

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 Signer³ and John Fieberg^{4*}

July 31, 2019

1 Running title: **Estimation of habitat-selection functions**

2

3 ¹ Institute of Evolutionary Biology and Environmental Studies, University of Zurich,
4 Winterthurerstrasse 190, 8057 Zurich, Switzerland

5

6 ² Department of Biostatistics, Epidemiology, Biostatistics and Prevention Institute, Uni-
7 versity of Zurich, Hirschengraben 84, 8001 Zurich, Switzerland

8

9 ³ Wildlife Sciences, Faculty of Forest Science and Forest Ecology, University of Goettin-
10 gen, Büsgenweg 3, 37077 Göttingen

11

12 ⁴ Department of Fisheries, Wildlife, and Conservation Biology, University of Minnesota,
13 St. Paul, MN, USA

14

15 * Corresponding authors. E-mail: stefanie.muff@uzh.ch, jfieberg@umn.edu

16

Abstract

17
18
19 **1.** Popular frameworks for studying habitat selection include resource-selection
20 functions (RSFs) and step-selection functions (SSFs), estimated using logistic
21 and conditional logistic regression, respectively. Both frameworks compare
22 environmental covariates associated with locations animals visit with envi-
23 ronmental covariates at a set of locations assumed available to the animals.
24 Conceptually, slopes that vary by individual, that is, random coefficient mod-
25 els, could be used to accommodate inter-individual heterogeneity with either
26 approach. While fitting such models for RSFs is possible with standard soft-
27 ware for generalized linear mixed effects models (GLMMs), straightforward
28 and efficient one-step procedures for fitting SSFs with random coefficients are
29 currently lacking.

30 **2.** To close this gap, we take advantage of the fact that the conditional logistic
31 regression model (*i. e.*, the SSF) is likelihood-equivalent to a Poisson model
32 with stratum-specific fixed intercepts. By interpreting the intercepts as a ran-
33 dom effect with a large (fixed) variance, inference for random-slope models
34 becomes feasible with standard Bayesian techniques, or with frequentist meth-
35 ods that allow one to fix the variance of a random effect. We compare this
36 approach to other commonly applied alternatives, including models without
37 random slopes and mixed conditional regression models fit using a two-step
38 algorithm.

39 **3.** Using data from mountain goats (*Oreamnos americanus*) and Eurasian
40 otters (*Lutra lutra*), we illustrate that our models lead to valid and feasible
41 inference. In addition, we conduct a simulation study to compare different
42 estimation approaches for SSFs, and to demonstrate the importance of includ-
43 ing individual-specific slopes when estimating individual- and population-level
44 habitat-selection parameters.

45 4. By providing coded examples using integrated nested Laplace approxima-
46 tions (INLA) and Template Model Builder (TMB) for Bayesian and frequen-
47 tist analysis via the R packages R-INLA and glmmTMB, we hope to make effi-
48 cient estimation of RSFs and SSFs with random effects accessible to anyone in
49 the field. SSFs with individual-specific coefficients are particularly attractive
50 since they can provide insights into movement and habitat-selection processes
51 at fine-spatial and temporal scales, but these models had previously been very
52 challenging to fit.

53 **Keywords:** Conditional logistic regression, glmmTMB, integrated nested Laplace
54 approximations (INLA), multinomial regression, random effects, resource-selection func-
55 tions, step-selection functions

56 1 Introduction

57 Ecologists have long been interested in understanding how animals select habitat and the
58 resulting fitness consequences from different space-use strategies (Gaillard et al., 2010).
59 Importantly, optimal behavioral strategies may depend on intrinsic factors specific to the
60 individual, such as its age, sex, and body condition (Lesmerises and St-Laurent, 2017),
61 as well as extrinsic factors, including climatic conditions (Raynor et al., 2017), local
62 predator communities (Heithaus, 2001), competition for resources (Rosenzweig, 1991), or
63 local availability of different habitat types (Mysterud and Ims, 1998). Interestingly, indi-
64 viduals from the same species often adopt different habitat-use strategies (*e. g.*, Leclerc
65 et al., 2016), suggestive of behavioral phenotypes or “personalities” (Stamps, 2007); these
66 differences may also have a strong genetic component (Jaenike and Holt, 1991). Thus,
67 understanding the causes and consequences of among-animal variation in habitat selec-
68 tion is key to addressing fundamental questions in ecology and evolution, including the
69 extent to which individuals develop specialized, individual niches that are narrower than
70 that of the population (Bolnick et al., 2002; Sheppard et al., 2018).

71 Modern biotelemetry devices have made it possible to monitor habitat use of multiple
72 animals at finer temporal and spatial scales, providing unique opportunities to study
73 variation in individual behaviors and habitat-selection strategies (Cagnacci et al., 2010).
74 Popular publications by Gillies et al. (2006) and Dingemanse and Dochtermann (2013)
75 have argued for the use of random effects (*i. e.*, random parameters drawn from a common
76 statistical distribution) to model individual variation in the context of habitat-selection
77 and behavioral studies, respectively. Random effects allow individual coefficients to be
78 “tied together” via an assumption that they come from a common population. These
79 models offer a powerful approach to studying inter-individual variability, because they
80 explicitly allow intercepts and/or slopes in the respective regression models to vary by
81 individual, while at the same time taking advantage of the shared information that is
82 present in the data from different individuals (Fieberg et al., 2009). An added benefit of
83 these models is that they can accommodate non-independent data arising from having
84 multiple observations on the same individual. By contrast, treating all observations as
85 though they are independent would result in optimistic standard errors and confidence
86 intervals, leading to what Hurlbert (1984) referred to as *pseudoreplication*.

87 **1.1 Habitat-Selection Analyses Using Resource-Selection and Step-** 88 **Selection Functions**

89 Habitat-selection analyses typically compare environmental covariates at locations vis-
90 ited by an animal to environmental covariates at a set of locations assumed available to
91 the animal (Manly et al., 2002). Historically, most analyses of animal telemetry data
92 focused on what Johnson (1980) called 3rd order selection, with available points sam-
93 pled randomly or systematically from within an animal’s estimated home range. In the
94 wildlife literature, the combined observed and available locations are typically analyzed
95 using *logistic regression*, with specific focus on estimating the exponential of the linear
96 predictor (with the intercept removed) referred to as a resource-selection function (RSF).

97 Warton and Shepherd (2010) provided context for interpreting RSFs by showing that
98 slope parameters in logistic regression models are asymptotically equivalent to slope pa-
99 rameters in an inhomogeneous Poisson point process (IPP) model. The IPP assumes
100 that the intensity function is a log-linear function of the covariates, thus regression pa-
101 rameters describe relationships between environmental covariates and the relative density
102 of observed locations in space, assuming all locations in the spatial domain are equally
103 accessible or available to the animal. Fithian and Hastie (2013) further showed that
104 equivalence between logistic regression and an IPP only holds when the model is cor-
105 rectly specified or when available points are “infinitely” weighted. Interestingly, several
106 other modelling approaches, including the maximum entropy method (Maxent, Phillips
107 et al., 2006), weighted distribution theory (Lele and Keim, 2006), and resource utilization
108 functions (Millsbaugh et al., 2006) have also been shown to be equivalent to fitting an
109 IPP model (Aarts et al., 2012; Fithian and Hastie, 2013; Hooten et al., 2013; Renner and
110 Warton, 2013).

111 Recent methodological development has focused on modelling habitat selection at
112 finer temporal and spatial scales, in part driven by concerns associated with serial auto-
113 correlation of animal locations, as points close in time are also expected to be close in
114 space (Arthur et al., 1996; Rhodes et al., 2005; Fortin et al., 2005). Recognizing that not
115 all areas of the availability domain (typically the home range of the animal) are equally
116 available at all time points, Fortin et al. (2005) suggested resampling step lengths (dis-
117 tances between successive observed locations) and turn angles (deviations from previous
118 bearings) to generate random movements and hence available points conditional on the
119 previously observed locations. This process results in stratified datasets with a different
120 set of available points associated with each observed location. The combined (stratified)
121 observed and available location data are typically analyzed using *conditional logistic re-*
122 *gression*, with the exponential of the linear predictor referred to as a step-selection func-
123 tion (SSF). Forester et al. (2009), Duchesne et al. (2015) and Avgar et al. (2016) further

124 refined this approach and demonstrated the utility of using common statistical distri-
125 butions to model and simulate step lengths and turn angles. Specifically, they showed
126 that it was possible to fit the equivalent of a biased random walk model when random
127 points were generated using specific statistical distributions and when movement-related
128 covariates (*e. g.*, turn angles, step length, log-step-length) were included in conditional
129 logistic regression models. These methods have recently been implemented in the `amt`
130 R package (Signer et al., 2019), making SSFs an exciting and accessible approach for
131 studying habitat selection at the scale of the movement step.

132 **1.2 Use of Random Effects in Resource-Selection and Step-Selection** 133 **Functions**

134 Gillies et al. (2006) recommended using logistic regression models with individual-specific
135 random intercepts to account for unequal sample sizes, and individual-specific random
136 coefficients (briefly denoted as *random coefficients* or *random slopes*) to account for
137 individual-specific differences when fitting RSFs. Similarly, Hebblewhite and Merrill
138 (2008) recommended random intercepts to account for correlation within nested group-
139 ings of locations from socially-structured populations (*e. g.*, repeated observations from
140 individual wolves and observations from wolves in the same pack). Gillies et al. (2006)
141 and Hebblewhite and Merrill (2008) further emphasized that random coefficients could
142 be used to model variation in habitat selection attributable to differences in habitat
143 availability, referred to as a *functional response* (Mysterud and Ims, 1998; Beyer et al.,
144 2010). Soon thereafter, Matthiopoulos et al. (2011) and Aarts et al. (2013) developed a
145 formal framework for modelling functional responses using a combination of random ef-
146 fects and fixed effects constructed from the first few moments (mean, variance) of habitat
147 covariates.

