(integrated) Step-Selection Functions

Johannes Signer

January 2024

Availability matters

- Habitat-selection function assume that each point in the availability domain (e.g., the home range) is available to the animal at any given point in time.

- This assumption might be met with relatively coarse data (large sampling rates), but the assumption becomes unrealistic for data sets with a higher temporal resolution.

Good overviews are

ECOLOGICAL
APPLICATIONS
ECOLOGICAL SOCIETY OF AMERICA

Conceptual and methodological advances in habitat-selection modeling: guidelines for ecology and evolution

Joseph M. Northrup ✉, Eric Vander Wal, Maegwin Bonar, John Fieberg, Michel P. Laforge, Martin Leclerc, Christina M. Prokopenko, Brian D. Gerber

Good overviews are

ECOLOGICAL APPLICATIONS
ECOLOGICAL SOCIETY OF AMERICA

Article | 🔓 Open Access | (cc) (i)

Conceptual and methodological advances in habitat-selection modeling: guidelines for ecology and evolution

Joseph M. Northrup ✉, Eric Vander Wal, Maegwin Bonar, John Fieberg, Michel P. Laforge, Martin Leclerc, Christina M. Prokopenko, Brian D. Gerber

First published: 09 October 2021 | https://doi.org/10.1002/eap.2470 | Citations: 7

Corresponding Editor: Jacob R. Goheen.

Movement Ecology

Home   About   Articles   Submission Guidelines

Review | Open Access | Published: 07 February 2014

Applications of step-selection functions in ecology and conservation

Henrik Thurfjell ✉, Simone Ciuti ✉ & Mark S Boyce ✉

Movement Ecology **2**, Article number: 4 (2014) | Cite this article

26k Accesses | **259** Citations | **2** Altmetric | Metrics

For a more statistical review on Step-Selection Functions (SSF) also see

The basic idea behind an SSF

- Habitat selection is constrained by the movement capacity (i.e., how far can an animal displace within a single time step) and
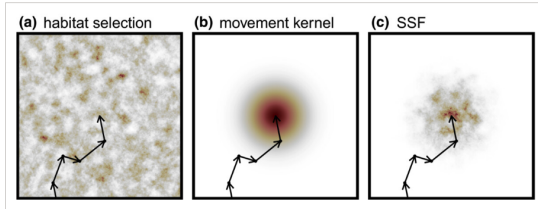
- the current position of the animal.

Figure 1: From Michelot et al. 2024

Step-selection function (SSF)

Fortin et al. (2005) introduced the step-selection function as:

$$u(s, t + \Delta t)|u(s', t) = \frac{w(X(s); \beta(\Delta t))}{\int_{\tilde{s} \in G} w(X(\tilde{s}, s'); \beta(\Delta t))d\tilde{s}}$$

- $u(s, t + \Delta t)|u(s', t)$ is conditional probability of finding the individual at location s at time $t + \Delta t$ given it was at location s' at time t.
- $w(X(s); \beta(\Delta t))$ is the exponential selection function, which again is $w(X(s); \beta(\Delta t)) = \exp(\beta_1 x_1(s) + \cdots + \beta_k x_k(s))$.
- The denominator is the full redistribution kernel that is usually approximated through random steps. But see Michelot et al. 2024 for other approaches.

- We can estimate the β's using a conditional logistic regression with the R function `clogit()` from the `survival` package.
- The `amt` package provides a wrapper to this called `fit_ssf()`.

Creating random steps for SSF

We will cover this in much more detail in the R walkthrough. The workflow to fit an SSF in most cases follows the same steps:

1. Make sure the data are in bursts of regular sampling rates (Note, the Δt in the equation).

Creating random steps for SSF

We will cover this in much more detail in the R walkthrough. The workflow to fit an SSF in most cases follows the same steps:

