Received: 12 January 2024 | Accepted: 21 May 2024

DOI: 10.1111/2041-210X.14367

Methods in Ecology and Evolution

RESEARCH ARTICLE

Step selection functions with non-linear and random effects

Natasha J. Klappstein[1] | Théo Michelot[1] | John Fieberg[2] | Eric J. Pedersen[3] | Joanna Mills Flemming[1]

[1]Department of Mathematics & Statistics, Dalhousie University, Halifax, Nova Scotia, Canada

[2]Department of Fisheries, Wildlife, and Conservation Biology, University of Minnesota, Minneapolis, Minnesota, USA

[3]Department of Biology, Concordia University, Montreal, Quebec, Canada

Correspondence
Natasha J. Klappstein
Email: nklappstein@dal.ca

Funding information
National Sciences and Engineering Research Council of Canada; National Aeronautics and Space Administration, Grant/Award Number: 80NSSC21K1182; Minnesota Agricultural Experimental Station

Handling Editor: Theoni Photopoulou

Abstract

1. Step selection functions (SSFs) are used to jointly describe animal movement patterns and habitat preferences. Recent work has extended this framework to model inter-individual differences, account for unexplained structure in animals' space use and capture temporally varying patterns of movement and habitat selection.
2. In this paper, we formulate SSFs with penalised smooths (similar to generalised additive models) to unify new and existing extensions, and conveniently implement the models in the popular, open-source `mgcv` R package.
3. We explore non-linear patterns of movement and habitat selection, and use the equivalence between penalised smoothing splines and random effects to implement individual-level and spatial random effects. This framework can also be used to fit varying-coefficient models to account for temporally or spatially heterogeneous patterns of selection (e.g. resulting from behavioural variation), or any other non-linear interactions between drivers of the animal's movement decisions.
4. We provide the necessary technical details to understand several key special cases of smooths and their implementation in `mgcv`, showcase the ecological relevance using two illustrative examples and provide R code to facilitate the adoption of these methods. This paper offers a broad overview of how smooth effects can be applied to increase the flexibility and biological realism of SSFs.

KEYWORDS
animal movement, generalised additive models, habitat selection, integrated step selection analysis, `mgcv`, penalised splines, random effects, step selection functions

1 | INTRODUCTION

Step selection functions (SSFs) are increasingly used to describe animal movement and habitat selection (Avgar et al., 2016; Forester et al., 2009; Rhodes et al., 2005). An SSF measures how an animal selects habitat at the scale of the observed movement step, while simultaneously estimating distributions of step lengths and turning angles. Compared to landscape-level habitat selection models (e.g. resource selection functions; RSFs), the temporal structure of SSFs better accounts for autocorrelation in animal tracking data. SSFs also make it possible to assess time-varying patterns of selection (Forester et al., 2009; Richter et al., 2020) as well as movement-habitat interactions that may be associated with behavioural patterns or landscape resistivity (Avgar et al., 2016). Their dynamic formulation makes them an important tool for predicting animal space use, with methods recently developed to scale SSF results to large-scale distributions (through simulations and mathematical methods; Potts & Börger, 2022; Signer et al., 2017, 2024). Therefore, to best understand animal spatial ecology, many methodological extensions of SSFs have been proposed to improve their biological realism and flexibility.

One important focus of SSF research has been to relax the common assumption that the parameters driving movement and habitat selection are constant through time, to account for behavioural variation. This can for example be done with hidden Markov models, where animals switch between behavioural states, each characterised by different patterns of movement and habitat preferences (Klappstein et al., 2023; Nicosia et al., 2017; Pohle et al., 2023). Another approach is to include interaction effects within an SSF, to specify covariate-dependent selection parameters. For example, habitat selection can vary with time (Forester et al., 2009; Richter et al., 2020), and movement speed can change with environmental resistivity (Avgar et al., 2016). Although this approach does not explicitly model behavioural switching, it provides a very flexible framework to capture time-varying movement and selection patterns. This method could be further extended to allow for non-linear habitat preferences and non-linear interaction effects (e.g. seasonal selection patterns). These ideas have been explored for RSFs (e.g. Dejeante, Valeix, & Chamaillé-Jammes, 2024; McCabe et al., 2021), but not yet for SSFs, which have the ability to assess variability in both habitat selection and movement.

Another direction of recent research has been to include random effects in SSFs, to capture variability otherwise unexplained by the model. Random slopes have been explored to account for inter-individual variability in movement and habitat selection patterns, attributable to animal 'personality' and affected by phenotypic plasticity (Chatterjee et al., 2024; Duchesne et al., 2010; Muff et al., 2020). Similarly, spatial random effects can help account for spatial pattern or variation not captured by other environmental covariates (Arce Guillen et al., 2023). This unexplained spatial variation could reflect an unmeasured driver of movement (e.g. predation risk, conspecific interactions) or centres of attraction (e.g. kill sites, dens), which if unaccounted for, may bias selection parameters for other covariates. In this paper, we will use a very general definition of random effects that includes penalised smooths, which have great flexibility in capturing spatial, group-level and temporal variation in movement and selection patterns (Hodges, 2014; Wood, 2017).

It is most common to implement SSFs using software for fitting conditional logistic regression models (Signer et al., 2019; Therneau, 2022). Typically, this approach has been limited to relatively simple log-linear models (although see `survival` for non-linear p-spline implementation; Therneau, 2022). The extensions mentioned above have been implemented using custom code or various software packages, such as `inlabru` for spatial random effects (Arce Guillen et al., 2023) and `glmmTMB` for random slope models (Muff et al., 2020). However, the increasing diversity of methodological extensions and implementation methods can present a challenge for practitioners, and it may not be easy to explore model formulations in the same inferential framework.

In this paper, we explain how an SSF can be formulated as a 'general smooth model' (a generalised additive model (GAM) without an exponential family distribution; Wood et al., 2016), and can conveniently be implemented in the popular R package `mgcv`. In `mgcv`, a wide range of smooth models can be specified using a convenient formula syntax, and fitted via likelihood-based methods such as restricted maximum likelihood. This framework greatly increases the flexibility of current SSF models, and we will focus on the following: (i) modelling non-linear patterns of selection and non-parametric movement kernels, (ii) capturing inter-group variability (i.e. random slopes and hierarchical smooths), (iii) accounting for unexplained spatial variation via spatial random effects and (iv) assessing selection patterns that change through time and/or space using varying-coefficient models. We will show that this implementation provides a unifying framework to fit, compare and contrast complex SSF formulations with non-linear and random effects.

2 | GENERAL MODEL FORMULATION AND IMPLEMENTATION

2.1 | Step selection functions

Consider a track of two-dimensional animal locations $\{\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_T\}$ collected at regular time intervals. In an SSF, the likelihood of a step ending at location $\mathbf{s}_{t+1}$ given the previous locations $\mathbf{s}_{1:t} = \{\mathbf{s}_1, \mathbf{s}_2, \dots \mathbf{s}_t\}$ is

$$p(\mathbf{s}_{t+1} \mid \mathbf{s}_{1:t}) = \frac{w(\mathbf{s}_t, \mathbf{s}_{t+1}) \phi(\mathbf{s}_{t+1} \mid \mathbf{s}_{1:t})}{\int_{\mathbf{r} \in \Omega} w(\mathbf{s}_t, \mathbf{r}) \phi(\mathbf{r} \mid \mathbf{s}_{1:t}) d\mathbf{r}}, \tag{1}$$

where w measures habitat selection, ϕ describes the movement patterns of the animal, and Ω is the study area (Forester et al., 2009; Rhodes et al., 2005). The denominator is a normalising constant, such that Equation (1) is a probability density function with respect to the endpoint of the step, $\mathbf{s}_{t+1}$. Note that some SSFs do not estimate movement jointly with habitat selection (e.g. Fortin et al., 2005), but we focus on SSFs where both ϕ and w are estimated (sometimes termed an 'integrated' SSF; Avgar et al., 2016; Forester et al., 2009). In this case, it is most common to define both components as log-linear models. Then, the habitat selection function is defined as $w(\mathbf{s}_t, \mathbf{s}_{t+1}) = \exp\{\boldsymbol{\beta}_h^\top \mathbf{c}_h(\mathbf{s}_t, \mathbf{s}_{t+1})\}$ where $\mathbf{c}_h(\mathbf{s}_t, \mathbf{s}_{t+1})$ is a vector of habitat covariates with associated coefficients $\boldsymbol{\beta}_h$. Likewise, the movement kernel is commonly written as $\phi(\mathbf{s}_{t+1} \mid \mathbf{s}_{1:t}) = \exp\{\boldsymbol{\beta}_m^\top \mathbf{c}_m(\mathbf{s}_{1:t}, \mathbf{s}_{t+1})\}$ with movement covariates $\mathbf{c}_m(\mathbf{s}_{1:t}, \mathbf{s}_{t+1})$ and coefficients $\boldsymbol{\beta}_m$. The two exponential functions can be factorised, such that Equation (1) becomes

