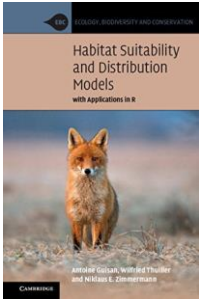
- Designate protected areas
- Estimate anthropogenic impacts
- Assess risks from invasive species or disease transmission from wildlife to humans.

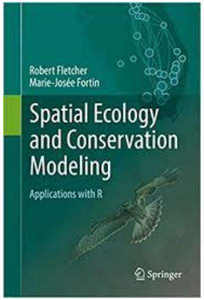
Species Distribution Models

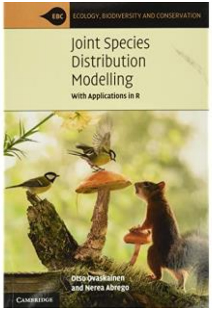
- ‘ISI’s Essential Science Indicators identifies species distribution modeling as the top ranked research front in ecology and the environmental sciences.’ (Renner and Warton 2013)
- Compare environments at observed and background/pseudo-absence locations
- MaxEnt, Generalized linear models (GLMs), Generalized Additive Models (GAMs), Boosted regression trees











Different ways of modeling the distribution of locations as a function of spatial predictors (resources, risks, and conditions)

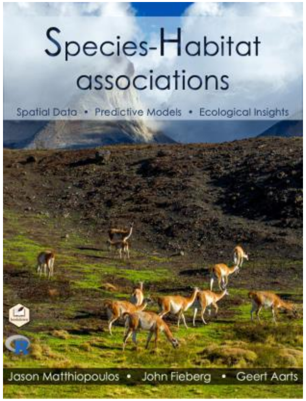
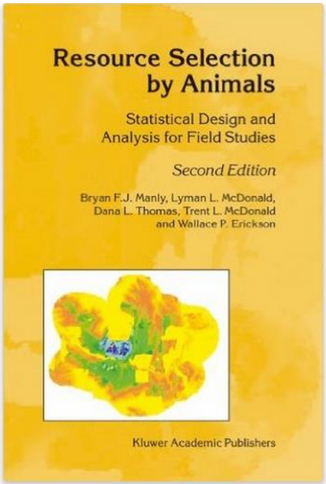


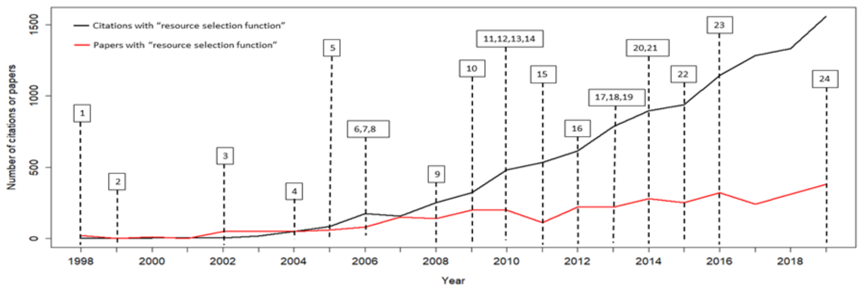
Table 3.1: Different types of habitat-use data used to parameterize species-habitat-association (SHA) models.

Data Collection Method	Sample Unit/Sampling Effort in Space and Time	Typical Sample Selection
Presence-only surveys	Spatial and temporal dimensions are often not well defined	Opportunistic
Presence-absence or count-based surveys	Spatial areas at fixed points in time	Random
Camera trap	Areas within the camera's field of view, followed over time	Systematic, Random
Telemetry	Individuals followed over time	Opportunistic

Habitat-Selection Models

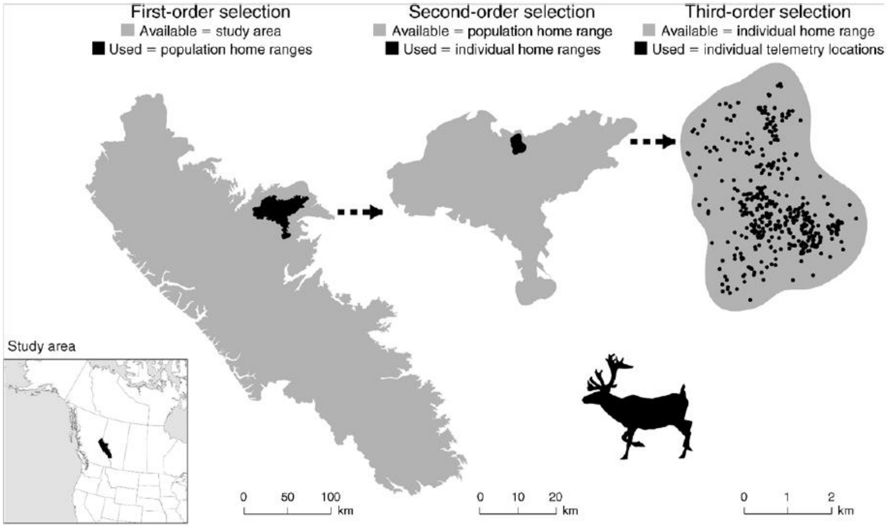
Quantify habitat preference by comparing used and available habitat





Northup, J.M, E. Vander Wal, M. Bonar, J. Fieberg, M.P. Laforge, M. Leclerc, C.M. Prokopenko, and B. D. Gerber. In review. Conceptual and methodological advances in habitat-selection modeling: guidelines for ecology and evolution. Ecological Applications.

Johnson, D. 1980. The comparison of usage and availability measurements for evaluating resource preference. Ecology 61:65-71.



Fourth order: local selection (e.g., within a feeding site)

DeCesare, et al. 2012. Transcending scale dependence in identifying habitat with resource selection functions. Ecological Applications 22(4):1068- 1083.