1. Make sure the data are in bursts of regular sampling rates (Note, the Δt in the equation).

2. Create steps (that is converting two consecutive points to a step with a start and end point).
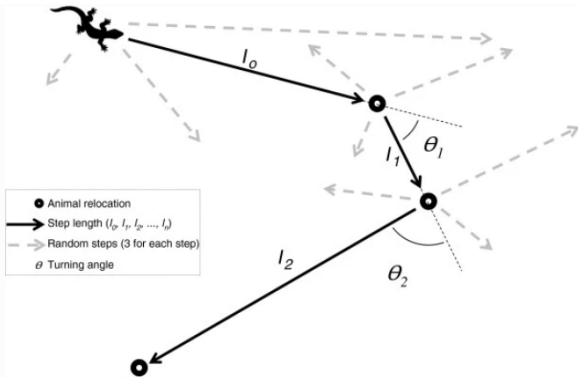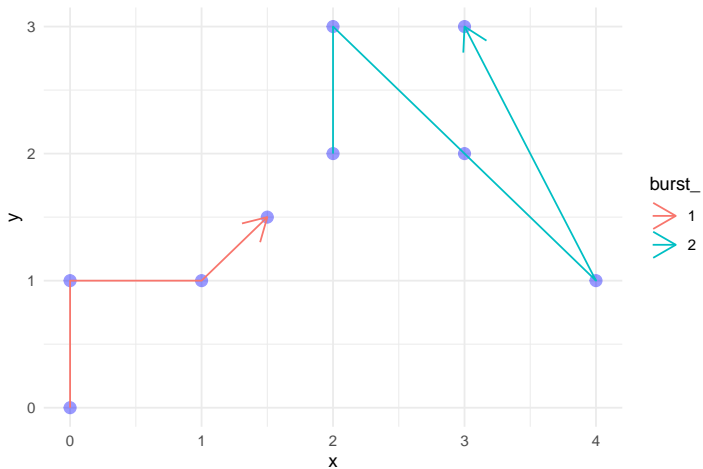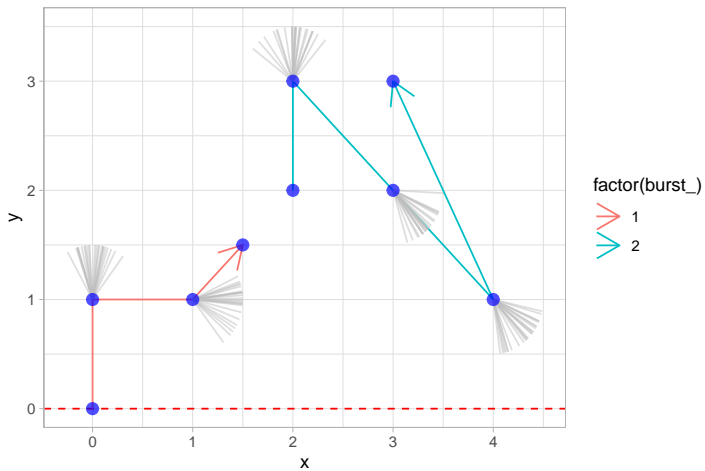
Figure 2: From Thurfjell et al. 2014

3. Sample for each observed step k random steps using the distribution from the observed steps and turn angles.

4. Extract covariates at the **end** of steps, and

5. Fit a conditional logistic regression.

- Interpretation of parameters can be done exactly the same way as for HSF (previous module).

- The relative selection strength (RSS) is how many times more probable it is, to select one step over an other step, when the two steps differ in one covariate at the end of the step.

We fitted a simple model to one deer: `fit_ssf(case_ ~ forest + strata(step_id_))`

```
# A tibble: 1 x 5
  term    estimate std.error statistic      p.value
  <chr>      <dbl>     <dbl>     <dbl>        <dbl>
1 forest     0.568     0.111      5.11 0.000000326
```

Think-pair-share: Try to give an interpretation on the estimated coefficient for forest.

We fitted a simple model to one deer: `fit_ssf(case_ ~ forest + strata(step_id_))`

```
# A tibble: 1 x 5
  term    estimate std.error statistic    p.value
  <chr>      <dbl>     <dbl>     <dbl>      <dbl>
1 forest     0.568     0.111      5.11 0.000000326
```

Think-pair-share: Try to give an interpretation on the estimated coefficient for forest.

- `forest` indicates if the end of a step is located inside a forest or not.

- $\exp(\beta_1) = \exp(0.302) = 1.35$. This indicates that it is 1.35 times more probably for a step to end in a forest patch than in a non-forest patch.

Take-home points

- The availability is usually constraint and SSFs allow to model this.

- Each observed step is paired with random k steps.

- An observed step and k random steps form a strata.

- We can use conditional logistic regression to estimate the selection coefficients

The integrated Step-Selection Function (iSSF)

- Forester et al. 2009 suggested to include the step length as an additional covariate to account for the bias that is introduced, because we are sampling from the observed steps. These steps are not independent of habitat selection.

- Avgar et al. 2016 coined the term **integrated** SSF and showed how commonly used statistical distributions can be used to model the selection-free movement kernel.

