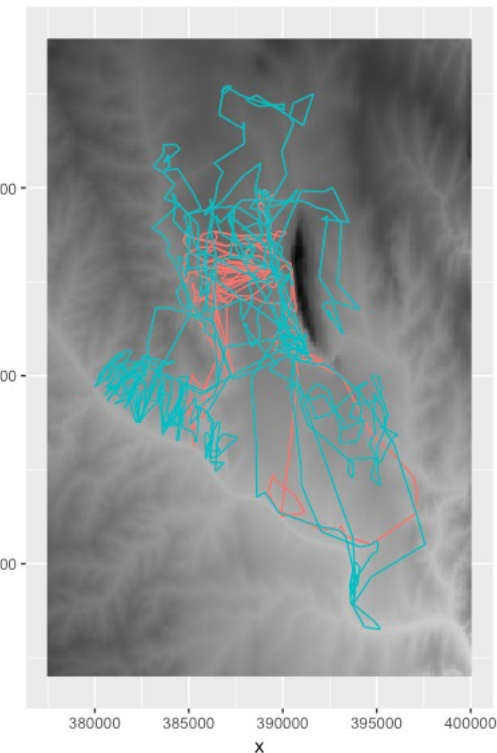


Linking movement and environment: Resource selection (RSF) and step selection functions (SSF)

Björn Reineking
LESSEM, Grenoble

Environmental effects on movement and utilization distributions

Foto: S. Rösner



- **Predict** (possibly for new environmental conditions)
 - Utilization distributions
 - Movement paths (only SSF)
- **Understand**
 - How the environment (resources such as food, risks such as predation, conditions such as temperature) shapes movement paths and long-term utilization distributions (as a function of animal characteristics such as size, sex, diet)
 - Quantify effect of drivers

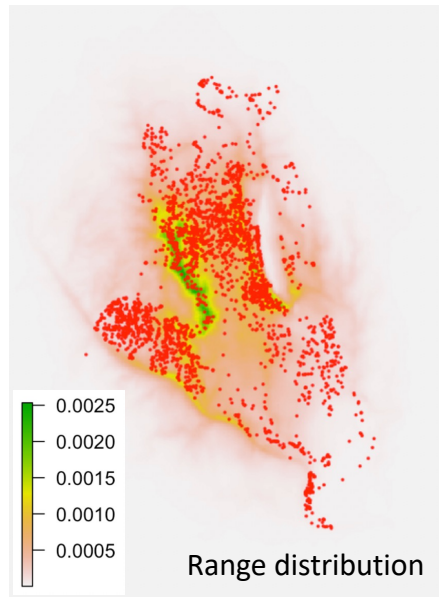
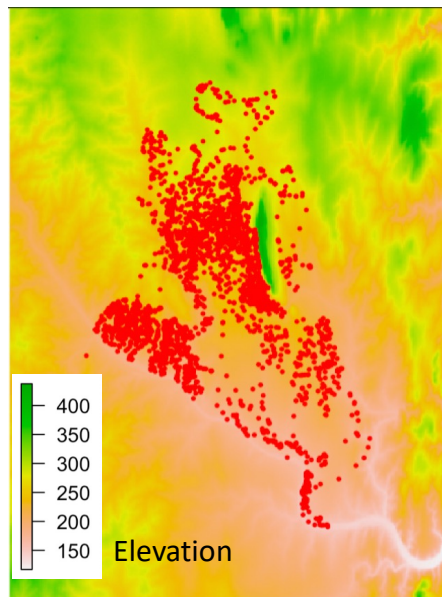
Model the probability of presence in a given area

Foto: S. Rösner

Both RSF and SSF model the probability of presence in a given area based on the environment and spatial constraints

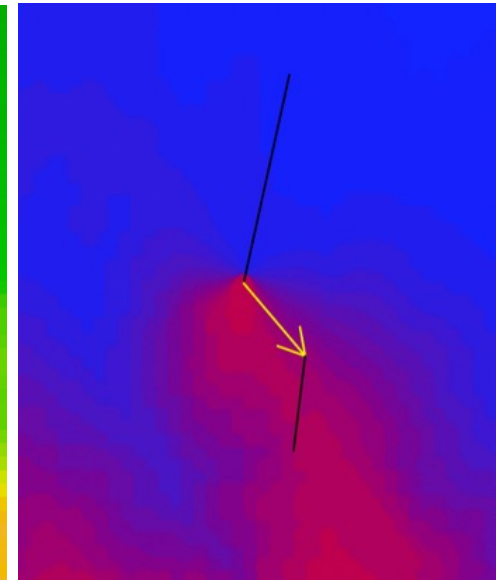
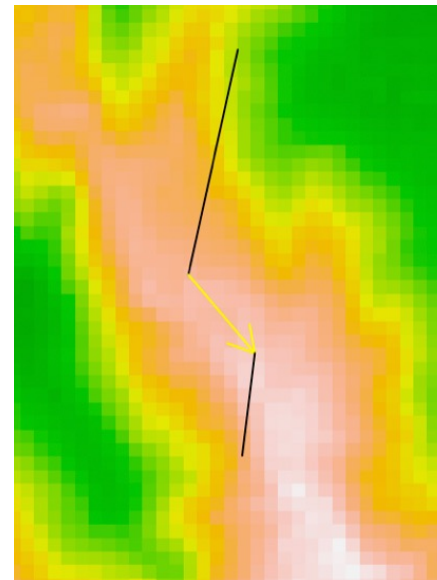
RSF

- Predicts range distribution (fraction of time the animal spends at a given site in the long term)



SSF

- Predicts position at time $t + \Delta t$ based on current position



Model the probability of presence in a given area

Foto: S. Rösner

Both RSF and SSF model the probability of presence in a given area based on the environment and spatial constraints

RSF

- Predicts occurrence distribution (fraction of time the animal spends at a given site)
- Assumes positions are statistically independent

SSF

- Predicts position at time $t + \Delta t$ based on current position
- Assumes velocities of successive steps are statistically independent

Model the probability of presence in a given area

Foto: S. Rösner

Both RSF and SSF model the probability of presence in a given area based on the environment and spatial constraints

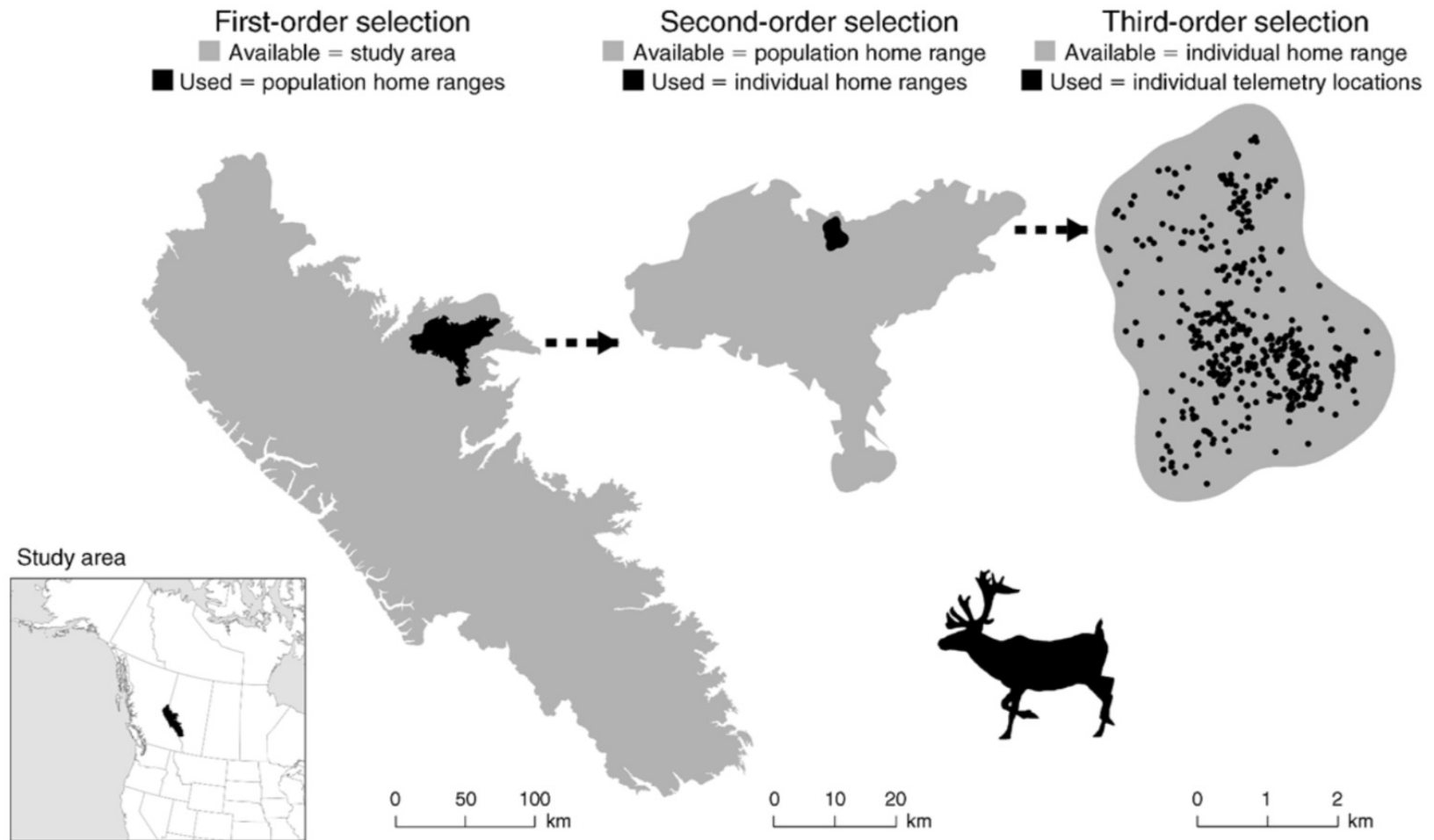
RSF

- Predicts occurrence distribution (fraction of time the animal spends at a given site)
- Assumes positions are statistically independent

SSF

- Predicts position at time $t + \Delta t$ based on current position
- Assumes velocities of successive steps are statistically independent
- Can simulate trajectories
- More difficult to get predictions of range distributions
- More efficient use of data

In general, selection parameters of RSF and SSF of a certain habitat are not identical



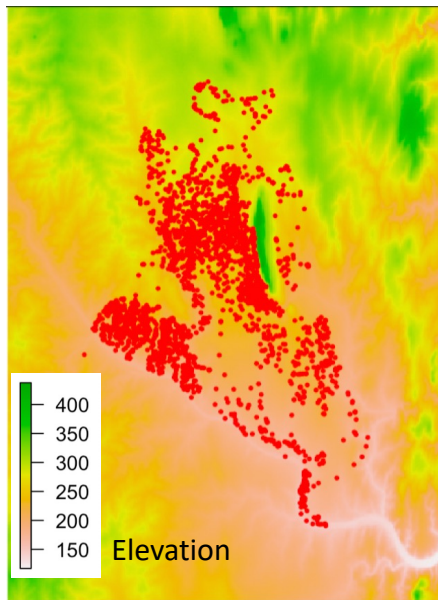
Fourth order: local selection (e.g., within a feeding site)

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

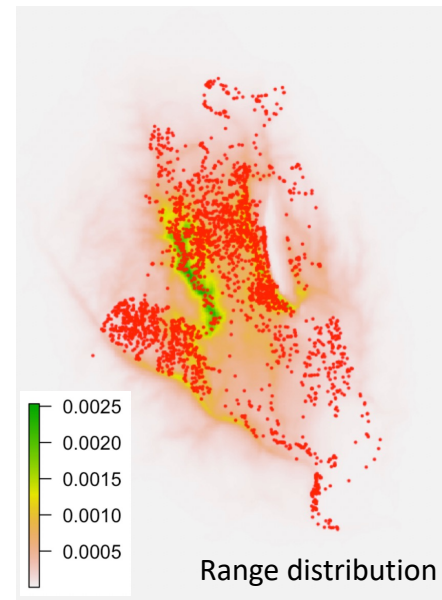
Slide from John Fieberg <https://movebankworkshopraleighnc.netlify.com>

What determines the range distribution ?

Foto: S. Rösner



?



Inhomogeneous Poisson process

Foto: S. Rösner

Given a point, the probability density of its location s is

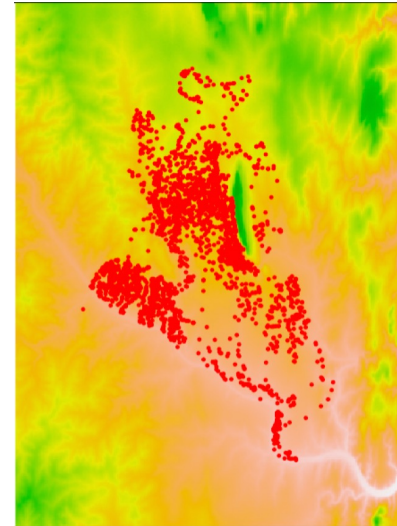
$$p(s|\lambda) = \frac{\lambda(s)}{\int_{\text{Area}} \lambda(S) dS}$$

The intensity function $\lambda(s)$ is ≥ 0 for all locations.

The integral over all potential positions S in Area ensures that the probability density integrates to 1.

The intensity function $\lambda(s)$ is typically constructed as an exponential model of k spatial covariates $h_i(s)$, where i ranges from 1 to k , and the β_i are the coefficients to be estimated.

$$\lambda(s) = \underbrace{\exp \left(\sum_{i=1}^k \beta_i h_i(s) \right)}_{\text{habitat selection } w(h(s))}$$



Inhomogeneous Poisson process

Foto: S. Rösner

Given a point, the probability density of its location s is

$$p(s|\lambda) = \frac{\lambda(s)}{\int_{\text{Area}} \lambda(S) dS}$$

The intensity function $\lambda(s)$ is ≥ 0 for all locations.

The integral over all potential positions S in Area ensures that the probability density integrates to 1.

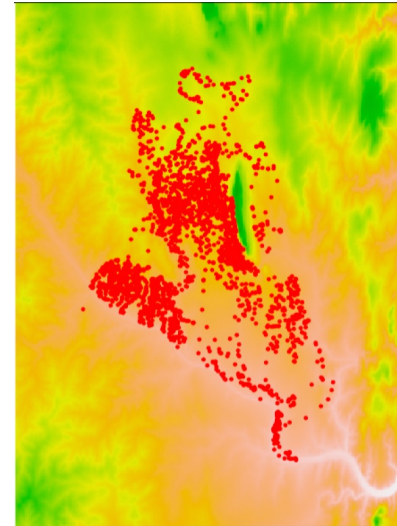
The intensity function $\lambda(s)$ is typically constructed as an exponential model of k spatial covariates $h_i(s)$, where i ranges from 1 to k , and the β_i are the coefficients to be estimated.

$$\lambda(s) = \exp \left(\underbrace{\sum_{i=1}^k \beta_i h_i(s)}_{\text{habitat selection } w(h(s))} \right)$$

There are different ways to calculate the normalisation constant $\int_{\text{Area}} \lambda(S) dS$

Two typical ways are

- Riemann: Sum over all cells of a regular grid
- Monte Carlo: Sum over a number K of random positions (drawn e.g. from a 2D normal distribution)



RSF and scale - What is the available area?

