

# Methods for modeling variability among-animals in habitat-selection studies

Received: 14 September 2018 | Accepted: 22 July 2019

DOI: 10.1111/1365-2656.13087

**BIOLOGGING**

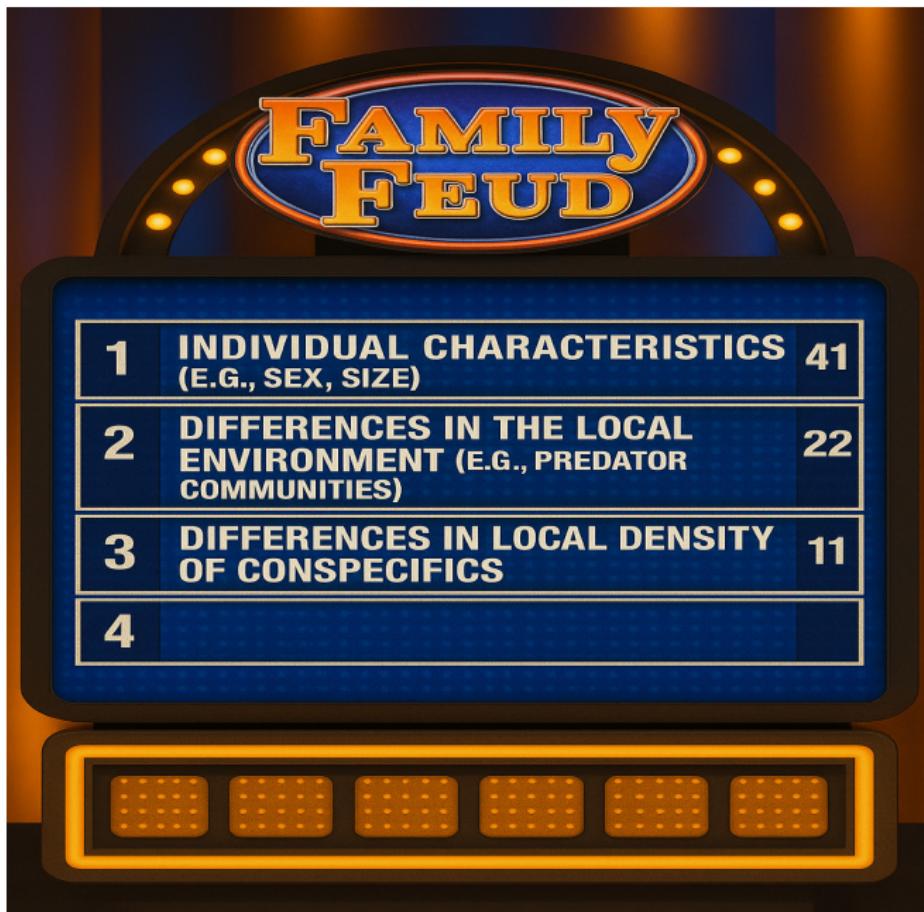
Journal of Animal Ecology  
 NATURE  
ECOLOGICAL  
SOCIETY

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

Stefanie Muff<sup>1,2</sup>  | Johannes Signer<sup>3</sup>  | John Fieberg<sup>4</sup> 

Why might animals have variable habitat selection parameters?

Why might animals have variable habitat selection parameters?



# Functional Responses

# Functional Responses

Habitat selection parameters,  $\beta$ , are informed by habitat use relative to habitat availability.

- ▶ when  $f^u(x) > f^a(x) \Rightarrow$  “preference” for habitat  $x$ .

## Functional Responses

Habitat selection parameters,  $\beta$ , are informed by habitat use relative to habitat availability.

- ▶ when  $f^u(x) > f^a(x) \Rightarrow$  “preference” for habitat  $x$ .

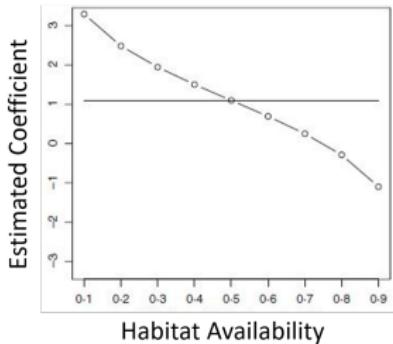
Some resources, like water, may be needed in small and relatively constant amounts

# Functional Responses

Habitat selection parameters,  $\beta$ , are informed by habitat use relative to habitat availability.