$$p(\mathbf{s}_{t+1} \mid \mathbf{s}_{1:t}) = \frac{\exp\{\boldsymbol{\beta}^\top \mathbf{c}(\mathbf{s}_{1:t}, \mathbf{s}_{t+1})\}}{\int_{\mathbf{r} \in \Omega} \exp\{\boldsymbol{\beta}^\top \mathbf{c}(\mathbf{s}_{1:t}, \mathbf{r})\} d\mathbf{r}}, \tag{2}$$

where $\mathbf{c}(\mathbf{s}_{1:t}, \mathbf{s}_{t+1}) = (\mathbf{c}_h(\mathbf{s}_t, \mathbf{s}_{t+1}), \mathbf{c}_m(\mathbf{s}_{1:t}, \mathbf{s}_{t+1}))$ and $\boldsymbol{\beta} = (\boldsymbol{\beta}_h, \boldsymbol{\beta}_m)$. Note that this model does not include an intercept, as this would not be identifiable (Manly et al., 2002). Typically $\mathbf{c}_h(\mathbf{s}_t, \mathbf{s}_{t+1})$ will include spatial covariates (e.g. foraging resources, proxies of risk, etc.), and the parameters $\boldsymbol{\beta}_h$ measure the animal's selection for (positive coefficients) or avoidance of (negative coefficients) these environmental features. A common movement model is the correlated random walk, and it can be implemented within the SSF framework by including specific functions of the step length and turning angle as covariates in $\mathbf{c}_m$ (Avgar et al., 2016; Duchesne et al., 2015; Forester et al., 2009). The coefficients $\boldsymbol{\beta}_m$ are then related to the parameters of the estimated distributions of the movement variables. This approach can be used to model turning angles with a von Mises distribution (Duchesne et al., 2015), and there are several options for step lengths, including the exponential, gamma, log-normal and Weibull distributions (Avgar et al., 2016; Forester et al., 2009). SSFs can also include interaction terms (i.e. covariate-dependent coefficients) to capture spatiotemporal variation in movement or habitat selection patterns (Avgar et al., 2016; Forester et al., 2009; Richter et al., 2020). To emphasise flexibility, we write the SSF conditional on all previous locations $\mathbf{s}_{1:t}$; in practice, it usually only depends on the two previous locations $\mathbf{s}_{t-1}$ and $\mathbf{s}_t$, which are used to derive the step length and turning angle.

The normalisation constant of the SSF (i.e. the integral in the denominator) is analytically intractable, and a common approach is to approximate the integral as the sum of function evaluations at random points from some distribution h (Michelot et al., 2024). This approximation is nearly likelihood-equivalent to conditional logisitic regression (CLR), and it can be convenient to re-write the SSF as such. That is, in this approach, we restrict the animal's movement choices to the random and observed points, and we can write the model in terms of the covariate values at these discrete choices. We define a stratum as the observation and random points at each time t. Then $\mathbf{X}_t$ is the design matrix, with columns for evaluations of $\mathbf{c}$ (i.e. habitat and movement variables) and one row for each location in the t-th stratum (i.e. the observation and the N random locations). Assuming the number of random points N is constant over all strata, we can approximate the SSF by

$$\Pr(y_{it} = 1 \mid \mathbf{X}_t) = \frac{\exp(\boldsymbol{\beta}^\top \mathbf{x}_{it})}{\sum_{n=0}^{N} \exp(\boldsymbol{\beta}^\top \mathbf{x}_{nt})}, \tag{3}$$

where $\mathbf{x}_{it}$ is the row of $\mathbf{X}_t$ indexed by i, and

$$y_{it} = \begin{cases} 1 & \text{if the } i\text{-¬th point in stratum } t \text{ is the observed location,} \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

Note that in Equation (3), we include the observation in the denominator, and therefore the sum is over $N+1$ points (Forester et al., 2009). Also, we temporarily ignored the distribution of random points h, and we must account for our sampling via a post hoc correction or 'update' to the parameters (Avgar et al., 2016; Forester et al., 2009). Typically random points will be sampled from step length and turning angle distributions derived from the empirical data (see Section 3.1 for details on sampling and corrections). This approach allows the SSF to be fitted using existing software for CLR (Avgar et al., 2016; Forester et al., 2009), and there are several options for log-linear SSFs, including the R packages `survival` (Therneau, 2022) and `glmmTMB` (Muff et al., 2020). Here, we recognise that Equation (3) is also likelihood-equivalent to a special case of a Cox proportional hazards (Cox PH) model (see Appendix A for details) and can therefore be fitted in the `mgcv` package. This is the same equivalence utilised in `survival`, where CLR can be fitted with `clogit` as a wrapper to `cox.ph` function. Implementation in `mgcv` as a GAM-like general smooth model allows for more flexible formulations with both non-linear and random effects, with model complexity controlled by a data-driven penalty. We describe this modelling framework next.

2.2 | SSFs with smooth effects

Equation (3) can be extended to include both 'conventional' random effects (e.g. random slopes) and non-linear functions, which we will collectively call 'smooth' terms following the terminology used for GAMs (Wood, 2017). We write the general model as

$$\Pr(y_{it} = 1 \mid \mathbf{X}_t, \mathbf{Z}_t) = \frac{\exp(\boldsymbol{\beta}^\top \mathbf{x}_{it} + \boldsymbol{\gamma}^\top \mathbf{z}_{it})}{\sum_{n=0}^{N} \exp(\boldsymbol{\beta}^\top \mathbf{x}_{nt} + \boldsymbol{\gamma}^\top \mathbf{z}_{nt})} \text{ and } \boldsymbol{\gamma} \sim N(\mathbf{0}, \boldsymbol{\Sigma}), \tag{5}$$

where $\mathbf{X}_t$ is the design matrix of fixed effect (unpenalised) terms and $\mathbf{Z}_t$ is the design matrix of smooth terms for the stratum t (with i-th row $\mathbf{z}_{it}$). In this mixed effect formulation, $\boldsymbol{\beta}$ is the vector of fixed effects and $\boldsymbol{\gamma}$ is the vector of random effects, assumed to follow a multivariate normal distribution with covariance matrix $\boldsymbol{\Sigma}$. The choice of a normally distributed $\boldsymbol{\gamma}$ is sometimes described as a prior, reflecting our belief about the random effect variance or function smoothness (which unperpins a Bayesian view of smoothing and uncertainty quantification; Kimeldorf & Wahba, 1970; Miller, 2021). If $\mathbf{Z}_t$ represents basis functions from a single smooth term, then the covariance matrix in Equation (5) is given by $\boldsymbol{\Sigma} = \mathbf{S}^- / \lambda$, where $\mathbf{S}$ is called the penalty matrix of the smoother (with pseudo-inverse $\mathbf{S}^-$) and λ is a smoothing parameter. If $\mathbf{Z}_t$ includes more than one smoother or smoothers are penalised by multiple terms, then $\boldsymbol{\Sigma}^-$ will be a block-diagonal matrix, with the penalty matrices multiplied by penalty-specific smoothing parameters on the matrix diagonal (Wood, 2017).