148 Most modern statistical software platforms provide methods for fitting generalized
149 linear mixed effects models (*e. g.*, logistic regression with random intercepts and slopes),

150 and therefore, allow for the possibility of studying individual-specific variation in studies
151 focused on 3rd order habitat selection. However, a literature search that we carried
152 out on all papers that cited Gillies et al. (2006), published between January 2016 and
153 May 2018 and claiming to use random effects in an RSF ($n = 69$), revealed that less
154 than 20% of all publications included individual-specific random slopes in their models,
155 while the majority of them only specified an individual-specific random intercept. This is
156 interesting because random intercept-only models are often not sufficient to account for
157 pseudoreplication (Schielzeth and Forstmeier, 2009). Further, in applications of RSFs,
158 the variability in the intercepts is largely driven by differences in the ratio of used to
159 available points, which is under control of the analyst (Fieberg et al., 2010). We will
160 come back to this point later in the paper.

161 In the context of SSFs, Duchesne et al. (2010) argued for incorporating individual-
162 specific slopes to allow the influence of habitat covariates to depend on what is locally
163 available to the animal (*i. e.*, for functional responses). Unfortunately, conditional logistic
164 regression models that include individual-specific random slopes are extremely challenging
165 to fit, especially with large numbers of strata (Craiu et al., 2011). To circumvent this
166 problem Craiu et al. (2011) developed a two-step estimation approach to fitting mixed-
167 effects models. This approach works well when the number of strata per individual
168 is large, but frequently fails (or leads to numerical instabilities) when one or several
169 individuals do not have enough variability in the encountered locations. As one of its
170 main limitations, it is not possible to use this two-step approach in cases where one or
171 more individuals do not encounter all factor levels of a categorical predictor.

172 **1.3 Objectives**

173 Our overarching goal of this paper is to provide both new and established users of RSFs
174 and SSFs with a coherent framework to formulate and fit the respective statistical mod-
175 els. In particular, the objectives are to: 1) reiterate the importance of including random

176 slopes in habitat-selection models, both for RSFs and SSFs; 2) reiterate the importance
177 of weighting available points when fitting logistic regression models to estimate RSFs;
178 and 3) present computationally efficient and consistent methods for fitting both RSFs
179 and SSFs with random effects. To allow fitting of SSFs, we propose to reformulate the
180 conditional logistic regression model as a (likelihood-equivalent) Poisson model, where
181 stratum-specific intercepts are included and efficiently modeled as a random effect with a
182 fixed large prior variance. We will explain why fixing the variance is important, and why,
183 for the same reason, random intercept variances in RSFs should also be fixed at a large
184 value, instead of estimated. We illustrate how all models discussed here can easily be fit
185 using R (R Core Team, 2018), either employing a Bayesian approach via integrated nested
186 Laplace approximations (INLA, Rue et al., 2009) using the R-interface R-INLA, or in a
187 frequentist approach using Template Model Builder (TMB) via the `glmmTMB` R-package
188 (Brooks et al., 2017; Magnusson et al., 2017). To illustrate the efficiency and accuracy
189 of these methods, we reanalyzed data from a study on mountain goats (*Oreamnos amer-*
190 *icanus*; Lele and Keim, 2006) and Eurasian otters (*Lutra lutra*; Weinberger et al., 2016),
191 and carried out a simulation study to compare the reformulated Poisson model for SSFs
192 to existing two-step procedures. We provide ready-to-use R code to replicate all of our
193 analyses (Muff et al., 2019).

194 **2 Background on analyzing RSFs and SSFs**

195 Both RSFs and SSFs quantify habitat selection by comparing environmental covariates
196 associated with locations that animals visit (encoded as $y = 1$) with environmental
197 covariates at a set of locations assumed available to the animal (encoded as $y = 0$). The
198 main difference between the RSF and the SSF approach is that the latter conditions (*i. e.*,
199 “matches”) the set of available points on the current location of the animal, resulting in a
200 stratified dataset, whereas RSFs use a single set of (pooled) available locations for each

201 animal, with these locations usually generated by sampling randomly or systematically
202 from within an animal’s home range (*e. g.*, Manly et al., 2002). The sampling scheme
203 used to generate available points dictates how the respective data should be analyzed
204 (Warton and Aarts, 2013): While RSFs can be estimated by fitting a standard logistic
205 regression model, SSFs need to account for the fact that a unique set of available points
206 is chosen for (or “matched to”) each observed location, which can be accomplished by
207 fitting a conditional logistic regression model. In the latter case, each observed location
208 thus forms a stratum along with its set of matched available locations. We give a short
209 overview of the two different regression models that are typically used to estimate RSFs
210 and SSFs.

211 **2.1 RSFs: logistic regression model**

212 Assume we have $n = 1, \dots, N$ individuals and a set $j = 1, \dots, J_n$ of used and available
213 locations for animal n . In the absence of any random effects, the probability that a point
214 y_{nj} with covariate vector \mathbf{x}_{nj} is used, $\Pr(y_{nj} = 1 | \mathbf{x}_{nj}) = \pi_{nj}$, can then be modeled as

$$\text{logit}(\pi_{nj}) = \boldsymbol{\beta}^\top \mathbf{x}_{nj}, \quad y_{nj} \sim \text{Bern}(\pi_{nj}), \quad (1)$$

215 with logistic (logit) link and covariate vector $\boldsymbol{\beta}$ that is the target of interest (Warton and
216 Shepherd, 2010). Standard generalized linear model (GLM) software, such as the `glm()`
217 function in R, can be used to estimate $\boldsymbol{\beta}$. An extension of model (1) to include individual-
218 specific random effects is conceptually straightforward, and the respective mixed model
219 can for example be fit by the `glmer()` function from the `lme4` package (Bates et al.,
220 2015).

221 It is important to note that, unlike prospective sampling designs involving a binary
222 response variable, the y_{nj} in RSF designs are not Bernoulli random variables. Rather, the
223 Bernoulli likelihood formed by (1) results in a set of estimating equations that produce

224 consistent estimators of β in an equivalent log-linear IPP model (Warton and Shepherd,
225 2010). This equivalence holds whenever the RSF model is correctly specified, or when
226 the number of available points is sufficiently large (Warton and Shepherd, 2010). How-
227 ever, using a sufficiently large number of available points is computationally inefficient,
228 and in fact Fithian and Hastie (2013) elegantly show that the same convergence limit
229 is obtained when instead infinite weights are assigned to all available points. For the
230 respective weighted logistic regression approach, the likelihood for the available “back-
231 ground” samples (*i. e.*, $y = 0$) is weighted with a weight W , while the used points ($y = 1$)
232 keep weight 1. Fithian and Hastie (2013) demonstrated how, for $W \rightarrow \infty$, the likeli-
233 hood converges to the IPP likelihood. In our experience values of $W = 1000$ typically
234 lead to good approximations, but larger values may be tried to check for convergence.
235 Weights are easily incorporated into most GLM software (*e. g.*, `glm()` or `glmer()`). We
236 do not reiterate the logistic regression likelihood here, but refer the reader to Hosmer and
237 Lemeshow (2000) for more on logistic regression, and to Warton and Shepherd (2010)
238 and Fithian and Hastie (2013) for a description and justification of its use for studying
239 habitat selection.

240 **2.2 SSFs: conditional logistic regression model**

241 Assume we have $n = 1, \dots, N$ individuals with realized steps at time points $t = 1, \dots, T_n$,
242 with $j = 1, \dots, J_{n,t}$ locations that were either used or available to animal n at time step
243 t . Note that, for notational simplicity, we may replace $J_{n,t}$ by J , because it is common
244 practice to match a constant number of available points to each observed location (*e. g.*,
245 10 available for 1 used location). Used and available locations associated with each step
246 form a choice set or *stratum*. This implies that the probability the n^{th} animal selects
247 the j^{th} unit with habitat-specific covariates \mathbf{x}_{ntj} at time point t , given the set of possible

248 choices $\mathbf{x}_{nt.} = \{\mathbf{x}_{nt1}, \dots, \mathbf{x}_{ntJ}\}$, is

$$\Pr(y_{ntj} = 1 | \mathbf{x}_{nt.}) = \pi_{ntj} = \frac{\exp(\boldsymbol{\beta}^\top \mathbf{x}_{ntj})}{\sum_{i=1}^J \exp(\boldsymbol{\beta}^\top \mathbf{x}_{nti})}, \quad (2)$$

249 with covariate vector $\boldsymbol{\beta}$ that is the target of estimation. A popular and computationally
250 efficient way to fit the discrete choice model (2) in the context of habitat-selection studies
251 is by interpreting it as a specific version of the stratified proportional hazards model
252 (Manly et al., 2002; McDonald et al., 2006). In the absence of random effects, this “Cox
253 trick” provides a framework for efficient inference using Maximum Likelihood (ML), for
254 instance by using the `clogit()` function from the `survival` package in R (Therneau,
255 2015b), although any function to fit the Cox proportional hazard survival model can be
256 used.

257 **3 Mixed effects modelling of RSFs and SSFs**

258 **3.1 The importance of random slopes**

259 Virtually all habitat-selection studies monitor multiple animals, and the respective data
260 are combined and modeled jointly. However, it is well known that such a sampling design
261 generally leads to pseudoreplication due to non-independence among the data points from
262 the same individual (see *e. g.*, Gillies et al., 2006; Duchesne et al., 2010; Fieberg et al.,
263 2010). Generalized linear mixed models (GLMM) offer a powerful approach to prop-
264 erly account for correlated measurements taken on the same animal, while also allowing
265 for differences in the intercepts and/or slopes among animals due to individual-specific
266 preferences and/or differences in habitat availability that induce functional responses in
267 habitat selection (Mysterud and Ims, 1998; Hebblewhite and Merrill, 2008; Matthiopoulos
268 et al., 2011; Aarts et al., 2013; Matthiopoulos et al., 2015).