- ▶ when  $f^u(x) > f^a(x) \Rightarrow$  “preference” for habitat  $x$ .

Some resources, like water, may be needed in small and relatively constant amounts



Aarts, G., Fieberg, J., Brasseur, S., & Matthiopoulos, J. (2013). Quantifying the effect of habitat availability on species distributions. *Journal of animal ecology*, 82(6), 1135-1145.

# Generalized Functional Responses

*Ecology*, 92(3), 2011, pp. 583–589  
© 2011 by the Ecological Society of America

## Generalized functional responses for species distributions

JASON MATTHIOPoulos,<sup>1,2,7</sup> MARK HEBBLEWHITE,<sup>3</sup> GEERT AARTS,<sup>4,5</sup> AND JOHN FIEBERG<sup>6</sup>

<sup>1</sup>Scottish Oceans Institute, School of Biology, University of St. Andrews, East Sands, St. Andrews, Fife KY16 8LB Scotland, United Kingdom

<sup>2</sup>Centre for Research into Environmental and Ecological Modeling, University of St Andrews, The Observatory, Buchanan Gardens, St Andrews, Fife KY16 9LZ Scotland, United Kingdom

<sup>3</sup>Wildlife Biology Program, Department of Ecosystem and Conservation Sciences, College of Forestry and Conservation, University of Montana, Missoula, Montana 59812 USA

<sup>4</sup>Royal Netherlands Institute for Sea Research (NIOZ), P.O. Box 59, 1790 AB Den Burg, The Netherlands

<sup>5</sup>IMARES Wageningen UR, Institute for Marine Resources and Ecosystem Studies, P.O. Box 167, 1790 AD Den Burg, The Netherlands

<sup>6</sup>Biometrics Unit, Minnesota Department of Natural Resources (Minnesota DNR), 5463-C West Broadway, Forest Lake, Minnesota 55025 USA

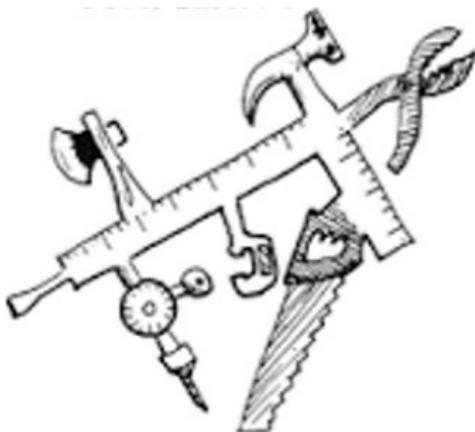
Capture variability in habitat-selection parameters (in RSFs) across multiple sampling instances using:

- ▶ moments of availability (mean, variance)
- ▶ interactions between these moments
- ▶ random coefficients

## Methods for modeling among-animal variability

## Methods for modeling among-animal variability

1. Fit models to individual animals (or sampling instances) and treat the estimates as data (two-step approach)
2. Mixed models, hierarchical models, random-effect models: allow parameters to vary by animal



# Two-Step Approach

Step 1: fit models to individuals

$$f_i^u(s) \propto \exp(x_1(s)\beta_{1i} + x_2(s)\beta_{2i} + \dots x_p(s)\beta_{pi})$$

# Two-Step Approach

Step 1: fit models to individuals

$$f_i^u(s) \propto \exp(x_1(s)\beta_{1i} + x_2(s)\beta_{2i} + \dots x_p(s)\beta_{pi})$$

Step 2: Do statistics on  $(\hat{\beta}_{1i}, \dots, \hat{\beta}_{pi})$

# Two-Step Approach

Step 1: fit models to individuals

$$f_i^u(s) \propto \exp(x_1(s)\beta_{1i} + x_2(s)\beta_{2i} + \dots x_p(s)\beta_{pi})$$

Step 2: Do statistics on  $(\hat{\beta}_{1i}, \dots, \hat{\beta}_{pi})$



Covariates:

- ▶ elevation
- ▶ population density
- ▶ forest (yes/no)

Fieberg J, Bohrer G, Davidson SC, Kays R (2018) Short course on analyzing animal tracking data. Presented at the North Carolina Museum of Natural Sciences, Raleigh, NC, USA. May 21-23, 2018.  
<https://movebankworkshopraleighnc.netlify.com/>

LaPoint, S., Gallery, P., Wikelski, M., and Kays, R. 2013. Animal behavior, cost-based corridor models, and real corridors. *Landscape Ecology*, 28, 1615-1630.