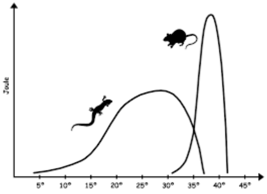
Species Distribution and Habitat-Selection Models

Both compare:

- Locations where individuals are found
- Random/available/background /pseudo-absence locations

Model the distribution of locations as a function of:

- Resources (more is better)
- Risks (less is better)
- Conditions (not too much or too little)



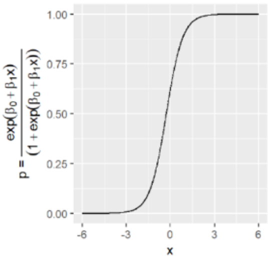
Logistic regression

- Monitor n habitat patches
- y_i = 1 if detected and 0 otherwise, i=1, 2, …, n
- Spatial predictors: $X_{i1}, X_{i2}, \dots, X_{ip}$

$$y_i \sim Bernoulli(p_i)$$

$$\log([p_i/(1-p_i)] = \beta_0 + X_{i1}\beta_1 + \ \dots + X_{ip}\beta_p$$

$$p_i = \frac{\exp\left(\beta_0 + X_{i1}\beta_1 + \ \dots + X_{ip}\beta_p\right)}{1+\exp\left(\beta_0 + X_{i1}\beta_1 + \ \dots + X_{ip}\beta_p\right)}$$



Resource-Selection Functions

- Observed locations: y_i= 1, background/available locations: y_i = 0

$$y_i \sim Bernoulli(p_i)$$
$$\log([p_i/(1-p_i)] = \beta_0 + X_1\beta_1 + \ \ldots + X_p\beta_p$$

- Used and background points may not be mutually exclusive
- Modeling points in continuous space rather than discrete sample units
- p_i depends on the number of background points
- How many random points should we include?
- How to interpret parameters in Resource-Selection Functions (RSF)

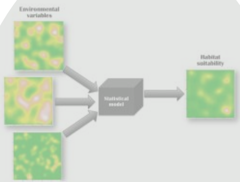
(RSF) = exp($X_1\beta_1 + \ \ldots + X_p\beta_p$)

Traditionally, **resource-selection functions** were described as measuring “relative probabilities of use.”

- Probability of using a site depends on the size of the site and how long individuals are monitored.
- Probability of using a point in space = 0 (to ensure integration over space = 1 for continuous probability distributions).

Better to think of modeling hazards (rates of use), which can be integrated over time or space to estimate utilization distributions.

Outline



Species-Habitat-Association Models

Species Distribution Models

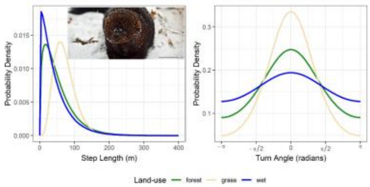
Habitat-Selection Models

$$f_u(x) = \frac{w(x,\beta)f_a(x)}{\int_{z\in E} w(z,\beta)f_a(z)dz}$$

Frameworks for Interpreting Parameters

Inhomogeneous Spatial Point Process

Weighted Distribution Theory



Modeling Animal Movement

Integrated Step-Selection Functions

Frameworks for Interpreting Parameters

HOW TO

A 'How-to' Guide for Interpreting Parameters in Habitat-Selection Analyses

John Fieberg, Johannes Signer, Brian Smith, Tal Avgar

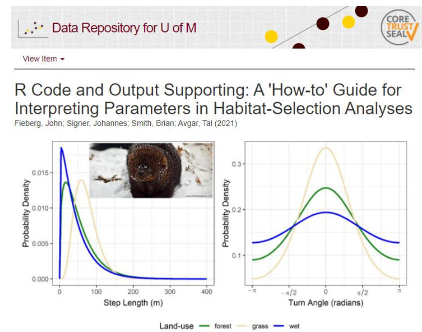
First Published: 14 February 2021

Abstract | PDF

Johannes Signer, University of Goettingen, Germany

Tal Avgar, Utah State University

Brian Smith, Grad Student, Utah State University



The Annals of Applied Statistics
2010, Vol. 4, No. 3, 1383–1402
DOI: 10.1214/10-AOAS331

POISSON POINT PROCESS MODELS SOLVE THE "PSEUDO-ABSENCE PROBLEM" FOR PRESENCE-ONLY DATA IN ECOLOGY[1]

BY DAVID I. WARTON AND LEAH C. SHEPHERD

University of New South Wales

Methods in Ecology and Evolution BRITISH ECOLOGICAL SOCIETY

Free Access

Comparative interpretation of count, presence–absence and point methods for species distribution models

Geert Aarts, John Fieberg, Jason Matthiopoulos

First published: 04 August 2011 | https://doi.org/10.1111/j.2041-210X.2011.00141.x | Citations: 140

The Annals of Applied Statistics
2013, Vol. 7, No. 4, 1917–1939
DOI: 10.1214/13-AOAS667

FINITE-SAMPLE EQUIVALENCE IN STATISTICAL MODELS FOR PRESENCE-ONLY DATA

BY WILLIAM FITHIAN[1] AND TREVOR HASTIE[2]

Stanford University

IBS Biometrics JOURNAL OF THE INTERNATIONAL BIOMETRIC SOCIETY

Original Article

Equivalence of MAXENT and Poisson Point Process Models for Species Distribution Modeling in Ecology

Ian W. Renner, David I. Warton

Inhomogeneous Poisson Point Process (IPP)

- Model density as a function of spatial predictors

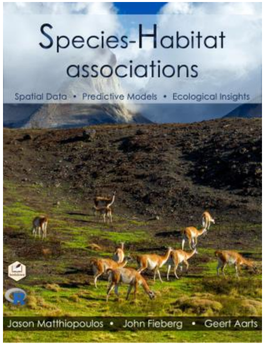
$$\log[\lambda(s)] = \beta_0 + \beta_1 X_1(s) + \cdots \beta_p X_p(s).$$

- $\lambda(s)$ is the intensity function and describes the expected density in a small sphere around s.