269 Our literature review on the RSF papers mentioned in the introduction suggests that

270 it is common practice to include individual-specific random intercepts, but not random
271 slopes when modelling habitat selection. This is remarkable for three reasons: First
272 and most importantly, random intercept-only models cannot (by definition) account for
273 among-individual variation in the regression *slopes*, that is, they cannot account for func-
274 tional responses. Further, the slope estimator from a logistic model that omits random
275 effects is a biased estimator of the mean slope in the population, a fact that has been
276 discussed repeatedly in the statistical and ecological literature (*e. g.*, Fieberg et al., 2009;
277 Muff et al., 2016). Second, omitting individual-specific random slopes when they actually
278 do vary between individuals induces too little uncertainty in the estimated parameters
279 (*e. g.*, Schielzeth and Forstmeier, 2009). Consequently, it is possible that researchers end
280 up with too high confidence in their potentially biased estimators of effect sizes. The
281 problem is particularly acute when there are lots of observations for each animal, which
282 is typically the case in telemetry studies. And third, the intercept in RSF models reflects
283 the probability of a location being used when all covariates are set equal to 0, and is
284 thus heavily influenced by the ratio of used versus available points (Fieberg et al., 2010).
285 Given that it is common to use a predefined, constant ratio of used to available points
286 for all animals (for example 10 available points per used point), it is not surprising that
287 random intercept estimators will sometimes return an among-animal variance component
288 of 0. We demonstrate all of these issues by comparing RSF models with and without
289 random slopes that we fit to data from mountain goats in Section 4.1. Moreover, the
290 first two issues are also relevant for SSF models, as pointed out by *e. g.*, Duchesne et al.
291 (2010), and are illustrated here with an analysis of Eurasian otters and with a simulation
292 study in Sections 4.2 and 4.3, respectively.

293 **3.2 Computational challenges for SSFs**

294 Fitting a GLMM is generally known to be a difficult and computationally demanding task,
295 and the user can choose among various model fitting procedures (an overview is given

296 by *e. g.*, Bolker et al., 2009, Table I). Note, however, that while standard logistic mixed
297 models (*i. e.*, RSFs) can be fit with several available software packages and functions (such
298 as `lme4::glmer()`), random effects modelling is even more challenging for SSFs, that is,
299 for conditional logistic regression, especially when the number of cases per stratum is
300 greater than 1, or when the strata are unbalanced (Craiu et al., 2011). Given that
301 proportional hazard (*i. e.*, survival) models are commonly used to analyse SSFs with only
302 fixed effects, it seems natural to interpret random-effects SSF models as survival models
303 with random effects (denoted as *frailty models*), for which R solutions, for example `coxme`
304 or `mclogit`, exist (*e. g.* Therneau, 2015a; Elff, 2016). Unfortunately, computation quickly
305 becomes prohibitive for telemetry data with large numbers of strata.

306 To address these challenges, several approaches to circumvent direct random effects es-
307 timation have been proposed, such as the use of generalized estimating equations (GEEs,
308 Craiu et al., 2008) or a two-step estimation approach (Craiu et al., 2011). GEEs, how-
309 ever, provide marginal parameter estimates that are analogous to those obtained from
310 models without random effects, which are known to underestimate the true effect sizes
311 experienced by individual animals (Lee and Nelder, 2004; Fieberg et al., 2009; Muff
312 et al., 2016); thus, we do not generally recommend them for habitat-selection studies.
313 The two-step approach is an efficient alternative that combines estimates of individual-
314 specific regression parameters from standard ML methods for independent data with an
315 expectation-maximization algorithm in conjunction with conditional restricted maximum
316 likelihood (REML). It is available via the `Ts.estim()` function from the `TwoStepCLogit`
317 package in R (Craiu et al., 2016). This approach is an approximate method that works
318 best when the number of strata per animal is large (Craiu et al., 2011). However, the
319 data must fulfill certain regularity conditions, namely all animals must have encountered
320 all levels of a categorical covariate, as it is otherwise not possible to obtain the individual-
321 specific estimates from the first step that are needed for the second step of the procedure.
322 Despite this major limitation of the two-step estimation method, it is one of the most

323 popular approaches for fitting SSFs with random effects, while fitting such models is
 324 currently considered unfeasible with standard GLM or GLMM software.

325 **3.3 An efficient alternative for SSFs**

326 We will now illustrate how relatively simple model reformulations allow one to fit mixed
 327 conditional logistic regression models in a standard GLMM. Starting (for notational
 328 simplicity) with the fixed effects-only model introduced in equation (2), we take advantage
 329 of the fact that the conditional logistic regression model is a special case of a multinomial
 330 model (*e. g.*, McCullagh and Nelder, 1989), and that as such it is likelihood-equivalent to
 331 the Poisson model

$$\mathbb{E}(y_{ntj}) = \mu_{ntj} = \exp(\alpha_{nt} + \boldsymbol{\beta}^\top \mathbf{x}_{ntj}) , \quad \text{with } y_{ntj} \sim \text{Po}(\mu_{ntj}) \quad (3)$$

332 (Whitehead, 1980; McCullagh and Nelder, 1989; Chen and Kuo, 2001), where α_{nt} is the
 333 stratum-specific intercept of animal n at time point t . Since a predefined fixed number of
 334 used points (usually one) is allowed within a stratum, the probability of use, conditional
 335 on the used and available locations in the stratum, is

$$\Pr(y_{ntj} = 1 | \mathbf{x}_{nt\cdot}) = \pi_{ntj} = \frac{\exp(\alpha_{nt} + \boldsymbol{\beta}^\top \mathbf{x}_{ntj})}{\sum_{i=1}^J \exp(\alpha_{nt} + \boldsymbol{\beta}^\top \mathbf{x}_{nti})} = \frac{\exp(\boldsymbol{\beta}^\top \mathbf{x}_{ntj})}{\sum_{i=1}^J \exp(\boldsymbol{\beta}^\top \mathbf{x}_{nti})} , \quad (4)$$

336 where the second equality holds because the stratum-specific intercepts α_{nt} cancel out.
 337 This illustrates that model (3) is maximizing the same likelihood-kernel as the condi-
 338 tional logistic model given in (2). Thus model (3), which is sometimes denoted as the
 339 *conditional Poisson* model, and conditional logistic regression models give equivalent pa-
 340 rameter estimates, $\hat{\boldsymbol{\beta}}$, and also the same standard errors (for a mathematical derivation
 341 see *e. g.*, McCullagh and Nelder, 1989, Chapter 6.4.2). Note that these considerations are
 342 not limited to the presence of only one used point per stratum, but are valid for multi-
 343 nomial data with *any* number of cases per stratum, and even hold when the different

344 strata in a dataset contain an unequal number of cases. In addition, the reformulation
345 also works when random effects are added to the linear predictors in (3), in which case
346 *any* convenient GLMM software can be used to fit the resulting mixed Poisson model.
347 This option to fit SSFs has already been pointed out by Duchesne et al. (2010), but it
348 has only rarely been used to analyze mixed conditional logistic regression models that
349 arise from habitat-selection studies (but see Bruun and Smith, 2003).

350 The obvious disadvantage of formulation (3) – and a potential reason why the ap-
351 proach is rarely used – is that a large number of stratum-specific fixed intercepts α_{nt}
352 must be estimated, which might again make the procedure prohibitive for movement
353 data with tens of thousands of realized steps, given that each step induces a stratum and
354 thus a separate intercept. Luckily, the α_{nt} are not actually of interest, and it is computa-
355 tionally more convenient and efficient to interpret them as a random effect $\alpha_{nt} \sim \mathbf{N}(0, \sigma_\alpha^2)$.
356 However, it is well known that estimates of random effects will, on average, be too small
357 in absolute terms, a phenomenon that is known as “shrinkage towards an overall mean” in
358 the statistics literature (*e. g.*, Robinson, 1991; Snijders and Bosker, 1999). While shrink-
359 age has, in general, many desirable properties, it would introduce a bias into the SSF
360 analysis (see illustration in Section 4.3). The trick to avoid shrinkage in the α_{nt} values,
361 while still taking advantage of the efficiency of a random effects model, is to not allow
362 the variance σ_α^2 to be freely estimated, but instead to fix it at a large value to ensure
363 that stratum-specific intercepts are not pulled towards 0, but are estimated essentially
364 like fixed-effects parameters.

365 This idea is easy to implement in a Bayesian approach, where such information can be
366 specified in the priors. In fact, exactly such models with fixed intercept variance have been
367 previously implemented in a Bayesian setting under the multinomial modelling framework
368 see *e. g.*, the WinBUGS manual section 9.7 (Lunn et al., 2000). Adding random effects

369 to the linear predictor leads to the mixed Poisson model

$$E(y_{ntj}) = \mu_{ntj} = \exp(\alpha_{nt} + \boldsymbol{\beta}^\top \mathbf{x}_{ntj} + \mathbf{u}_n^\top \mathbf{z}_{ntj}), \quad \text{with } y_{ntj} \sim \text{Po}(\mu_{ntj}), \quad (5)$$

370 with individual-specific random slopes \mathbf{u}_n^\top , design vector \mathbf{z}_{ntj} (typically a sub-vector of
371 \mathbf{x}_{ntj}), and $\alpha_{nt} \sim N(0, \sigma_\alpha^2)$ with σ_α^2 fixed at a large value, for example 10^6 . It may be
372 prudent to verify that the results are robust when even larger values of σ_α^2 are used.