# amt: Individual-Specific Coefficients

Quick and easy using the `amt` package in conjunction with `tidyverse` in R

Fit models to individuals:

```
rsffits <- dat_rsf %>% nest(data = !c(id, sex)) %>%
  mutate(mod = map(data, function(x)
    glm(case_ ~ elevation + popden + forest,
        data = x, weight=w, family = binomial)))
```

Pull off coefficients:

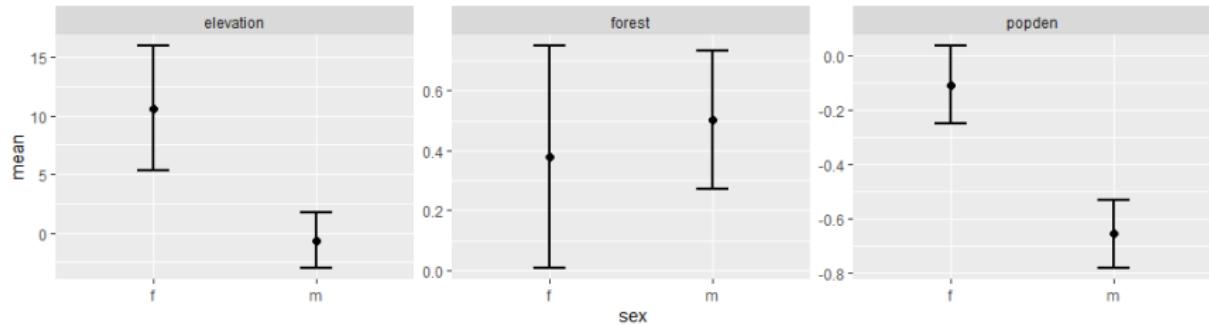
```
rsffits2 <- rsffits %>%
  dplyr::mutate(tidy = purrr::map(mod, broom::tidy),
    n = purrr::map(data, nrow) %>% simplify()) %>%
  unnest(tidy) %>% select(-c(data, mod))
```

# Data Frame: Individual-Specific Coefficients

```
> rsffits2
# A tibble: 24 x 8
  id sex term      estimate std.error statistic p.value     n
  <dbl> <chr> <chr>      <dbl>     <dbl>     <dbl>    <dbl> <dbl>
1 2   f   (Intercept) -8.36      0.489    -17.1  1.33e-65 14828
2 2   f   forest       1.02      0.0602    16.9  2.04e-64 14828
3 2   f   elevation    1.64      0.524     3.13  1.75e- 3 14828
4 2   f   popden      0.142     0.0772    1.83  6.73e- 2 14828
5 4   f   (Intercept)  1.79      0.950     1.88  5.95e- 2 11132
6 4   f   forest       0.160     0.0667    2.39  1.67e- 2 11132
7 4   f   elevation    13.0     1.12      11.7  2.25e-31 11132
8 4   f   popden     -0.260     0.0723    -3.60  3.22e- 4 11132
9 5   f   (Intercept)  6.69      2.85      2.35  1.90e- 2 5599
10 5   f   forest      -0.0470   0.114     -0.412 6.80e- 1 5599
# ... with 14 more rows
> |
```

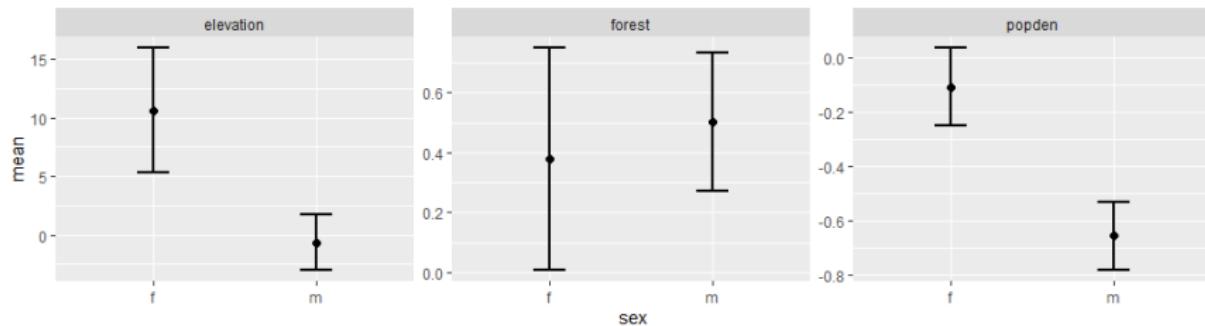
Signer, J., J. Fieberg, and T. Avgar (2019). Animal movement tools (amt): R package for managing tracking data and conducting habitat selection analyses. Ecology and Evolution 9:880-890.