Foto: S. Rösner

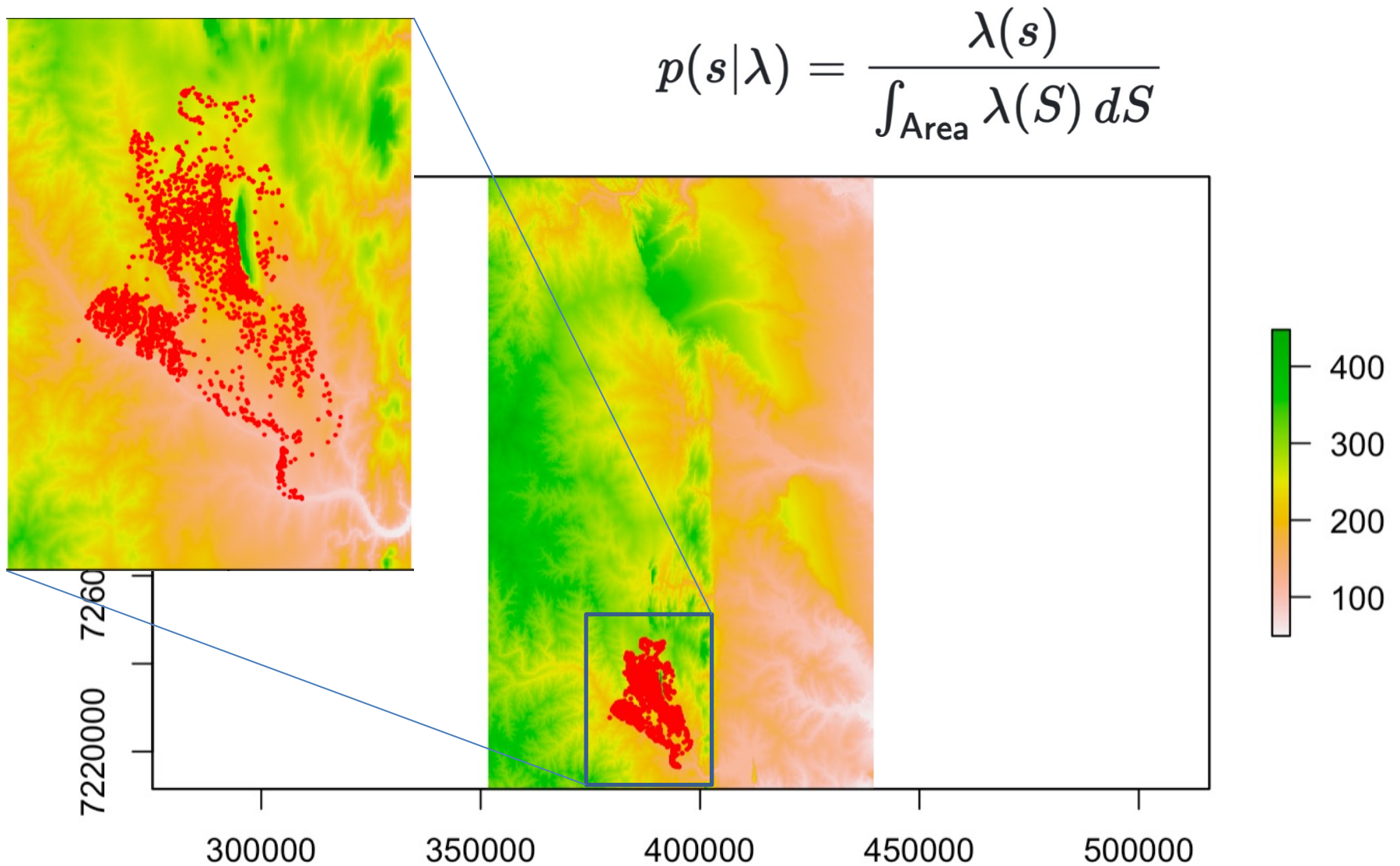




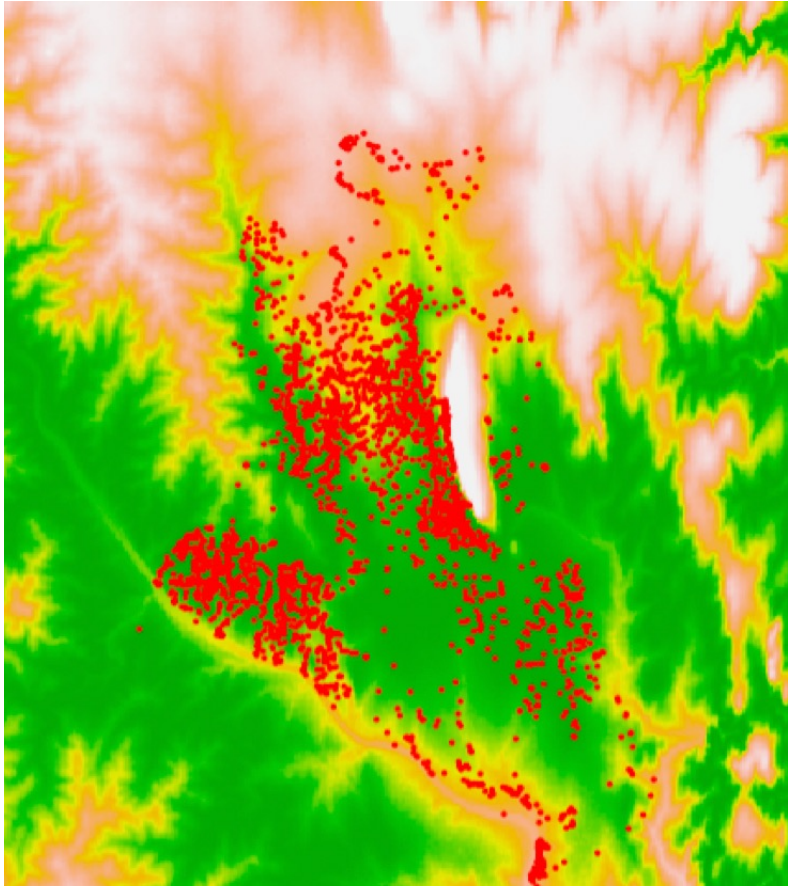
Foto: S. Rösner

RSF and scale – Include home ranging behaviour

Represents behavioural choices. Reduces dependence on definition of available area.

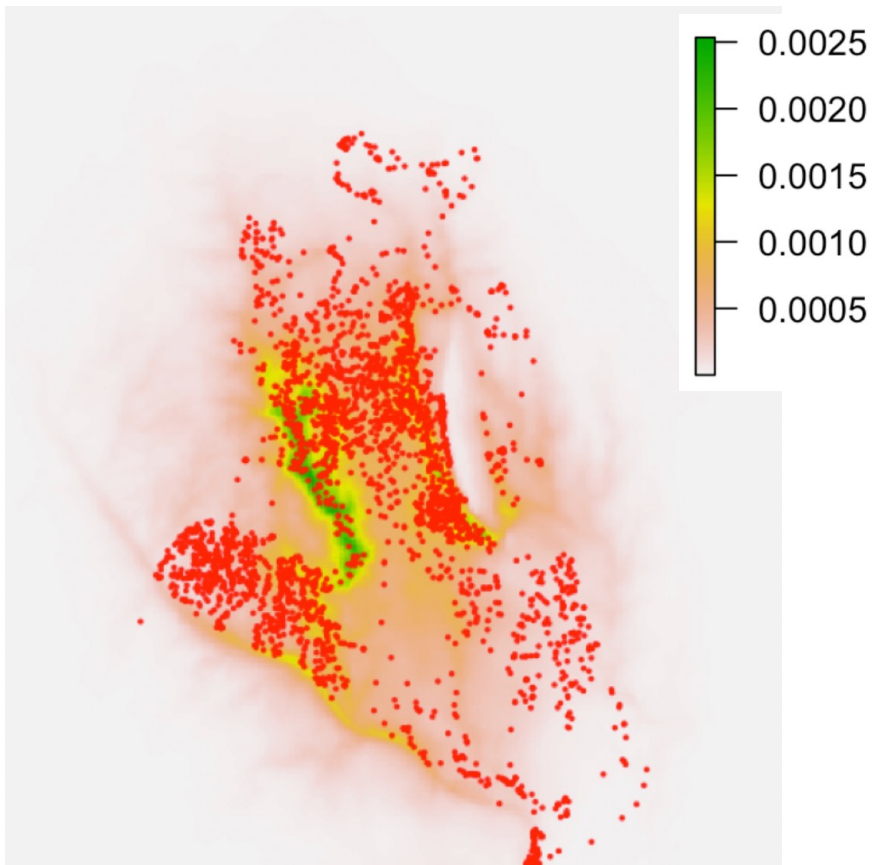
Include $x, y, I(x^2 + y^2)$ in model

RSF Selection



$$\lambda(s) = \underbrace{\exp \left(\sum_{i=1}^k \beta_i h_i(s) \right)}_{\text{habitat selection } w(h(s))}$$

RSF Range distribution



$$\lambda(s) = \underbrace{\exp \left(\sum_{i=1}^k \beta_i h_i(s) \right)}_{\text{habitat selection } w(h(s))} \underbrace{\exp \left(-\beta_{rr} [(x - \mu_x)^2 + (y - \mu_y)^2] \right)}_{\text{home ranging } \phi(s)}$$

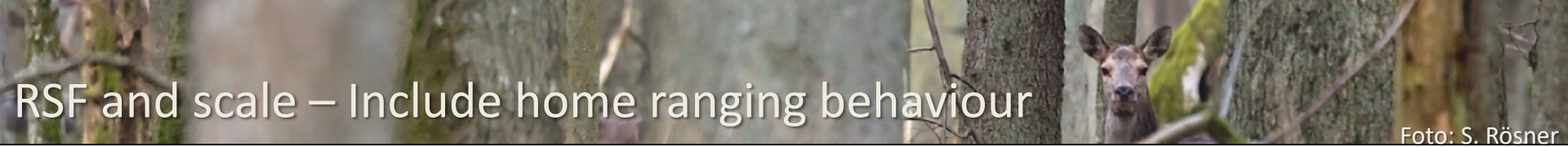


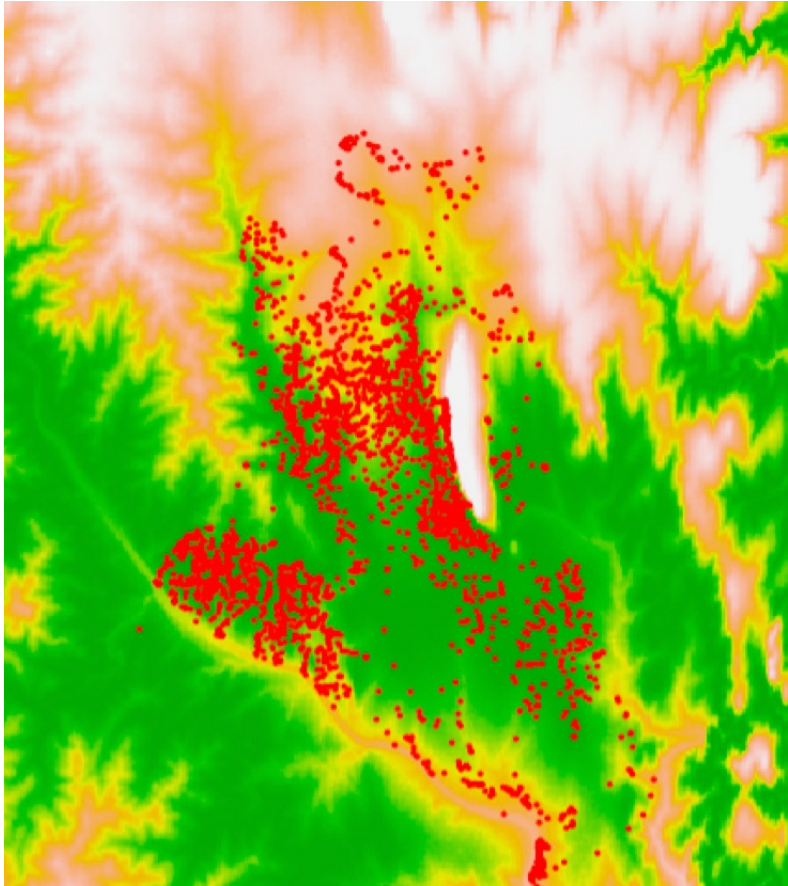
Foto: S. Rösner

RSF and scale – Include home ranging behaviour

Represents behavioural choices. Reduces dependence on definition of available area.

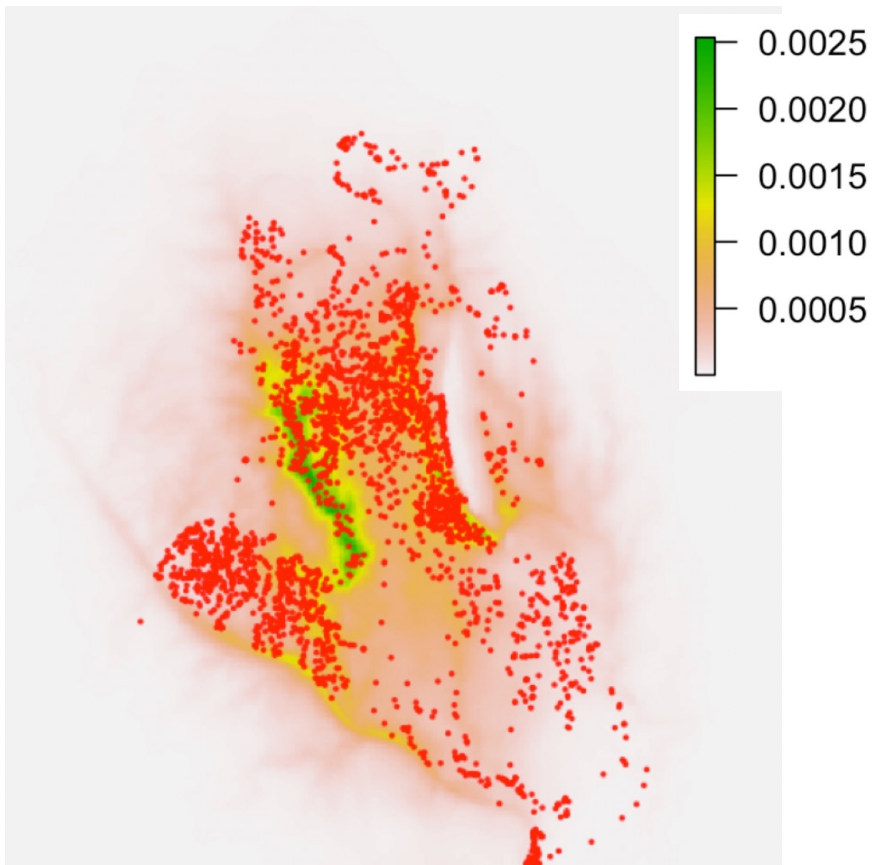
Include $x, y, I(x^2 + y^2)$ in model

RSF Selection



$$\lambda(s) = \underbrace{\exp \left(\sum_{i=1}^k \beta_i h_i(s) \right)}_{\text{habitat selection } w(h(s))}$$

RSF Range distribution



$$\lambda(s) = \underbrace{\exp \left(\sum_{i=1}^k \beta_i h_i(s) \right)}_{\text{habitat selection } w(h(s))} \underbrace{\exp \left(-\beta_{rr} [(x - \mu_x)^2 + (y - \mu_y)^2] \right)}_{\text{home ranging } \phi(s)}$$

Model of availability

$$p(s|\lambda) = \frac{\lambda(s)}{\int_{Area} \lambda(S) dS}$$

When the n observed localisations are independent, the likelihood l of the data is

$$l(\text{data}|\lambda) = \sum_{i=1}^n \log \lambda(s(t_i)) - n \log \int_{Area} \lambda(S) dS$$



Foto: S. Rösner

RSF and autocorrelation: downweight locations

- RSF assumes that successive positions are independent.
- Good news: autocorrelation can be accounted for by weighing individual positions (Alston et al. 2023 MEE)
- `ctmm::rsf.fit()` or « by hand » using weights from `akde` (see R code in exercises)
- But: effective sample sizes are in general much lower than the number of tracking points (for buffalo Cilla: 3500 points at 1 hour interval translate to effective sample size of < 20).

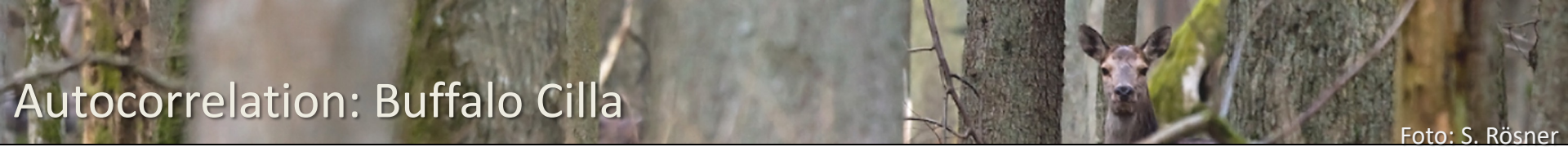
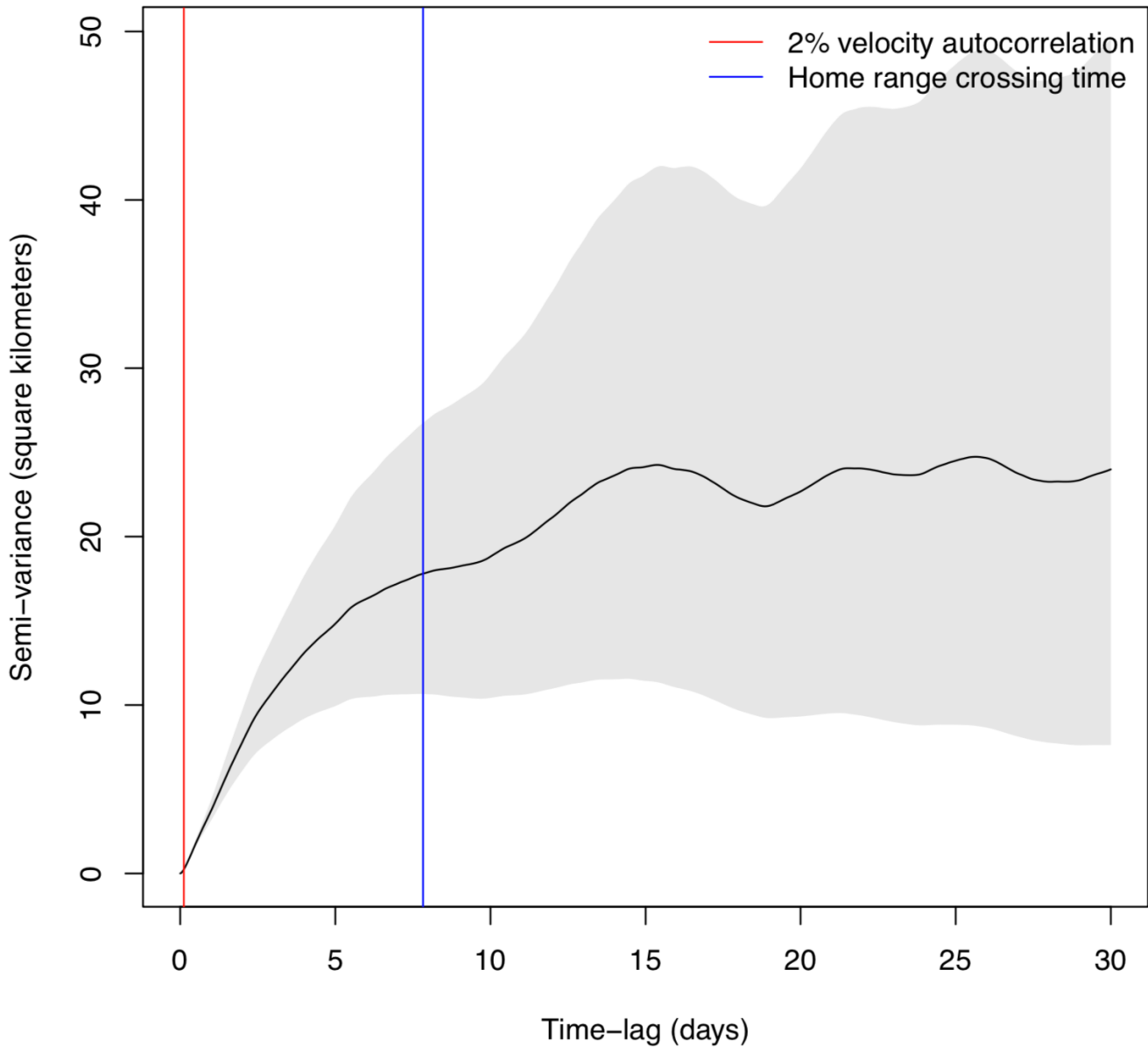


Foto: S. Rösner

Semi-variance: In 2D, a quarter of the mean squared distance between points separated by a given time lag



$$p(s|\lambda) = \frac{\lambda(s)}{\int_{Area} \lambda(S) dS}$$

When the n observed localisations are independent, the likelihood l of the data is

$$l(\text{data}|\lambda) = \sum_{i=1}^n \log \lambda(s(t_i)) - n \log \int_{Area} \lambda(S) dS$$

When they are **not** independent, we can weight the localisations in the likelihood, using the weights from the akde