Characteristics/Assumptions:

- The number of points in any area *G*, is a Poisson random variable with mean $\int_G \lambda(s) ds$.
- Locations are independent (any clustering can be explained by spatial covariates)

IPP as a Data Integrator



- Counts in grid cells
- Presence-absence data in grid cells

$$y_i \sim Bernoulli(p_i)$$
$$\log(-\log(1-p_i)) = \beta_0 + \beta_1 X_1 + \cdots \beta_p X_p$$

- Presence-only data (as a thinned point process)
- Distance sampling (as a thinned point process)
- Can add spatial random effects

Logistic Regression and IPP model

- Logistic regression provides unbiased estimates of parameters in an Inhomogeneous point-process model:
  - as the number of available points increases to infinity (Warton and Shepherd 2010)
  - if we assign large ("infinite") weights to the available points (Fithian and Hastie 2013)
  - in practice, we can assign large finite weights (e.g., 1000 or 5000 to available points, and a 1 to used points).

Warton, D.I. and Shepherd, L.C., 2010. Poisson point process models solve the "pseudo-absence problem" for presence-only data in ecology. The Annals of Applied Statistics, 4(3), pp.1383-1402.

Fithian, W. and T. Hastie (2013). Finite-sample equivalence in statistical models for presence-only data. Annals of Applied Statistics 7, 1917-1939.

Implementation

1. Observed locations ($Y_i = 1$)

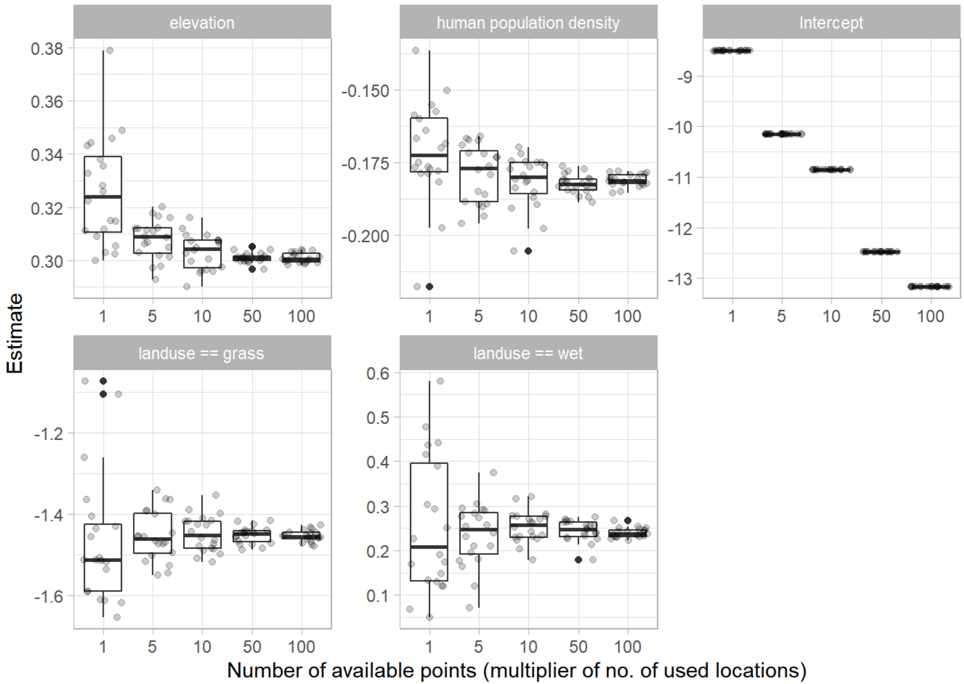
2. Sample available locations (randomly, systematically) from an area A (e.g., the animal's home range; set $Y_i = 0$).

3. Assign weights (1 to used points, and a large number – say 1000 - to available points)

4. Fit logistic regression model using the weights (throw away the intercept).

5. Increase the number of available points until slope parameters stabilize.

$$L(\beta|s_1, s_2, \dots, s_n) = \prod_{i=1}^{n} \frac{\exp(\beta_0 + \beta_1 X_1(s_i) + \cdots \beta_p X_p(s_i))}{\int_{s \in G} \exp(\beta_0 + \beta_1 X_1(s) + \cdots \beta_p X_p(s)) ds}$$



Does a Point Process Framework Make Sense for Telemetry Data?

Inhomogeneous Point Process Model

- Useful for interpreting parameters:
  - Species Distribution Models
  - Resource-Selection Functions

Methods in Ecology and Evolution

Methods in Ecology and Evolution 2015, 6, 366–379 doi: 10.1111/2041-210X.12352

SPECIAL FEATURE – REVIEW
NEW OPPORTUNITIES AT THE INTERFACE BETWEEN ECOLOGY AND STATISTICS

Point process models for presence-only analysis

Ian W. Renner[1]*, Jane Elith[2], Adrian Baddeley[3], William Fithian[4], Trevor Hastie[4], Steven J. Phillips[5], Gordana Popovic[6] and David I. Warton[6]

Telemetry Data

- Historical emphasis on behavior (selection of habitat)
- Repeated observations on the same individuals
- Temporal data dimension
- Autocorrelation

Weighted Distribution Theory

Weighted Distribution Theory

$$u(X) = \frac{w(X;\beta)a(X)}{\int_z w(Z;\beta)a(Z)dZ}$$

Environmental Space:

- $u(X)$ = the frequency distribution of habitat covariates, X, at locations used by our study animals.
- $a(X)$ = the frequency distribution of habitat covariates, X, at locations assumed to be available to our study animals.

Lele, S. R., & Keim, J. L. (2006). Weighted distributions and estimation of resource selection probability functions. *Ecology*, *87*(12), 3021–3028.