For background see

Integrated step selection analysis: bridging the gap between resource selection and animal movement

❶ Correction(s) for this article ⌄

Tal Avgar ✉, Jonathan R. Potts, Mark A. Lewis, Mark S. Boyce

For background see

Methods in Ecology and Evolution | BRITISH ECOLOGICAL SOCIETY

Standard Paper | 🔒 Open Access | (cc) (i) (=) (\$)

Integrated step selection analysis: bridging the gap between resource selection and animal movement

❗ Correction(s) for this article ⌄

Tal Avgar ✉, Jonathan R. Potts, Mark A. Lewis, Mark S. Boyce

Journal of Animal Ecology | BRITISH ECOLOGICAL SOCIETY

HOW TO… | 🔒 Open Access | (cc) (i) (\$)

A 'How to' guide for interpreting parameters in habitat-selection analyses

John Fieberg ✉, Johannes Signer, Brian Smith, Tal Avgar

iSSF in a bit more detail

- This adds a further term to the equation we saw before:

$$u(s, t + \Delta t) | u(s', t) = \frac{w(X(s); \beta(\Delta t))\phi(s, s'; \gamma(\Delta t))}{\int_{\tilde{s} \in G} w(X(\tilde{s}, s'); \beta(\Delta t))\phi(\tilde{s}, s'; \gamma(\Delta t))d\tilde{s}}$$

- $\phi(s, s'; \gamma(\Delta t))$ is a selection-independent movement kernel that describes how the animal would move in homogeneous habitat.

# Characterizing animal movement (the movement kernel)

- In order to use statistical models to describe the movement process.

- This is most easily done using statistical distributions for step lengths and turn angles.

Distributions in statistics

- Random variables, are variables where each outcome has a probability and these probabilities are often mathematically summarized with functions that are characterized with one or more parameters (=**distributions**).

- A distribution translates the a possible outcome in a probability, given some parameters.

For example, a normal distribution is characterized by the μ (mean) and σ (standard deviation). If we know μ and σ we can plot the distribution

```
curve(dnorm(x, mean = 10, sd = 2), from = 5, to = 15)
```
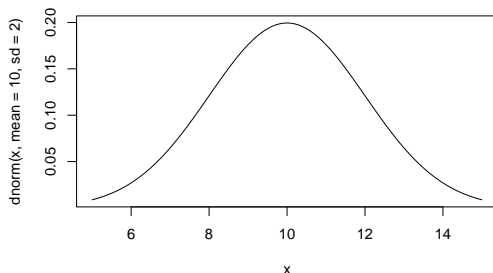

Note, dnorm() is a built-in function for the normal density function.

- The normal distribution is very widely used in statistics, however, it is not useful to characterize step lengths and turn angles of animals, because values can range from $-\infty$ to ∞.

Think-pair-share: what properties do distributions need to have, in order to describe animal movement?

- The normal distribution is very widely used in statistics, however, it is not useful to characterize step lengths and turn angles of animals, because values can range from $-\infty$ to ∞.

Think-pair-share: what properties do distributions need to have, in order to describe animal movement?

- For step lengths we need a distribution with a support of positive values.

- For turn angles we need a wrapped distribution.

Distributions for step lengths

- Suitable distribution for step lengths: Gamma, Exponential, Half Normal, Weibull and possibly others.

- We will use the Gamma distribution.

- The Gamma distribution has two parameters: shape and scale.
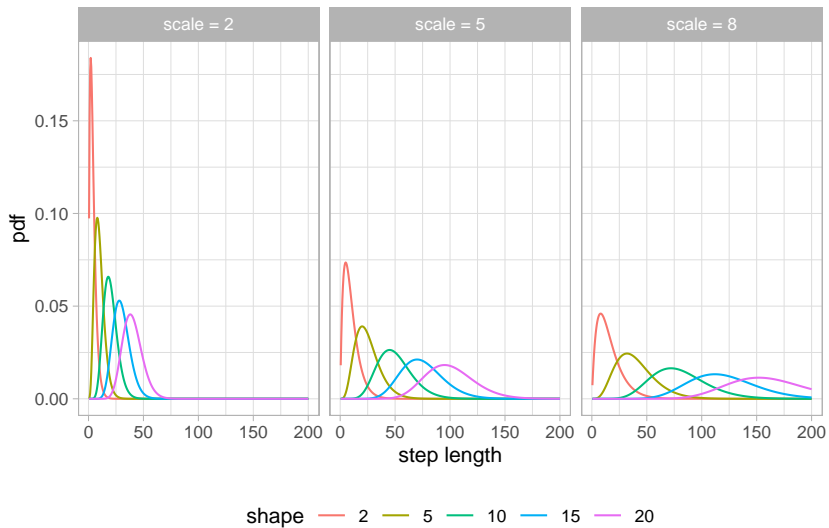
Distributions for turn angles

- For turn angles we need a wrapped distribution.

- A typical distributions are the von Mises or the wrapped Cauchy distribution.