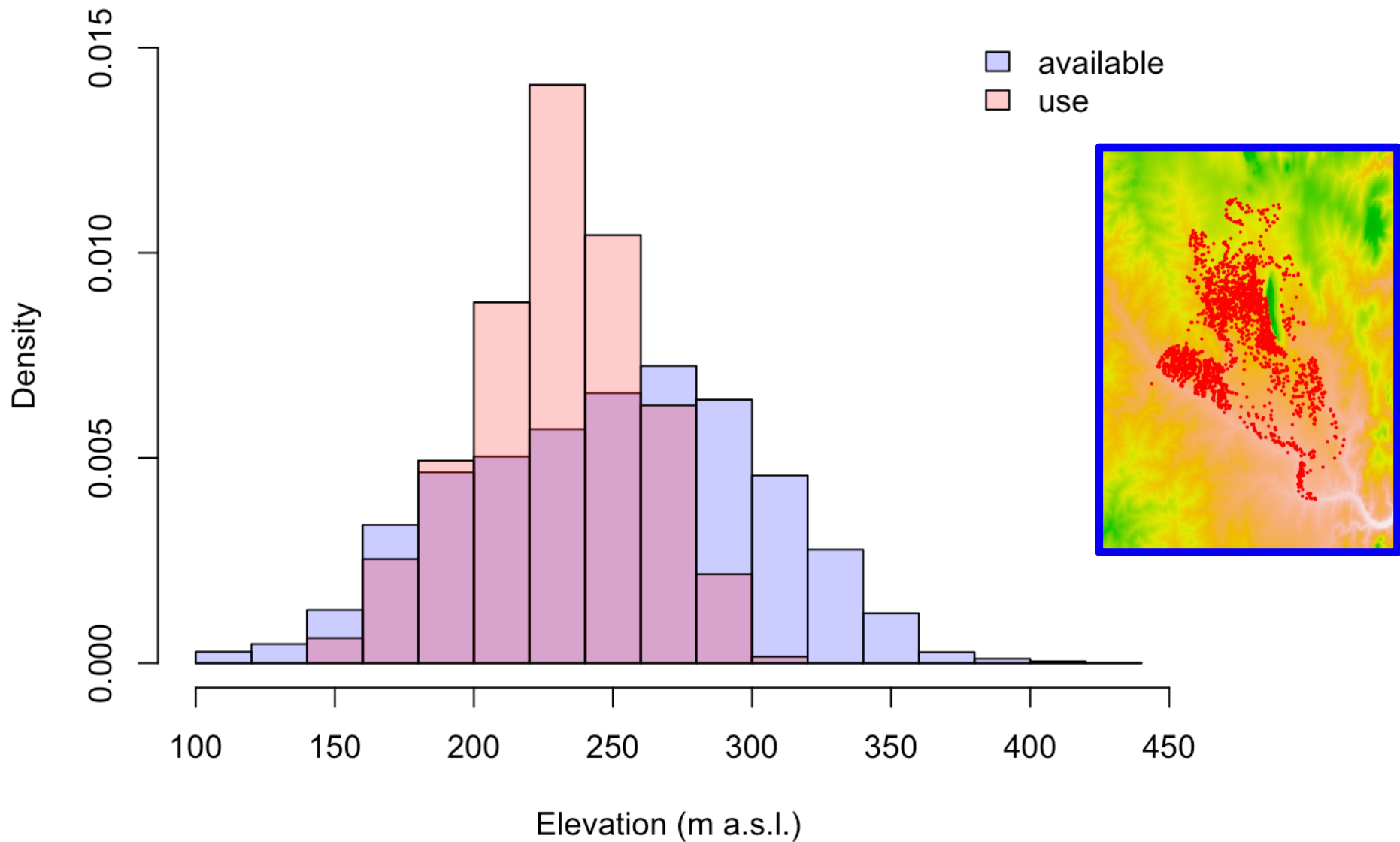
$$l(\text{data}|\lambda) = \sum_{i=1}^n w(t_i) \log \lambda(s(t_i)) - N \log \int_{Area} \lambda(S) dS$$

$$\sum_{i=1}^n w(t_i) = N \quad N \leq n$$

What do we mean by selection? Use vs availability in environmental space

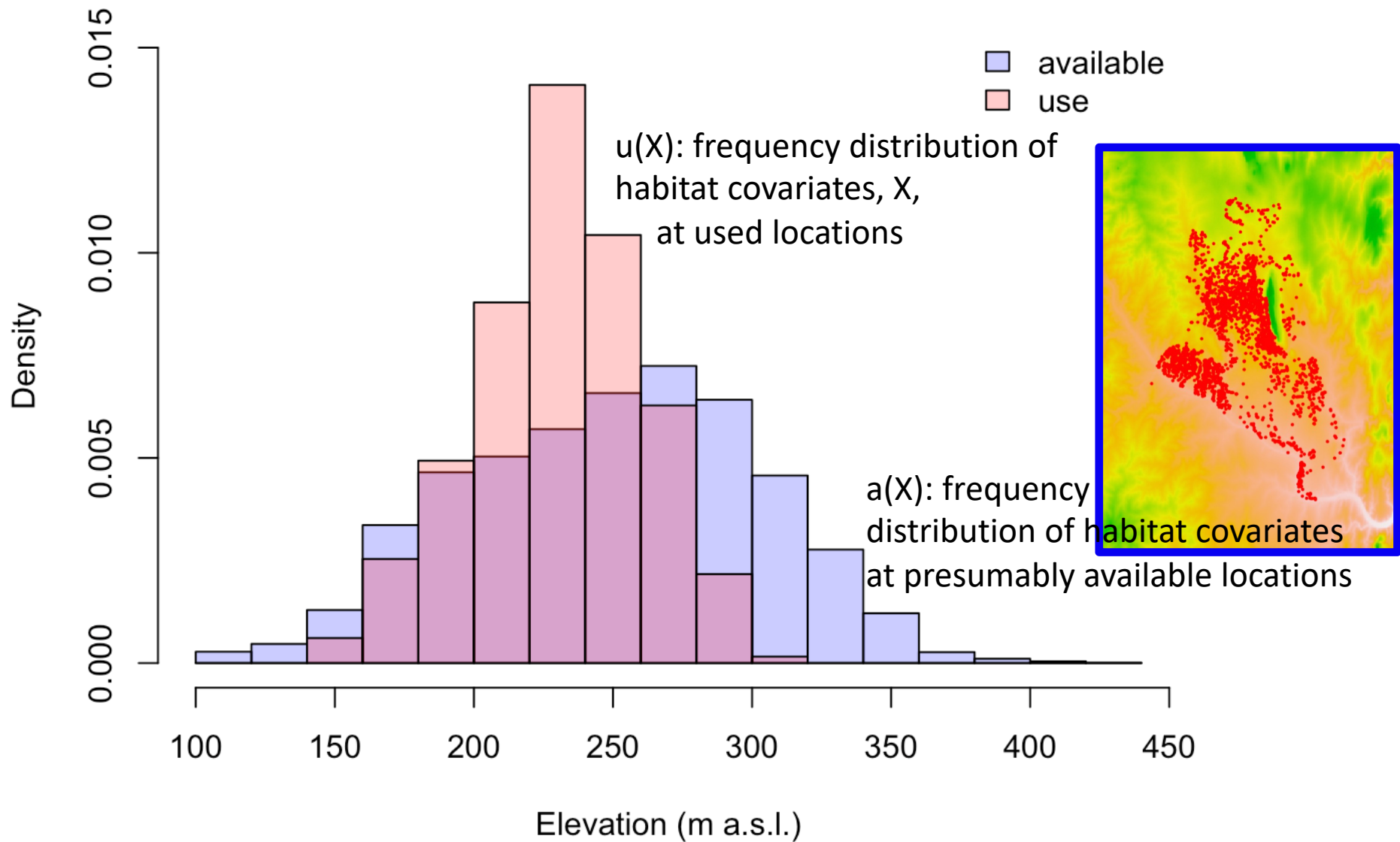
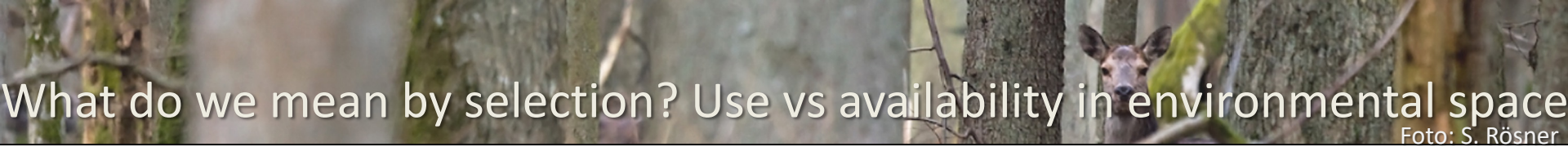
Foto: S. Rösner

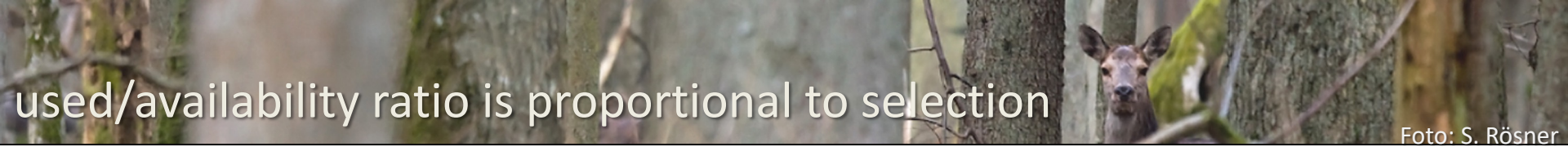
Here we assume that every point in geographic space is equally available



What do we mean by selection? Use vs availability in environmental space

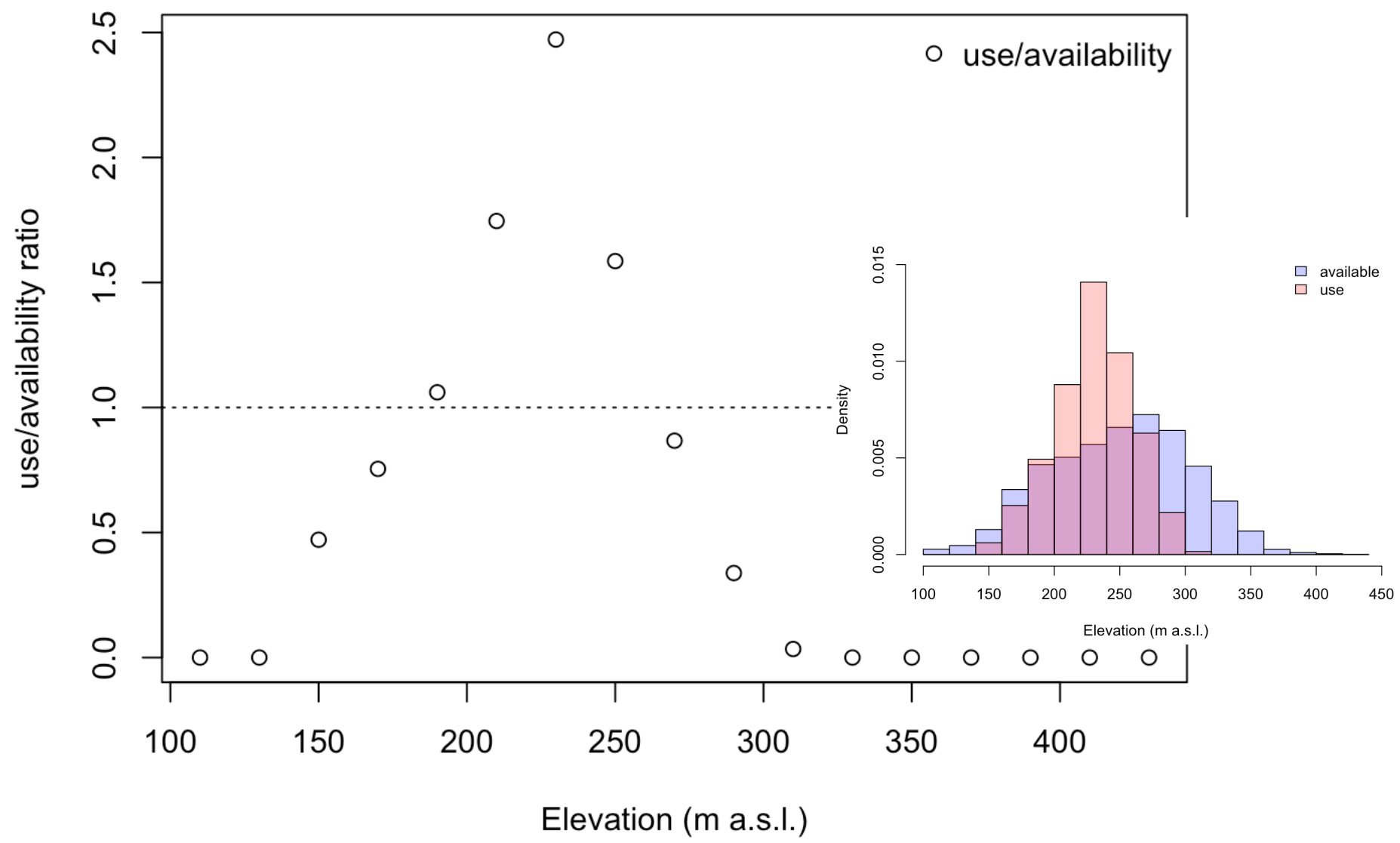
Foto: S. Rösner





used/availability ratio is proportional to selection

Foto: S. Rösner



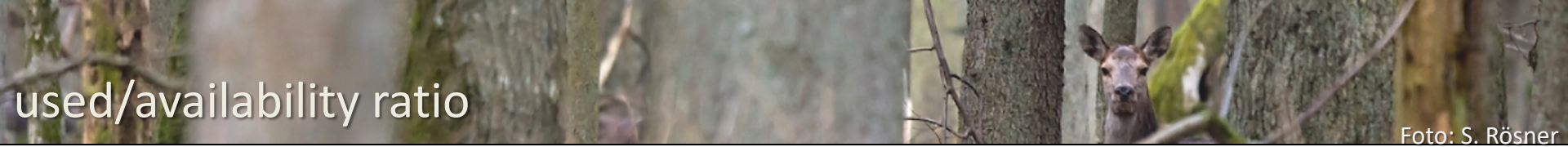
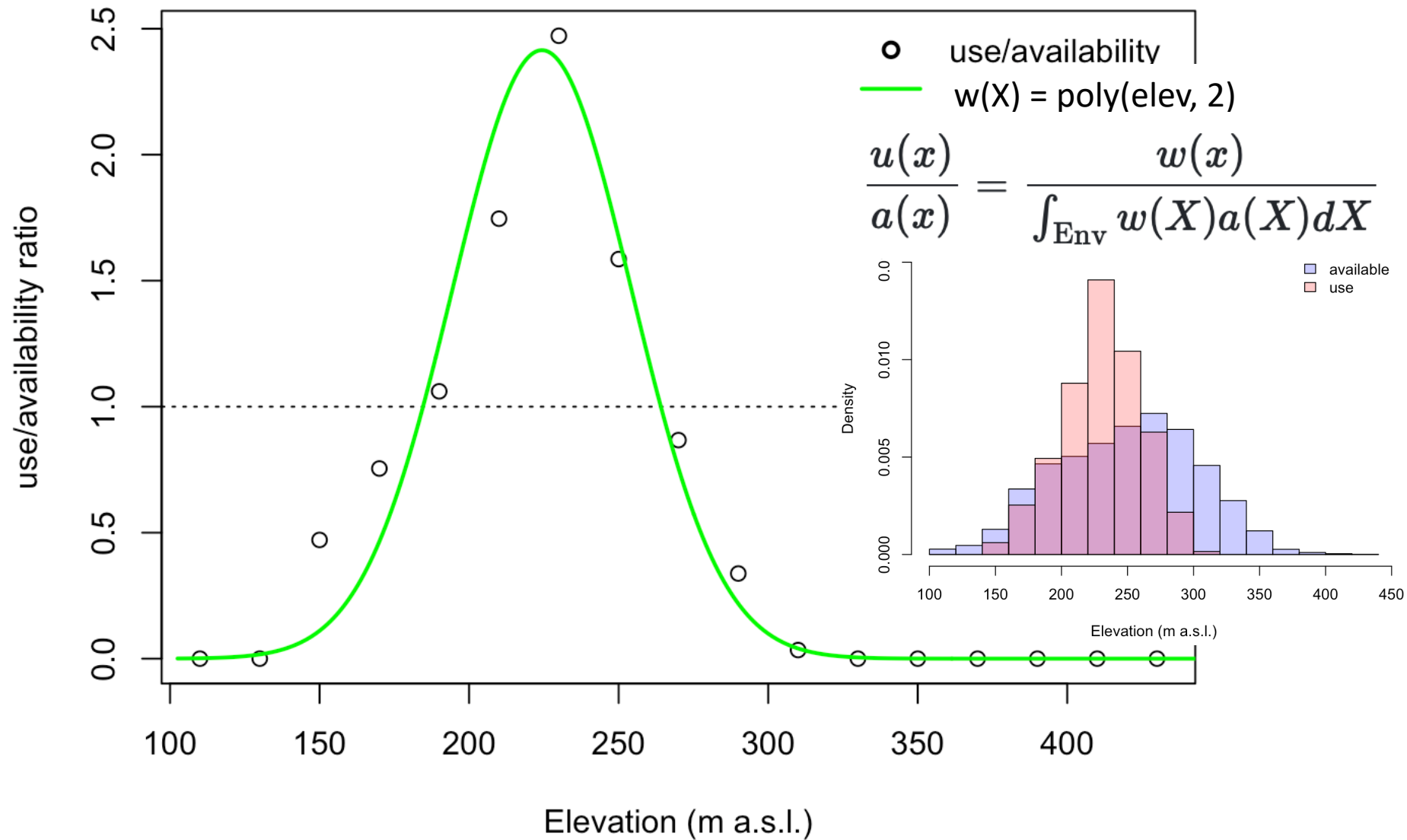


