Interpreting and Predicting (exponential) HSFs & (integrated) SSFs

Tal Avgar

Interpreting eHSF: what's in a β?

The natural logarithm of the selection strength for location x_1 relative to location x_2:

$$ln\left(\frac{exp[\beta_0 + \sum_{i=1}^{n} \beta_i \cdot h_i(x_1)]}{exp[\beta_0 + \sum_{i=1}^{n} \beta_i \cdot h_i(x_2)]}\right) = \sum_{i=1}^{n} \beta_i \cdot [h_i(x_1) - h_i(x_2)]$$

- The selection coefficient for habitat covariate i, β_i, is thus the lnRSS for x_1 vs x_2, conditional on $[h_i(x_1) - h_i(x_2)] = 1$ and all other covariates being equal
- The selection coefficient for habitat covariate i, β_i, is thus the conditional effect size of h_i

Interpreting eHSF: what's in a β?

The natural logarithm of the **selection strength** for location x_1 **relative** to location x_2:

$$ln\left(\frac{exp[\beta_0 + \sum_{i=1}^{n} \beta_i \cdot h_i(x_1)]}{exp[\beta_0 + \sum_{i=1}^{n} \beta_i \cdot h_i(x_2)]}\right) = \sum_{i=1}^{n} \beta_i \cdot [h_i(x_1) - h_i(x_2)]$$

- The selection coefficient for habitat covariate i, β_i, is thus the lnRSS for x_1 vs x_2, conditional on $[h_i(x_1) - h_i(x_2)] = 1$ and all other covariates being equal
- The selection coefficient for habitat covariate i, β_i, is the conditional effect size of h_i

Interpreting eHSF: what's in a β?

The natural logarithm of the **selection strength** for location x_1 **relative** to location x_2:

$$ln\left(\frac{exp[\beta_0 + \sum_{i=1}^{n} \beta_i \cdot h_i(x_1)]}{exp[\beta_0 + \sum_{i=1}^{n} \beta_i \cdot h_i(x_2)]}\right) = \sum_{i=1}^{n} \beta_i \cdot [h_i(x_1) - h_i(x_2)]$$

- The selection coefficient for habitat covariate i, β_i, is thus the ***lnRSS*** for x_1 vs x_2, conditional on $[h_i(x_1) - h_i(x_2)] = 1$ and all other covariates being equal
- The selection coefficient for habitat covariate i, β_i, is the **conditional effect size** of h_i

Avgar, T., Lele, S.R., Keim, J.L. & Boyce, M.S. (2017). Relative Selection Strength: Quantifying effect size in habitat- and step-selection inference. *Ecol. Evol.*, 7, 5322–5330.

Interpreting iSSF: what's in a β?

For m step attributes (e.g., habitat quality or step length), $z_j(x)$:

$$p[x_t = x_i] = \frac{exp\left[\sum_{j=1}^{m} \beta_j \cdot z_j(x_i)\right]}{\int exp\left[\sum_{j=1}^{m} \beta_j \cdot z_j(x)\right] dx}$$

If x_1 and x_2 are endpoints of two alternative steps originating from $x_{t-\tau}$:

$$\frac{p[x_t = x_1]}{p[x_t = x_2]} = \frac{exp\left[\sum_{j=1}^{m} \beta_j \cdot z_j(x_1)\right]}{exp\left[\sum_{j=1}^{m} \beta_j \cdot z_j(x_2)\right]} = exp\left[\sum_{j=1}^{m} \beta_j \cdot \left[z_j(x_1) - z_j(x_2)\right]\right]$$

Avgar, T., Lele, S.R., Keim, J.L. & Boyce, M.S. (2017). Relative Selection Strength: Quantifying effect size in habitat- and step-selection inference. *Ecol. Evol.*, 7, 5322–5330.

Interpreting iSSF: what's in a β?