Equation (5) is very flexible, as mixed effect models provide a convenient framework to describe not only inter-group heterogeneity (e.g. using random slopes) but also non-linear effects using penalised splines and spatial dependence via random fields (Hodges, 2014; Michelot, 2023; Wood, 2017). These formulations correspond to different choices of the covariates in the model matrix $\mathbf{Z}_t$ and of the penalty matrix $\mathbf{S}$ (and therefore the covariance matrix $\boldsymbol{\Sigma}$). We describe random slopes and random intercepts as 'conventional' random effects, but note that throughout the paper we use the term random effects more generally to also describe other smooth effects. Next, we will formalise this equivalence between conventional random effects and penalised smooths, and explain their uses within the SSF framework (see Figure 1 for a summary of key examples). Throughout, we denote a linear predictor for the i-th location of stratum t as $\eta_{it} = \boldsymbol{\beta}^\top \mathbf{x}_{it} + \boldsymbol{\gamma}^\top \mathbf{z}_{it}$ (as in Equation 5).

2.2.1 | Conventional random effects

'Conventional' random effects, which are used to capture differences between individuals or groups, are perhaps the simplest special case of the mixed effect model in Equation (5). SSFs lack an intercept, and so we will focus on defining a model with random slopes to capture variability in habitat selection or movement patterns (Chatterjee

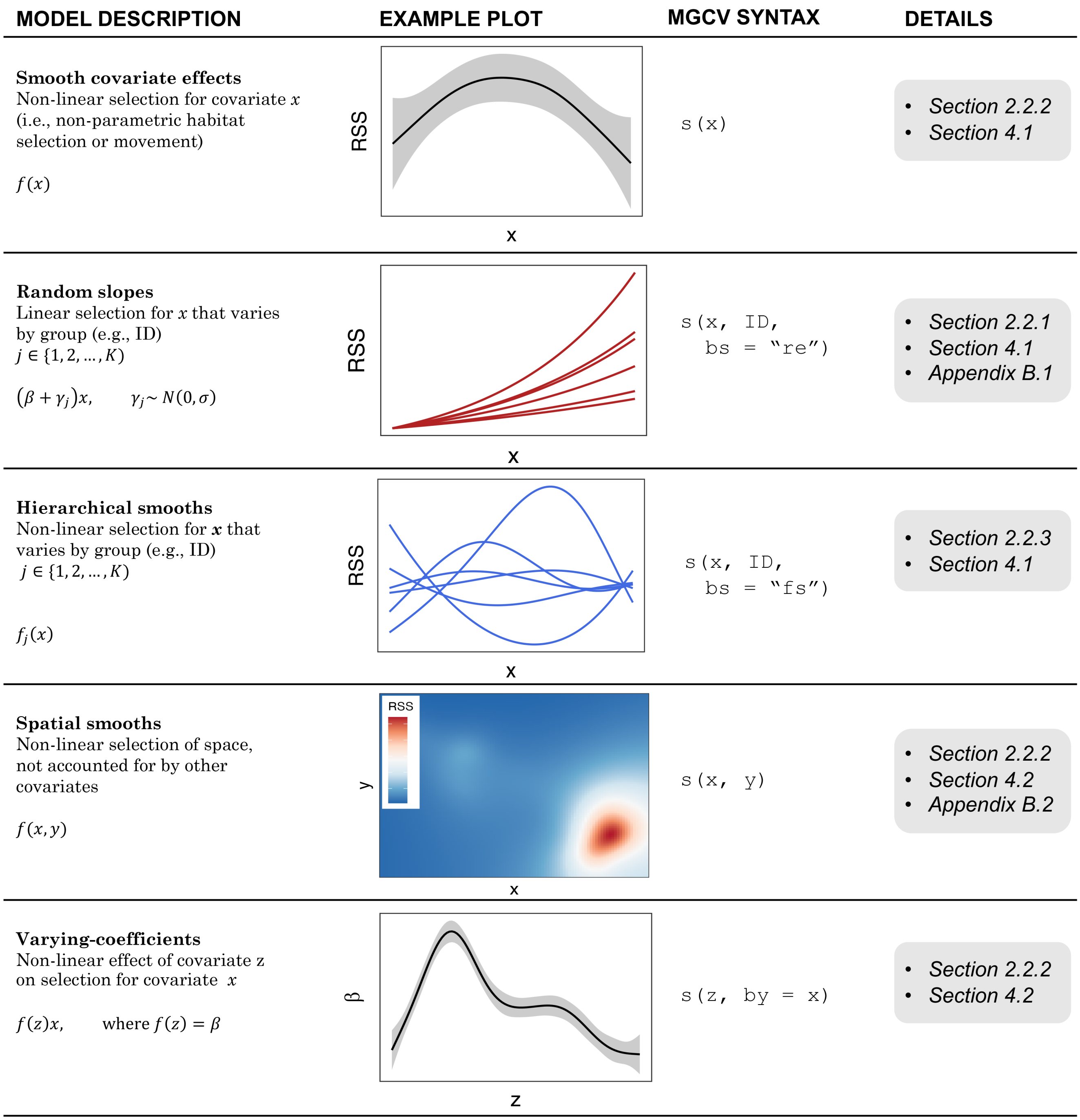

FIGURE 1 Summary of smooth effects discussed in this paper. 'Model description' explains the form of the smooth effect, which is demonstrated visually in the example plot (in which RSS is the relative selection strength). The `mgcv` syntax only refers to the relevant term in the model formula, where `s()` denotes a smooth function, `bs` is the basis function type, and `by` is used to specify interactions. Further details of each effect can be found elsewhere in the paper and appendices.

et al., 2024; Muff et al., 2020). The linear predictor of an SSF with independent random slopes for a covariate x takes the form,

$$\eta_{itj} = (\beta + \gamma_j) x_{itj} \tag{6}$$

with random effect levels $j \in 1, 2, \ldots, K$. In this case, each z_{it} in Equation (5) is zero when the observation is not part of the j-th random effect level (and $z_{it} = x_{it}$ when it is). The penalty matrix of the i.i.d. random effects is the identity matrix ($\mathbf{S} = \mathbf{I}$), leading to the random effect distribution $\boldsymbol{\gamma} \sim N\left(\mathbf{0},\ \sigma_\gamma^2 \mathbf{I}\right)$, where $\sigma_\gamma^2 = 1/\lambda$. That is, the smoothing parameter λ is inversely related to the variance of the random effects; a large value of λ corresponds to a small value of σ_γ^2 and thus, shrinkage of individual or group-specific parameters toward the population mean.

2.2.2 | Smooth effects

To include a wide range of non-parametric movement and habitat terms, it is possible to model non-linear functions (e.g. penalised splines, Gaussian processes) as random effects. We define a single-penalty smooth function of some covariate u as

$$f(u) = \sum_{k=1}^{K} \gamma_k \Psi_k(u), \qquad \boldsymbol{\gamma} \sim N\left(\mathbf{0}, \frac{1}{\lambda}\mathbf{S}^-\right), \tag{7}$$

where the Ψ_k are simple mathematical functions called 'basis functions', weighted by the coefficients γ_k (see Figure 2 for an example; Hefley et al., 2017; Pedersen et al., 2019; Wood, 2017). In Equation (7), the penalty matrix is determined by the choice of smoothing basis, and the smoothness of the function (i.e. the degree of penalisation) is controlled by λ. When λ is high, this indicates low covariance of the basis coefficients, such that adjacent basis function evaluations are more similar, resulting in a smoother function (Figure 2). In contrast to simpler polynomials and regression splines with complexity specified a priori, we estimate λ from the data to control the trade-off between model fit and complexity.

This approach is also extensible to smooths with $m \geq 2$ penalties (e.g. adaptive smoothers with variable smoothness along the covariate range) by replacing $\mathbf{S}^- / \lambda$ in Equation (7) with the pseudo-inverse of a sum of the m penalty matrices for a given smoother multiplied by specific penalty terms: $\left(\sum_{j=1}^{m} \lambda_j \mathbf{S}_j\right)^-$. This is important because it allows for modelling non-linear interactions between covariates with different units as smooth terms via tensor-product smoothers (Wood et al., 2013).