373 For the Bayesian analyses presented here, we will take advantage of INLA via its R
374 interface R-INLA. INLA avoids sampling by accurately approximating posterior marginal
375 distributions (Rue et al., 2009), and it has therefore become a popular and efficient alter-
376 native to Markov chain Monte Carlo (MCMC) or likelihood-based inference, in particular
377 for GLMMs (Fong et al., 2010). Importantly, while fixing a variance in a Bayesian anal-
378 ysis is straightforward and natural, it is of course also possible in a likelihood framework.
379 Model (5) can therefore also be fit with a frequentist GLMM software, provided that
380 there is an option to constrain σ_α^2 to a fixed, large value to avoid shrinkage of the in-
381 tercepts. To our knowledge, this is currently not implemented in `glmer()` in the `lme4`
382 package in R, but it is possible with the `glmmTMB` package (Brooks et al., 2017; Magnusson
383 et al., 2017). Consequently, we will fit frequentist GLMMs using `glmmTMB::glmmTMB()`
384 to estimate SSFs according to model (5).

385 It may seem a logical consequence to suggest infinitely weighted Poisson regression to
386 estimate the model parameters of equation (5) for SSFs, given that infinitely weighted
387 logistic regression is recommended for RSFs. However, the assumptions that hold for
388 RSFs are violated because strata (which are the sampling-units of SSFs) only contain
389 very few available points ($y = 0$), thus the large-sample properties of RSFs do not apply
390 to the case of SSFs, and convergence to the IPP is therefore not guaranteed (see *e. g.*,
391 assumptions of Theorems 3.2 and 3.3 in Warton and Shepherd, 2010). As a consequence,
392 weighting introduces a bias, unless the use to availability ratio is very small. We will
393 illustrate this point with a simulation (see Section 4.3 and Figure S1 in the Appendix).

3.4 Individual-specific intercepts in RSFs

As mentioned in Section 3.1, the (individual-specific) intercept term in an RSF is largely determined by the sampling ratio of used and available points for each individual (Warton and Shepherd, 2010, Theorem 3.2). However, the intercept is also influenced by the probability that a point is used (versus available) for the case when all covariates are set equal to 0. If all covariates \boldsymbol{x} in equation (1) have been mean-centered, for example, this reflects an “average” point in the habitat ensemble of all individuals. Thus, even in the presence of equal sampling ratios for all individuals, individual-specific intercepts may still vary due to differences in the distribution of habitat covariates within each individual’s home range (*e. g.*, varying availability of woodland). Importantly, in the same way that the intercept is used to *condition on* habitat availability at the current position of an individual in an SSF, the intercept conditions on the habitat availability in the home range of the respective individual in an RSF. As a consequence, we recommend that individual-specific intercepts should *not* be shrunk towards an overall mean, but instead should also be given a large, fixed prior variance just like the stratum-specific intercepts in SSF models in Section 3.3.

4 Applications

The code and data for all analyses in this Section are available at the Data Repository of the University of Minnesota (Muff et al., 2019).

4.1 Habitat selection of mountain goats: an RSF analysis

To reiterate the problems with fitting random intercept-only models, we considered data collected from GPS-collared mountain goats in British Columbia, previously analyzed by Lele and Keim (2006) with fixed-effects-only models, and available in the `Resource-Selection` R package (Lele et al., 2017). This dataset consists of use and availability

418 locations for each of 10 different mountain goats, with a use to available ratio of 1:2 for
419 each goat, and a total number of 6338 used points. Although such a low use to available
420 ratio is generally considered inadequate (see *e. g.*, Northrup et al., 2013), we employ the
421 example here purely for illustration purposes. We first fit a RSF containing a single
422 predictor, elevation (centered and scaled to have mean 0 and sd 1) along with a random
423 intercept (variance not fixed) for each goat. The model was fit with an unweighted lo-
424 gistic regression using `glmmTMB::glmmTMB()`, and returned a variance estimate for the
425 among-animal variability in intercepts very close to 0 (Table 1, model M1), reflecting
426 that the differences in the intercepts are mainly determined by the use to available ratio,
427 as pointed out in Sections 3.1 and 3.4. Interestingly, a variance estimate of exactly 0 was
428 obtained when using default settings in the `lme4::glmer()` function (results not shown),
429 reflecting the challenge of estimating such a small variance.

430 We next considered RSFs that included elevation plus a centered and scaled measure
431 of aspect, and compared the estimates from a random intercept-only model (model M2)
432 to those from a model containing independent random intercepts and slopes (model M3),
433 both fit with `glmmTMB()`. In model M3, the standard errors associated with the slope
434 coefficients for aspect and elevation were an order of magnitude larger than when they
435 were not allowed to vary by individual in model M2. These results clearly demonstrate
436 the problems noted by Schielzeth and Forstmeier (2009), namely that random intercept-
437 only models tend to underestimate standard errors of (potentially biased) fixed effects
438 parameters. Finally, we fit the weighted logistic regression model (using $W = 1000$) with
439 random intercept and slopes, with fixed intercept variance at 10^6 (model M4), because
440 this is the procedure we recommend. Weighting the likelihood and fixing the variance of
441 the intercepts in M4 led to a noticeable increase in the estimate of β_{ele} and a decrease in
442 the estimate of σ_{ele}^2 with respect to the unweighted model, while it had little effect on the
443 estimated values of β_{asp} and σ_{asp}^2 . Very similar results to model M4 were obtained when
444 we carried out a Bayesian analysis using R-INLA, and also when the model was fit with

445 an intercept variance that could be freely estimated (results therefore not shown, but see
446 data and code for all analyses).

447 **4.2 Habitat selection of otters: an SSF analysis**

448 We reanalyzed data collected and presented by Weinberger et al. (2016) involving nine
449 radio-collared otters that were tracked between six months and three years in the Eu-
450 ropean Alps. To fit SSFs to these data, each observed location was matched with nine
451 random (available) points generated by resampling step lengths and turning angles from
452 their empirical distribution (Fortin et al., 2005). Due to the absence of an efficient alter-
453 native, the original analysis was performed with a two-step estimation method provided
454 by the `TwoStepCLogit::Ts.estim()` function. The original model included 12 covariates
455 and random effects for all of them. Here, however, we only included the variables of main
456 interest, namely the factorial covariate *habitat type* (with levels *main discharge*, *reservoir*
457 and *residual water*), and the continuous variable *river width*. Moreover, because Forester
458 et al. (2009) showed that the addition of a distance function to the linear predictor is
459 required to reduce bias in the parameter estimators, we included *step length* as an ad-
460 ditional covariate. The data contained a total of 41 670 data points with 4 167 realized
461 steps, where the latter thus corresponds to the number of strata.

462 For illustration, we started by fitting fixed effects-only models. To this end, the well
463 established stratified Cox model was fit via the `survival::clogit()` function. The
464 respective results were compared to the outcome from the conditional Poisson model as
465 given by equation (3), where the stratum-specific intercepts are implicitly estimated by
466 modelling them as a random intercept with a fixed variance $\alpha_{nt} \sim \mathbf{N}(0, 10^6)$; we also
467 re-ran the models with $\alpha_{nt} \sim \mathbf{N}(0, 10^{12})$ to verify that results were robust to this choice.
468 We estimated the parameters both with the frequentist approach using `glmmTMB`, and
469 with the Bayesian approach using `R-INLA`, with independent $\beta \sim \mathbf{N}(0, 10^4)$ priors for all
470 components in the vector of slope parameters. This led to parameter estimates that were

471 essentially indistinguishable from those obtained via the stratified Cox model (Table 2),
472 illustrating that the conditional Poisson model is equivalent to the conditional logistic
473 model, and that we can circumvent the estimation of the stratum-specific fixed intercepts
474 by a random effect with large fixed variance. Note that this equivalence does not hold
475 when σ_α^2 is freely estimated instead, and that this would lead to invalid results, as will be
476 illustrated in the simulation below (Section 4.3). Computation times were on the order
477 of a few seconds for all procedures.

478 Next, we included independent individual-specific random slopes for all covariates
479 (except for step length). We again estimated parameters with `glmmTMB` and `R-INLA`, using
480 the conditional Poisson model (5). For the Bayesian model, the same priors as above were
481 used for the fixed effects and the intercept α_{nt} . In addition, penalized complexity (PC)
482 priors $\text{PC}(3, 0.05)$ were assigned to the precisions of the remaining random slopes (note
483 that priors in the Bayesian framework are typically given to precisions, not variances),
484 but results were insensitive to this choice. PC priors were recently proposed as robust
485 and intuitive alternatives to inverse gamma priors, and were shown to have excellent
486 robustness properties with respect to the choice of their hyperprior parameters (Simpson
487 et al., 2017). PC priors are parameterized as $\text{PC}(u, \alpha)$, where the interpretation of the
488 parameters (u, α) is that $\Pr(\sigma > u) = \alpha$ for the standard deviation σ , thus the user can
489 specify how likely it is ($0 < \alpha < 1$) that σ is larger than a specific value $u > 0$.