## Step 2: Statistics on Statistics



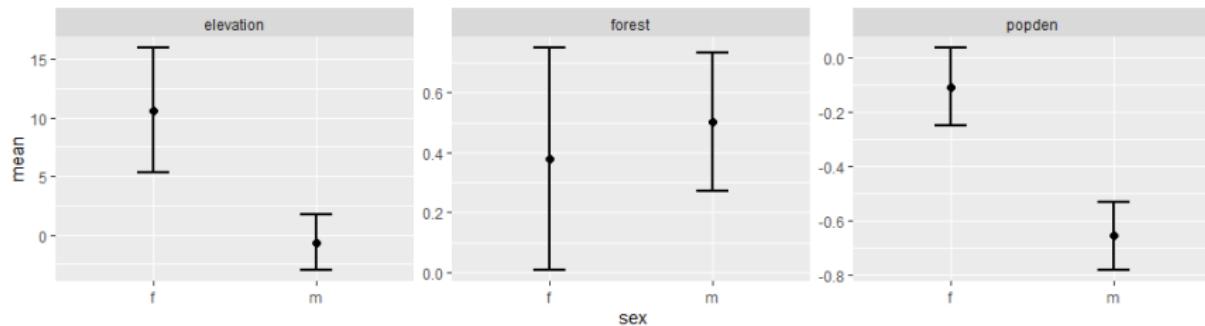
- ▶ relate coefficients to animal-specific characteristics (e.g., age, sex) using say `lm`

## Step 2: Statistics on Statistics



- ▶ relate coefficients to animal-specific characteristics (e.g., age, sex) using say `lm`
- ▶ calculate variance/covariance (biased high due to sampling variability)

## Step 2: Statistics on Statistics



- ▶ relate coefficients to animal-specific characteristics (e.g., age, sex) using say `lm`
- ▶ calculate variance/covariance (biased high due to sampling variability)
- ▶ plot coefficients against availability to explore functional responses

## Two-Step Approach

- ▶ Often a useful starting point (exploratory data analysis)

## Two-Step Approach

- ▶ Often a useful starting point (exploratory data analysis)
- ▶ Fewer assumptions than mixed-effect models (no distributional assumptions about random effects)

## Two-Step Approach

- ▶ Often a useful starting point (exploratory data analysis)
- ▶ Fewer assumptions than mixed-effect models (no distributional assumptions about random effects)
- ▶ Can use bootstrap for inference (assuming individuals are independent)

## Two-Step Approach

- ▶ Often a useful starting point (exploratory data analysis)
- ▶ Fewer assumptions than mixed-effect models (no distributional assumptions about random effects)
- ▶ Can use bootstrap for inference (assuming individuals are independent)

For justification of 2-step approach, see: Murtaugh, P. A. (2007).  
Simplicity and complexity in ecological data analysis. *Ecology*, 88(1),  
56-62.

## Two-Step Approach

- ▶ Often a useful starting point (exploratory data analysis)
- ▶ Fewer assumptions than mixed-effect models (no distributional assumptions about random effects)
- ▶ Can use bootstrap for inference (assuming individuals are independent)

For justification of 2-step approach, see: Murtaugh, P. A. (2007). Simplicity and complexity in ecological data analysis. *Ecology*, 88(1), 56-62.

In the context of step-selection functions:

- ▶ Craiu, R. V., T. Duchesne, D. Fortin, and S. Baillargeon (2011). Conditional logistic regression with longitudinal follow-up and individual-level random coefficients: A stable and efficient two-step estimation method. *Journal of Computational and Graphical Statistics* 20, 767-784.
- ▶ Craiu, R. V., T. Duchesne, D. Fortin, and S. Baillargeon (2016). TwoStepCLogit: Conditional Logistic Regression: A Two-Step Estimation Method. R package version 1.2.5.

# Mixed-Effect or Hierarchical Models

$$f_i^u(s) \propto \exp(x_1(s)\beta_{1i} + x_2(s)\beta_{2i} + \dots x_p(s)\beta_{pi})$$

# Mixed-Effect or Hierarchical Models

$$f_i^u(s) \propto \exp(x_1(s)\beta_{1i} + x_2(s)\beta_{2i} + \dots x_p(s)\beta_{pi})$$

Further assume:

$$(\beta_{1i}, \dots, \beta_{2i}) \sim N(\mu, \psi)$$

Similar 2-step approach, but assume the regression parameters come from a common normal distribution.