There are several useful applications of smoothing splines, including modelling non-linear covariate effects, random fields and formulating interaction models with smooth effects (Wood, 2017). Here, we briefly summarise three important uses of smooth functions in SSFs (which we explore in more detail in Sections 3 and 4):

1. Including smooth functions of covariates can capture non-linear patterns of habitat selection and non-parametric movement kernels. The linear predictor of an SSF with a smooth of covariate x (e.g. step length, turning angle or a spatial feature) takes the form $\eta_{it} = f(x_{it})$. This can relax the assumption that animals have a constant rate of selection (see Section 4.1) and simple unimodal distributions of step lengths and turning angles.
2. Incorporating a spatial smooth (i.e. random effect or random field) can account for spatial variation that is not explained by the other environmental covariates in the model (Arce Guillen et al., 2023). An SSF with a spatial smooth of the spatial coordinates (here, denoted u and v) has the linear predictor $\eta_{it} = f(u_{it},\ v_{it})$ with smooth function $f(u, v) = \sum_{k=1}^{K} \gamma_k \Psi_k(u, v)$, where Ψ_k is a two-dimensional basis function (Hefley et al., 2017; Wood, 2017). In Appendix B.2, we show how an unknown centre of attraction can be captured with a spatial smooth, and in Section 4.2, we use a spatial smooth

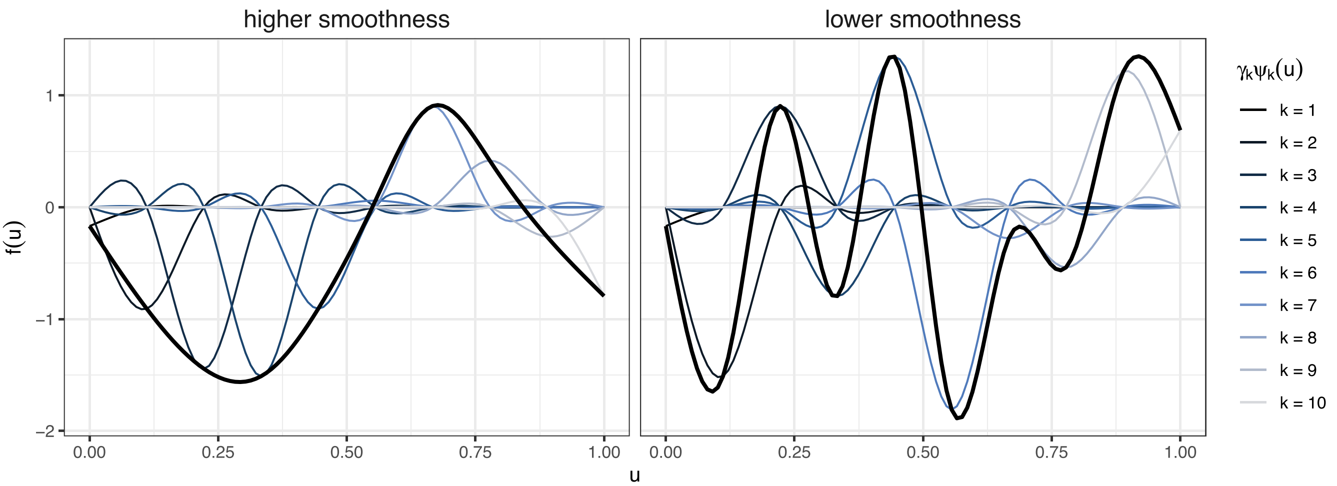

FIGURE 2 Example of how to derive cubic regression splines: the blue lines are the basis functions Ψ multiplied by their coefficients $\boldsymbol{\gamma}$, which are then summed to obtain the smooth functions (black lines). The panels show two examples with different smoothing penalties (i.e. λ).

to model patterns of a zebra's space use that cannot be explained by preference for a measured habitat variable.

3. Formulating a model with varying coefficients (i.e. where the linear effect β of a covariate varies smoothly with another covariate; Wood, 2017) greatly increases the flexibility of SSFs. An interaction between the covariate x and a non-linear effect of u can be written as $\eta_{it} = f(u_{it})x_{it}$, where $f(u)$ is given by Equation (7). That is, the selection coefficient for x is specified as a non-linear function of u. This approach is highly flexible, as it allows for any specification of non-linear covariate-dependent habitat selection or movement. For example, in Section 4.2, we show how zebra movement speed varies cyclically with time of day, reflecting underlying behavioural variation (Klappstein et al., 2023).

2.2.3 | Hierarchical smooths

We can formulate step selection models with multiple levels of smooth effects via hierarchical smooths (Pedersen et al., 2019). In this framework, a separate smooth relationship is estimated for each random effect level (e.g. each individual). This is conceptually similar to a random slope model, as it allows individuals or groups to differ in their selection or movement patterns, but in a non-linear manner. Consider a model with a hierarchical smooth with K groups for a continuous covariate x. The linear predictor for the j-th group is $\eta_{itj} = f_j(x_{itj})$, where each smooth f_j takes the general form presented in Equation (7). The same general formulation could be extended to account for inter-group variability in any smooth effect (e.g. spatial smooths). Further, f_j is often modelled as the sum of a population-level smooth relationship and individual-specific smooth deviations from the population. This gives great flexibility to specify which model components are shared across individuals and which are not (e.g. shape or smoothness of the relationship), and Pedersen et al. (2019) describe several important possible formulations. The `sz` smoother has also recently been added to `mgcv` as a type of hierarchical smoother with the main effect factored out of the smooth term, leaving only differences between individuals in the smoother. This allows for estimation of both an overall main effect (via a non-hierarchical smoother) and individual-level effects (via an `sz` smoother). This addresses some of the issues identified in Pedersen et al. (2019) with estimating hierarchical GAMs that include both a global and group-level smoother for the same term.

3 | PRACTICAL GUIDANCE

In this section, we provide some guidance for practitioners interested in using `mgcv` to fit SSFs with smooth effects. We cover how to sample random points for different movement models, how to choose appropriate settings in `mgcv`, as well as model selection, interpretation, and diagnostics (Figure 3 presents a summary of the workflow).

3.1 | Sampling random points

As described in Section 2.1, the implementation of SSFs requires sampling random points from a two-dimensional distribution h to obtain an approximation of the likelihood (Michelot et al., 2024). Note that, when the movement kernel ϕ is estimated, the random points are not assumed to represent movement or habitat availability, as this is estimated during model fitting. Within the CLR framework, it is most convenient to specify h with the same general form as ϕ, as this simplifies post hoc corrections to the parametric movement kernels (Table 1; Forester et al., 2009; Avgar et al., 2016). Therefore, when the movement kernel is chosen as a parametric model, we suggest that random points be sampled from the same families of distributions of step lengths and turning angles (Figure 3). The estimated coefficients associated with movement covariates (e.g. $\cos(\alpha)$, L) from CLR then represent deviations from the parameters in h. As such, the parameters of the movement kernel ϕ can be derived using simple formulas (i.e. 'corrections' to the tentative CLR coefficients). For common step length and turning angle distributions, Table 1 shows the terms that need to be included in the linear predictor of the SSF, and the corrections required to recover movement parameters (see also Avgar et al., 2016).

Alternatively, spatially uniform random points can be used and require no post hoc corrections (as h is constant), although this comes at the cost of reduced computational efficiency (i.e. more random points may be needed for accurate parameter estimation; Michelot et al., 2024). Despite this cost, spatially uniform points are a sensible choice when estimating non-parametric movement kernels, for which simple corrections are not possible. In this case, we recommend sampling points $\{\mathbf{r}_{t1}, \dots, \mathbf{r}_{tN}\}$ uniformly over a disc centred on the previous observed location $\mathbf{s}_{t-1}$, with a radius R close to the maximum observed step length for adequate spatial coverage (Klappstein et al., 2022; Michelot et al., 2024). In practice, this can be done by generating each turning angle from $\mathrm{Unif}(-\pi, \pi)$ and each step length as the square root of a random draw from $\mathrm{Unif}(0, R^2)$.