490 Results from the conditional Poisson models were compared to the outcome of the
491 two-step procedure via `Ts.estim()`, where it was also assumed that the random effects
492 were independent. These results (Table 2) illustrate two important points: First, the
493 inclusion of individual-specific random slopes in the Poisson regression model leads to
494 different parameter estimates and to much larger standard errors for the slope estimates
495 than when fixed effects-only models are used, which again confirms that fixed effects-only
496 models tend to give overly precise standard errors and biased estimators of regression pa-
497 rameters in the presence of inter-individual heterogeneity. And second, the reformulation

498 of the conditional logistic regression model as a Poisson model with random stratum-
499 specific intercept, as given in (5), leads to feasible estimation of mixed effects parameters
500 in a single modelling step. While computations with other single-step R procedures,
501 such as adding random effects (frailties) to survival models using `coxme::coxme()`, were
502 unfeasible even when only 1 000 out of the more than 4 000 strata were used (we inter-
503 rupted the sessions after 24h of non-convergence), `glimmTMB()` terminated in roughly 5
504 seconds and `R-INLA` in 70 seconds on an Intel Core i7-6500U 4 x 2.50GHz processor for
505 the full dataset. On the other hand the `Ts.estim()` procedure was still considerably
506 faster (about 0.5 seconds), but we note that the parameter estimates from the approxi-
507 mate two-step procedure are not in very good agreement with those from the (correctly
508 specified) Poisson model, especially for β_{REST} and σ_{REST}^2 . Finally, to illustrate that the
509 two-step procedure fails when at least one individual does not encounter all levels of a
510 factorial variable, we artificially removed all strata that contained either used or available
511 points falling in residual water for the individual that had the fewest visits to this habitat
512 type (a total of 12 strata were removed). As expected, the `Ts.estim()` procedure could
513 not be run, while stable results were obtained from fitting the Poisson model.

514 **4.3 Simulation analysis of an SSF design**

515 To more systematically compare different estimation approaches for SSFs, we simulated
516 and analyzed data with known true coefficient values. The simulation of movement tracks
517 involved two continuous covariates: *elevation* and *habitat*. We simulated elevation and
518 habitat as independent unconditional Gaussian Random Fields (GRF; as implemented
519 in Ribeiro Jr and Diggle, 2016) with range $\sigma^2 = 0.1$ and a partial sill of $\phi = 50$ to obtain
520 smooth and realistic surfaces for the two covariates. Each setup was replicated 500 times
521 to obtain a sampling distribution of the estimated coefficients and to investigate bias and
522 variance of the different estimators.

523 We simulated movements of 20 animals according to a biased random walk starting at

524 the center of the landscape at time $t = 0$. To find the position at time $t + 1$, each animal n
 525 was given 200 candidate locations, where the coordinates for each candidate location were
 526 determined by drawing a random step length from an exponential distribution with rate
 527 parameter $\lambda = 1$, and a random turning-angle from a uniform distribution. One candidate
 528 location was then selected at random with probability proportional to $\exp(\boldsymbol{\beta}^\top \mathbf{x})$, where
 529 \mathbf{x} are the covariate values at the end point of each potential step, and $\boldsymbol{\beta}^\top = (-4, 4)$
 530 was the vector of selection coefficients. Animals were assigned individual-specific slopes
 531 for both variables, generated from uncorrelated Gaussian distributions with mean $\boldsymbol{\beta}$ and
 532 variances $\sigma_{ele}^2 = 10$ and $\sigma_{hab}^2 = 5$. For each animal, we simulated 200 time steps, and each
 533 observed step was paired with 9 random (control) steps. Following Forester et al. (2009),
 534 we generated random steps with step lengths from an exponential distribution with rate
 535 $\lambda = 1/(2\bar{l})$, with \bar{l} equal to the mean realized step length, and with the direction of
 536 random steps drawn from a uniform distribution distribution of turning angles between
 537 $-\pi$ and π . We then included step length (l) in the linear predictor to correct for the
 538 bias due to the way we generated random step lengths (*i. e.*, exponential with $\lambda = 1/(2\bar{l})$)
 539 rather than $\lambda = 1$).

540 These data were analyzed with the mixed conditional Poisson model of equation
 541 (5) using R-INLA and `glmmTMB` including random slopes for elevation and habitat. The
 542 variance of the stratum-specific intercept was fixed to $\sigma_\alpha^2 = 10^6$. To illustrate that fixing
 543 this variance is important, we also fit the same model with σ_α^2 estimated instead (only
 544 with `glmmTMB` to avoid redundancy). For INLA we used $\mathbf{N}(0, 10^3)$ priors on the fixed
 545 effects, and $1/\sigma_{ele}^2 \sim \text{PC}(10, 0.01)$ and $1/\sigma_{hab}^2 \sim \text{PC}(5, 0.01)$ priors on the precisions of
 546 the random effects. As a comparison, we also estimated regression parameters using
 547 the two-step approach implemented in `Ts.estim()` assuming independent slopes, and fit
 548 fixed-effects models with Cox models using the `clogit()` function.

549 The Poisson models with fixed σ_α^2 fit with R-INLA and `glmmTMB` retrieved consistent
 550 estimators of the fixed-effects parameters, and the two-step estimator was also nearly

551 unbiased (Figure 1). This was not true, however, when the stratum-specific intercept
552 variance was estimated by the model rather than fixed at 10^6 , in which case all estima-
553 tors were heavily biased. Importantly, we also observe that ignoring random effects leads
554 to biased estimators of fixed-effects parameters when, like here, there is inter-individual
555 heterogeneity in the slopes. All variance estimators were slightly underestimated for all
556 methods, namely because the step-length variable in the predictor absorbs some of the
557 variability in the selection coefficients. In fact, we were able to obtain less biased variance
558 estimators when we omitted the step-length variable (see Figure S2 in the Appendix).
559 The impact of including step-length in the linear predictor on the variance estimators
560 is interesting and unexpected, and it is an apparent contrast to Forester et al. (2009),
561 where the inclusion of step-length is recommended to avoid bias in fixed-effect param-
562 eters. This trade-off between bias in the estimators of fixed effect parameters and the
563 variance parameters deserves more attention in future research. Finally, as pointed out in
564 Section 3.3, weighted regression models resulted in biased estimators except for very large
565 numbers of random steps per stratum (Figure S1 in the Appendix); therefore, weighted
566 alternatives were not further investigated here.

567 **5 Discussion**

568 Recent technological advances have made it possible track a wider range of species for
569 longer durations, leading to an explosion of high-temporal resolution location data (Kays
570 et al., 2015). For example, Movebank, an online platform for storing, managing, and shar-
571 ing data now includes about 1.2 billion locations from over 5500 studies of 850 different
572 taxa (Kranstauber et al., 2011; Wikelski and Kays, 2018). The widespread availability
573 of fine-scale temporal data is fueling the development of new statistical approaches for
574 modelling animal movement data (*e. g.*, Hooten et al., 2017; Jonsen et al., 2018) and also
575 provides unique opportunities to study among-individual variability in movement and

576 habitat-selection patterns. Step-selection functions are appealing because they provide
577 an objective approach to determining habitat availability based on movement character-
578 istics of the study species (Fortin et al., 2005; Thurfjell et al., 2014). Although fitting
579 step-selection models to individual animals is straightforward, efficient estimation proce-
580 dures for models fit to multiple animals have been lacking, hindering our ability to quan-
581 tify among-animal variability in their habitat-selection patterns. Mixed-effects models
582 are an attractive option, but these models are well acknowledged to be computationally
583 challenging to fit in this context (Duchesne et al., 2010).

584 We proposed to fit RSFs and SSFs in a unified, standard GLMM framework, which
585 is possible by combining three statistical results. First, we make use of the fact that the
586 conditional logistic regression model, which needs to be fit to derive SSFs, is actually
587 a multinomial model, and as such it is likelihood-equivalent to a Poisson model. This
588 renders mixed-effects modelling for SSFs equivalent to fitting any Poisson GLMM, which
589 implies that incorporating individual-specific variation in SSFs is no more challenging
590 than doing so for RSFs. Second, because individual- or stratum-specific intercepts are not
591 actually of interest in RSFs or SSFs, and because they are determined by sampling ratios
592 and habitat availability, these intercepts should be treated as fixed effects, or equivalently
593 and more efficiently, as random effects with large, fixed variance. Doing so prevents these
594 intercept parameters from being shrunk towards the overall mean. The magnitude of the
595 shrinkage, and hence bias, may be minimal for RSFs that include many observations for
596 each individual (as in the goat example of Section 4.1), but can be substantial for SSFs
597 which tend to include only a few observations in each stratum (Figure 1). And third,
598 we reiterated that the logistic regression likelihood to estimate RSFs should always be
599 weighted with a large weight W on the available points, in order to ensure convergence
600 to the IPP likelihood which is guaranteed for $W \rightarrow \infty$ (Fithian and Hastie, 2013).

601 Fixing the individual- or stratum-specific intercept variance is particularly straightfor-
602 ward in a Bayesian framework, where the user is required to specify priors on all unknown

603 parameters. To ensure efficient Bayesian inference we have relied on the INLA approach
604 via the R-INLA interface. Of course, all models discussed here can also be approached via
605 MCMC sampling, although this may be very inefficient. We include an MCMC imple-
606 mentation of an SSF analysis to fisher (*Pekania pennanti*) data using the Stan language
607 (Carpenter et al., 2017) in the data repository that accompanies this article. For that
608 example, Stan required an order of magnitude more time to converge than INLA (Stan
609 ≈ 38 min for two parallel chains with 2000 iterations each, INLA ≈ 1 min). Users that
610 prefer frequentist inference should choose a software package that allows to fix a random
611 effect variance to a prespecified value. Here, we fit these models using `glmmTMB`, which
612 provides fast inference, and has previously proven useful for analyzing large telemetry
613 data sets (Jonsen et al., 2018). Table 3 gives an overview of models and procedures that
614 we recommend for efficient and accurate inference on either fixed-effects or random-effects
615 RSFs and SSFs.