# Mixed-Effect or Hierarchical Models

$$f_i^u(s) \propto \exp(x_1(s)\beta_{1i} + x_2(s)\beta_{2i} + \dots x_p(s)\beta_{pi})$$

Further assume:

$$(\beta_{1i}, \dots, \beta_{2i}) \sim N(\mu, \psi)$$

Similar 2-step approach, but assume the regression parameters come from a common normal distribution.

Advantages:

- ▶ inference at individual- and population-level with single model

# Mixed-Effect or Hierarchical Models

$$f_i^u(s) \propto \exp(x_1(s)\beta_{1i} + x_2(s)\beta_{2i} + \dots x_p(s)\beta_{pi})$$

Further assume:

$$(\beta_{1i}, \dots, \beta_{2i}) \sim N(\mu, \psi)$$

Similar 2-step approach, but assume the regression parameters come from a common normal distribution.

Advantages:

- ▶ inference at individual- and population-level with single model
- ▶ can “borrow strength” across individuals when estimating  $(\beta_{1i}, \dots, \beta_{2i})$

# Mixed-Effect or Hierarchical Models

$$f_i^u(s) \propto \exp(x_1(s)\beta_{1i} + x_2(s)\beta_{2i} + \dots x_p(s)\beta_{pi})$$

Further assume:

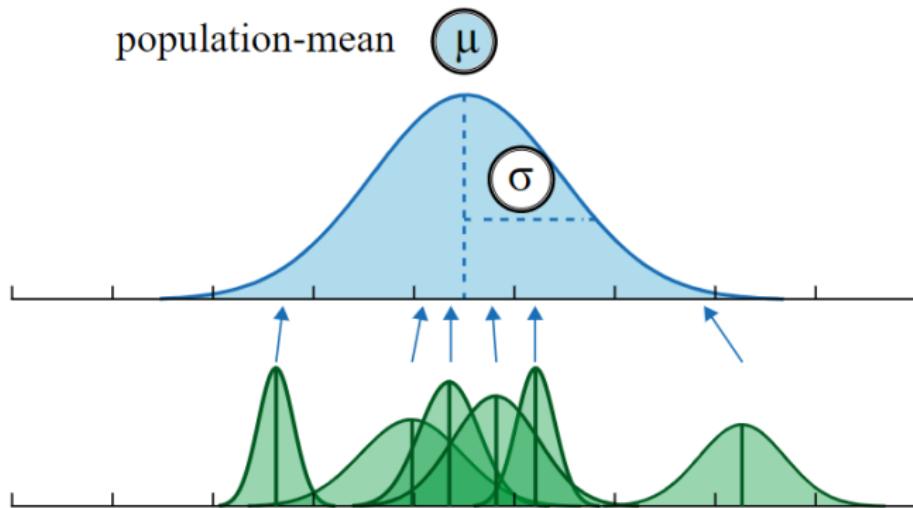
$$(\beta_{1i}, \dots, \beta_{2i}) \sim N(\mu, \psi)$$

Similar 2-step approach, but assume the regression parameters come from a common normal distribution.

Advantages:

- ▶ inference at individual- and population-level with single model
- ▶ can “borrow strength” across individuals when estimating  $(\beta_{1i}, \dots, \beta_{2i})$
- ▶ can estimate  $\mu, \psi$  while accounting for sampling variability

# Individual Estimates Borrow Strength (Exhibit Shrinkage)



## Downside of Random Effects

- ▶ More assumptions (parameters are normally distributed)
- ▶ Added complexity (requires numerical integration to calculate the likelihood), computationally challenging to fit
- ▶ Potentially more difficult for practitioners to understand, correctly specify, and communicate

# Random Effects and Habitat-selection Models

Resource-selection functions:

- ▶ logistic regression with random effects, lots of options (e.g.,  
`glmer` in `lme4` package)

# Random Effects and Habitat-selection Models

Resource-selection functions:

- ▶ logistic regression with random effects, lots of options (e.g., `glmer` in `lme4` package)

Step-selection functions:

- ▶ conditional logistic regression, few good options
- ▶ `coxme` for small numbers of strata (SLOW!)
- ▶ `TwoStepCLogit::Ts.estim()`, a formal two-step approach (will fail if some individuals do not experience all levels of a categorical variable)

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

Stefanie Muff<sup>1,2</sup>  | Johannes Signer<sup>3</sup>  | John Fieberg<sup>4</sup> 



Steffanie Muff, Norwegian University of Science and Technology, Johannes Signer, University of Goettingen

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

Stefanie Muff<sup>1,2</sup>  | Johannes Signer<sup>3</sup>  | John Fieberg<sup>4</sup> 



Steffanie Muff, Norwegian University of Science and Technology, Johannes Signer, University of Goettingen

- ▶ How are random effects used in applications of RSFs?