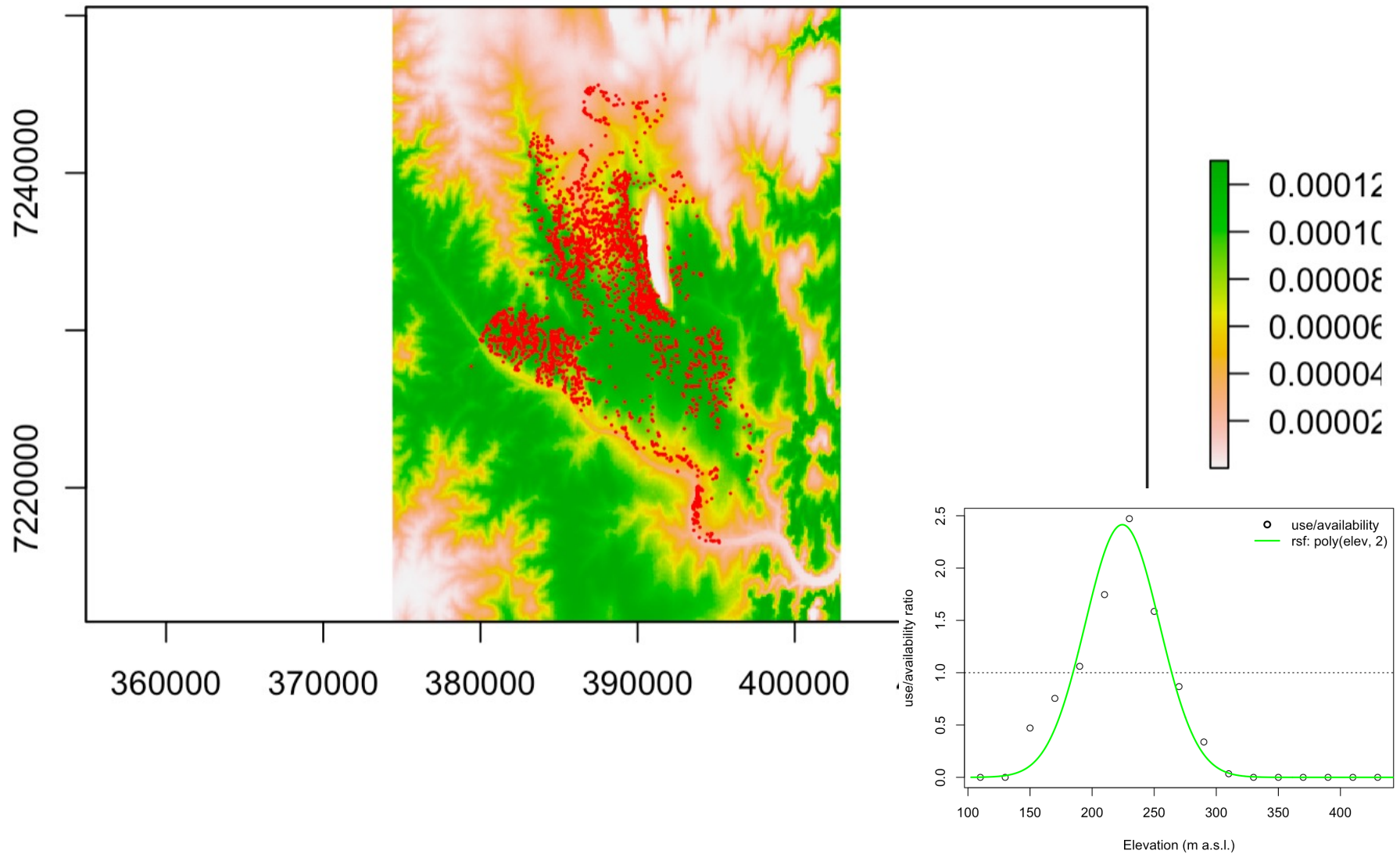
Foto: S. Rösner

used/availability ratio



RSF prediction as quadratic function of elevation

Foto: S. Rösner



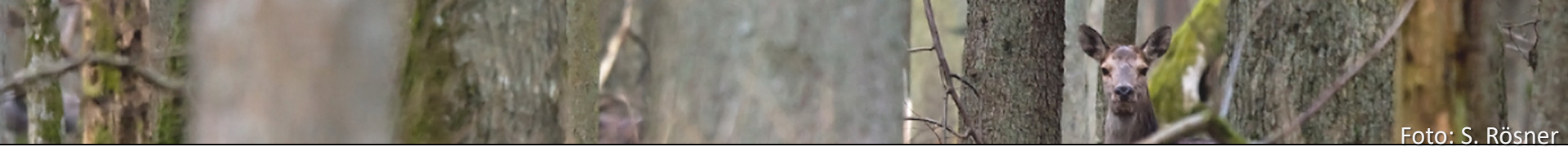


Foto: S. Rösner

Geographic space

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

Environmental space

$$u(X) = \frac{w(X, \beta)a(X)}{\int_{Z \in E} w(Z, \beta)a(Z)dZ}.$$

s: location in geographic space

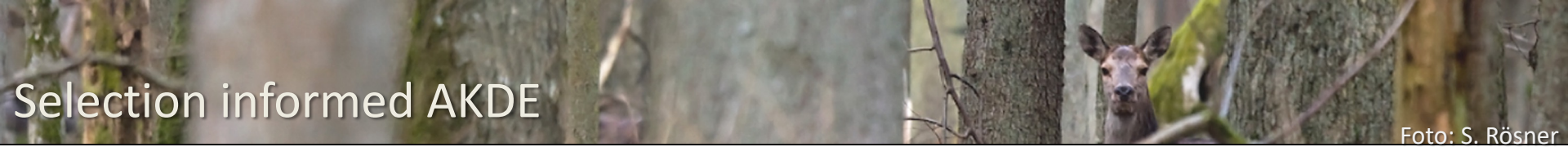
$u(s)$ = frequency distribution of locations used by animals (aka ranging distribution)

- $u(X)$ = 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.

The way we describe availability of locations in geographic space will affect our estimates of habitat selection

Availability of X in environmental space depends on

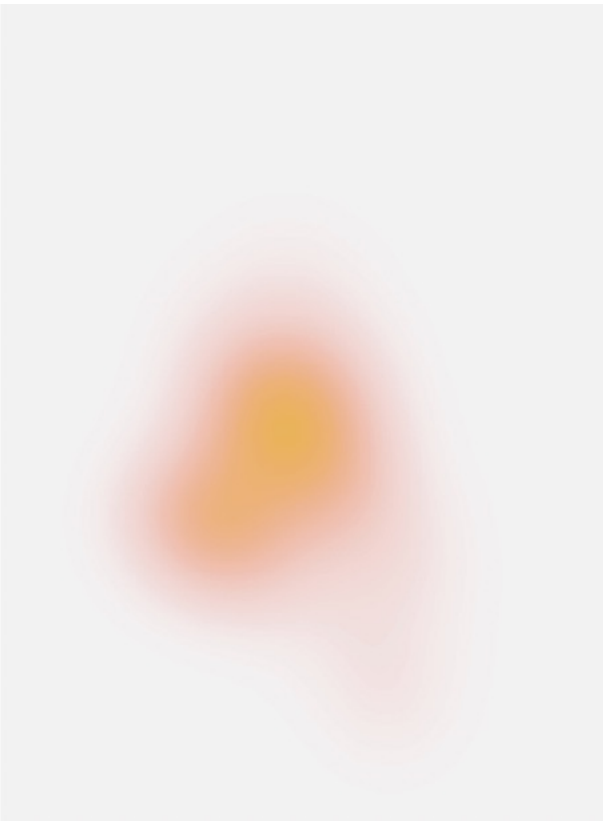
- the distribution of X in geographic space (i.e. the map of X)
- The availability of locations in geographic space for the animal



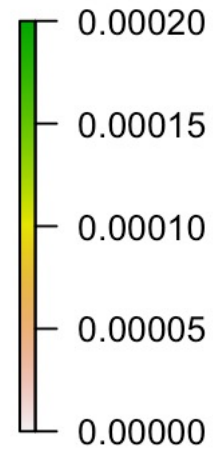
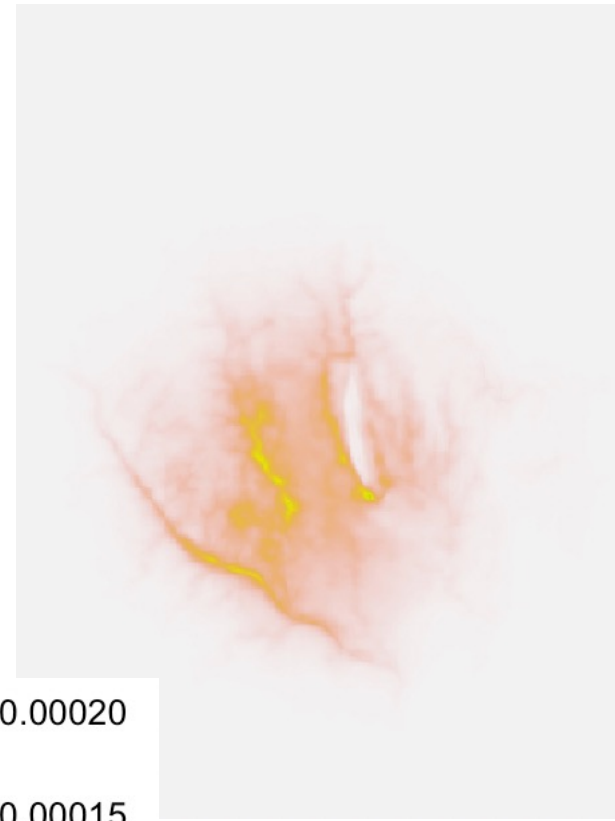
Selection informed AKDE

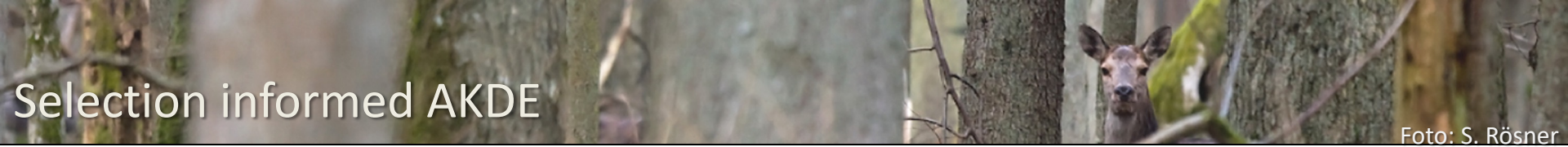
Foto: S. Rösner

akde



rsf.fit





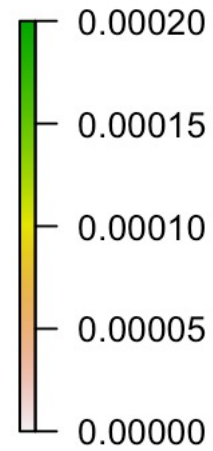
Selection informed AKDE

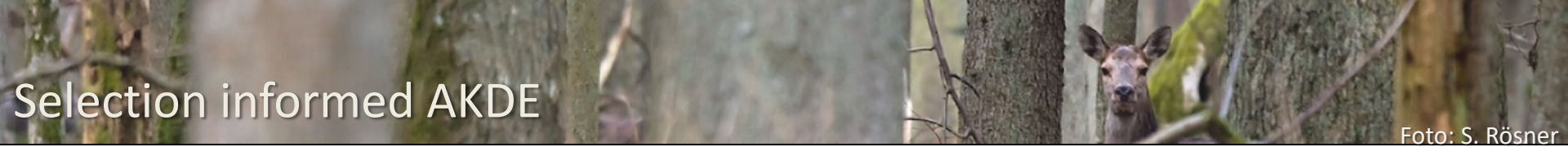
Foto: S. Rösner

akde

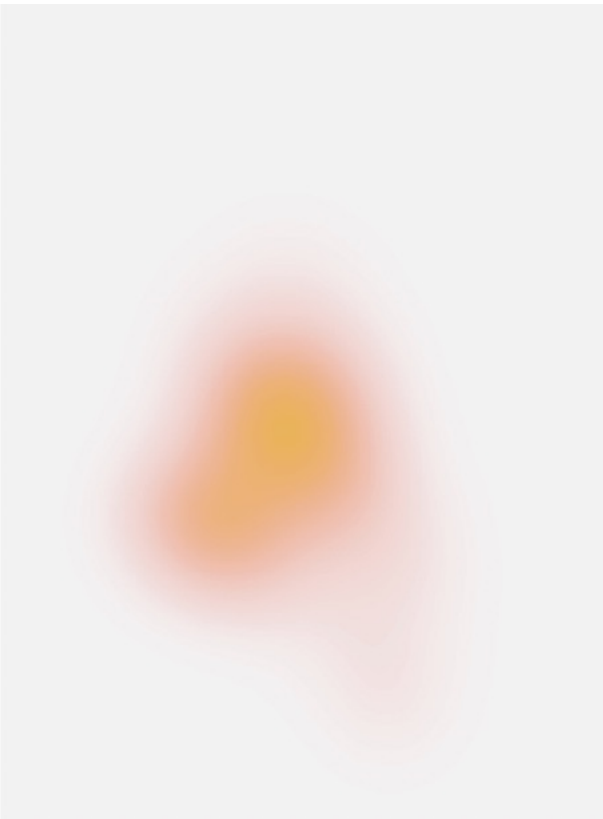


rsf.fit

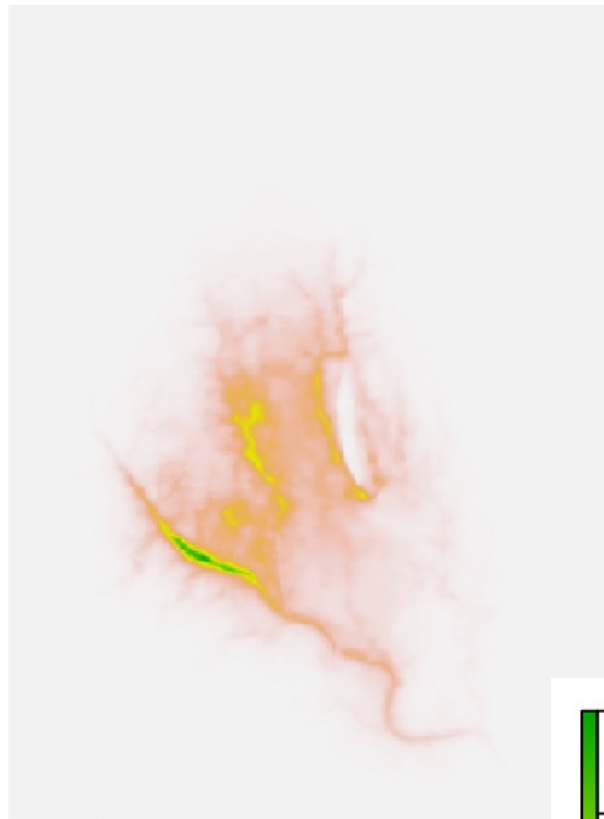




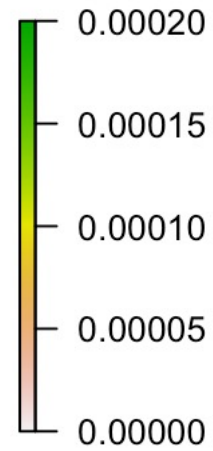
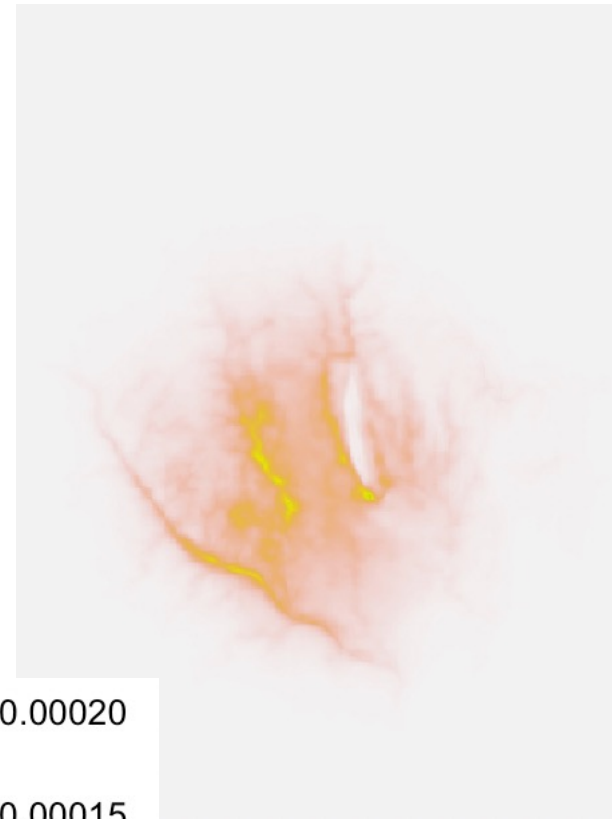
akde

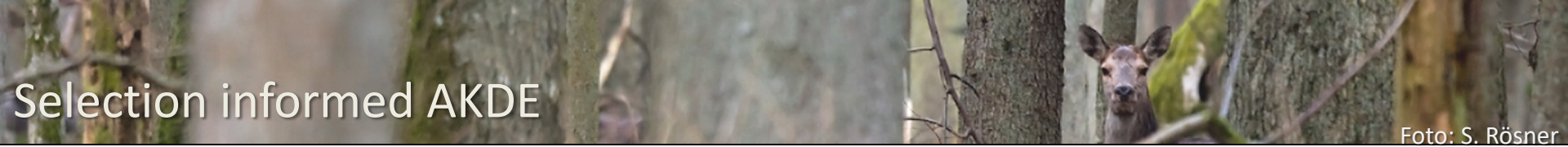


akde + rsf.fit



rsf.fit





Selection informed AKDE

Foto: S. Rösner

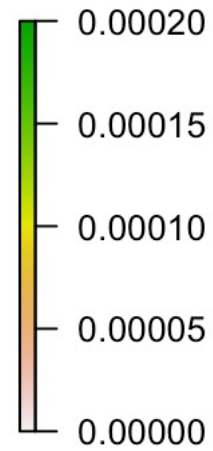
akde



akde + rsf.fit



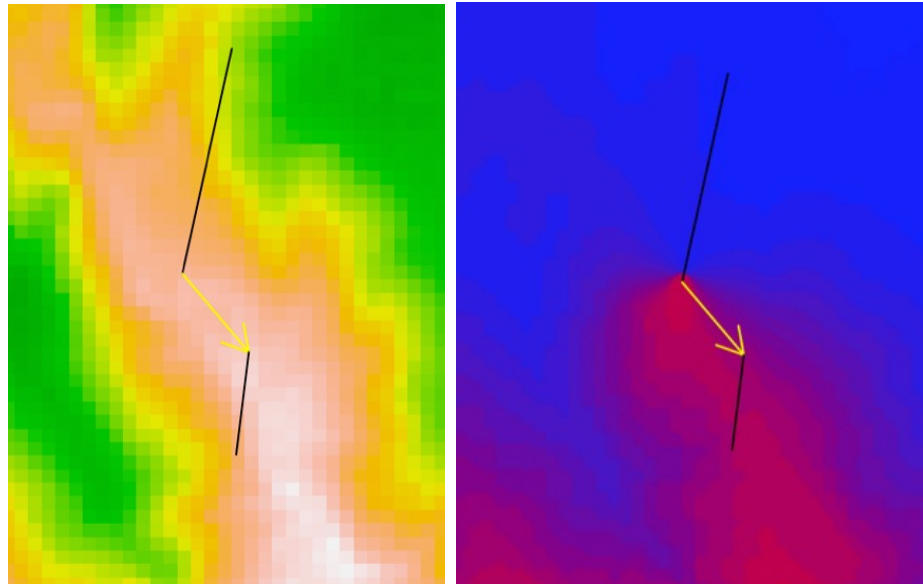
rsf.fit



Model the probability of presence in a given area

Foto: S. Rösner

SSF predicts the probability of the animal's next position based on its current position (and direction of movement) and environmental conditions