616 Prior to now, fitting random coefficient SSFs was often only computationally feasible
617 via two-step procedures that combine estimates of individual-specific habitat-selection in-
618 ference (Craiu et al., 2011), a strategy what was proposed for habitat-selection inference
619 more generally (*e. g.*, Fieberg et al., 2010; Hooten et al., 2016). An advantage of using
620 `Ts.estim` is that it is typically much faster than `glmmTMB` or R-INLA, as illustrated by
621 the computation times of the otter data analysis in Section 4.2. However, it must be kept
622 in mind that `Ts.estim` is an approximate procedure that does not guarantee consistent
623 results, and that it may fail to converge or even does not run, for example when at least
624 one animal does not encounter all habitat types. Moreover, it might be worth noting that
625 `Ts.estim` does not return any information-theoretic measures like AIC, BIC, or DIC to
626 help guide model selection. Still, for very large datasets and models, where GLMMs may
627 demand too much computational power, it certainly remains a convenient and efficient al-
628 ternative. Regarding efficiency, we have also seen that frequentist analyses with `glmmTMB`
629 can be considerably faster than the Bayesian route using R-INLA. In fact, efficiency gain

630 will rarely be the reason to choose Bayesian over likelihood inference. An interesting
631 benefit of Bayesian procedures is that they give (marginal) posterior distributions of
632 all parameters, whereas frequentist approaches usually only return point estimates and
633 standard errors for fixed effect parameters, but no measures of uncertainty for variance
634 parameters (although `glmmTMB` is a notable exception). In addition, various modelling
635 extensions, such as spatial or temporal dependencies (*e. g.*, Lindgren et al., 2011) or mea-
636 surement error in covariates (*e. g.*, Muff et al., 2015) are often much more straightforward
637 to incorporate, or even only computationally feasible, in a Bayesian setup.

638 Although the importance of including random coefficients in regression models of
639 habitat-selection studies has been stressed repeatedly (Gillies et al., 2006; Duchesne et al.,
640 2010), our literature review suggests that random-effects models are often understood as
641 models that merely include a random intercept. Here we have reiterated and illustrated
642 that such practice may lead to too high confidence in results that are potentially biased.
643 By providing coded examples using `R-INLA` and `glmmTMB`, we hope to make efficient esti-
644 mation of RSFs and SSFs with random effects accessible to anyone in the field. SSFs with
645 individual-specific coefficients are particularly attractive since they can provide insights
646 into movement and habitat-selection processes at fine-spatial and temporal scales (Avgar
647 et al., 2016; Signer et al., 2019), but these models had previously been very challenging
648 to fit.

649 **Acknowledgements**

650 JF received partial support from the Minnesota Agricultural Experimental Station and
651 the McKnight Foundation and SM was supported by the Faculty of Science of the Uni-
652 versity of Zurich. The authors would like to thank Irene Weinberger for the permission
653 to use the Eurasian otter dataset, and Benedikt Gehr for comments on an earlier version
654 of the manuscript. We also thank three reviewers for their thoughtful input that greatly

655 helped to improve this article.

656 **Author's contributions**

657 JF, SM and JS conceived the research idea, SM developed the statistical framework,
658 SM, JF and JS conceived of the design and analysis of the data, JS developed and ran
659 the simulations, JS conducted the literature review, SM and JF led the writing of the
660 manuscript. All authors contributed critically to the drafts and gave final approval for
661 publication.

662 **Data accessibility**

663 Data and code for all examples and simulations presented here will be archived with the
664 Data Repository of the University of Minnesota (<https://www.lib.umn.edu/datamanagement/drum>)
665 upon acceptance of the paper (DOI: 10.13020/8bhv-dz98, see Muff et al., 2019). The
666 repository also includes a short tutorial including data not presented here.

Model	$\hat{\beta}_{\text{ele}}$	$\hat{\beta}_{\text{asp}}$	$\hat{\sigma}_{\text{intercept}}^2$	$\hat{\sigma}_{\text{ele}}^2$	$\hat{\sigma}_{\text{asp}}^2$
M1 (Random intercept)	0.12 (0.05)		0.008		
M2 (Random intercept)	0.14 (0.03)	0.52 (0.02)	0.013		
M3 (Random intercept + slopes)	0.07 (0.38)	0.66 (0.11)	0.96	1.40	0.10
M4 (Random intercept + slopes)	0.12 (0.31)	0.65 (0.11)		0.93	0.12

Table 1: Estimates for the slopes of elevation ($\hat{\beta}_{\text{ele}}$) and aspect ($\hat{\beta}_{\text{asp}}$), and for the variances of the random effects ($\hat{\sigma}_{\text{intercept}}^2$, $\hat{\sigma}_{\text{ele}}^2$, $\hat{\sigma}_{\text{asp}}^2$) from four models fit to data from 10 GPS-collared mountain goats. Models M1 – M3 were fit with an unweighted likelihood. Model M4, which is the recommended model, was fit with weighted logistic regression ($W = 1000$) and fixed intercept variance ($\sigma_{\text{int}}^2 = 10^6$). All models were fit using `glmmTMB()`.

Slope estimates	$\hat{\beta}_{\text{RESE}}$	$\hat{\beta}_{\text{REST}}$	$\hat{\beta}_{\text{Width}}$
I. Fixed effects models			
clogit	-0.07 (0.07)	-0.38 (0.10)	0.16 (0.04)
cPois (INLA)	-0.07 (0.07)	-0.38 (0.10)	0.16 (0.04)
cPois (glmmTMB)	-0.07 (0.07)	-0.39 (0.10)	0.16 (0.04)
II. Mixed effects models (random intercept & slopes)			
Two-step	0.04 (0.17)	-0.24 (0.24)	0.10 (0.12)
cPois (INLA)	0.02 (0.18)	-0.33 (0.22)	0.11 (0.14)
cPois (glmmTMB)	-0.004 (0.14)	-0.35 (0.16)	0.12 (0.11)
Variance estimates (Mixed models only)	$\hat{\sigma}_{\text{RESE}}^2$	$\hat{\sigma}_{\text{REST}}^2$	$\hat{\sigma}_{\text{Width}}^2$
Two-step	0.17	0.35	0.08
cPois (INLA)	0.08 (0.02,0.78)	0.10 (0.03,1.02)	0.05 (0.02,0.47)
cPois (glmmTMB)	0.07 (0.01,0.64)	0.10 (0.01,1.12)	0.07 (0.02,0.28)

Table 2: Estimated slopes for reservoir ($\hat{\beta}_{\text{RESE}}$), residual water ($\hat{\beta}_{\text{REST}}$) and river width ($\hat{\beta}_{\text{width}}$) and for the corresponding variance parameters of the Eurasian otter example when using the Cox model (clogit), the Poisson model with stratum-specific intercept (cPois) fit with R-INLA or glmmTMB(), and the two-step procedure `Ts.estim()` (Two-step). For the INLA output, posterior means are given for the slope estimates, and posterior modes for the variances. Values in parentheses are standard errors (for the slope estimates) and 95% credible intervals (for the variances); `Ts.estim()` does not provide measures of uncertainty for variance parameters.

	RSF designs	SSF designs
Example	Mountain goats (sec. 6.1)	Eurasian otters (sec. 6.2)
Fixed effects	Models: Logistic regression	Conditional Poisson regression (model (3) in text)
	R procedures: <code>inla()</code> , <code>glm()</code> , <code>glmmTMB()</code>	<code>clogit()</code> function or <code>inla()/glmmTMB()</code> for Poisson models with stratum-specific random effect and large fixed variance σ_α^2 .
Mixed effects	Models: Mixed logistic regression	Mixed conditional Poisson regression (model (5) in text)
	R procedures: <code>inla()</code> , <code>glmer()</code> , <code>glmmTMB()</code>	<code>inla()</code> , <code>glmmTMB()</code> , <code>Ts.estim()</code>

Table 3: Overview of sampling designs and procedures in R that we recommend for efficient computation. Note that we recommend to carry out RSF analyses using the *infinitely weighted* version, while unweighted regression is recommended for SSFs.