For m step attributes (e.g., habitat quality or step length), $z_j(x)$:

$$p[x_t = x_i] = \frac{exp\left[\sum_{j=1}^{m} \beta_j \cdot z_j(x_i)\right]}{\int exp\left[\sum_{j=1}^{m} \beta_j \cdot z_j(x)\right] dx}$$

If x_1 and x_2 are endpoints of two alternative steps originating from $x_{t-\tau}$:

$$\frac{p[x_t = x_1]}{p[x_t = x_2]} = \frac{exp\left[\sum_{j=1}^{m} \beta_j \cdot z_j(x_1)\right]}{exp\left[\sum_{j=1}^{m} \beta_j \cdot z_j(x_2)\right]} = exp\left[\sum_{j=1}^{m} \beta_j \cdot \left[z_j(x_1) - z_j(x_2)\right]\right]$$

Avgar, T., Lele, S.R., Keim, J.L. & Boyce, M.S. (2017). Relative Selection Strength: Quantifying effect size in habitat- and step-selection inference. *Ecol. Evol.*, 7, 5322–5330.

Interpreting iSSF: what's in a β?

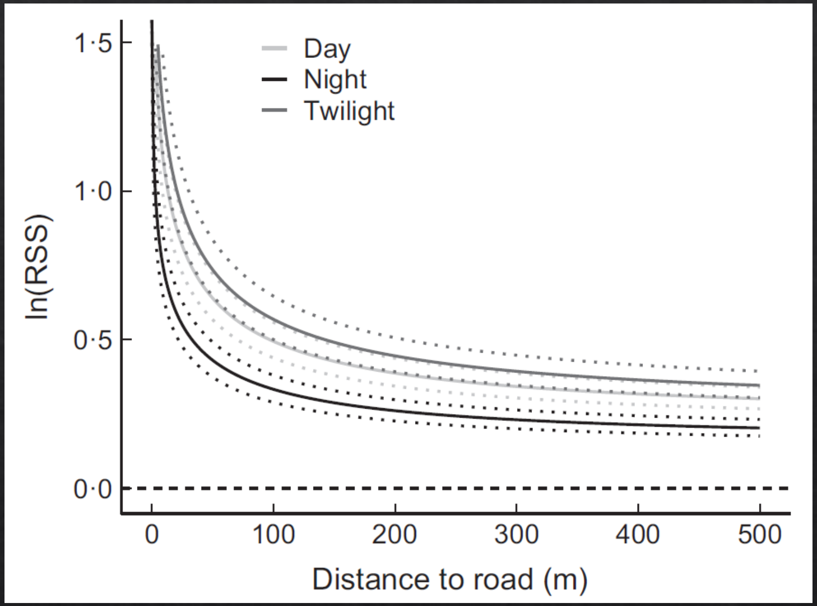
$$ln\left(\frac{p[x_t = x_1]}{p[x_t = x_2]}\right) = \sum_{j=1}^{m} \beta_j \cdot [z_j(x_1) - z_j(x_2)]$$

β_j is thus the ***lnRSS*** for a step ending at x_1 relative to a step ending at x_2 given the two steps are identical in all attributes but z_j, which is one unit larger for the step ending at x_1 than for the step ending at x_2

Avgar, T., Lele, S.R., Keim, J.L. & Boyce, M.S. (2017). Relative Selection Strength: Quantifying effect size in habitat- and step-selection inference. *Ecol. Evol.*, 7, 5322–5330.