3.2 | Implementation in `mgcv`

Using `mgcv` requires implementing an SSF as a Cox PH model via the `cox.ph` family. To do so, we define a constant 'event time' variable (`times`; all observations set to the same value) in conjunction with a stratification variable (`stratum`; an identifier for each grouping of an observed location and its random points) as our response and use the `weights` argument to distinguish the observed location from the random points (i.e. via an indicator variable; `obs`). Ultimately, this defines a CLR model, and the linear predictor can be defined using standard formula syntax in `mgcv`. For example, a model with non-parametric distributions of step lengths and turning angles, and a linear effect of habitat covariate x is specified as

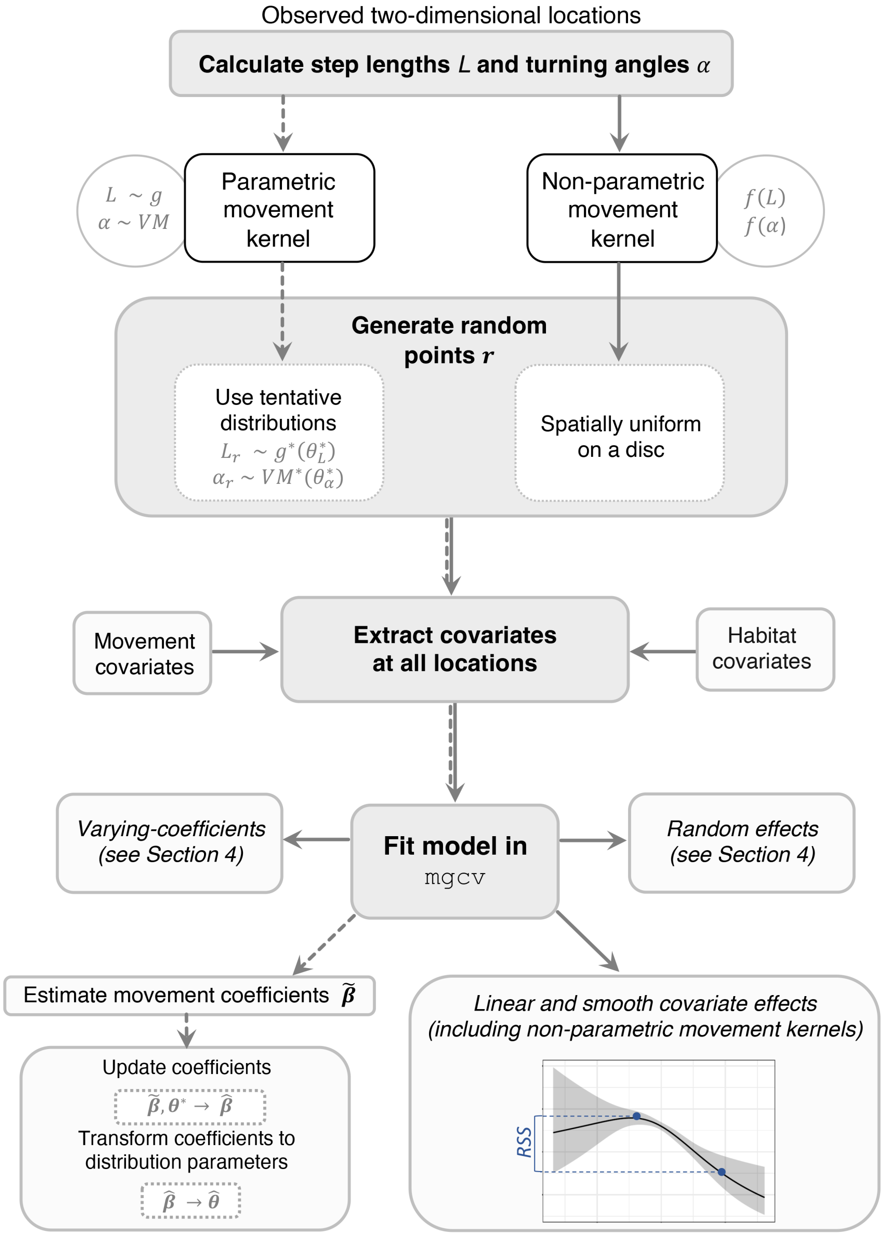

FIGURE 3 Summary of workflow to implement SSFs described in this paper (dashed lines indicate the most common workflow for parametric movement kernels). L denotes step length, α denotes turning angles, and the distributions g^* (any step length distribution from the exponential family) and VM^* (von Mises) used to generate random points are typically estimated from empirical data. RSS is the relative selection strength calculated as the ratio between the two blue points, where the y-axis is the exponential of the partial effect.

```
fit <- gam(cbind(times, stratum) ~ s(step) + s(angle) + x,
      data = dataset,
      family = cox.ph,
      weights = obs)
```

where `times` is a vector of 1's, and `obs` is a vector of 0's and 1's. The model parameters are estimated via restricted maximum likelihood (REML).

3.2.1 | Choice of basis functions and basis dimension

The choice of a smoother defines the form of the basis functions and the penalty term that controls the function smoothness (Pedersen et al., 2019; Wood, 2017). Therefore, various random effect and smooth models can be fitted simply by choosing the appropriate basis functions. By default, `mgcv` uses thin plate regression splines for both one- and multi-dimensional smooths. Although these generally perform well, there are several other general smoother options available (see Wood, 2017, and `mgcv` documentation for full descriptions). For example, a spatial random field could be specified as a Gaussian process with several options for the covariance function. As outlined by Miller et al. (2020), this approach is similar to the stochastic partial differential equation approach, which has been previously implemented for spatial random effects in SSFs via `inlabru` (Arce Guillen et al., 2023). Specialised smoothers can be used to better account for boundary behaviour (e.g. soap film smoothers, Duchon splines) and non-isotropic coordinates (e.g. splines on the sphere). Additionally, cyclic cubic regression splines can be used to capture cyclic patterns, which are ubiquitous in ecology (e.g. effects of time of day;

TABLE 1 Common step length (L) and turning angle (α) distributions used in SSFs. If the random points are generated from tentative distributions of step lengths and turning angles, we obtain tentative estimates of the coefficients (denoted by $\tilde{\beta}$). These estimates do not correspond to the parameters of the SSF linear predictor, and therefore, we must correct these parameters (i.e., account for the random point distribution with parameters $\boldsymbol{\theta}^*$) to obtain the SSF parameters (denoted as $\hat{\boldsymbol{\beta}}$), which can then transformed to the distribution parameters of interest $\hat{\boldsymbol{\theta}}$.

Step length or angle distribution		Distribution parameters	SSF linear predictor	Correction[a] $(\tilde{\beta}, \theta^* \rightarrow \hat{\beta})$	Transformation $(\hat{\beta} \rightarrow \hat{\theta})$
Step length L	*Exponential*	Rate λ	$\beta_1 L - \log(L)$	$\hat{\beta}_1 = \tilde{\beta}_1 - \lambda^*$	$\hat{\lambda} = -\hat{\beta}_1$
	Gamma	Shape a	$\beta_1 L + \beta_2 \log(L)$	$\hat{\beta}_1 = \tilde{\beta}_1 - \frac{1}{b^*}$	$\hat{a} = \hat{\beta}_2 + 2$
		Scale b		$\hat{\beta}_2 = \tilde{\beta}_2 + a^* - 2$	$\hat{b} = -\frac{1}{\hat{\beta}_1}$
	Log-normal	Mean μ	$\beta_1 \log(L) + \beta_2 \log(L)^2$	$\hat{\beta}_1 = \tilde{\beta}_1 + \frac{\mu^*}{(\sigma^*)^2} - 2$	$\hat{\mu} = -\frac{\hat{\beta}_1 + 2}{2\hat{\beta}_2}$
		SD σ		$\hat{\beta}_2 = \tilde{\beta}_2 - \frac{1}{2(\sigma^*)^2}$	$\hat{\sigma} = \sqrt{\frac{-1}{2\hat{\beta}_2}}$
Turning angle α	*von Mises*	Concentration κ	$\beta_1 \cos(\alpha)$	$\hat{\beta}_1 = \tilde{\beta}_1 - \kappa^*$	$\hat{\kappa} = \hat{\beta}_1$
		Mean μ			$\hat{\mu} = \begin{cases} 0 \text{ if } \hat{\beta}_1 \geq 0 \\ \pi \text{ if } \hat{\beta}_1 < 0 \end{cases}$

Abbreviation: SSF, step selection function.

[a]The correction step is only necessary when the SSF is being implemented with non-uniform random locations. These corrections assume that the random points are sampled from a joint distribution of step lengths and turning angles that matches the general form of the modelled distribution ϕ.