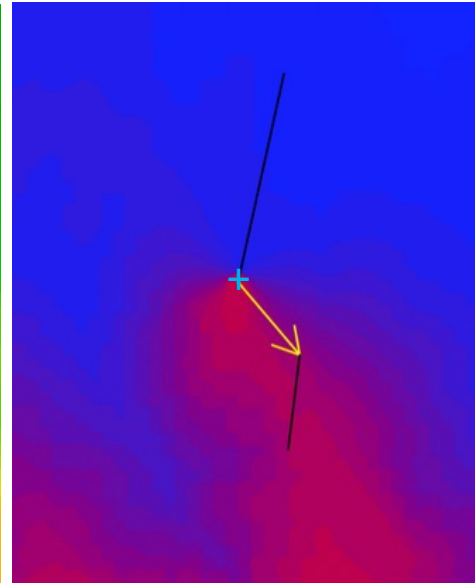
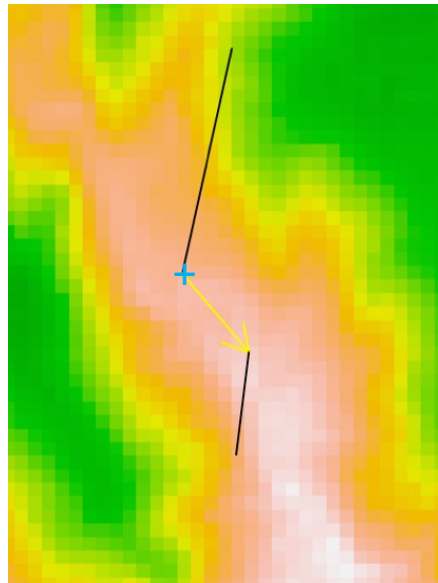


Model the probability of presence in a given area

Foto: S. Rösner

SSF predicts the probability of the animal's next position based on its current position (and direction of movement) and environmental conditions

Realised trajectory overlaid on DEM

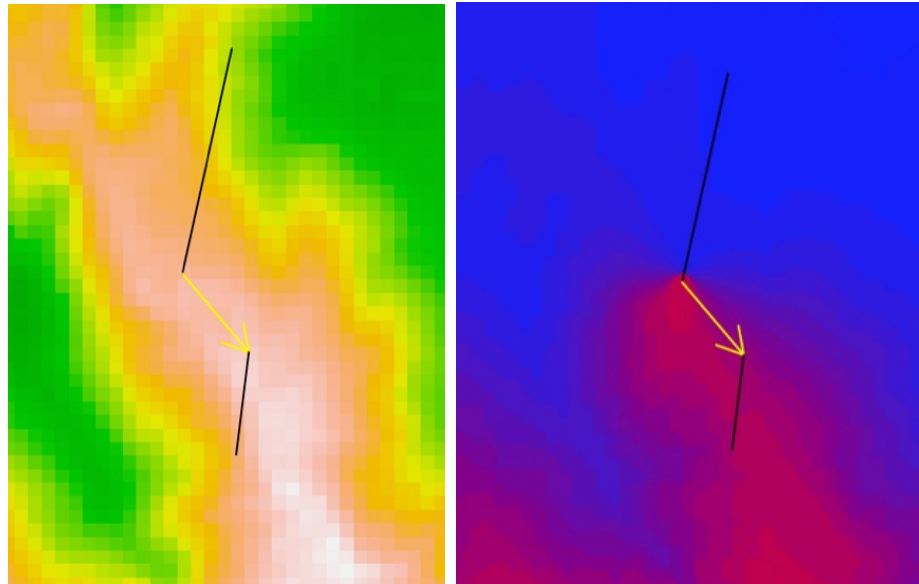


Modelled probability for next location from position « + », red indicating higher probability.
Realised step in yellow

Model the probability of presence in a given area

Foto: S. Rösner

SSF predicts the probability of the animal's next position based on its current position (and direction of movement) and environmental conditions



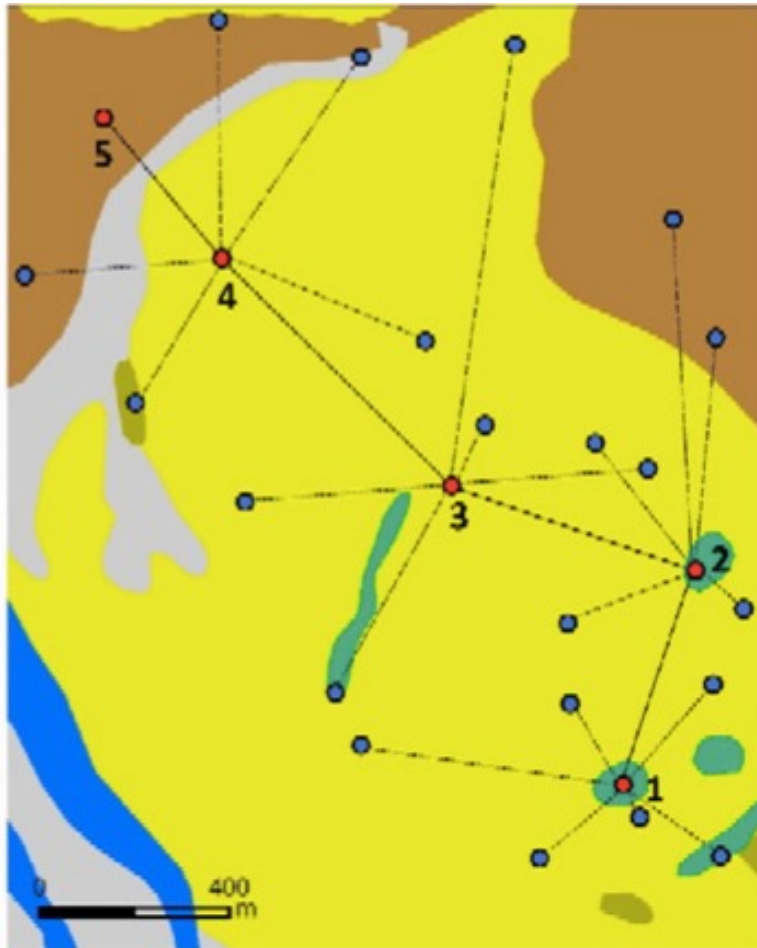
- Movement is **discrete** in time, with **a single fixed*** time step length (e.g. 4 hours)
- Successive steps are stochastically **independent** (potentially accounting for direction of last step) (no autocorrelation in velocities)

* But see Hofman et al. 2024 Movement ecology and Eisaguirre et al. 2024 Movement ecology

Step selection function

Foto: S. Rösner

$$p(s_{t+\Delta t}) = \frac{w(h(s_{t+\Delta t}))\phi(s_{t+\Delta t}, s_t, s_{t-\Delta t}, h(s_t))}{\int_{Area} w(h(S))\phi(S, s_t, s_{t-\Delta t}, h(s_t))dS}$$



$s_{t+\Delta t}$: next position

s_t : current position

$s_{t-\Delta t}$: previous position

$w()$: selection kernel

$h(s)$: spatial predictors

$\phi()$: movement kernel

Ω : area reachable from current position

There are different ways to calculate the integration constant.

Typically, we calculate the sum over a number K of randomly selected alternative step positions (that follow a distribution of the selection-free movement kernel).

Higher K reduce numerical error in the integration, but is more costly in terms of memory and computing. I recommend $K \geq 200$.

Signer et al. 2024 MEE

Step selection function

Foto: S. Rösner

$$p(s_{t+\Delta t}) = \frac{w(h(s_{t+\Delta t}))\phi(s_{t+\Delta t}, s_t, s_{t-\Delta t}, h(s_t))}{\int_{Area} w(h(S))\phi(S, s_t, s_{t-\Delta t}, h(s_t))dS}$$

The two parts, i.e. movement kernel ϕ and environmental selection kernel w are modelled as linear combinations of predictors on a log scale.

For example:

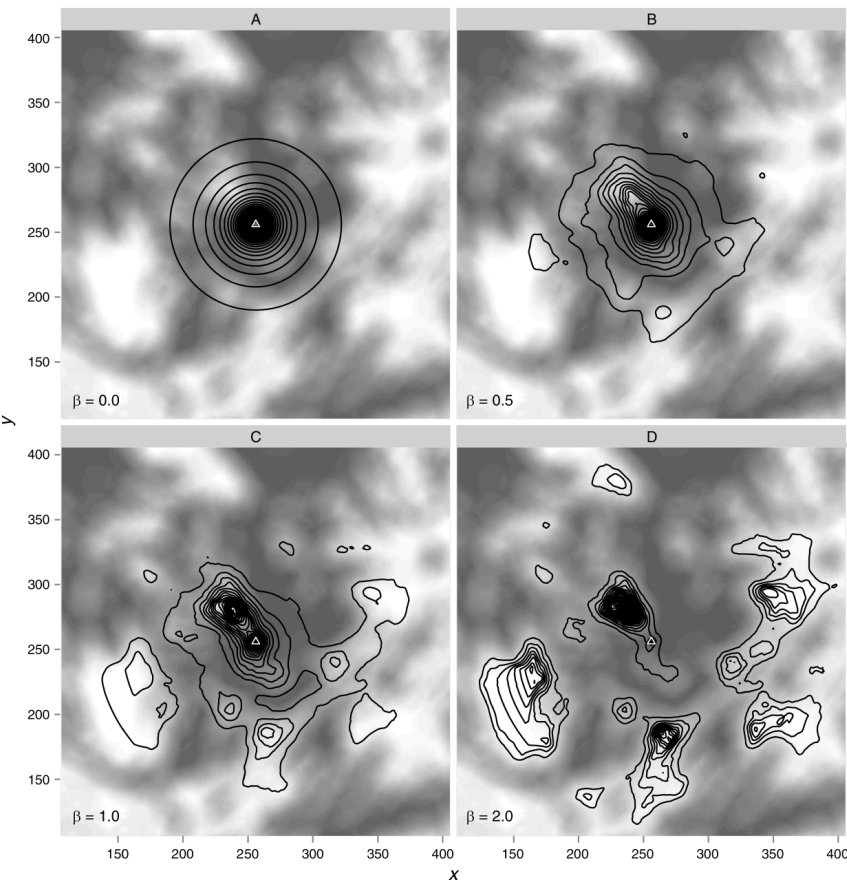
$$\phi = \exp(\beta_0 * \text{step_length}(s(t+\Delta t), s(t)) + \beta_1 * \log(\text{step_length}(s(t+\Delta t), s(t))) + \beta_2 * \cos(\text{turning_angle}(s(t+\Delta t), s(t), s(t-\Delta t))))$$

$$w = \exp(\sum \beta_i * h_i(s(t+\Delta t)))$$



Foto: S. Rösner

$$p(s_{t+\Delta t}) = \frac{w(h(s_{t+\Delta t}))\phi(s_{t+\Delta t}, s_t, s_{t-\Delta t}, h(s_t))}{\int_{Area} w(h(S))\phi(S, s_t, s_{t-\Delta t}, h(s_t))dS}$$



Typically, two parts are considered in modelling the probability p to move from $s(t)$ to $s(t+\Delta t)$ given direction of movement

w : Environmental selection

ϕ : Movement kernel (without environmental selection)

Potts et al. 2014 Ecology & Evolution

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

$$p(s_{t+\Delta t}) = \frac{w(h(s_{t+\Delta t}))\phi(s_{t+\Delta t}, s_t, s_{t-\Delta t}, h(s_t))}{\int_{Area} w(h(S))\phi(S, s_t, s_{t-\Delta t}, h(s_t))dS}$$

Differences to RSF

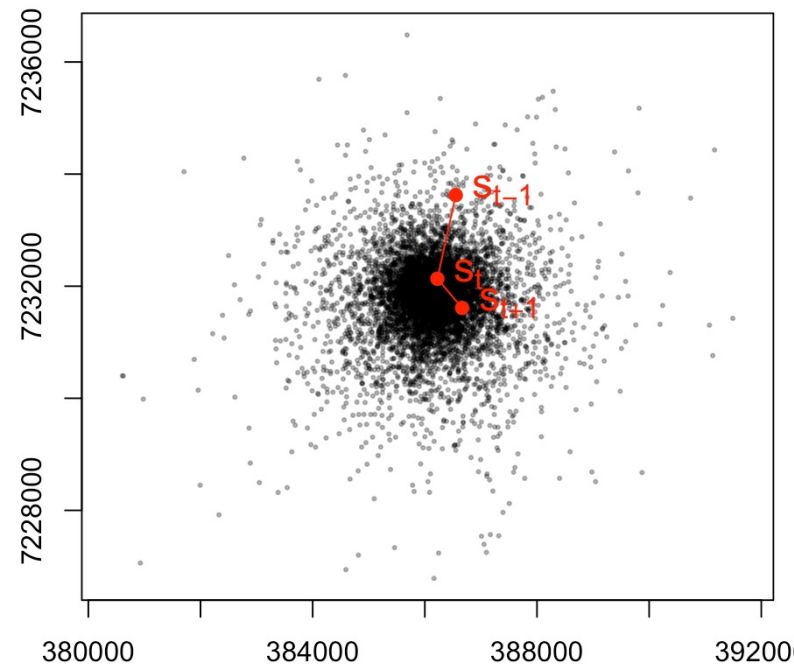
The integration constant is different for each observed step.

For a given step, we have to integrate over a relatively small area around the animal's current position.

Typical approach: stochastic integration

Approximate the integral with a (weighted) sum of values calculated at randomly chosen locations

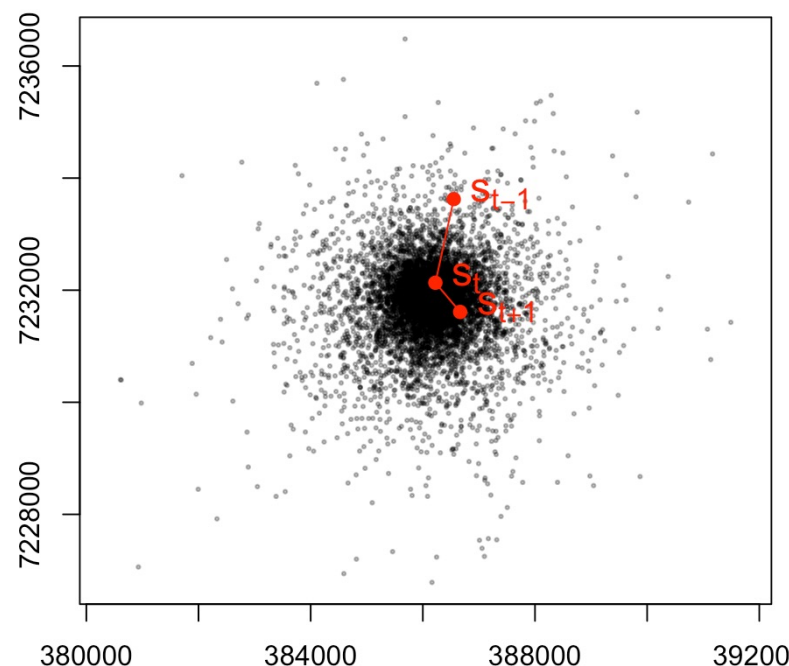
- “alternative steps”, “background points”, “pseudo-absences”, “quadrature points”
- *Importance sampling*: Sample more locations in “interesting” areas



$$p(s_{t+\Delta t}) = \frac{w(h(s_{t+\Delta t}))\phi(s_{t+\Delta t}, s_t, s_{t-\Delta t}, h(s_t))}{\int_{Area} w(h(S))\phi(S, s_t, s_{t-\Delta t}, h(s_t))dS}$$

Quadrature points S_j are drawn from a proposal distribution g (e.g. gamma distribution for step lengths and von Mises distribution for turn angles, with distribution parameters estimated from the observed movement track)

$$S_j \sim g$$



$$p(s_{t+\Delta t}) = \frac{w(h(s_{t+\Delta t}))\phi(s_{t+\Delta t}, s_t, s_{t-\Delta t}, h(s_t))/g(s_{t+\Delta t})}{\sum_{j=1}^Q w(h(S_j))\phi(S_j, s_t, s_{t-\Delta t}, h(s_t)) / g(S_j)}$$