Weighted Distribution Theory

$$u(s) = \frac{w(X(s);\beta)a(s)}{\int_{g \in G} w(Z(g);\beta)a(g)dg}$$

Geographic Space:

- $u(s)$ = utilization distribution in geographic space (lat, long)
- $a(s)$ = availability in geographic space (usually assumed to be constant)

Equivalent to IPP if:

- $w(X(s);\beta) = \exp(X_1\beta_1 + \ ... + X_p\beta_p)$
- $a(s)$ = constant

Methods in Ecology and Evolution BRITISH ECOLOGICAL SOCIETY

Free Access

Comparative interpretation of count, presence–absence and point methods for species distribution models

Geert Aarts, John Fieberg, Jason Matthiopoulos

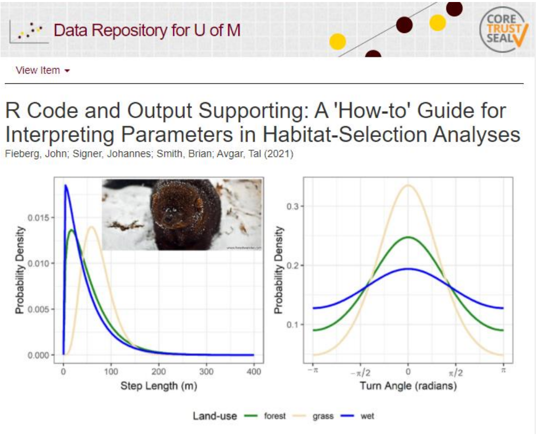
First published: 04 August 2011 | https://doi.org/10.1111/j.2041-210X.2011.00141.x |

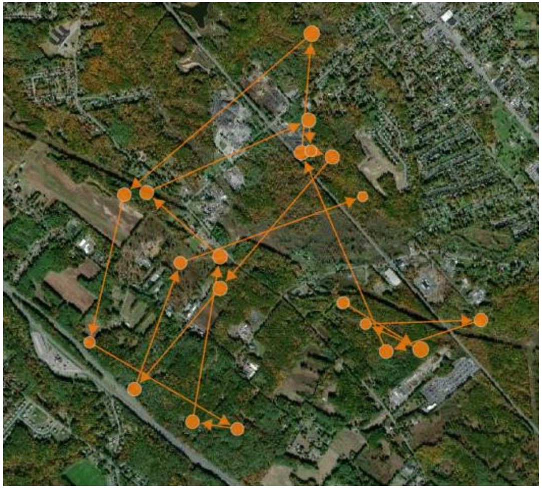
Exemplification

Logistic regression model with predictors:

- Landcover class (forest, grass, wet)
- Human Population Density
- Elevation

Lupe, a fisher from Upstate NY (Lapoint et al. 2013)





AppA_HSF_examples.Rmd (html)

```
rsf_dat <- dat %>%
  random_points() %>%
  extract_covariates(landuse) %>%
  extract_covariates(elevation) %>%
  extract_covariates(popden) %>%
  mutate(elevation = scale(elevation)[, 1],
         popden = scale(popden)[, 1],
         landuseC = reclass_landuse(landuse),
         forest = landuseC == "forest",
         weight = ifelse(case_, 1, 1e3)
  )
```

To find out more!

?random_points

```
Lupe.dat$w <- ifelse(Lupe.dat$case_, 1, 5000)
HSF.Lupe1 <- glm(case_ ~ elevation + popden + landuseC,
                 data = Lupe.dat, weight = w,
                 family = binomial(link = "logit"))
```

```r
summary(HSF.Lupe1)
```

```
##
## Call:
## glm(formula = case_ ~ elevation + popden + landuseC, family = binomial(link = "logit"),
##     data = Lupe.dat, weights = w)
##
## Deviance Residuals:
##    Min      1Q  Median      3Q     Max
## -0.277  -0.152  -0.136  -0.123   5.469
##
## Coefficients:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -13.1678     0.0195 -676.15  < 2e-16 ***
## elevation       0.3023     0.0165   18.29  < 2e-16 ***
## popden         -0.1842     0.0211   -8.75  < 2e-16 ***
## landuseCgrass  -1.4632     0.2780   -5.26  1.4e-07 ***
## landuseCwet     0.2397     0.1085    2.21    0.027 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 84847  on 303403  degrees of freedom
## Residual deviance: 84437  on 303399  degrees of freedom
## AIC: 84447
##
## Number of Fisher Scoring iterations: 9
```

Quantitative Predictors

- Location s1: elevation = 3, popden = 1.5, landC = “wet”
- Location s2: elevation = 2, popden = 1.5, landC = “wet”

$$\frac{u(s_1)}{u(s_2)} = \frac{\exp\left(3\beta_{elevation} + 1.5\beta_{popden} + 0\beta_{forest} + 0\beta_{grass}\right) a(s_1)}{\exp\left(2\beta_{elevation} + 1.5\beta_{popden} + 0\beta_{forest} + 0\beta_{grass}\right) a(s_2)}$$

$$\frac{u(s_1)}{u(s_2)} = \exp(\beta_{elevation}) \frac{a(s1)}{a(s_2)}$$

If the two points are equally available, then

$$\frac{u(s_1)}{u(s_2)} = \exp(\beta_{elevation}) = 1.35$$

Lupe would be 1.35 times more likely to use/select the location with the higher elevation.