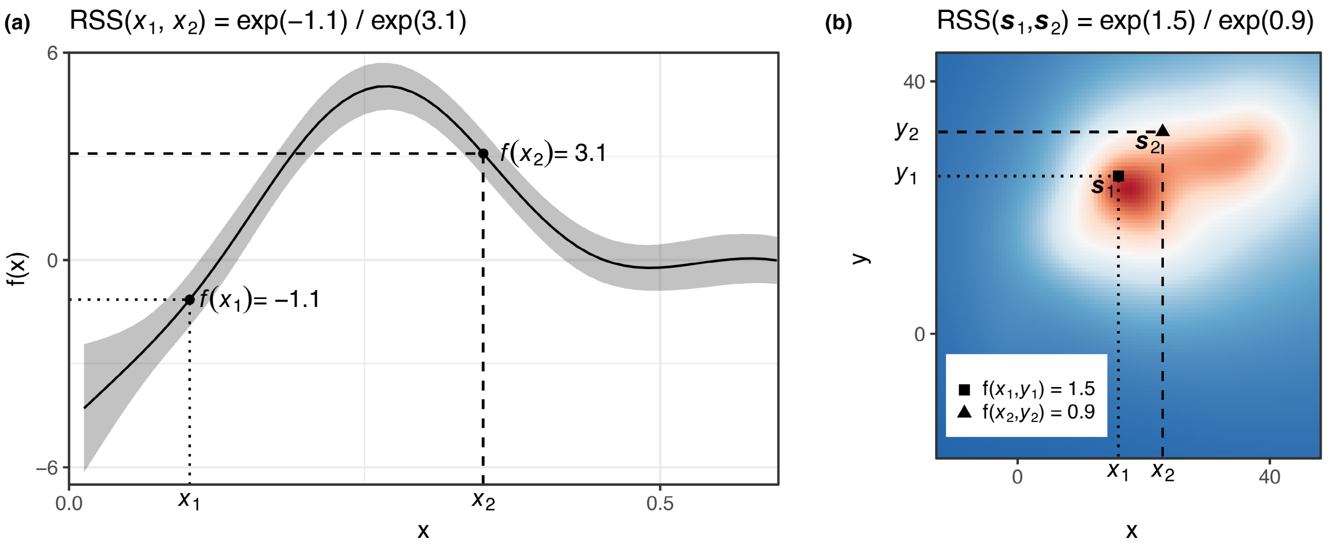

FIGURE 4 Relative selection strength (RSS) examples. (a) shows the RSS for the covariate x (all else held equal), where RSS(x_1, x_2) refers to the RSS of location with covariate value x_1 relative to x_2. (b) shows how to interpret a spatial smooth, where the RSS is based on spatial locations (i.e. different values of easting x and northing y), where RSS($\mathbf{s}_1$, $\mathbf{s}_2$) refers to the RSS of the first location $\mathbf{s}_1$ (with coordinates x_1, y_1) compared to the second $\mathbf{s}_2$ (with coordinates x_2, y_2).

Feldmann et al., 2023). Since we can view a random effect as a smooth (and vice versa), we can also specify a random slope model via the basis functions (see Figure 1 and coded examples for relevant settings in `mgcv`).

In principle, the basis dimension K (i.e. number of columns in the associated design matrix) only sets the upper limit to the wiggliness of the function. That is, for large enough K, the smoothness of the function is constrained by the penalty matrices and estimated smoothness parameters λ. In practice, the shape of the estimated function might therefore not be affected much by this choice, but a large K will be more computationally demanding (Wood, 2017). We suggest trying multiple choices of K to assess if model outputs change between fits, and checking how K compares to the effective degrees of freedom (if the values are similar, it may indicate that a higher K is required). Note that in the case of random slopes, K is the number of groups and cannot be changed.

3.2.2 | Comparison to glmmTMB

In Appendix B.1, we explore the performance of `mgcv` and `glmmTMB` for estimating individual-level random slopes in a simple SSF with a single habitat covariate and gamma distribution of step lengths. `glmmTMB` was more computationally-efficient, particularly for a large number of individuals. The function `gam()` is not structured to utilise the sparsity of most random effect structures, and is known be relatively inefficient for many random effect levels (Wood, 2017). An alternative could be to formulate the SSF as a conditional Poisson model (i.e. with stratum-specific intercepts; Muff et al., 2020), which should allow fitting with the mixed model function `gamm()` or to be run in parallel via `bam()`. However, we did not explore this, and it is unclear if this would increase efficiency. Our simulated example showed that, for both `mgcv` and `glmmTMB`, fixed effect estimators were negligibly biased, and estimators of the random effect variance were negatively biased only in `glmmTMB` (consistent with results from Muff et al., 2020).

3.3 | Model interpretation

Both smooth and parametric (main) effects can be interpreted in terms of relative selection strength (RSS; Avgar et al., 2017; Fieberg et al., 2021). For linear effects, there is an intuitive interpretation of the coefficient, where (holding all else equal) $\exp(\beta)$ is how many times more likely an animal is to take a step when the covariate increases by 1 (similarly to other regression coefficients). The relative selection strength between any two steps can be calculated as the ratio of the predicted selection $\exp(\beta x_i)$ for each step. For example, consider steps with covariate values x_1 and x_2; the RSS of the step with x_1 relative to the step with x_2 is $RSS(x_1, x_2) = \exp(\beta x_1) / \exp(\beta x_2)$. For a smooth term, we lose the straightforward interpretation of $\exp(\beta)$, but we can calculate the RSS in the same way. It may also be useful to plot the predicted selection for a grid of covariate values, but note that the y-axis can only be interpreted relatively (Figures 3–5). Note that this same interpretation applies to spatial smooths, but where the RSS is calculated between two spatial locations (i.e. the spatial smooth should not be confused with the spatial distribution of the animal; Figure 4b). The RSS (with uncertainty) for linear SSFs can be calculated using the `log_rss` function of the package `amt` (Signer et al., 2019); this is not possible for models fitted via `mgcv`, but similar calculations are possible using other software packages (e.g. `marginaleffects`; Arel-Bundock, 2023). Note that interaction terms (e.g. varying coefficients) and random effects have standard interpretations (as described in Wood, 2017).

Confidence intervals for estimated smooth curves should be interpreted as pointwise (rather than simultaneous) CIs for a given smooth curve (Wood, 2017). Confidence intervals for smooth curves estimated using REML have been shown to have good coverage properties when averaged across the function, but might have above- or below-nominal coverage at different points in the curve, so should be interpreted with caution. Visualising posterior simulations from the fitted curve provides a better visual guide to the degree of functional uncertainty in a fitted model, rather than just the estimated curve and confidence interval. These posterior simulations can be produced via the `smooth_samples` function of the `gratia` package (Simpson, 2023).

3.4 | Model selection and diagnostics

AIC is commonly used for variable selection in regression models. There are two types of AIC for models with random effects fitted with `mgcv`: (i) marginal AIC (which tends to favour complex models) and (ii) conditional AIC (favouring simpler models) (Wood, 2017). Wood et al. (2016) describes a 'corrected' conditional AIC, which better accounts for smoothing parameter uncertainty in the penalty term. This is the default when `AIC()` is applied to a `gam` model object, and it is one possible approach to choose among competing model formulations.

Another approach is to use penalisation to automatically exclude smooth terms that have no clear effect, by 'shrinking' the effect to zero. For most smooths, increasing λ to a large number only constrains the relationship to be any straight line, rather than zero (the set of straight lines is called the null space of the penalty). This problem can be resolved by also penalising non-zero straight lines. In `mgcv`, this can be done with specialised shrinkage bases, for which the smoothing parameter λ also penalises the null space (i.e. where $\lambda \to \infty$ effectively removes the corresponding term from the model) or by including an additional penalty with the `select=TRUE` option for any smooth term (i.e. the 'double-penalty' approach; Marra & Wood, 2011). Although AIC and shrinkage are convenient model selection techniques, we suggest they be used in conjunction with sensible candidate models formed from expert opinion.

In general, there are no simple residual-based checks for SSFs, and this is not solved by our proposed implementation in `mgcv`. Although `mgcv` has a convenient `gam.check` function, the residuals are specifically designed to assess the assumptions of the Cox PH model. For example, Schoenfield residuals test the proportional hazards assumption and deviance residuals are derived from Martingale residuals based on the cumulative hazard function (Wood, 2017). However, the proportional hazards assumption does not apply to SSFs, and therefore we advise against using these residuals for model checking. Diagnositics are notoriously difficult for SSFs, and although out of the scope of this paper, simulation-based methods could be promising for SSFs with smooths (DiRenzo et al., 2023; Fieberg et al., 2018, 2024).