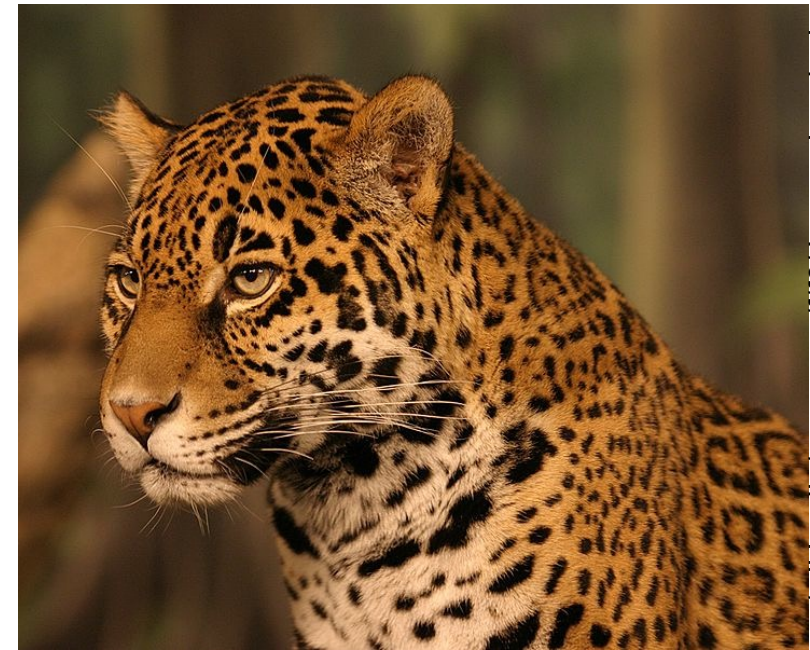
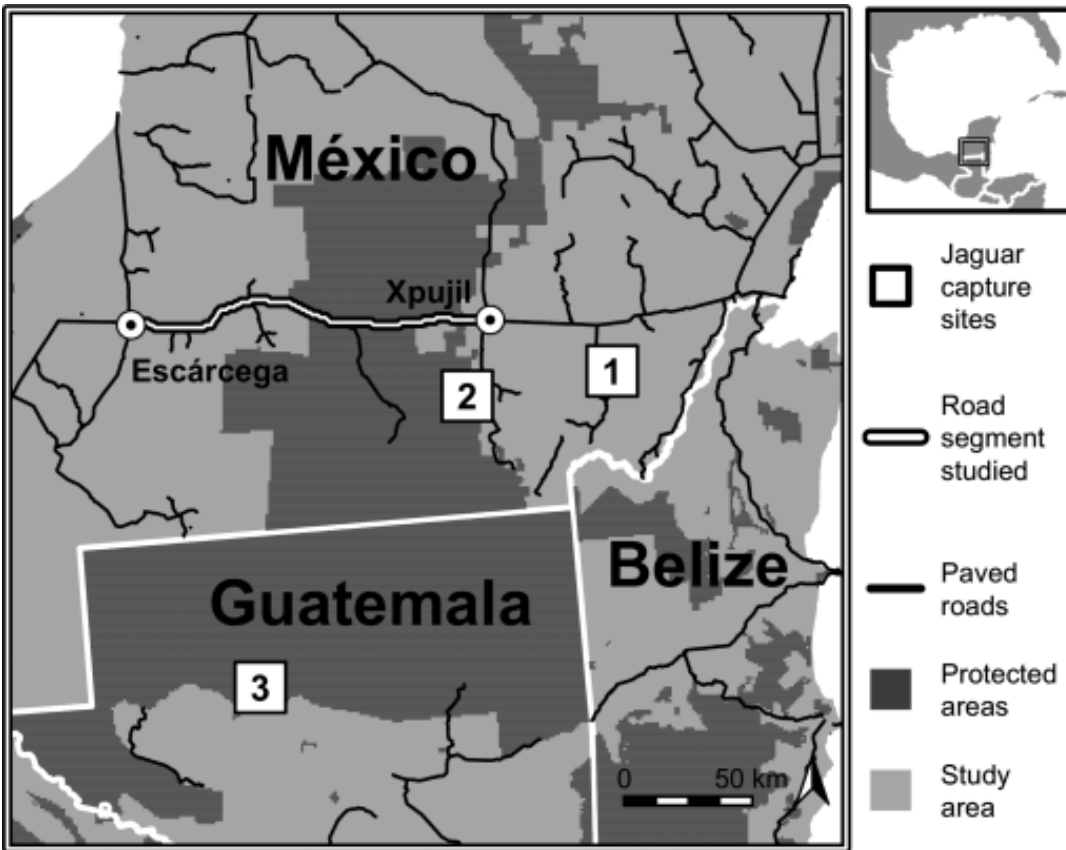
When using importance sampling in SSF (the default in `amt`), the parameter estimates for the movement kernel (e.g. step length, $\log(\text{step length})$, $\cos(\text{turn angle})$) represent differences to the parameter values of the proposal distribution.

Parameter values for the step length and turn angle distributions have to be calculated from the fitted `ssf` model and the proposal distribution.

See `?amt::update_gamma` or `?amt::update_vonmises`

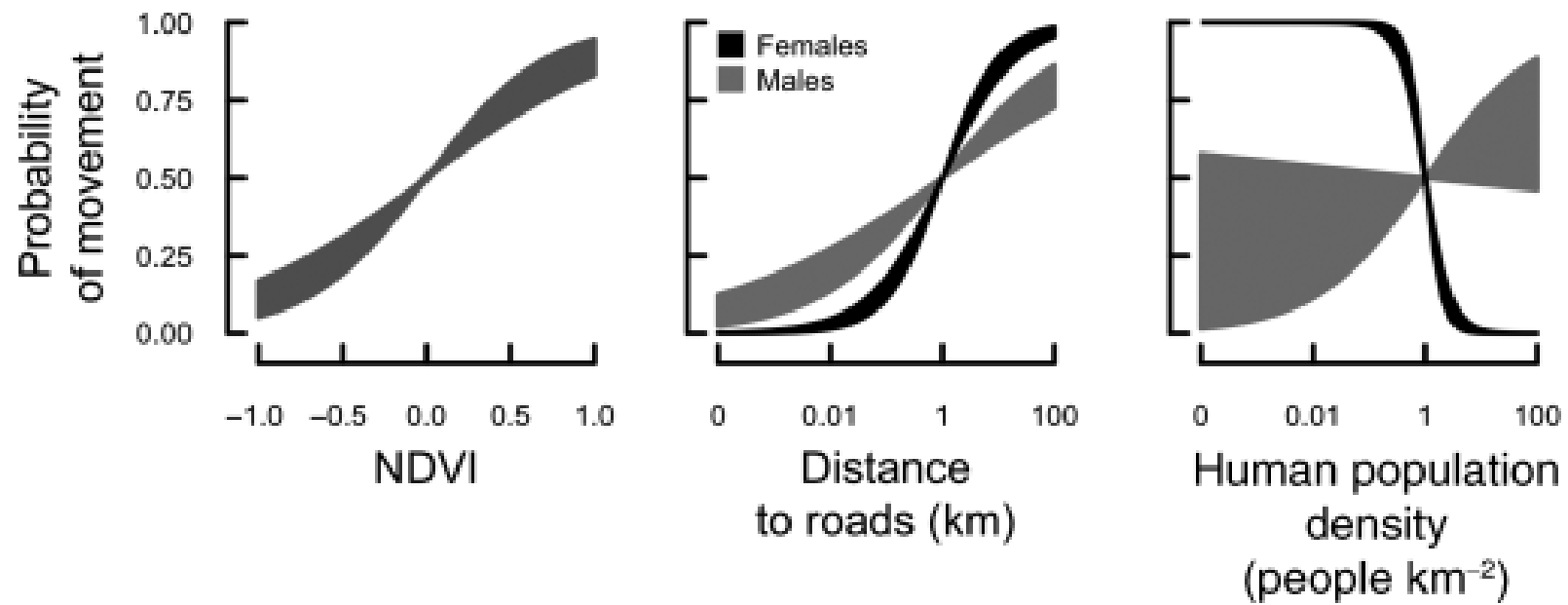
Modeling movement to mitigate fragmentation from road expansion in the Mayan Forest

Foto: S. Rösner



Habitat preferences

Foto: S. Rösner

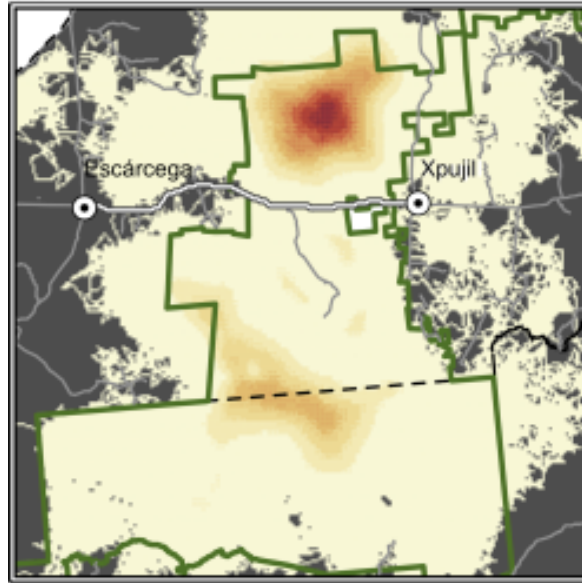


Simulated movement paths

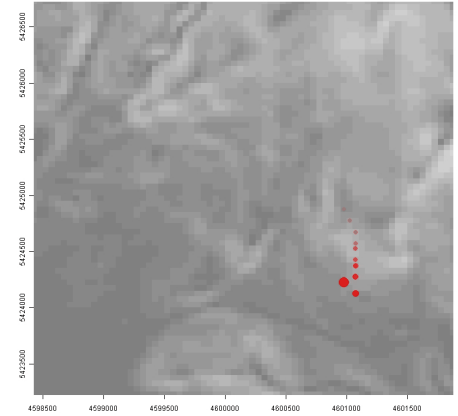
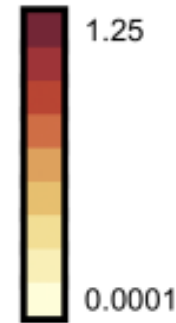


Female

(a)

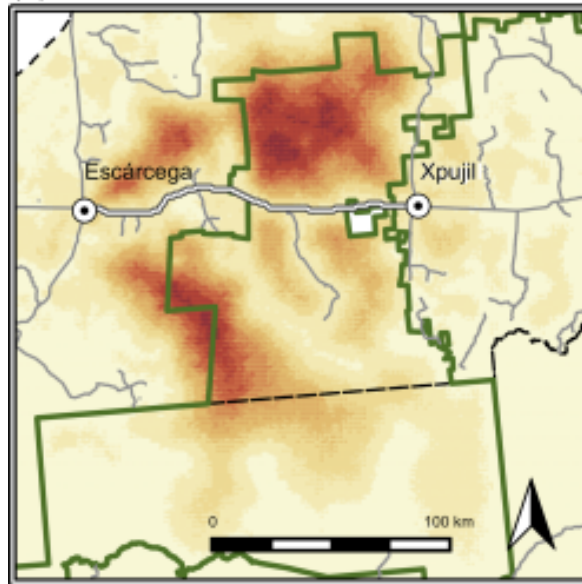


Average
cell use
per jaguar



Male

(b)

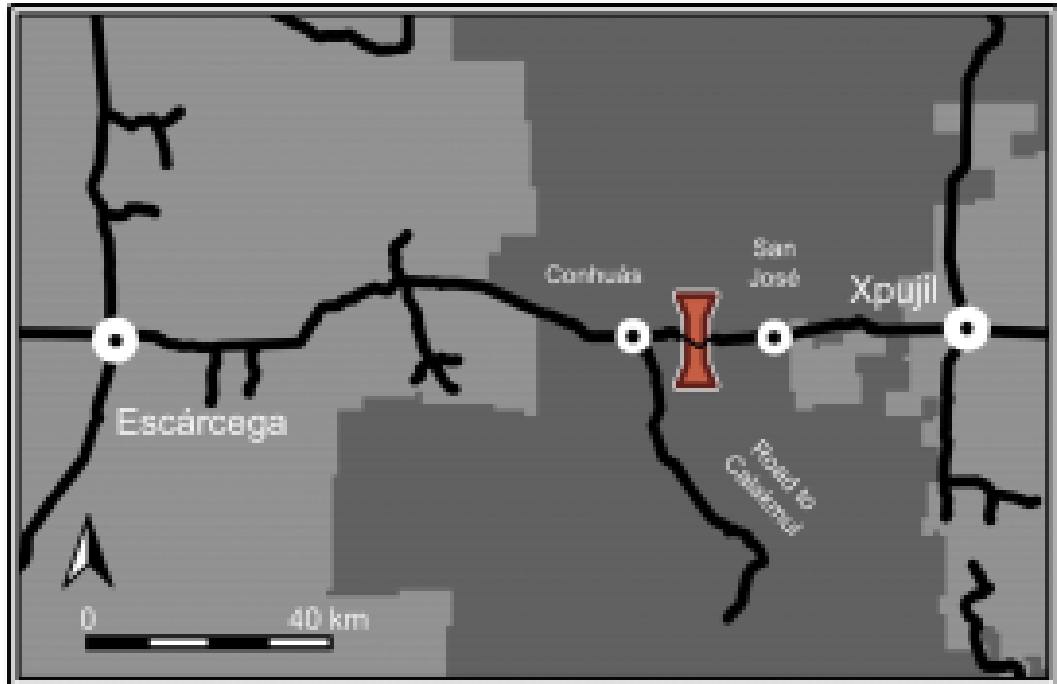
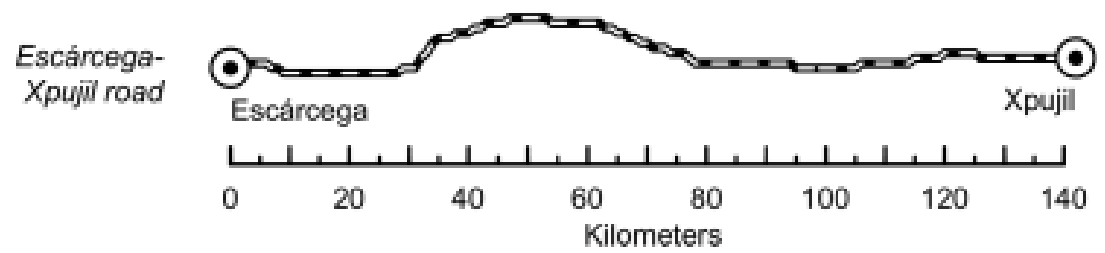
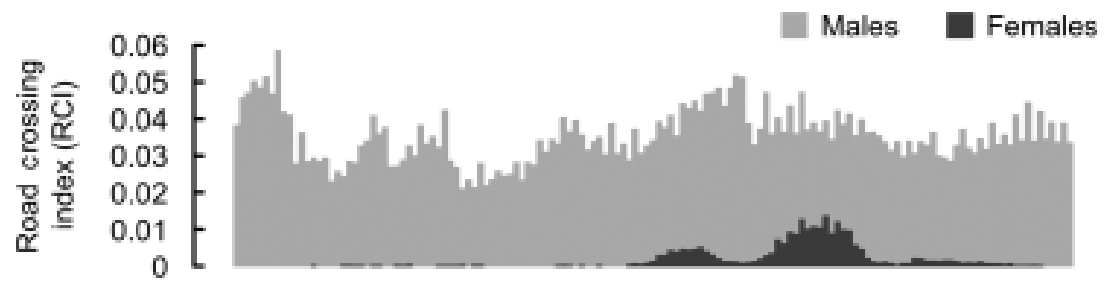


- Paved Roads
- Protected areas
- - Country boundaries

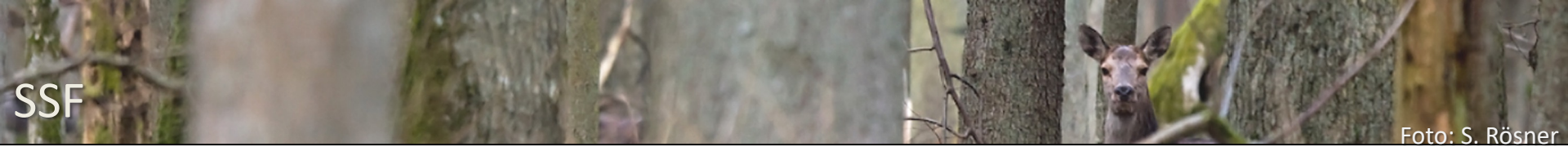


Foto: S. Rösner

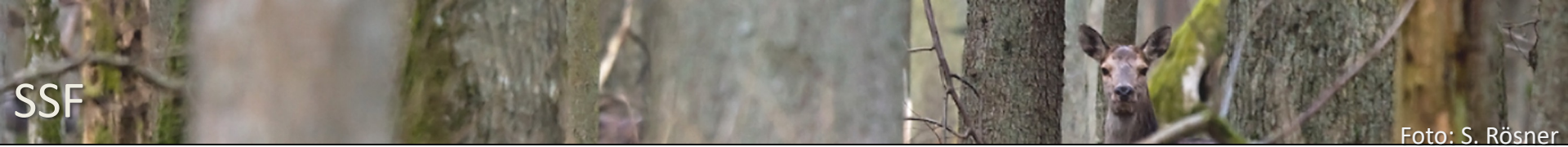
Location of Jaguar crossing suggested by simulations



- Jaguar crossing
- Protected areas
- Paved roads
- Towns



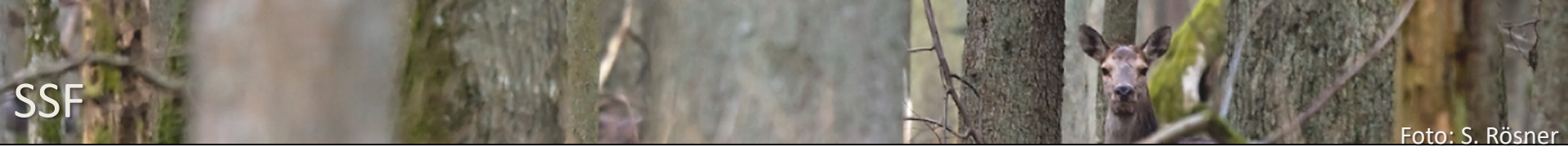
- Estimates effect of environment
- Generative model
 - Simulate paths and range distribution for current and future environmental conditions
- Standard software (conditional logistic regression; Poisson GLMM)
- Single fixed timescale
- “Beaming” of individuals



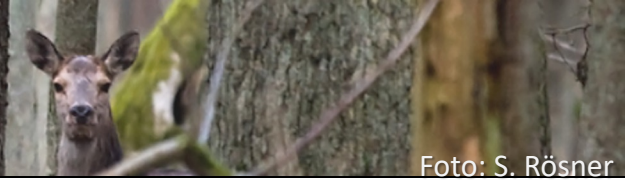
SSF

Foto: S. Rösner

- Estimates effect of environment
- Generative model
 - Simulate paths and range distribution for current and future environmental conditions
- Standard software (conditional logistic regression; Poisson GLMM)
- Single fixed timescale
- “Beaming” of individuals
- Not free in choosing timescale
 - Lower limit (uncorrelated velocities)
 - Scale dependence of parameters
- No hidden states
- Range distributions in general costly to calculate