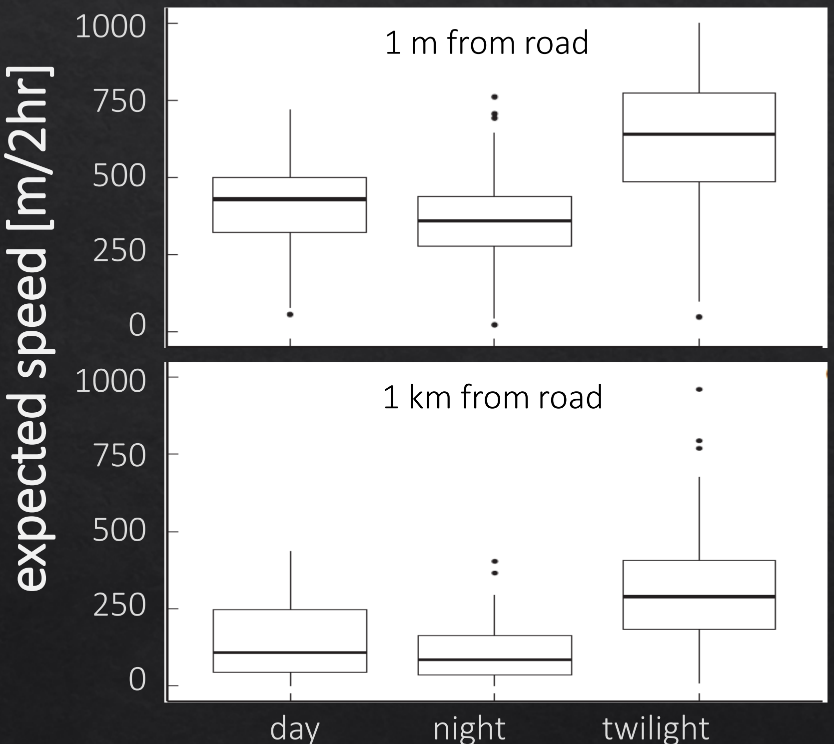
Characterizing wildlife behavioral responses to roads using iSSA

lnRSS for one location in relation to another, which is 250 m closer to a road, as a function of the distance of the second location from the road

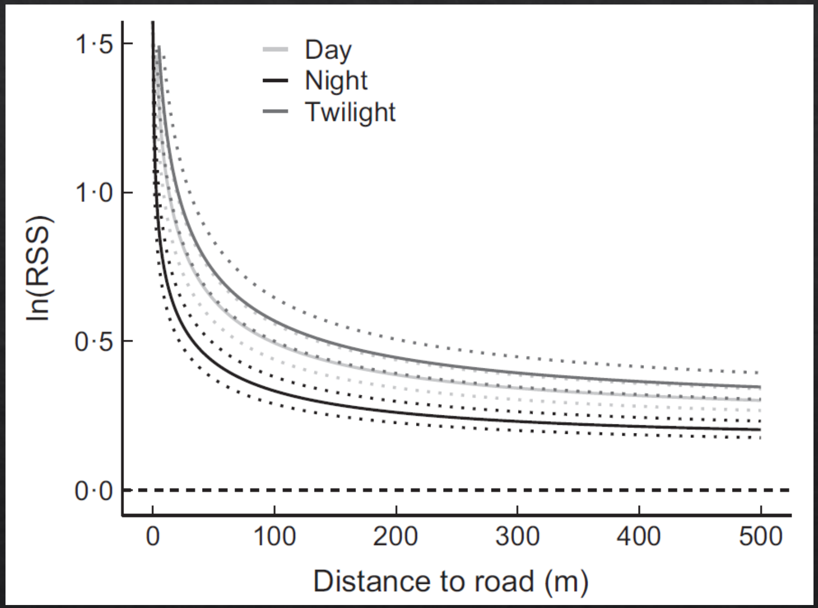


Prokopenko, C. M., Boyce, M.S., and Avgar, T. (2017). Characterizing wildlife behavioural responses to roads using integrated step selection analysis. *J Appl Ecol*, 54:470-479

Characterizing wildlife behavioral responses to roads using iSSA



lnRSS for one location in relation to another, which is 250 m closer to a road, as a function of the distance of the second location from the road



Prokopenko, C. M., Boyce, M.S., and Avgar, T. (2017). Characterizing wildlife behavioural responses to roads using integrated step selection analysis. *J Appl Ecol*, 54:470-479

Predicting what?

Predicting what?

The **Utilization Distribution (UD)**: the probability density of animal occurrence over space – a map of space-use intensity

- Steady-state UD: the expected UD after an infinitely long time, independent of the starting point
- Transient UD: the expected UD after a finite time, conditional on the starting point

Predicting what?