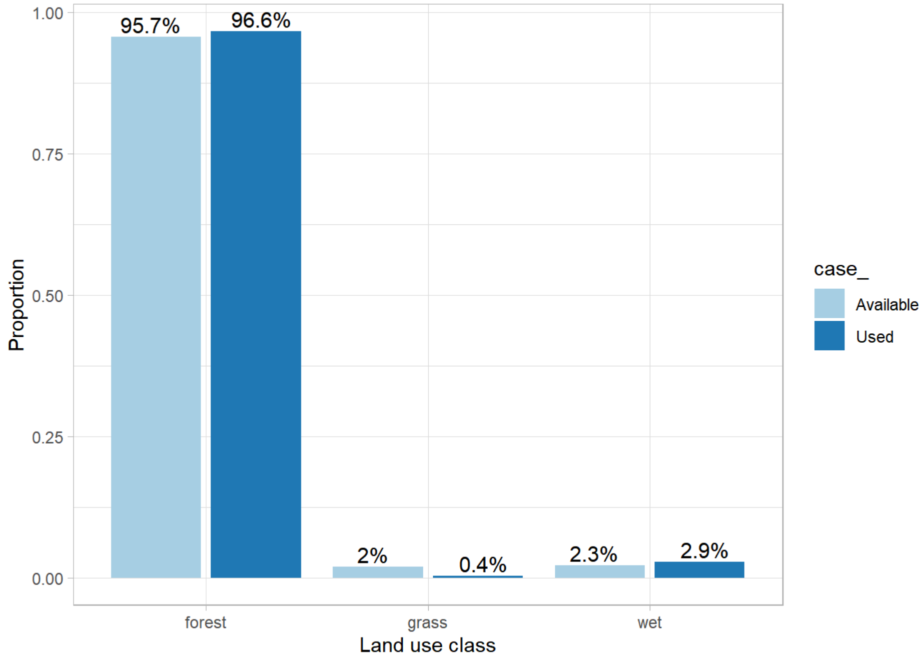
Quantitative Predictors

$\exp(\beta_{elevation})$ describes the relative intensity of use of any 2 locations that differ by 1 unit of elevation but are otherwise equivalent, i.e.:

- They have the same values for all other predictors
- The locations are equally available (or accessible)

Parameters

Predictor	β
(Intercept)	-12.918
elevation	0.303
popden	-0.183
landuseC1forest	-0.250
landuseC1grass	-1.727



Categorical Predictors

- Location s1: elevation = 2, popden = 1.5, landuseC = “wet”
- Location s2: elevation = 2, popden = 1.5, landuseC = “forest”

$$\frac{u(s_1)}{u(s_2)} = \frac{\exp\left(2\beta_{elevation} + 1.5\beta_{popden} + 0\beta_{forest} + 0\beta_{grass}\right) a(s_1)}{\exp\left(2\beta_{elevation} + 1.5\beta_{popden} + 1\beta_{forest} + 0\beta_{grass}\right) a(s_2)}$$

$$\frac{u(s_1)}{u(s_2)} = \exp(-\beta_{forest}) \frac{a(s1)}{a(s_2)}$$

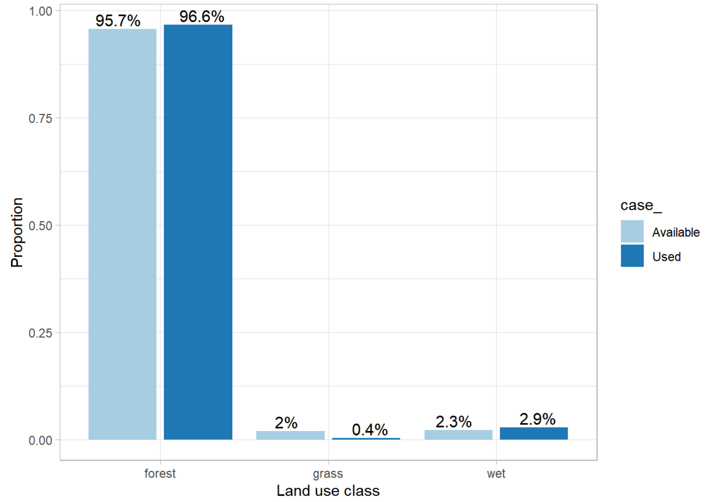
$$\frac{u(s_1)}{u(s_2)} = \exp\left(-\beta_{forest}\right) = 1.28$$

Given two equally available locations that differ only in landcover class, Lupe would be 1.28 times more likely to select the one in wet (versus forest).

Categorical Predictors

$$\frac{u(s_1)}{u(s_2)} = \exp(-\beta_{forest})\frac{a(s1)}{a(s_2)}$$

$$\frac{u(s_1)}{u(s_2)} = 1.28\,\frac{0.023}{0.957} = \frac{1}{33}$$



- Suggests Lupe is 33 times more likely to be found in forest than wet
- Assumes, naively, that forest and wet do not differ in terms of human population density and elevation

Integrated Spatial Intensities

$$\frac{u(s, s \in wet)}{u(s, s \in forest)} = \frac{\int_G u(s) I(s \in wet) ds}{\int_G u(s) I(s \in forest) ds}$$

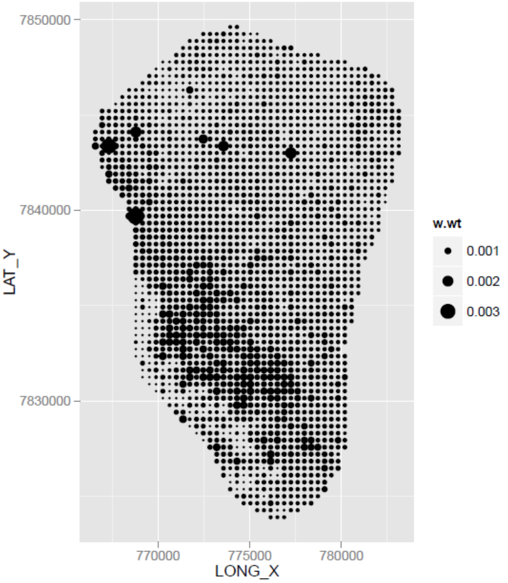
$$\frac{\hat{u}(s, s \in wet)}{\hat{u}(s, s \in forest)} = \frac{\sum_{i=1}^{n_a} \widehat{w}(X(s_i); \beta) I(s_i \in wet)}{\sum_{i=1}^{n_a} \widehat{w}(X(s_i); \beta) I(s_i \in forest)} = \frac{1}{33}$$

- Lupe is 33 times more likely to be found in forest versus wet areas.