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

Stefanie Muff<sup>1,2</sup>  | Johannes Signer<sup>3</sup>  | John Fieberg<sup>4</sup> 



Steffanie Muff, Norwegian University of Science and Technology, Johannes Signer, University of Goettingen

- ▶ How are random effects used in applications of RSFs?
- ▶ Develop computationally efficient method of fitting SSFs with random effects

# One Step: Random Effects

- ▶ Random effects were proposed for RSFs over 10 years ago<sup>1</sup>

The screenshot shows a Google Scholar search results page. The search query is "Application of random effects to the study of resource selection by animals". The results are filtered to "Articles" and show approximately 35,200 results found in 0.11 seconds. On the left, there are search filters for time range: "Any time", "Since 2018", "Since 2017", "Since 2014", and "Custom range...". The main result is a link to an article by CS Gillies, M Hebblewhite, and SE Nielsen, published in the Journal of Animal Ecology in 2006. The abstract discusses the application of random effects in resource selection studies. Below the abstract, there are options to "Cite by 479", "Related articles", and "All 23 versions". A red box highlights the "Cited by 479" button.

≡ Google Scholar | Search icon

Articles About 35,200 results (0.11 sec)

Any time  
Since 2018  
Since 2017  
Since 2014  
Custom range...

[HTML] Application of random effects to the study of resource selection by animals  
CS Gillies, M Hebblewhite, SE Nielsen... - Journal of Animal ..., 2006 - Wiley Online Library  
2Most frequently, individual animals are monitored and pooled to estimate population-level effects without regard to group or individual-level variation. Pooling assumes that both observations and their errors are independent, and resource selection is constant given individual variation ...  
☆ 99 Cited by 479 Related articles All 23 versions

<sup>1</sup> Gillies et al. "Application of random effects to the study of resource selection by animals." *Journal of Animal Ecology* 75.4 (2006): 887-898.

# One Step: Random Effects

- ▶ Random effects were proposed for RSFs over 10 years ago<sup>1</sup>

A screenshot of a Google Scholar search results page. The search bar at the top contains the query "random effects". Below the search bar, there are two tabs: "Articles" (selected) and "Books". The search results section shows a list of articles. The first result is a paper by CS Gillies, M Hebblewhite, and SE Nielsen, published in the Journal of Animal Ecology in 2006. The title is "[HTML] Application of random effects to the study of resource selection by animals". Below the title, it says "Cited by 479". A red box highlights this citation count.

≡ Google Scholar |  Search icon

Articles About 35,200 results (0.11 sec)

Any time  
Since 2018  
Since 2017  
Since 2014  
Custom range...

[HTML] Application of random effects to the study of resource selection by animals  
CS Gillies, M Hebblewhite, SE Nielsen... - Journal of Animal ..., 2006 - Wiley Online Library  
2Most frequently, individual animals are monitored and pooled to estimate population-level effects without regard to group or individual-level variation. Pooling assumes that both observations and their errors are independent, and resource selection is constant given individual variation ...  
☆ 99 Cited by 479 Related articles All 23 versions

- ▶ Majority of studies (80 % since 2016) only include random intercept and no random slope(s).

<sup>1</sup> Gillies et al. "Application of random effects to the study of resource selection by animals." Journal of Animal Ecology 75.4 (2006): 887-898.

## RSFs: Random Intercept-Only Models

1. Intercept in RSFs is not of interest and depends heavily on the sampling ratio of used versus available points

## RSFs: Random Intercept-Only Models

1. Intercept in RSFs is not of interest and depends heavily on the sampling ratio of used versus available points
2. Cannot (by definition) account for among-animal variation in the regression slopes (i.e., functional responses)!