4 | ILLUSTRATIVE EXAMPLES

In this section, we present two real data examples to illustrate the models and implementation described in Section 2: (i) we explore how to best account for inter-individual variability in polar bear (*Ursus maritimus*) habitat selection, and (ii) we formulate a model with a spatial smooth and a time-varying (parametric) movement kernel for a single plains zebra (*Equus quagga*). In both examples, we chose to include a parametric movement kernel, and therefore we sampled from tentative distributions of step lengths that matched the form of ϕ (following Figure 3). We matched each observed location with 25 random steps, generated from a gamma distribution of step lengths (with parameters derived from the empirical data) and

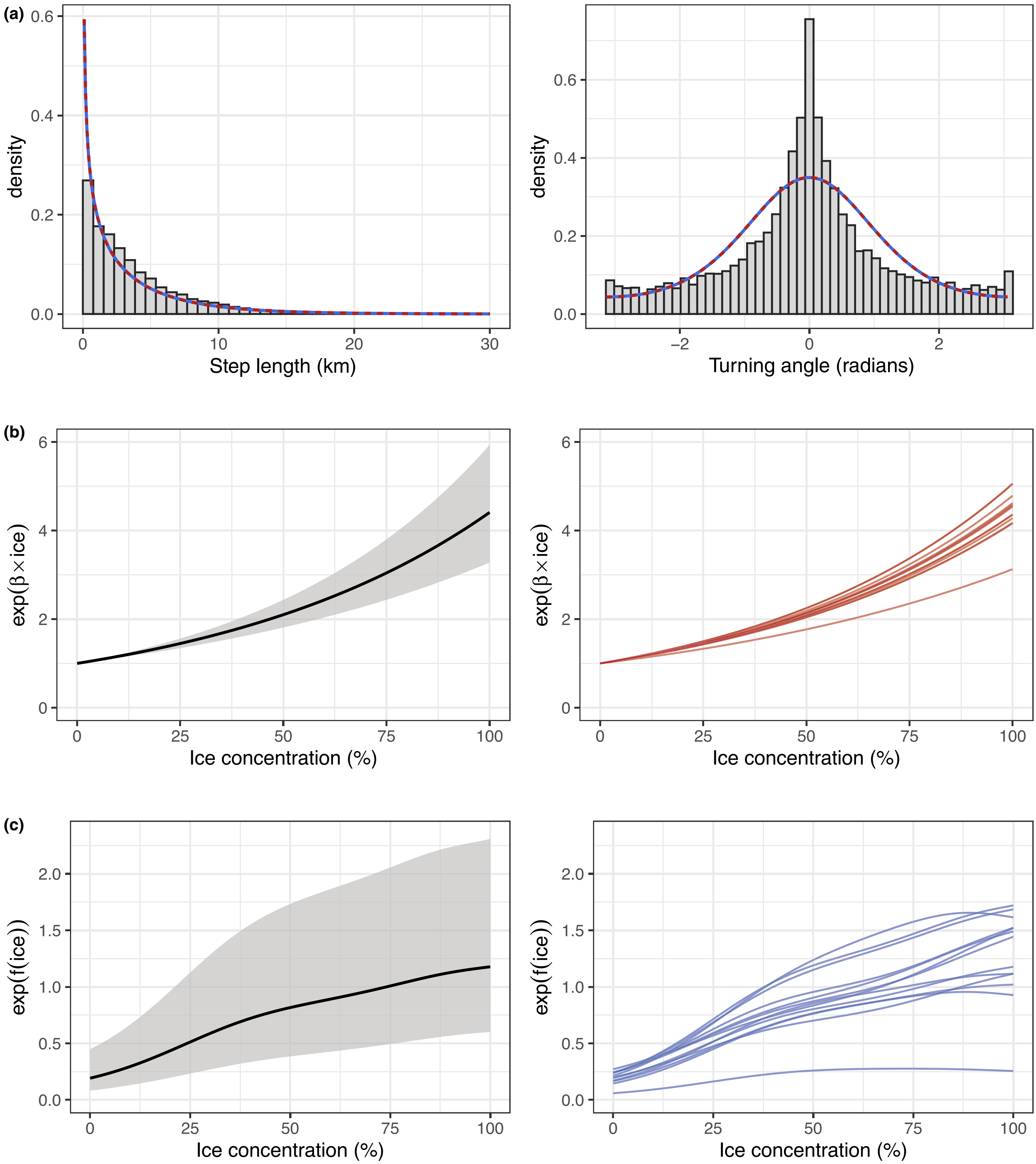

FIGURE 5 Results for polar bear example: (a) step length and turning angle distributions for both models (shown in dashed/solid lines), (b) random slopes for ice concentration, and (c) hierarchical smooths for ice concentration. In both (b) and (c), the left plot shows the population estimate with 95% CIs and right plot shows individual-level estimates.

uniform turning angles. In all models, we modelled step lengths L with a gamma distribution (by including L and $\log(L)$ as covariates) and turning angles α with a von Mises distribution (by including $\cos(\alpha)$ as a covariate). All models were fitted on an Apple MacBook Pro with an M1 Pro processor and 16GB of RAM.

4.1 | Capturing inter-individual variability of polar bear habitat selection

We obtained 4-h GPS locations from 13 polar bears in the Beaufort Sea. We regularised the movement tracks and interpolated spatial covariates (following Klappstein et al., 2020, 2022). We analysed a total of 14,927 locations (range: 862–1810 per individual). The main goal of inference was to assess inter-individual variability in selection for ice concentration, which is linked to polar bear prey distribution and energetic costs of travel (Klappstein et al., 2022; Pilfold et al., 2014). Therefore, we considered two models that allow for this inter-group variation: a model with random slopes and a model with hierarchical smooths. We assumed that the movement kernel ϕ (described at the beginning of Section 4) was shared across all individuals. The linear predictor for the i-th location of stratum t of the random slope model (for $j \in 1, 2, \dots 13$ individuals) takes the form

$$\eta_{itj} = \beta_1 L_{itj} + \beta_2 \log(L_{itj}) + \beta_3 \cos(\alpha_{itj}) + (\beta_4 + \gamma_j) x_{itj}, \quad (8)$$

where x is the ice concentration. The analogous hierarchical smooth model has the form,

$$\eta_{itj} = \beta_1 L_{itj} + \beta_2 \log(L_{itj}) + \beta_3 \cos(\alpha_{itj}) + f(x_{itj}) + f_j(x_{itj}), \quad (9)$$

where both smooths were constructed from $K = 5$ basis functions. Each model contained both a global term for the slope or smooth, as well as inter-individual deviations from that global mean, and the main difference was that selection for ice concentration was either modelled as a log-linear effect (Equation 8) or as a smooth (i.e. non-linear) effect (Equation 9).

It was faster to fit the random slope model (approximately 6 min), compared to the hierarchical smooth model (approximately 14 min). Estimates of movement parameters were the same (to two decimal places) for both models ($\hat{\beta}_1 = 0.19$, $\hat{\beta}_2 = -1.39$, $\hat{\beta}_3 = 1.04$), indicating that polar bear step lengths followed a gamma distribution with mean $\hat{\mu} = 3.22$ km and standard deviation $\hat{\sigma} = 4.12$ km, and turning angles followed a von Mises distribution with mean $\hat{\mu} = 0$ and concentration $\hat{\kappa} = 1.04$ (Figure 5a). Note that the animals' movement patterns are affected by habitat selection, so the observed and estimated distributions in Figure 5a are not necessarily expected to match completely. The models differed in their characterisation of the selection for ice concentration. The population mean of the random slopes model indicated that the 'typical' bear selected for higher values of ice concentration, and was $\exp(\hat{\beta}_4) = 1.015$ times more likely to take a step for every 1% increase in ice concentration. We cannot easily derive this relationship from the hierarchical smooth model, but we can assess relative selection between any two competing ice concentrations by comparing the estimated smooth function at each point. For example, consider if the average bear was presented with a choice of 0% or 50% ice concentration. The smooth and linear models predict that the bear would be 4.3 and 2.1 (respectively) times more likely to choose the step with 50% ice concentration. However, note that the smooth function varies across the covariate range and the RSS will not be the same for any two points with 50% difference (as is the case for a linear model).