How to create a pretty map?

Estimate $u(s) = \frac{exp(X(s);\beta)}{\int_{g \in G} exp(Z(g);\beta)dg}$ with

$$\frac{\exp(x_i\beta)}{\sum_{i=1}^{n_a} \exp(x_i\beta)}$$



HOW TO

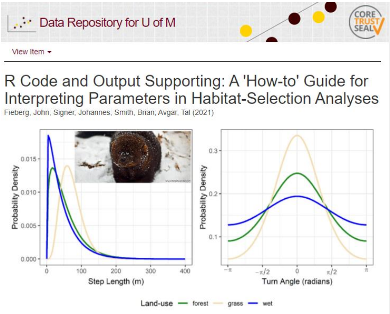
A 'How-to' Guide for Interpreting Parameters in Habitat-Selection Analyses

John Fieberg, Johannes Signer, Brian Smith, Tal Avgar

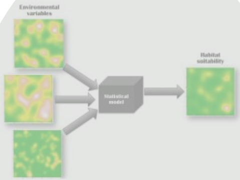
First Published: 14 February 2021

Abstract | **PDF**

- Interpret parameters when allowing for non-linear relationships
- Parameters in models with interactions



Outline



Species-Habitat-Association Models

Species Distribution Models

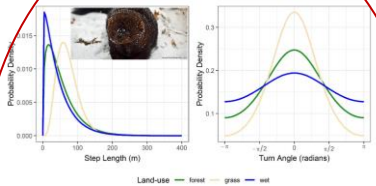
Habitat-Selection Models

$$f_u(x) = \frac{w(x,\beta)f_a(x)}{\int_{z \in E} w(z,\beta)f_a(z)dz}$$

Frameworks for Interpreting Parameters

Inhomogeneous Spatial Point Process

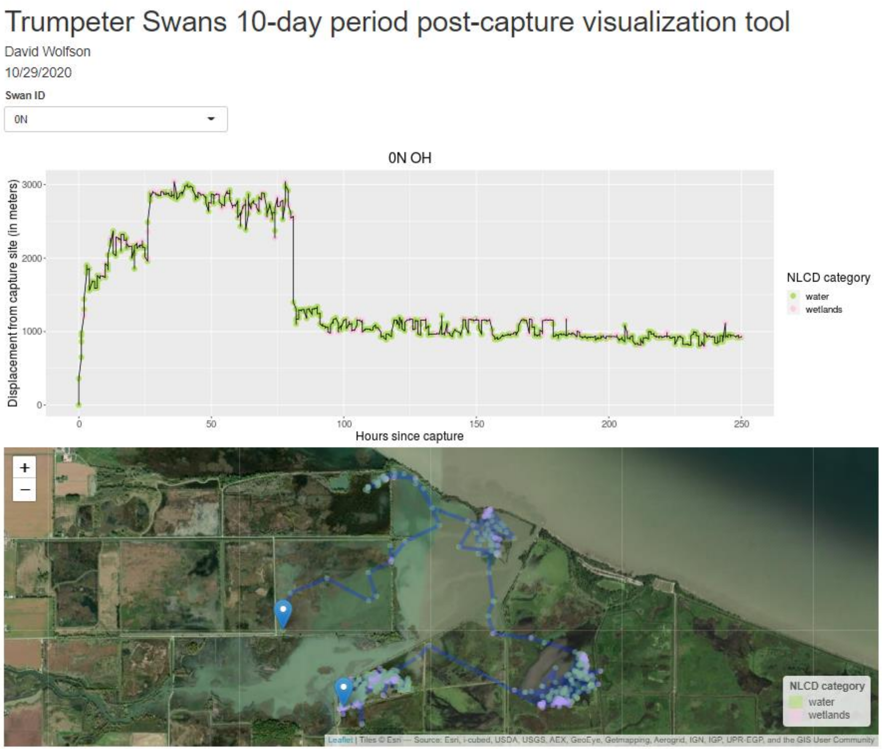
Weighted Distribution Theory



Modeling Animal Movement

Integrated Step-Selection Functions

Habitat Availability and Movement Constraints



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.

Mitigating pseudoreplication and bias in resource selection functions with autocorrelation-informed weighting

Jesse M. Alston[1,2*†], Christen H. Fleming[3,4†], Roland Kays[5,6], Jarryd P. Streicher[7], Colleen T. Downs[7], Tharmalingam Ramesh[7,8], and Justin M. Calabrese[1,2,9]

[1]Center for Advanced Systems Understanding, Görlitz, Germany

[2]Helmholtz-Zentrum Dresden Rossendorf (HZDR), Dresden, Germany

[3]Smithsonian Conservation Biology Institute, National Zoological Park, Front Royal, VA USA

[4]Department of Biology, University of Maryland, College Park, MD USA

[5]Department of Forestry and Environmental Resources, North Carolina State University, Raleigh, NC USA

[6]North Carolina Museum of Natural Sciences, Raleigh, NC USA

[7]Centre for Functional Biodiversity, School of Life Sciences, University of KwaZulu-Natal, Pietermaritzburg, South Africa

[8]Sálim Ali Centre for Ornithology and Natural History (SACON), Coimbatore, Tamil Nadu, India

[9]Department of Ecological Modelling, Helmholtz Centre for Environmental Research (UFZ), Leipzig, Germany