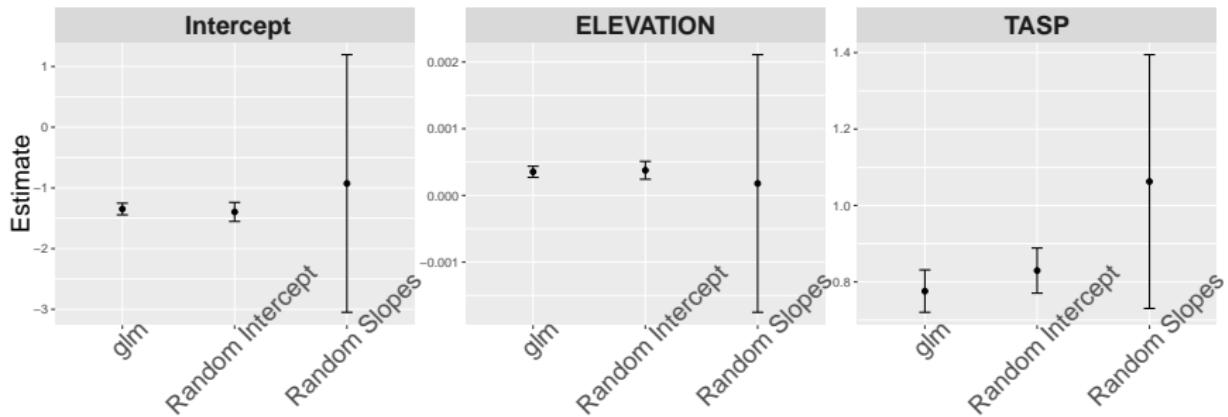
## RSFs: Random Intercept-Only Models

1. Intercept in RSFs is not of interest and depends heavily on the sampling ratio of used versus available points
2. Cannot (by definition) account for among-animal variation in the regression slopes (i.e., functional responses)!
3. SEs will be too small, particularly with lots of observations for each animal

Schielzeth, H. and W. Forstmeier (2009). Conclusions beyond support: Overconfident estimates in mixed models. Behavioral Ecology 20, 416-420.

# Example: Goat RSFs

```
glm <- glm(STATUS ~ TASP + ELEVATION, family=binomial(),  
           data = goats)  
glmer_int <- glmmTMB(STATUS ~ TASP + ELEVATION + (1|ID),  
                      family=binomial(), data = goats)  
glmer_randcoef <- glmmTMB(STATUS ~ TASP + ELEVATION +  
                           (1+TASP+ELEVATION|ID),  
                           family=binomial(), data = goats)
```



# Computationally Efficient Step-Selection Functions

How can we fit SSFs with random effects?

# Computationally Efficient Step-Selection Functions

How can we fit SSFs with random effects?

Reformulate SSFs as a Poisson model with stratum-specific intercepts  $\alpha_{nt}$  (Armstrong et al. 2014):

- ▶ Same likelihood kernel as condition logistic regression likelihood, same  $\hat{\beta}$ , same  $SE(\hat{\beta})$
- ▶ Treat intercepts as random with large fixed variance:  $\alpha_{nt} \sim N(0, 10^6)$  (avoids shrinkage)
- ▶ Options: `glmmTMB` or `INLA` (Bayesian)

Armstrong et al. 2014. Conditional Poisson models: a flexible alternative to conditional logistic case cross-over analysis. *BMC medical research methodology* 14:122.

# Applied Example



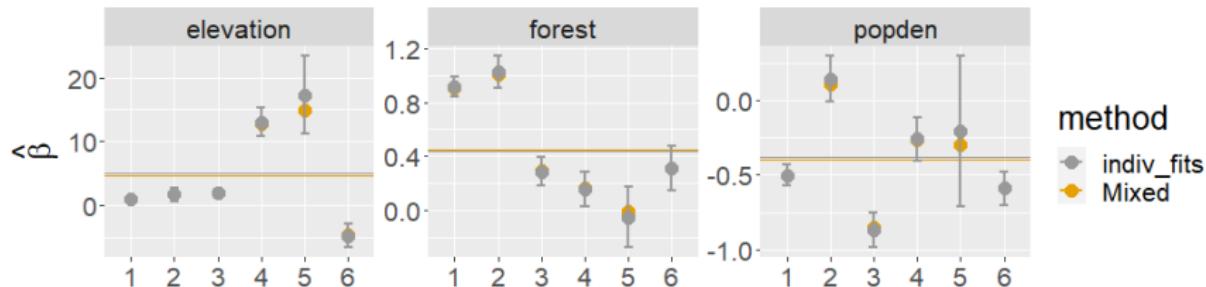
Covariates:

- ▶ elevation
- ▶ population density
- ▶ forest (yes/no)

Compare two-step approaches and mixed models:

- ▶ estimates of individual coefficients
- ▶ mean and variance of the coefficients across animals

# Resource-Selection Functions



Variance estimates	$sd(\hat{\beta}_{\text{elevation}})$	$sd(\hat{\beta}_{\text{popden}})$	$sd(\hat{\beta}_{\text{forest}})$
naive two-step (indiv_fits)	8.35	0.352	0.428
mixed model	7.07	0.315	0.382

## Resource-Selection Functions

Naive two-step and mixed-model estimates extremely similar

- ▶ large  $n \Rightarrow$  little shrinkage or borrowing of strength

## Resource-Selection Functions

Naive two-step and mixed-model estimates extremely similar

- ▶ large  $n \Rightarrow$  little shrinkage or borrowing of strength
- ▶ BUT, we assumed independence within an individual (will usually not appropriate)

## Resource-Selection Functions

Naive two-step and mixed-model estimates extremely similar

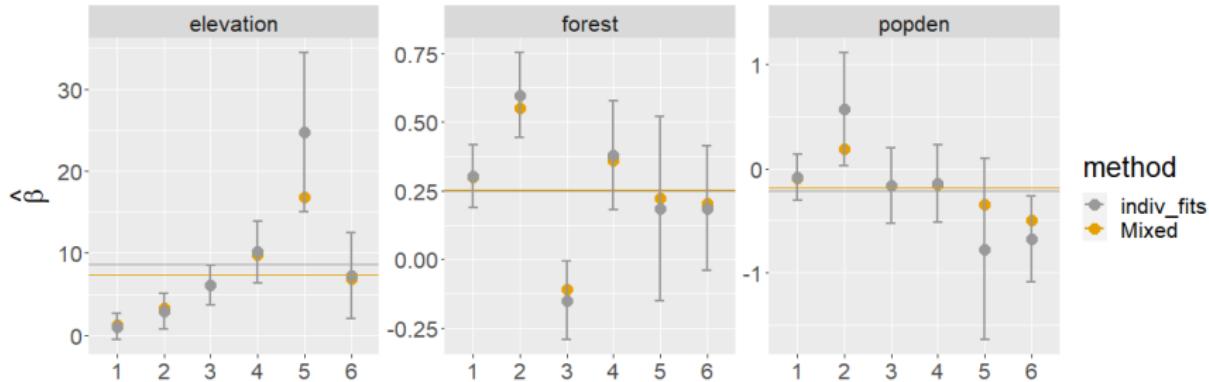
- ▶ large  $n \Rightarrow$  little shrinkage or borrowing of strength
- ▶ BUT, we assumed independence within an individual (will usually not appropriate)
- ▶ likelihood weighting using estimates of effective sample size has been suggested as a way to deal with within-individual autocorrelation (Alson et al. in review. Mitigating pseudoreplication and bias in resource selection functions with autocorrelation-informed weighting)

## Resource-Selection Functions

Naive two-step and mixed-model estimates extremely similar

- ▶ large  $n \Rightarrow$  little shrinkage or borrowing of strength
- ▶ BUT, we assumed independence within an individual (will usually not appropriate)
- ▶ likelihood weighting using estimates of effective sample size has been suggested as a way to deal with within-individual autocorrelation (Alson et al. in review. Mitigating pseudoreplication and bias in resource selection functions with autocorrelation-informed weighting)
- ▶ two-step approach + bootstrap (resampling individuals) could relax the assumption of independence within an individual

# Step-Selection Functions



Variance estimates	$sd(\beta_{elevation})$	$sd(\beta_{popden})$	$sd(\beta_{forest})$
naive two-step (indiv_fits)	7.67	0.465	0.282
mixed model	4.91	0.293	0.252

## Step-Selection Functions

Here, there are clear advantages to using mixed-model:

- ▶ beneficial shrinkage (due to smaller sample sizes since we are modeling transitions requiring equally-spaced time points)
- ▶ estimates of variance parameters are not biased by sampling variability
- ▶ BUT: steps may still be autocorrelated depending on the sampling frequency

## Next steps

We will explore code for:

- ▶ implementing a two-step approach using tidyverse principles
- ▶ fitting mixed SSFs using glmmTMB (frequentist)