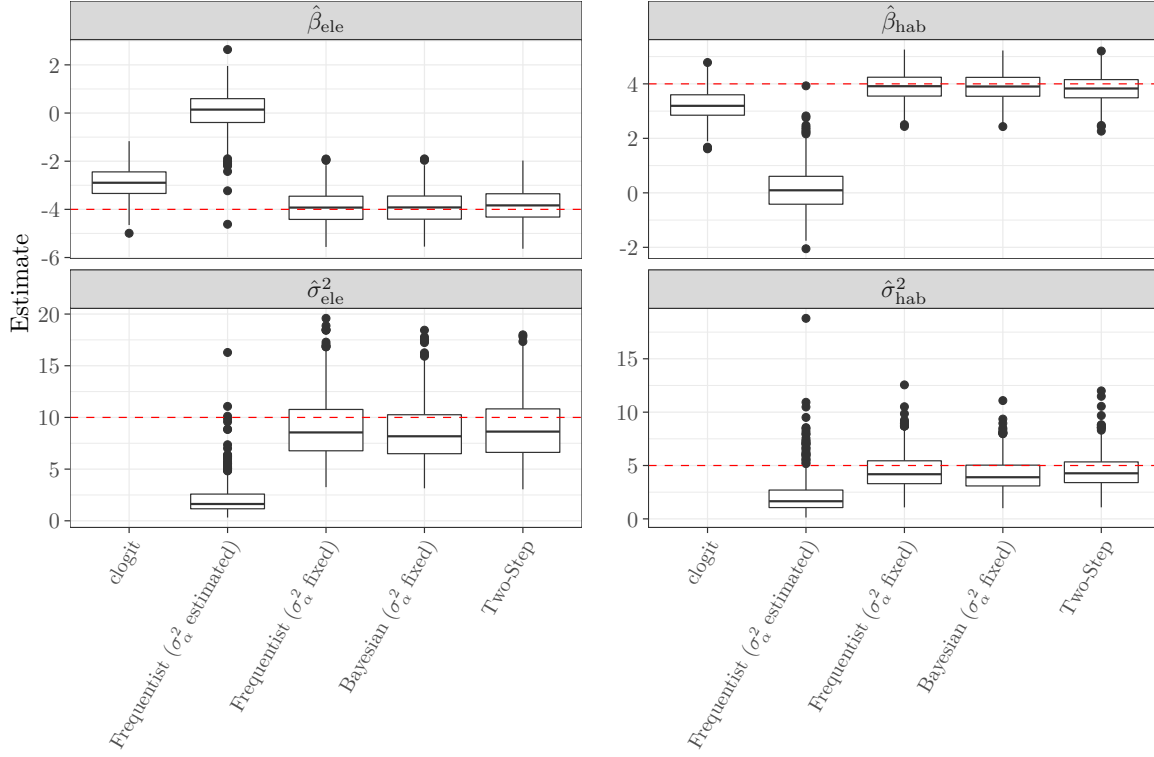


Figure 1: Sampling distribution for estimated SSF coefficients from conditional logistic regression without random effects using the `clogit()` function, from the conditional Poisson regression model with random coefficients using either a frequentist (`glmmTMB`) or a Bayesian approach (`R-INLA`), and from a two-step approach implemented in the `Ts.estim()` function. In the Bayesian case, the estimates are the posterior means for the fixed effects and the posterior modes for the variances. The frequentist approach was implemented both with $\sigma_\alpha^2 = 10^6$ fixed (as recommended) or by estimating σ_α^2 (for illustration). Boxplots show the distribution of the estimates from 500 replications. Variance estimates $\hat{\sigma}_{hab}^2 > 20$ were removed for better visibility (only affects frequentist with σ_α^2 estimated). The horizontal red dashed lines indicated the true value used for the simulations.

References

- 667
668 Aarts, G., J. Fieberg, S. Brasseur, and J. Matthiopoulos (2013). Quantifying the effect of
669 habitat availability on species distributions. *Journal of Animal Ecology* 82, 1135–1145.
- 670 Aarts, G., J. Fieberg, and J. Matthiopoulos (2012). Comparative interpretation of count,
671 presence–absence and point methods for species distribution models. *Methods in Ecology and Evolution* 3, 177–187.
- 672
- 673 Arthur, S. M., B. F. Manly, L. L. McDonald, and G. W. Garner (1996). Assessing habitat
674 selection when availability changes. *Ecology* 77, 215–227.
- 675 Avgar, T., J. R. Potts, M. A. Lewis, and M. S. Boyce (2016). Integrated step selection
676 analysis: bridging the gap between resource selection and animal movement. *Methods in Ecology and Evolution* 7, 619–630.
- 677
- 678 Bates, D., M. Mächler, B. Bolker, and S. Walker (2015). Fitting linear mixed-effects
679 models using lme4. *Journal of Statistical Software* 67, 1–48.
- 680 Beyer, H. L., D. T. Haydon, J. M. Morales, J. L. Frair, M. Hebblewhite, M. Mitchell,
681 and J. Matthiopoulos (2010). The interpretation of habitat preference metrics under
682 use–availability designs. *Philosophical Transactions of the Royal Society B: Biological Sciences* 365(1550), 2245–2254.
- 683
- 684 Bolker, B. M., M. E. Brooks, C. J. Clark, S. W. Geange, J. R. Poulsen, M. H. H. Stevens,
685 and J.-S. S. White (2009). Generalized linear mixed models: a practical guide for
686 ecology and evolution. *Trends in Ecology and Evolution* 24, 127–135.
- 687 Bolnick, D. I., R. Svanbäck, J. A. Fordyce, L. H. Yang, J. M. Davis, C. D. Hulsey,
688 and M. L. Forister (2002). The ecology of individuals: incidence and implications of
689 individual specialization. *The American Naturalist* 161(1), 1–28.

690 Brooks, M. E., K. Kristensen, K. J. van Benthem, A. Magnusson, C. W. Berg, A. Nielsen,
691 H. J. Skaug, M. Maechler, and B. M. Bolker (2017). Modeling zero-inflated count data
692 with glmmTMB. *bioRxiv preprint bioRxiv:132753*.

693 Bruun, M. and H. G. Smith (2003). Landscape composition affects habitat use and
694 foraging flight distances in breeding european starlings. *Biological Conservation 114*,
695 179 – 187.

696 Cagnacci, F., L. Boitani, R. A. Powell, and M. S. Boyce (2010). Animal ecology meets
697 gps-based radiotelemetry: a perfect storm of opportunities and challenges.

698 Carpenter, B., A. Gelman, M. Hoffman, D. Lee, B. Goodrich, M. Betancourt,
699 M. Brubaker, J. Guo, P. Li, and A. Riddell (2017). Stan: A probabilistic programming
700 language. *Journal of Statistical Software, Articles 76*, 1–32.

701 Chen, Z. and L. Kuo (2001). A note on the estimation of the multinomial logit model
702 with random effects. *The American Statistician 55*, 89–95.

703 Craiu, R. V., T. Duchesne, and Fortin (2008). Inference methods for the conditional
704 logistic regression model with longitudinal data. *Biometrical Journal 50*, 97–109.

705 Craiu, R. V., T. Duchesne, D. Fortin, and S. Baillargeon (2011). Conditional logistic
706 regression with longitudinal follow-up and individual-level random coefficients: A stable
707 and efficient two-step estimation method. *Journal of Computational and Graphical*
708 *Statistics 20*, 767–784.

709 Craiu, R. V., T. Duchesne, D. Fortin, and S. Baillargeon (2016). *TwoStepCLogit: Con-*
710 *ditional Logistic Regression: A Two-Step Estimation Method*. R package version 1.2.5.

711 Dingemanse, N. J. and N. A. Dochtermann (2013). Quantifying individual variation in
712 behaviour: mixed-effect modelling approaches. *Journal of Animal Ecology 82*, 39–54.

713 Duchesne, T., D. Fortin, and N. Courbin (2010). Mixed conditional logistic regression
714 for habitat selection. *Journal of Animal Ecology* 79, 548–555.

715 Duchesne, T., D. Fortin, and L.-P. Rivest (2015). Equivalence between step selection
716 functions and biased correlated random walks for statistical inference on animal move-
717 ment. *PLOS ONE* 10, e0122947.

718 Elff, M. (2016). *mclogit: Mixed Conditional Logit Models*. R package version 0.4.4.

719 Fieberg, J., J. Matthiopoulos, M. Hebblewhite, M. S. Boyce, and J. L. Frair (2010).
720 Correlation and studies of habitat selection: problem, red herring or opportunity?
721 *Phil. Trans. R. Soc. B* 365, 2233–2244.

722 Fieberg, J., R. H. Rieger, M. C. Zicus, and J. S. Schildcrout (2009). Regression mod-
723 elling of correlated data in ecology: subject-specific and population averaged response
724 patterns. *Journal of Applied Ecology* 46, 1018–1025.

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

727 Fong, Y., H. Rue, and J. Wakefield (2010). Bayesian inference for generalized linear
728 mixed models. *Biostatistics* 11, 397–412.

729 Forester, J. D., H. K. Im, and P. J. Rathouz (2009). Accounting for animal movement
730 in estimation of resource selection functions: sampling and data analysis. *Ecology* 90,
731 3554–3565.

732 Fortin, D., H. Beyer, M. Boyce, and D. Smith (2005). Wolves influence elk movements:
733 behavior shapes a trophic cascade in Yellowstone National Park. *Ecology* 86, 1320–
734 1330.

735 Gaillard, J.-M., M. Hebblewhite, A. Loison, M. Fuller, R. Powell, M. Basille, and
736 B. Van Moorter (2010). Habitat–performance relationships: finding the right metric

737 at a given spatial scale. *Philosophical Transactions of the Royal Society B: Biological*
738 *Sciences* 365(1550), 2255–2265.

739 Gillies, C. S., M. Hebblewhite, S. E. Nielsen, M. A. Krawchuk, C. L. Aldridge, J. L. Frair,
740 D. J. Saher, C. E. Stevens, and C. L. Jerde (2006). Application of random effects to
741 the study of resource selection by animals. *Journal of Animal Ecology* 75, 887–898.

742 Hebblewhite, M. and E. Merrill (2008). Modelling wildlife–human relationships for social
743 species with mixed-effects resource selection models. *Journal of Applied Ecology* 45,
744 834–844.

745 Heithaus, M. R. (2001). Habitat selection by predators and prey in communities with
746 asymmetrical intraguild predation. *Oikos* 92(3), 542–554.

747 Hooten, M. B., F. E. Buderman, B. M. Brost, E. M. Hanks, and J. S. Ivan (2016). Hi-
748 erarchical animal movement models for population-level inference. *Environmetrics* 27,
749 322–333.

750 Hooten, M. B., E. M. Hanks, D. S. Johnson, and M. W. Alldredge (2013). Reconciling
751 resource utilization and resource selection functions. *Journal of Animal Ecology* 82,
752 1146–1154.

753 Hooten, M. B., D. S. Johnson, B. T. McClintock, and J. M. Morales (2017). *Animal*
754 *Movement: Statistical Models for Telemetry Data*. Boca Raton: CRC Press.

755 Hosmer, D. W. and S. Lemeshow (2000). *Applied Logistic Regression (Wiley Series in*
756 *Probability and Statistics)* (2 ed.). Wiley-Interscience Publication.

757 Hurlbert, S. H. (1984). Pseudoreplication and the design of ecological field experiments.
758 *Ecological monographs* 54(2), 187–211.