The **Utilization Distribution (UD)**: the probability density of animal occurrence over space – a map of space-use intensity

- **Steady-state** (stationary) **UD**: the expected UD after an infinitely long time, independent of the starting point
- **Transient** (non-stationary) **UD**: the expected UD after a finite time, conditional on the starting point

Predicting the UD based on eHSF

The probability that location x will be used (occupied; have population density > 0) is given by:

$$w(x) = c \cdot exp[\sum_{i=1}^{n} \beta_i \cdot h_i(x)]$$

Predicting the UD based on eHSF

The probability that location x will be used (occupied; have population density > 0) is given by:

$$w(x) = \cancel{c}\ exp[\sum_{i=1}^{n} \beta_i \cdot h_i(x)]$$

Predicting the UD based on eHSF

The probability that location x will be used (occupied; have population density > 0) is given by:

$$w(x) = c \cdot exp[\sum_{i=1}^{n} \beta_i \cdot h_i(x)]$$

The UD is a probability density, and should thus sum to 1:

$$UD(x_j) = \frac{w(x_j)}{\sum_{j=1}^{\Omega} w(x_j)} = \frac{exp[\sum_{i=1}^{n} \beta_i \cdot h_i(x)]}{\sum_{j=1}^{\Omega} exp[\sum_{i=1}^{n} \beta_i \cdot h_i(x)]}$$

Predicting the UD based on eHSF

Exponential habitat-selection functions are parametrized under the assumption that used points were observed with a probability that is exactly proportional to the *steady-state UD*. Consequently, mapping of the parametrized HSF onto a habitat raster can be interpreted as the steady-state UD, but provides no information on transient dynamics

Predicting the UD based on iSSF

- Cannot be done by applying $exp\left[\sum_{j=1}^{m} \beta_j \cdot z_j(x_i)\right]$ to the habitat raster; the relative residence time in x depends on the conditional probability of all the steps leading to and from x
- An iSSF-based UD must be simulated: an iSSF is a fully parametrised individual-based simulation of a locally biased random walk
- Simulations are stochastic and must be repeated many (many) times
  - The steady-state UD is approximated by running the simulation for many (many) steps and/or starting it from various starting positions across the spatial domain
  - The transient UD is approximated by running the simulation for fewer steps, always starting from the same initial position within the spatial domain

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

Predicting the UD based on iSSF

- <u>Cannot be done</u> by applying $exp\left[\sum_{j=1}^{m} \beta_j \cdot z_j(x_i)\right]$ to the habitat raster; the relative residence time in x depends on the conditional probability of all the steps leading to and from x
- <u>An iSSF-based UD must be simulated</u>: an iSSF is a fully parametrised individual-based simulation of a ‘locally biased random walk’
- Simulations are stochastic and must be repeated many (many) times
  - The steady-state UD is approximated by running the simulation for many (many) steps and/or starting it from various starting positions across the spatial domain
  - The transient UD is approximated by running the simulation for fewer steps, always starting from the same initial position within the spatial domain

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

Predicting the UD based on iSSF

- <u>Cannot be done</u> by applying $exp\left[\sum_{j=1}^{m} \beta_j \cdot z_j(x_i)\right]$ to the habitat raster; the relative residence time in x depends on the conditional probability of all the steps leading to and from x
- <u>An iSSF-based UD must be simulated</u>: an iSSF is a fully parametrised individual-based simulation of a 'locally biased random walk'
- Simulations are stochastic and must be repeated many (many) times
  - The steady-state UD is approximated by running the simulation for many (many) steps and/or starting it from various starting positions across the spatial domain
  - The transient UD is approximated by running the simulation for fewer steps, always starting from the same initial position within the spatial domain

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

Concluding remarks

- Exponential habitat-selection models make strong assumptions about independence and stationarity, making UD estimation relatively simple (if you are willing to accept these assumptions)
- Integrated step-selection models relax these assumptions, making them more ecologically plausible, and inferentially flexible, but also more complex to use as a UD estimator
- Grain (resolution; scale) matters: UD's should always be estimated using the same spatial, and, in the case of iSSF, temporal, grain (pixel size, step durations) as the data's

Concluding remarks

- Exponential habitat-selection models make strong assumptions about independence and stationarity, making UD estimation relatively simple (if you are willing to accept these assumptions)
- Integrated step-selection models relax these assumptions, making them more ecologically plausible, and inferentially flexible, but also more complex to use as a UD estimator
- Grain (resolution; scale) matters: UD's should always be estimated using the same spatial, and, in the case of iSSF, temporal, grain (pixel size, step durations) as the data's