- Estimates effect of environment
- Generative model
 - Simulate paths and range distribution for current and future environmental conditions
- Standard software (conditional logistic regression; Poisson GLMM)
- Accommodates much ecological realism
- Single fixed timescale
- “Beaming” of individuals
- Not free in choosing timescale
 - Lower limit (uncorrelated velocities)
 - Scale dependence of parameters
- No hidden states
- Range distributions in general costly to calculate

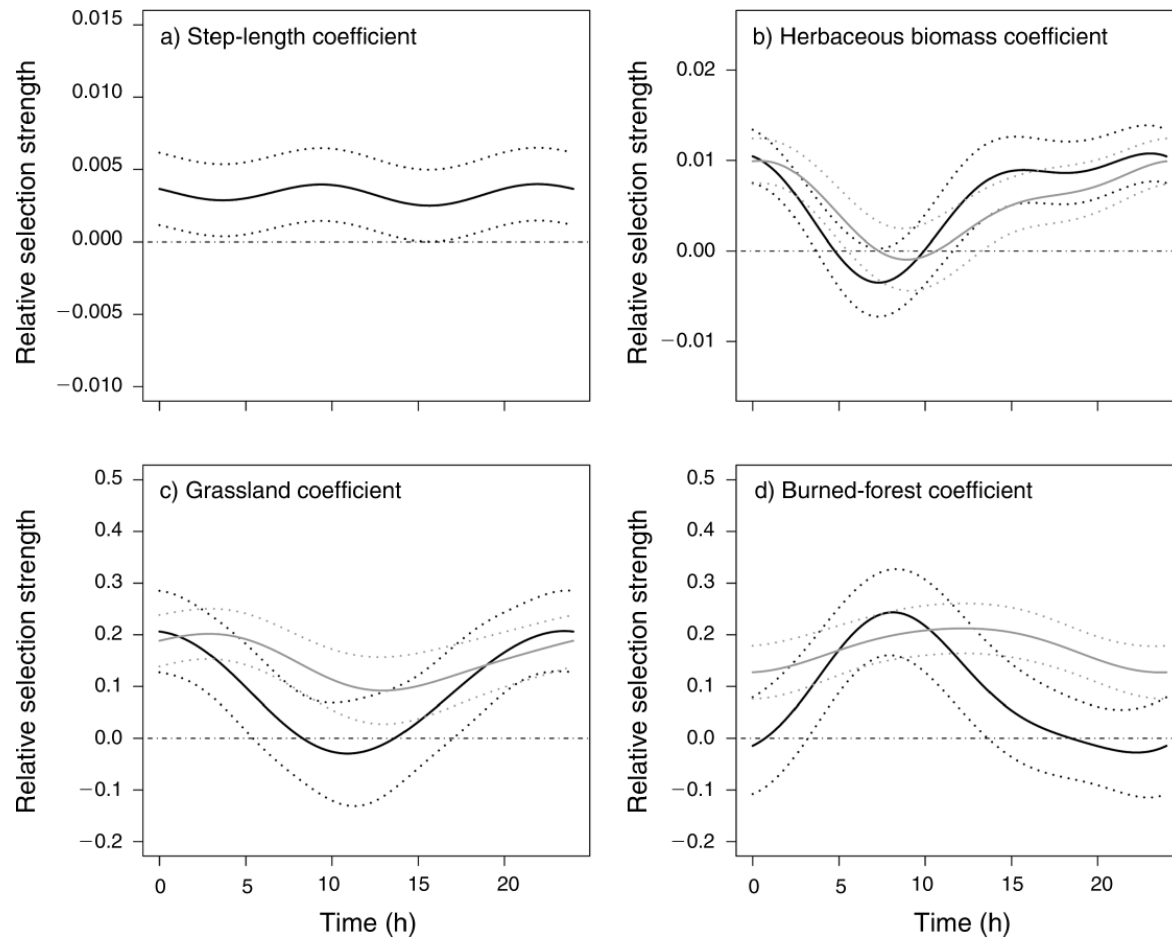


- Temporal variation in selection
- Home range
- Barriers
- Migration
- Slowing speed in the darkness vs. actively moving towards darkness
- Movement influenced by earlier space use of other individuals

Temporal variation in selection



Foto: S. Rösner

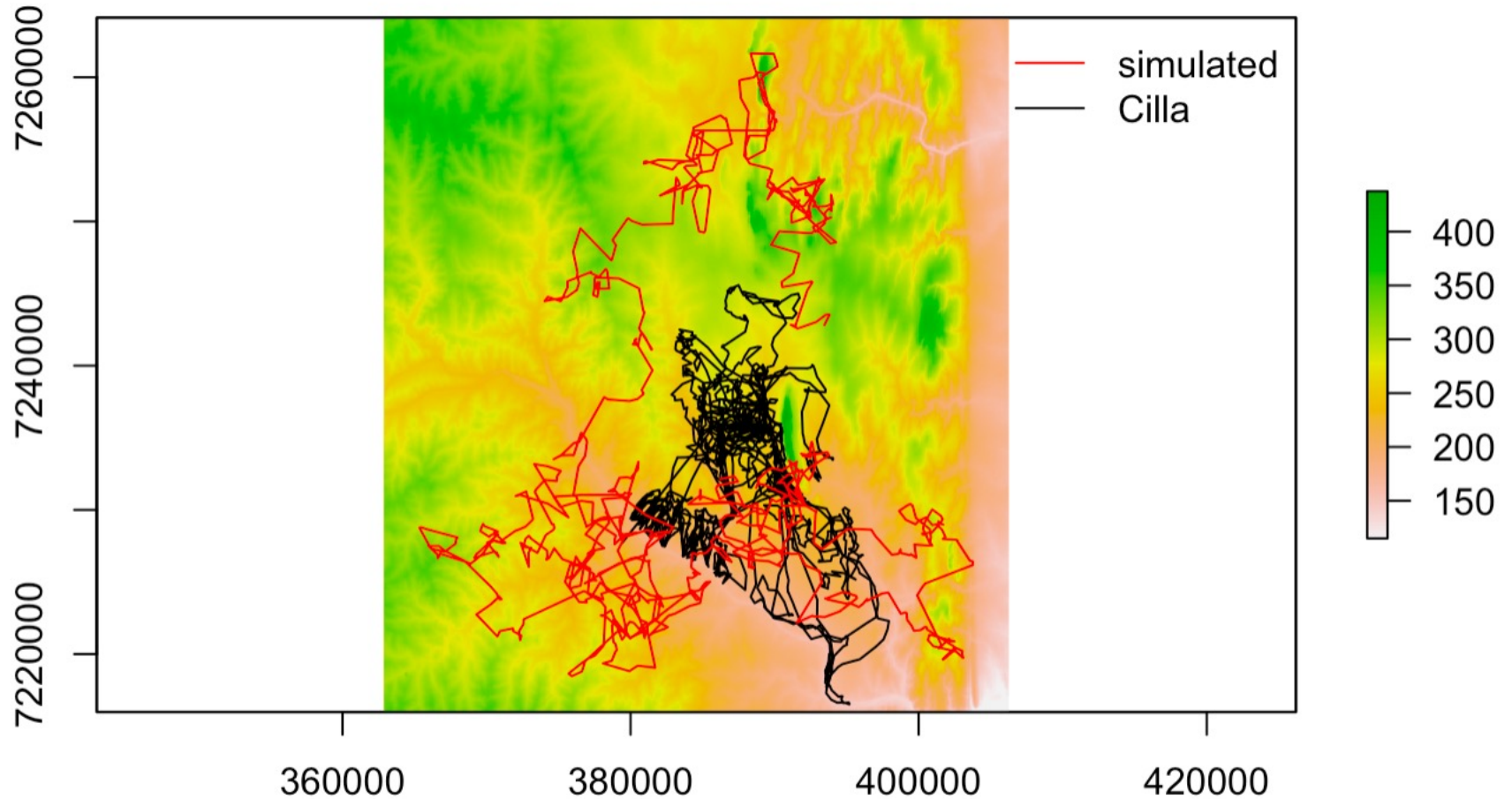


16 cow elk in
Yellowstone during
summer months, 5h
intervals

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

Distance to water as environmental predictor

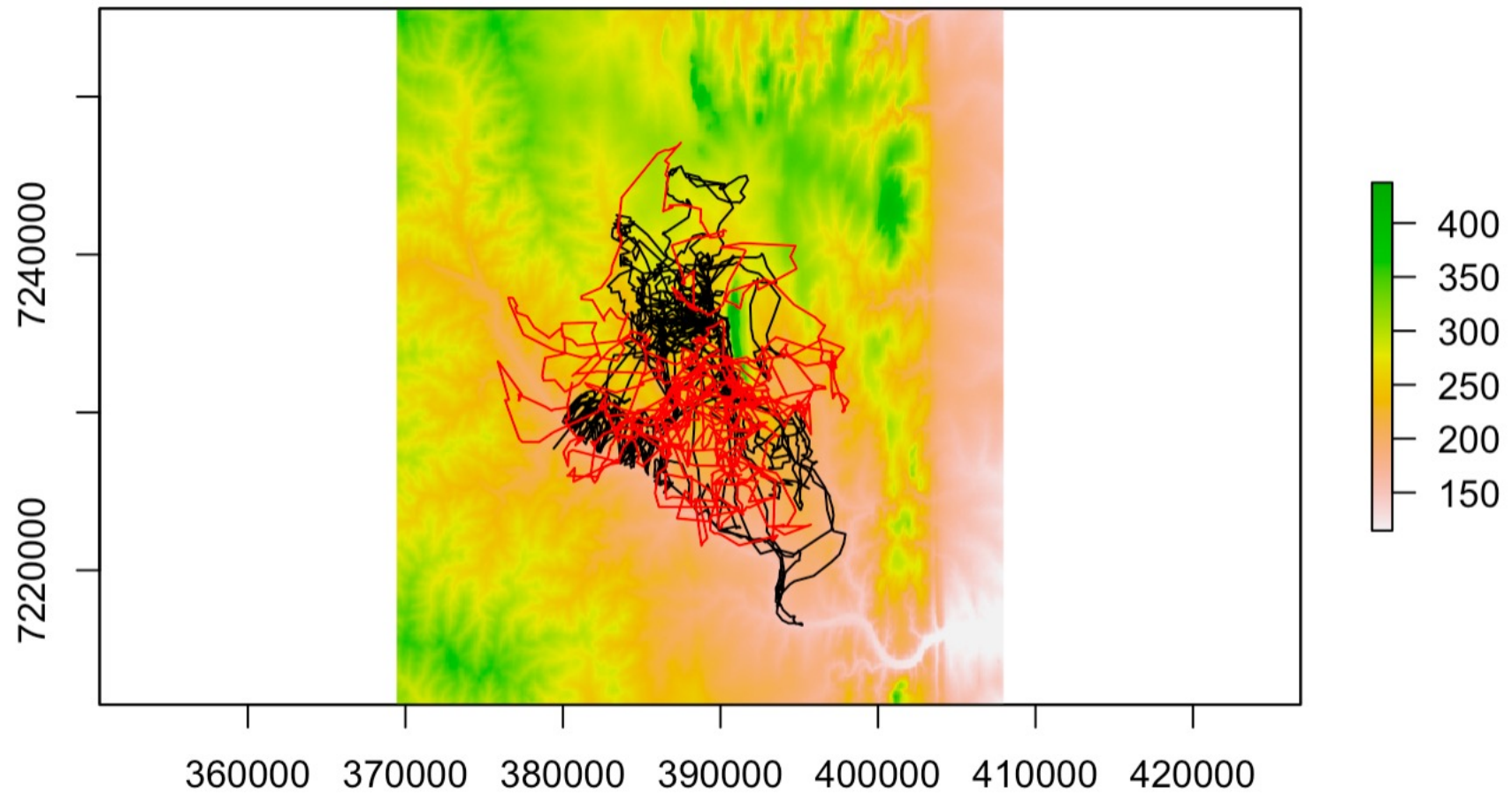
Foto: S. Rösner

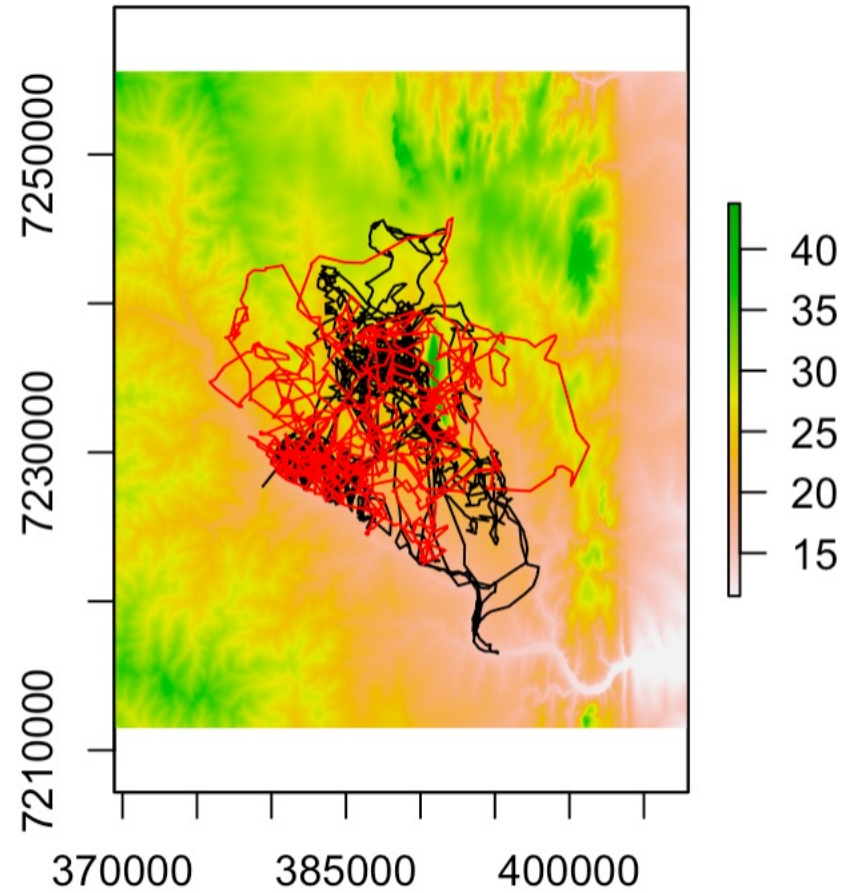
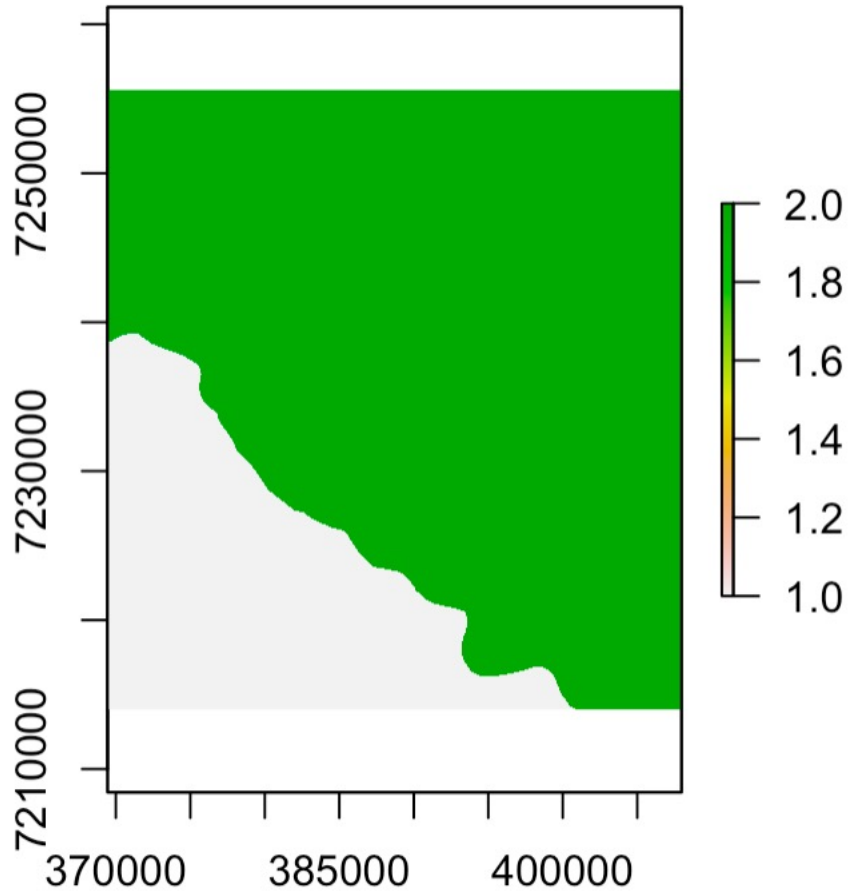