759 Jaenike, J. and R. D. Holt (1991). Genetic variation for habitat preference: evidence and
760 explanations. *The American Naturalist* 137, S67–S90.

761 Johnson, D. H. (1980). The comparison of usage and availability measurements for
762 evaluating resource preference. *Ecology* 61, 65–71.

763 Jonsen, I., C. McMahon, T. Patterson, M. Auger-Methe, R. Harcourt, M. Hindell, and
764 S. Bestley (2018). Movement behaviour responses to environment: fast inference of
765 individual variation with a mixed effects model. *bioRxiv preprint bioRxiv:314690*.

766 Kays, R., M. C. Crofoot, W. Jetz, and M. Wikelski (2015). Terrestrial animal tracking
767 as an eye of life and planet. *Science* 348, aaa2478.

768 Kranstauber, B., A. Cameron, R. Weinzerl, T. Fountain, S. Tilak, M. Wikelski, and
769 R. Kays (2011). The Movebank data model for animal tracking. *Environmental Mod-
770 elling & Software* 26, 834–835.

771 Leclerc, M., E. Vander Wal, A. Zedrosser, J. E. Swenson, J. Kindberg, and F. Pelletier
772 (2016). Quantifying consistent individual differences in habitat selection. *Oecologia* 180,
773 697–705.

774 Lee, Y. and J. A. Nelder (2004). Conditional and marginal models: another view. *Sta-
775 tistical Science* 19, 219–238.

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

778 Lele, S. R., J. L. Keim, and P. Solymos (2017). *ResourceSelection: Resource Selection
779 (Probability) Functions for Use-Availability Data*. R package version 0.3-2.

780 Lesmerises, R. and M.-H. St-Laurent (2017). Not accounting for interindividual variability
781 can mask habitat selection patterns: a case study on black bears. *Oecologia* 185(3),
782 415–425.

783 Lindgren, F., H. Rue, and J. Lindström (2011). An explicit link between Gaussian fields

784 and Gaussian Markov random fields: The SPDE approach (with discussion). *Journal*
785 *of the Royal Statistical Society. Series B (Statistical Methodology)* 73, 423 – 498.

786 Lunn, D., A. Thomas, N. Best, and D. Spiegelhalter (2000). WinBUGS – A Bayesian
787 modelling framework: concepts, structure, and extensibility. *Statistics and Comput-*
788 *ing* 10, 325–337.

789 Magnusson, A., H. J. Skaug, A. Nielsen, C. W. Berg, K. Kristensen, M. Maechler,
790 K. J. van Benthem, B. M. Bolker, and M. E. Brooks (2017). *glmmTMB: General-*
791 *ized Linear Mixed Models using Template Model Builder*. R package version 0.1.3.
792 <https://github.com/glmmTMB>.

793 Manly, B. F., L. L. McDonald, D. L. Thomas, T. L. McDonald, and W. P. Erickson (2002).
794 *Resource Selection by Animals: Statistical Design and Analysis for Field Studies* (2
795 ed.). Dordrecht, The Netherlands: Kluwer Academic Publishers.

796 Matthiopoulos, J., J. Fieberg, G. Aarts, H. L. Beyer, J. M. Morales, and D. T. Hay-
797 don (2015). Establishing the link between habitat selection and animal population
798 dynamics. *Ecological Monographs* 85, 413–436.

799 Matthiopoulos, J., M. Hebblewhite, G. Aarts, and J. Fieberg (2011). Generalized func-
800 tional response for species distributions. *Ecology* 92, 583–589.

801 McCullagh, P. and J. A. Nelder (1989). *Generalized Linear Models*. London: Chapman
802 and Hall.

803 McDonald, T. L., B. F. J. Manly, R. M. Nielson, and L. V. Diller (2006). Discrete-choice
804 modelling in wildlife studies exemplified by northern spotted owl nighttime habitat
805 selection. *The Journal of Wildlife Management* 70, 375–383.

806 Millsaugh, J. J., R. M. Nielson, L. McDonald, J. M. Marzluff, R. A. Gitzen, C. D.
807 Rittenhouse, M. W. Hubbard, and S. L. Sheriff (2006). Analysis of resource selection
808 using utilization distributions. *Journal of Wildlife Management* 70, 384–395.

- 809 Muff, S., L. Held, and L. F. Keller (2016). Marginal or conditional regression models for
810 correlated non-normal data? *Methods in Ecology and Evolution* 7, 1514–1524.
- 811 Muff, S., A. Riebler, L. Held, H. Rue, and P. Saner (2015). Bayesian analysis of mea-
812 surement error models using integrated nested Laplace approximations. *Journal of the*
813 *Royal Statistical Society, Applied Statistics Series C* 64, 231–252.
- 814 Muff, S., J. Signer, and J. Fieberg (2019). R code and output supporting: Accounting for
815 individual-specific variation in habitat-selection studies: Efficient estimation of mixed-
816 effects models using bayesian or frequentist computation. University of Minnesota
817 Digital Conservancy. 10.13020/8bhv-dz98.
- 818 Mysterud, A. and R. A. Ims (1998). Functional responses in habitat use: availability
819 influences relative use in trade-off situations. *Ecology* 79, 1435–1441.
- 820 Northrup, J. M., M. B. Hooten, C. R. Anderson, and G. Wittemeyer (2013). Practical
821 guidance on characterizing availability in resource selection functions under a use-
822 availability design. *Ecology* 94, 1456–1463.
- 823 Phillips, S. J., R. P. Anderson, and R. E. Schapire (2006). Maximum entropy modeling
824 of species geographic distributions. *Ecological Modelling* 190, 231 – 259.
- 825 R Core Team (2018). *R: A Language and Environment for Statistical Computing*. Vienna,
826 Austria: R Foundation for Statistical Computing.
- 827 Raynor, E. J., H. L. Beyer, J. M. Briggs, and A. Joern (2017). Complex variation in
828 habitat selection strategies among individuals driven by extrinsic factors. *Ecology and*
829 *evolution* 7(6), 1802–1822.
- 830 Renner, I. W. and D. I. Warton (2013). Equivalence of MAXENT and Poisson point
831 process models for species distribution modeling in ecology. *Biometrics* 69, 274–281.

- 832 Rhodes, J. R., C. A. McAlpine, D. Lunney, and H. P. Possingham (2005). A spatially
833 explicit habitat selection model incorporating home range behavior. *Ecology* 86, 1199–
834 1205.
- 835 Ribeiro Jr, P. J. and P. J. Diggle (2016). *geoR: Analysis of Geostatistical Data*. R package
836 version 1.7-5.2.
- 837 Robinson, G. K. (1991). That blup is a good thing: The estimation of random effects
838 (with discussion). *Statistical Science* 6, 15–51.
- 839 Rosenzweig, M. L. (1991). Habitat selection and population interactions: the search for
840 mechanism. *The American Naturalist* 137, S5–S28.
- 841 Rue, H., S. Martino, and N. Chopin (2009). Approximate Bayesian inference for latent
842 Gaussian models by using integrated nested Laplace approximations (with discussion).
843 *Journal of the Royal Statistical Society. Series B (Methodological)* 71, 319–392.
- 844 Schielzeth, H. and W. Forstmeier (2009). Conclusions beyond support: Overconfident
845 estimates in mixed models. *Behavioral Ecology* 20, 416 – 420.
- 846 Sheppard, C. E., R. Inger, R. A. McDonald, S. Barker, A. L. Jackson, F. J. Thompson,
847 E. I. Vitikainen, M. A. Cant, and H. H. Marshall (2018). Intragroup competition pre-
848 dict individual foraging specialisation in a group-living mammal. *Ecology letters* 21(5),
849 665–673.
- 850 Signer, J., J. Fieberg, and T. Avgar (2019). Animal movement tools (amt): R package
851 for managing tracking data and conducting habitat selection analyses. *Ecology and*
852 *Evolution* 9, 880 – 890.
- 853 Simpson, D., H. Rue, A. Riebler, T. G. Martins, and S. H. Sørbye (2017). Penalising
854 model component complexity: A principled, practical approach to constructing priors.
855 *Statistical Science* 32, 1–28.

- 856 Snijders, T. A. B. and R. J. Bosker (1999). *Multilevel Analysis*. London: Sage.
- 857 Stamps, J. A. (2007). Growth-mortality tradeoffs and ‘personality traits’ in animals.
858 *Ecology letters* 10(5), 355–363.
- 859 Therneau, T. M. (2015a). *coxme: Mixed Effects Cox Models*. R package version 2.2-5.
- 860 Therneau, T. M. (2015b). *A Package for Survival Analysis in S*. Version 2.38.
- 861 Thurfjell, H., S. Ciuti, and M. Boyce (2014). Applications of step-selection functions in
862 ecology and conservation. *Movement Ecology* 2, 4.
- 863 Warton, D. and G. Aarts (2013). Advancing our thinking in presence-only and used-
864 available analysis. *Journal of Animal Ecology* 82, 1125–1134.
- 865 Warton, D. I. and L. C. Shepherd (2010). Poisson point process models solve the “pseudo-
866 absence problem” for presence-only data in ecology. *Annals of Applied Statistics* 4,
867 1383–1402.
- 868 Weinberger, I. C., S. Muff, A. Kranz, and F. Bontadina (2016). Flexible habitat selection
869 paves the way for a recovery of otter populations in the European Alps. *Biological*
870 *Conservation* 199, 88–95.
- 871 Whitehead, J. (1980). Fitting Cox’s regression model to survival data using GLIM.
872 *Applied Statistics* 29, 268–275.
- 873 Wikelski, M. and R. Kays (2018). Movebank: archive, analysis and sharing of
874 animal movement data. Hosted by the Max Planck Institute for Ornithology,
875 <https://www.movebank.org>.