Parameters may change

- Parameters for step-length distribution and turn-angle distribution can change, resembling different behavioral states of the animal (e.g., foraging vs. resting).
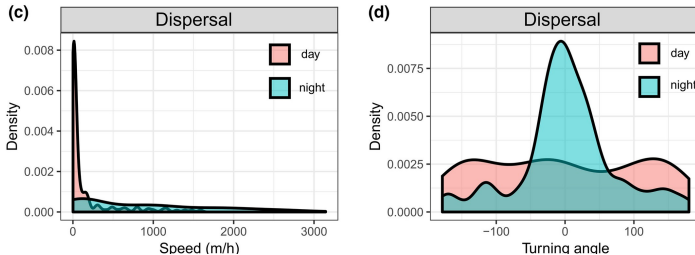

Figure 3: Figure from Moll et al. 2021

Take-home messages

1. We can use statistical distributions to characterize the movement of animals.

2. Statistical distributions are described using one or more parameters.

3. Different behavioral states of the animal can be described with different parameter sets.

- We can estimate the selection coefficients (β's) and movement related coefficients (γ's) again using a conditional logistic regression.

- But we are restricted to distributions from the exponential family.

- See Michelot et al. 2024 how this can be relaxed.

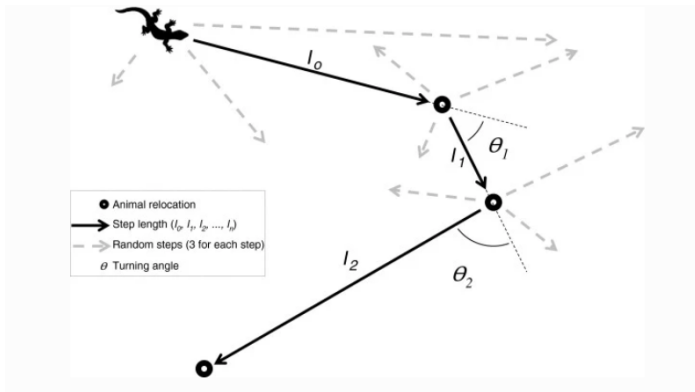Remember how we can characterize movement in discrete time:

Figure 4: From Thurfjell et al. 2014

- Before we sampled random steps from the observed step lengths and turn angles.

- We replace this step by using statistical distributions (most often Gamma distribution for step lengths and von Mises distribution for turn angles).

Creating random steps for iSSFs

We will cover this in much more detail in the R walkthrough. The workflow is usually as follows:

1. Make sure the data are in bursts of regular sampling rates (Note, the Δt in the equation).

2. Create steps (that is converting two consecutive points to a step with a start and end point).

3. *Sample for each observed step k random steps using the distribution from the observed steps and turn angles.*

4. Sample for each observed step *k* random steps **using a statistical distribution fit to the observed steps or turn angles**.

5. Extract covariates at the *end* of steps, and

6. Fit a conditional logistic regression.

How to include movement?

- Including the habitat independent movement kernel in the model is possible through functions for the step length and turn angles[1].

- For the Gamma distribution the coefficients of the
  - step length is linked to the scale parameter, and
  - the log of the step length is linked to the shape parameter.

- For the von Mises distribution the cosine of the turn angle is linked to the concentration parameter.

Once the model is fitted to the data, the tentative parameter estimates can be adjusted (more on this in module 7).

---

[1] see here for details:
https://conservancy.umn.edu/bitstream/handle/11299/218272/AppC_iSSA_movement.html

How many random steps do we need?

- There is no rule of thumb, generally the more random steps the better.

- Typical number of random steps range from 10 to 100.

- An easy way to check if the number of random steps is sufficient, is to rerun the analysis several times with increasing an increasing number of random steps and check when estimates stabilize.

Take-home points

- SSF can be extended by including movement (**iSSF**).

- We sample step length and turn angles from fitted statistical distributions.

- We then include functions of steps in the model.

Key resources/publications

- 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 in Ecology and Evolution, 7(5), 619-630.

- Fortin, D., Beyer, H. L., Boyce, M. S., Smith, D. W., Duchesne, T., & Mao, J. S. (2005). Wolves influence elk movements: behavior shapes a trophic cascade in Yellowstone National Park. Ecology, 86(5), 1320-1330.

- Michelot, T., Klappstein, N. J., Potts, J. R., & Fieberg, J. (2023). Understanding step selection analysis through numerical integration. Methods in Ecology and Evolution.

- 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.

- Northrup, J. M., Vander Wal, E., Bonar, M., Fieberg, J., Laforge, M. P., Leclerc, M., . . . & Gerber, B. D. (2022). Conceptual and methodological advances in habitat-selection modeling: guidelines for ecology and evolution. Ecological Applications, 32(1), e02470.

- Signer, J., Fieberg, J., & Avgar, T. (2019). Animal movement tools (amt):