Distance to water and home range

Foto: S. Rösner





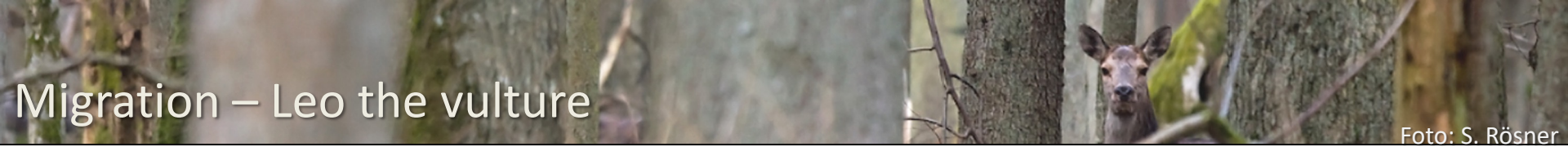
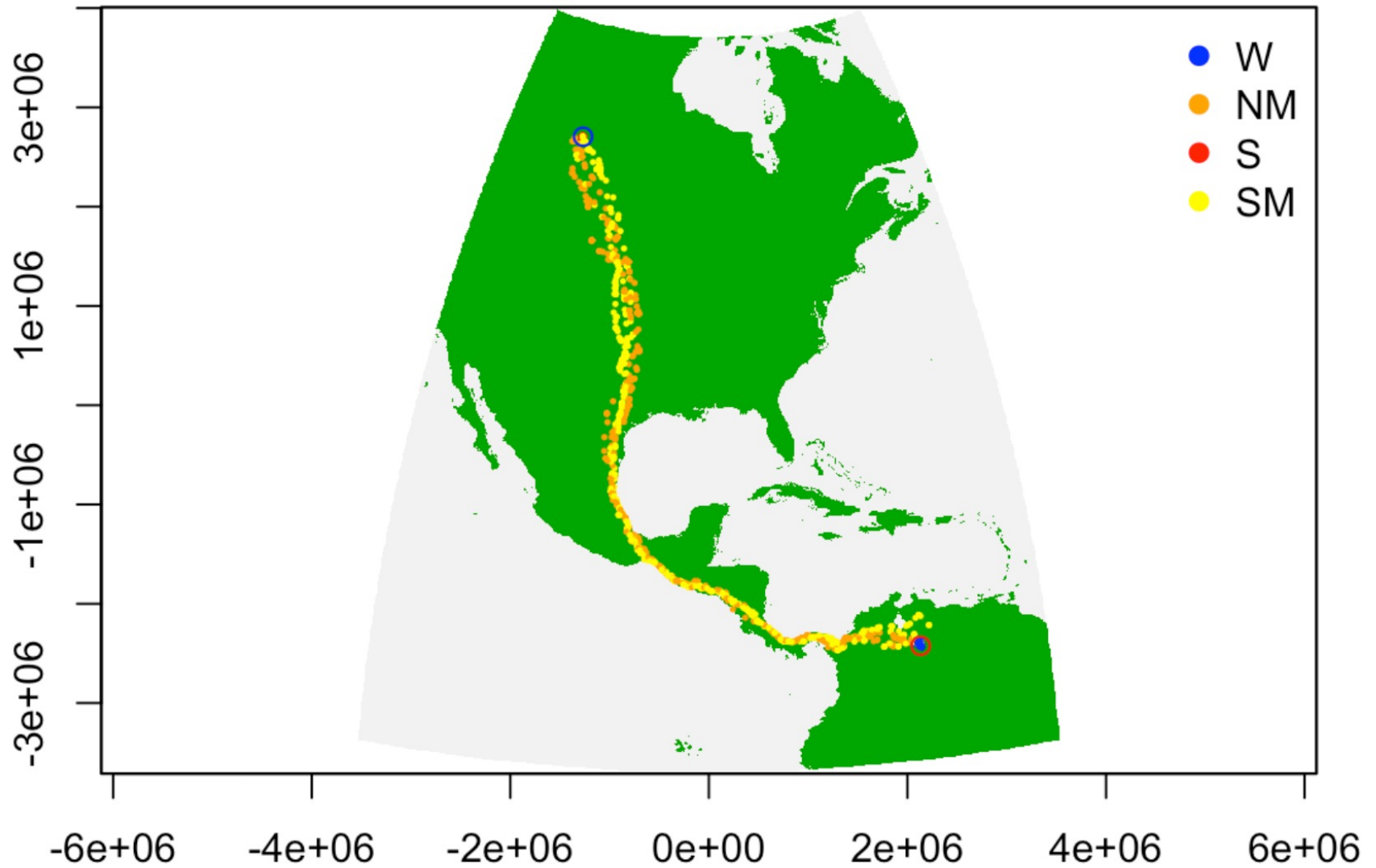


Foto: S. Rösner

Migration – Leo the vulture



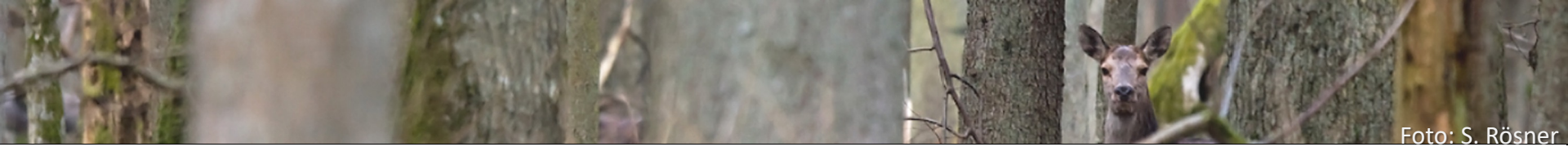
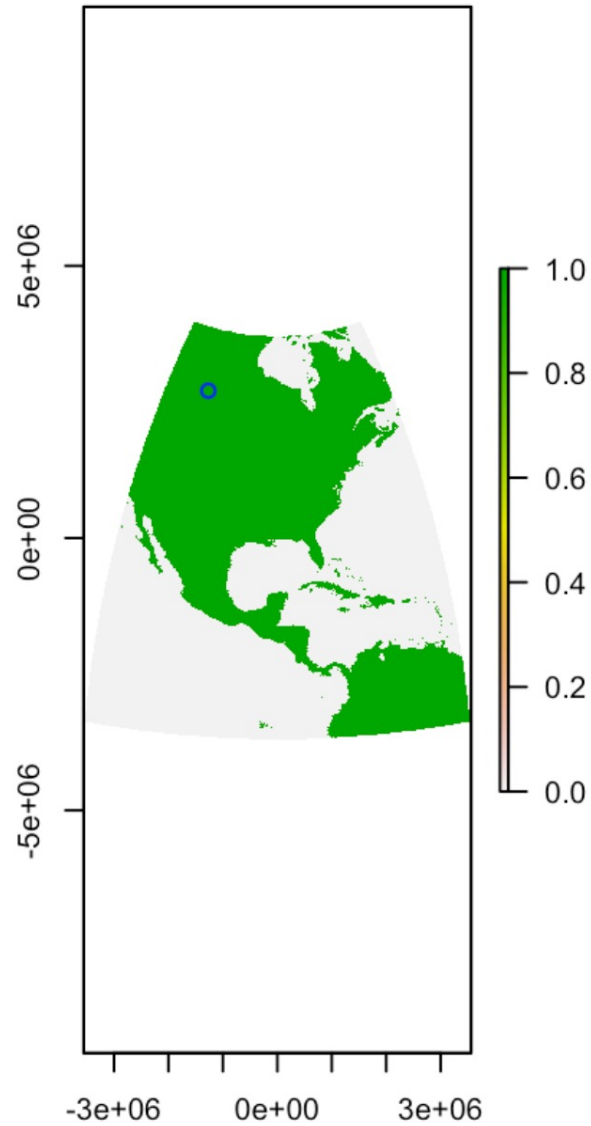
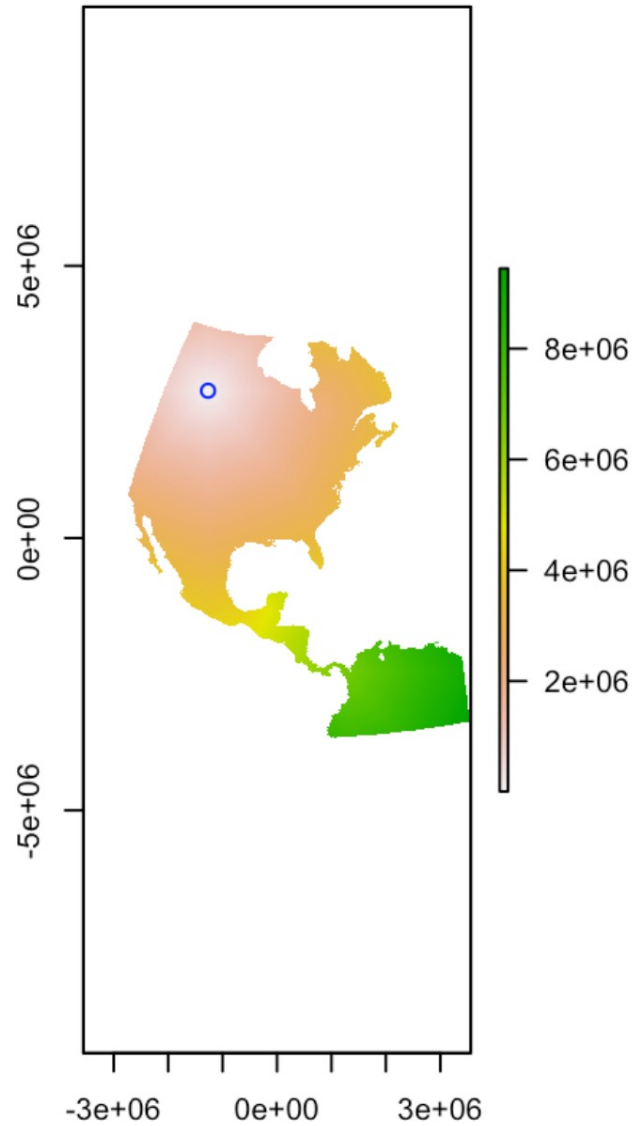


Foto: S. Rösner

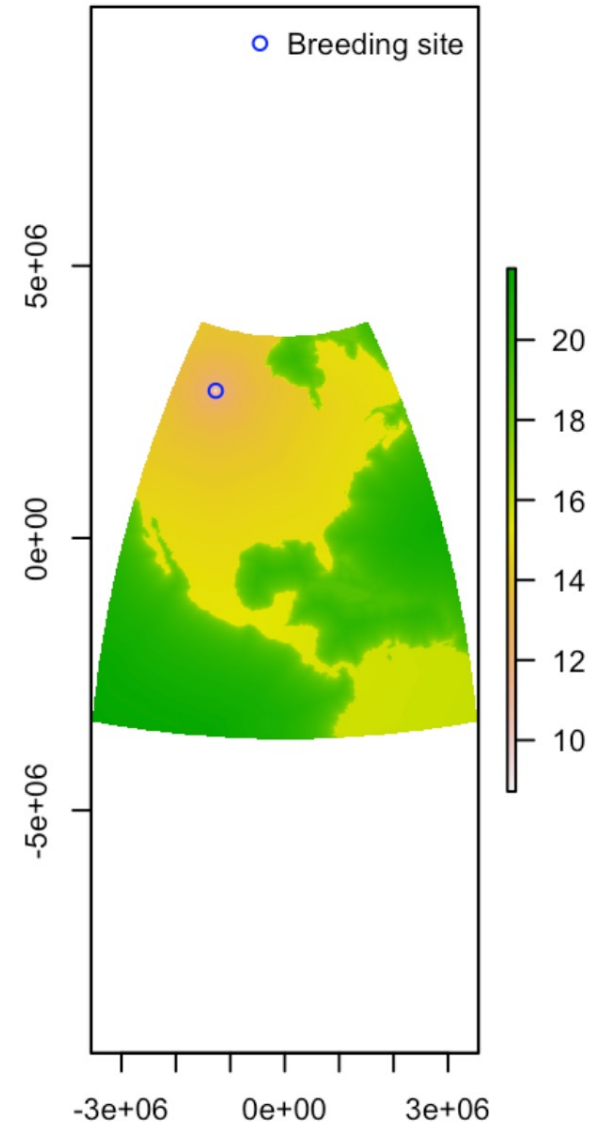
Land



Distance to breeding site

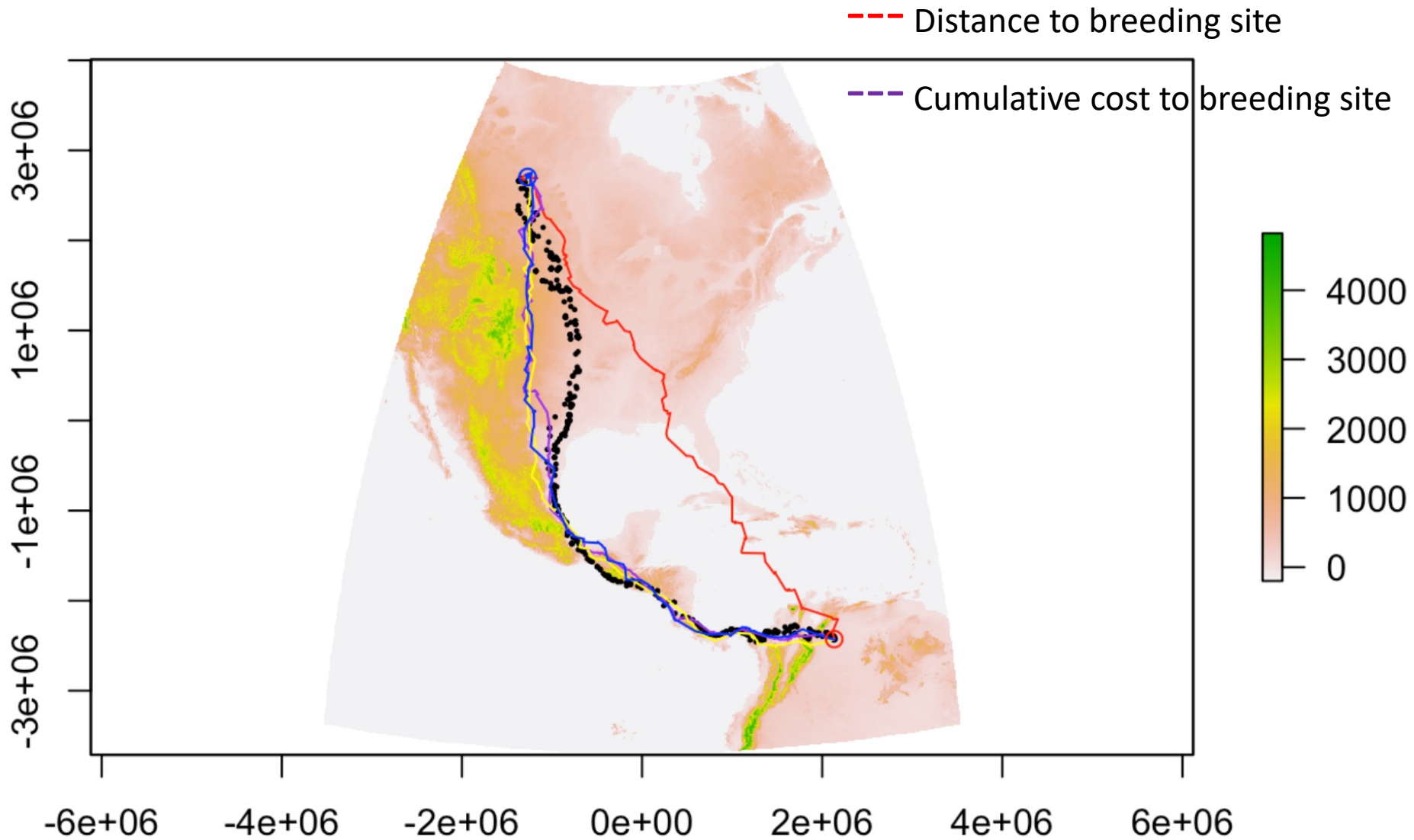


log(cost to breeding site)



Simulated Leo trajectories for northward migration

Foto: S. Rösner



- Functional response (amount of habitat/resource)
- Population density

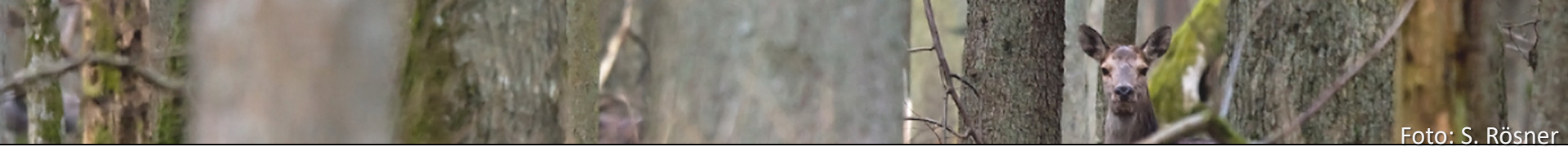


Foto: S. Rösner

- Estimates effect of environment + properties of the individual
- Generative model
 - Simulate paths and range distribution for current and future environmental conditions
- Standard software (conditional logistic regression; Poisson GLMM)
- Accommodates much ecological realism
- **Multiple animals**
- Single fixed timescale
- “Beaming” of individuals
- Not free in choosing timescale
 - Lower limit (uncorrelated velocities)
 - Scale dependence of parameters
- No hidden states



Foto: S. Rösner

Recommended literature for methods of RSF/SSF

Methods in Ecology and Evolution

Methods in Ecology and Evolution 2015, 6, 366–379 doi: 10.1111

SPECIAL FEATURE – REVIEW NEW OPPORTUNITIES AT THE INTERFACE BETWEEN ECOLOGY AND STATISTICS Point process models for presence-only analysis

Ian W. Renner^{1*}, Jane Elith², Adrian Baddeley³, William Fithian⁴, Trevor Hastie⁴, Steven J. Phillips⁵, Gordana Popovic⁶ and David I. Warton⁶

Received: 17 May 2022 | Accepted: 22 September 2022

DOI: 10.1111/2041-210X.14025

RESEARCH ARTICLE

Methods in Ecology and Evolution

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

Jesse M. Alston^{1,2,3} | Christen H. Fleming^{4,5} | Roland Kays^{6,7} | Jarryd P. St. Colleen T. Downs⁸ | Tharmalingam Ramesh^{8,9} | Björn Reineking¹⁰ | Justin M. Calabrese^{1,2,11}

Methods in Ecology and Evolution

Methods in Ecology and Evolution 2012, 3, 177–187 doi: 10.1111

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

Geert Aarts^{1,2*}, John Fieberg³ and Jason Matthiopoulos^{4,5}

Methods in Ecology and Evolution

Methods in Ecology and Evolution 2016, 7, 619–630

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

Tal Avgar^{1*}, Jonathan R. Potts², Mark A. Lewis^{1,3} and Mark S. Boyce¹

¹Department of Biological Sciences, University of Alberta, Edmonton, AB T6G 2G1, Canada; ²Statistical Modelling Unit, University of Sheffield, Sheffield S2 7PU, UK; and ³Department of Mathematics, University of Alberta, Edmonton, AB T6G 2G1, Canada

Ecology, 90(12), 2009, pp. 3554–3565
© 2009 by the Ecological Society of America

Accounting for animal movement in resource selection functions: sampling considerations

JAMES D. FORESTER,^{1,3} HAE KYUNG IM,¹ AND TAL AVGAR^{1*}

Received: 25 November 2020 | Accepted: 2 February 2021

DOI: 10.1111/1365-2656.13441

HOW TO...

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

John Fieberg¹ | Johannes Signer² | Brian Smith³ | Tal Avgar³

Journal of Animal Ecology

- Buffalo in Kruger
 - RSF (ctmm)
 - SSF (amt)

We will focus on one single animal (Cilla).

There is code on how to model multiple animals



Johannes
Signer