[*]Corresponding author: j.alston@hzdr.de

[†]Equal contributions

https://www.biorxiv.org/content/10.1101/2022.04.21.489059v1.full.pdf

Step-Selection Analyses

$$u(s) = \frac{w(X(s);\beta)a(s)}{\int_{g \in \mathrm{G}} w(Z(g);\beta)a(g)dg}$$

$$u(s, t + \Delta t)|u(s', t) = \frac{w(X(s);\beta)\phi(s, s';\theta)}{\int_z w(X(s);\beta)\phi(s, s';\theta)}$$

- $u(s, t + \Delta t)|u(s', t)$ = conditional probability of finding the individual at location s at time $t + \Delta t$, given it was at location s' at time t
- $\phi(s, s';\theta)$ is a "selection-free" movement kernel that describes the relative likelihood of moving from s to s'.
- $w(X(s);\beta)$ is referred to as a step-selection function

Models for $\phi(s, s'; \theta)$

- Discrete time movement model
  - Step lengths
  - Turn angles

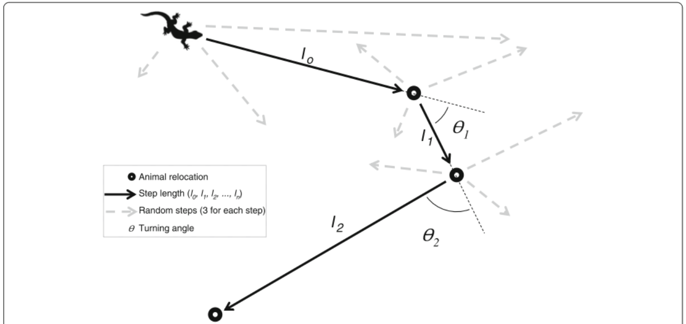


Figure 1 Example of movement pathway in SSFs. Example of how a movement pathway can be simplified into linear step lengths and turning angles occurring between successive locations in any type of animal tracked visually or using VHF or GPS devices. In this example, 3 random steps have been matched with actual steps walked by the lizard.

- Leads to time-dependent availability distributions
  - $a(s, t + \Delta t) | u(s', t) = \phi(s, s'; \theta)$

Thurfjell, H., Ciuti, S. and Boyce, M.S., 2014. Applications of step-selection functions in ecology and conservation. Movement ecology, 2(1), p.4.

- Requires observations to be equally spaced in time.

```
ssf_dat <- dat %>%
  track_resample(rate = minutes(2), tolerance = seconds(20))
```

Conditional Logistic Regression

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

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

- 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 $\phi(s, s'; \theta)$

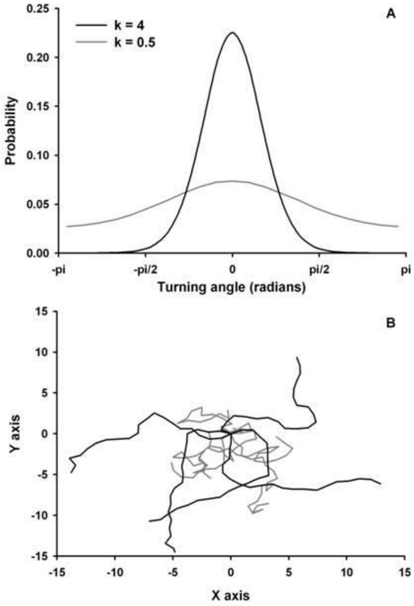
What can we do?

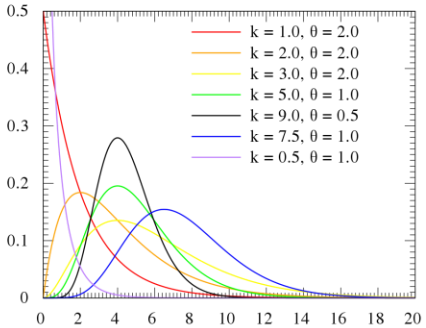
- Resample observed step lengths and turn angles
- Use step-lengths and turn angles to parameterize statistical distributions and sample from them

Distributions

- Step lengths are typically modeled using an exponential or gamma distribution (right skewed, continuous)

- Turn angles are often assumed to be uniform or are modeled using a von Mises distribution





amt Package

The random_steps() 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

```
ssf_dat <- dat %>%
  track_resample(rate = minutes(2), tolerance = seconds(20)) %>%
  steps_by_burst() %>%
  random_steps() %>%
  extract_covariates(landuse, where = "both") %>%
  extract_covariates(elevation, where = "both") %>%
  extract_covariates(popden, where = "both") %>%
  mutate(elevation_start = scale(elevation_start),
         elevation_end = scale(elevation_end),
         popden_start = scale(popden_start),
         popden_end = scale(popden_end),
         landuseC_start = reclass_landuse(landuse_start),
         landuseC_end = reclass_landuse(landuse_end),
         forest_start = landuseC_start == "forest",
         forest_end = landuseC_end == "forest",
         cos_ta_ = cos(ta_),
         log_sl_ = log(sl_)
  ) %>%
  filter(!is.na(ta_))
```

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.

Problem

$$u(s, t + \Delta t)|u(s', t) = \frac{w(X(s); \beta)\phi(s, s'; \theta)}{\int_z w(X(s); \beta)\phi(s, s'; \theta)}$$

- $\phi(s, s'; \theta)$ is a “selection-free” movement kernel that describes the relative likelihood of moving from s to s’.
- $w(X(s); \beta)$ is referred to as a step-selection function

We see: $u(s, t + \Delta t)|u(s', t)$ - i.e., the combined effect of movement and habitat selection, which we use to estimate $\phi(s, s'; \theta)$.

Problem and Potential Solution

- Parameterizing the movement model without accounting for habitat selection can lead to biased estimates of movement parameters
- With certain statistical distributions, we can remove the bias by including movement characteristics (e.g., step length) associated with used and available points.

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.

Opportunity

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

For the mathematical details, see papers below and AppC_iSSA_movement.html from Fieberg et al. 2021.

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.

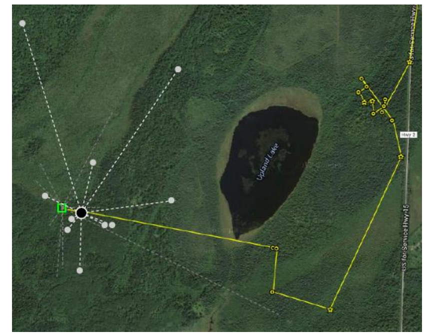
Duchesne, T., Fortin, D. and Rivest, L.P. (2015) Equivalence between step selection functions and biased correlated random walks for statistical inference on animal movement. PloS one, 10, e0122947.

Implementation

Received: 17 July 2018 | Accepted: 20 July 2018
DOI: 10.1002/ece3.4823
ORIGINAL RESEARCH
WILEY Ecology and Evolution

Animal movement tools (amt): R package for managing tracking data and conducting habitat selection analyses

Johannes Signer[1] | John Fieberg[2] | Tal Avgar[3]

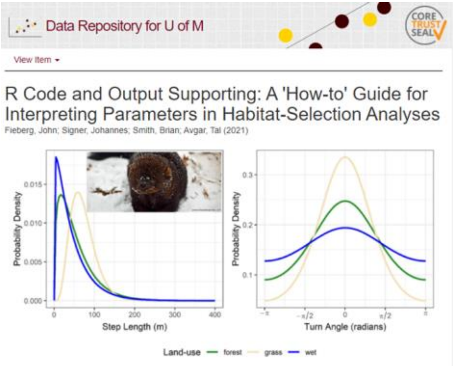


- Estimate a tentative selection-free movement kernel, using observed step-lengths and turn angles, $\hat{\phi}(s, s'; \hat{\theta})$
- Generate time-dependent available locations by simulating potential movements from the previously observed location
$$a(s, t + \Delta t) | u(s', t) = \phi(s, s'; \theta)$$
- Estimate β using conditional logistic regression, with strata formed by combining time-dependent used and available locations.
- Include movement characteristics (e.g., step-length, cos(turn angle) as covariates to correct for bias due to two-step approach (Forester et al. 2009)

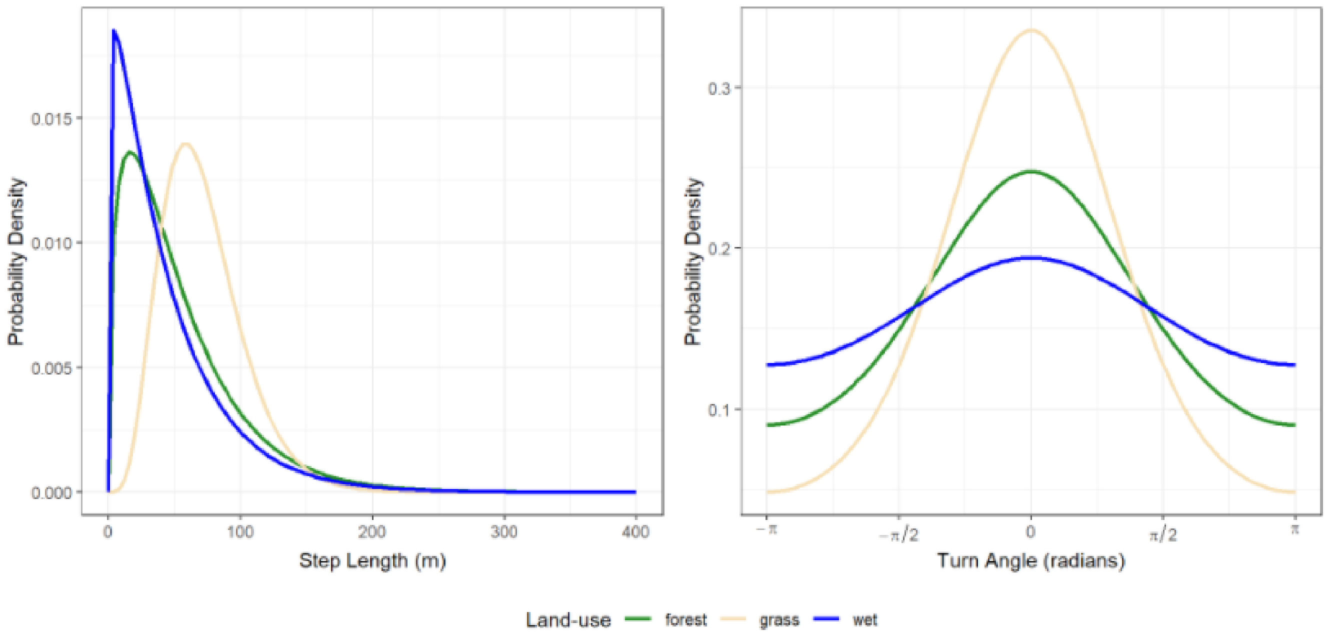
Forester, J. D., Im, H. K., & Rathouz, P. J. (2009). Accounting for animal movement in estimation of resource selection functions: sampling and data analysis. *Ecology*, *90*(12), 3554-3565.

Integrated Step-Selection Analyses

- Re-estimate movement parameters in $\hat{\phi}(s, s'; \hat{\theta})$ using regression coefficients (Duchesne et al 2015, Avgar et al. 2016)
- Include interactions to allow movement parameters to depend on current habitat.




```
m3 <- ssf_dat %>%
  fit_issf(case_ ~ popden_end + elevation_end +
             landuseC_end +
             sl_ + log_sl_ + cos_ta_ +
             landuseC_start:(sl_ + log_sl_ + cos_ta_) +
             strata(step_id_), model = TRUE)
```



We see that Lupe tends to take larger, more directed steps when in `grass` and slower and more tortuous steps in `wet` habitat. One possibility is that Lupe tends to travel through `grass` and forage in `forest` and `wet` habitats.