Both models captured inter-individual variability, but the smooths model also captured variability in the pattern of selection (i.e. individuals have functions with different shapes; Figure 5c). We compared the models with AIC (as described in Wood et al., 2016) and found that the hierarchical smooth model had the lower AIC (ΔAIC = 38.3). This is consistent with previous results on polar bear habitat selection, which indicated that individual polar bears select for an optimal ice concentration between 60 to 100% (rather than a log-linear relationship; Klappstein et al., 2022) corresponding with maximum prey biomass at 85% (Pilfold et al., 2014).

4.2 | Spatiotemporal variation of zebra movement and habitat selection

We analysed 7246 GPS locations from a single plains zebra, collected at a 30-min resolution in Hwange National Park in Zimbabwe from January–April 2014 (previously described in Klappstein et al., 2023; Michelot et al., 2020). The goal of the analysis was threefold: (i) assess habitat selection for different vegetation types, (ii) capture temporal variation in movement patterns and (iii) account for any remaining spatial variation using a spatial smooth.

We assessed habitat selection for a categorical vegetation variable (woodland, bushland, bushed grassland and grassland as the reference category). As in the previous example, we defined ϕ with a gamma step length distribution and a von Mises distribution of turning angles. We allowed the scale parameter of the gamma distribution (i.e. the step length coefficient) to vary based on the time of day τ via a cyclic spline (with $K = 15$), to capture temporal patterns in movement speed that repeat or occur each day. Finally, we included a spatial smooth (i.e. an isometric smooth interaction between the easting u and northing v) using two-dimensional thin plate regression splines (with $K = 30$). The model linear predictor was

$$\begin{aligned}\eta_{it} &= f_1(\tau_t) L_{it} + \beta_1 \log(L_{it}) + \beta_2 \cos(\alpha_{it}) + \beta_3 \delta_{it}^{w} + \beta_4 \delta_{it}^{b} \\ &+ \beta_5 \delta_{it}^{bg} + f_2(u_{it}, v_{it}),\end{aligned} \quad (10)$$

where δ^{w}, δ^{b}, δ^{bg} are indicator variables for woodland, bushland and bushed grassland.

It took approximately 15 min to fit the model. The zebra's mean step length changed throughout the day (captured by variability in the associated coefficient; Figure 6a), ranging from approximately 80–100 m in the evening to 200–250 m in the early morning and mid-afternoon (Figure 6b). Consistent with previous findings, our results indicate that the zebra selected for grassland over all other habitat types (Figure 6c; Klappstein et al., 2023; Michelot et al., 2020). The spatial smooth showed remaining spatial pattern, once other covariates had been considered. In particular, there was one area at the northern edge of the map that the zebra disproportionately selected for (Figure 6d), which could indicate an important spatial feature that was not accounted for in the model. Note that the estimated spatial smooth should be interpreted in terms of RSS (see Figure 4b) and cannot be interpreted as a spatial distribution.

5 | DISCUSSION

In this paper, we explained how smooth (i.e. random) effects provide a unifying framework for many SSF extensions, which can

Methods in Ecology and Evolution

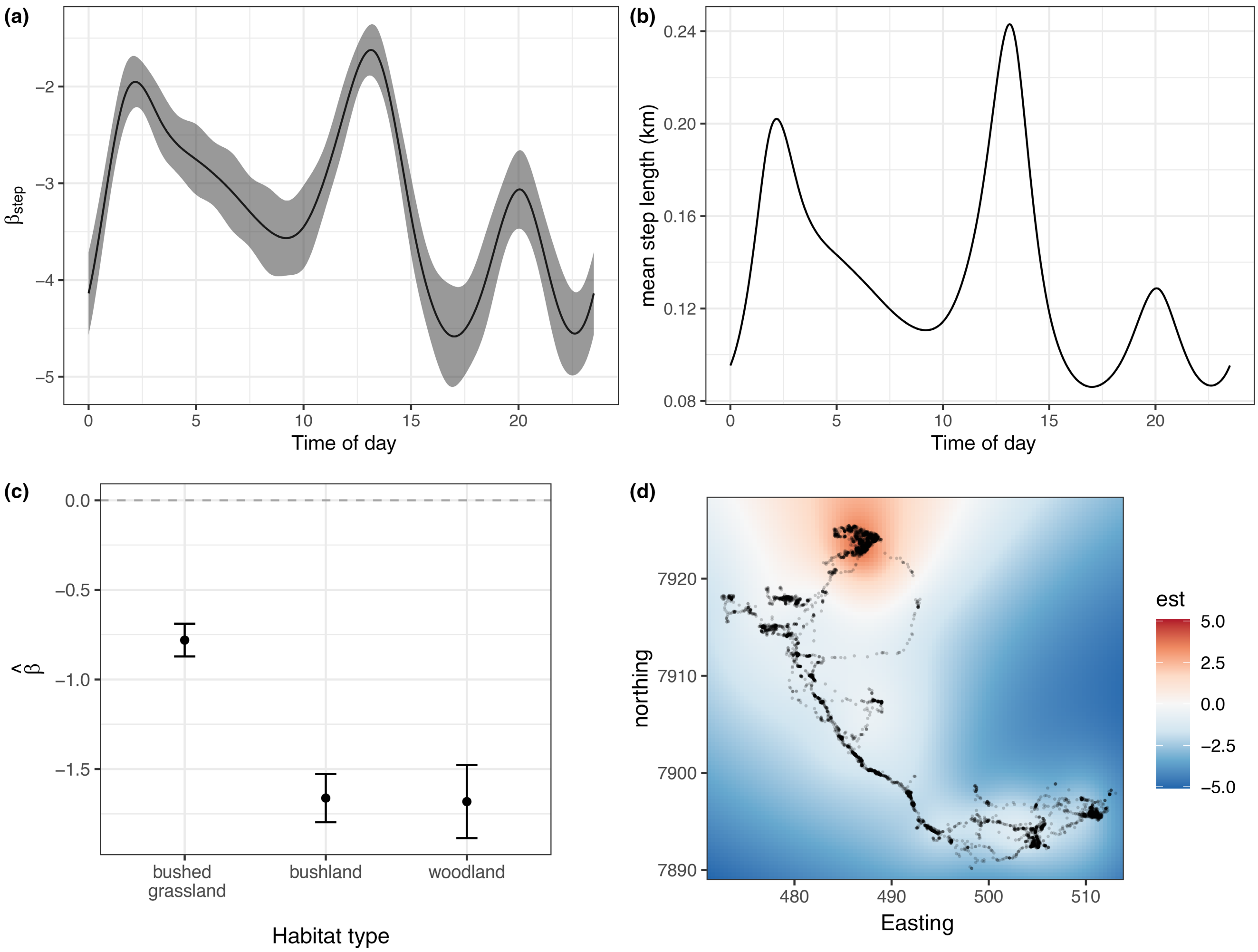

FIGURE 6 Results of the zebra model. (a) and (b) show how step length varies as a function of time of day. (a) shows how the step length coefficient varied throughout the day, and (b) is the same relationship, translated to the scale of the mean step length. (c) Selection coefficients with 95% CIs for habitat types (with grassland as a reference category, i.e. corresponding to zero). (d) Partial effect of the spatial smooth (on the log scale).

be easily implemented in the widely used R package `mgcv`. Using smooths, we can account for non-linearity, inter-individual variability and other spatial and temporal dependencies in habitat selection and movement patterns. Although we focused on SSF analyses where the movement is estimated, this approach could also be used when the movement kernel ϕ is assumed to be known (Fortin et al., 2005). In that case, the flexible covariate effects that we described could only be included in the habitat selection component of the model (and not on ϕ, which is chosen by the analyst prior to model fitting; Michelot et al., 2024). Smooth effects are also highly relevant for RSFs (e.g. spatial smoothing, hierarchical smooths), which can also be implemented in `mgcv` with a binomial distribution (e.g. McCabe et al., 2021). We gave a general overview of the flexible nature of smooth effects, focusing on several key cases we believe will be particularly relevant, but there are many additional possibilities that could further extend the capabilities of habitat selection models.

5.1 | Utility of smooth effects

In general, smooth interactions are powerful tools in SSFs, as they can model complex covariate-dependence in habitat selection and movement. We showed how varying-coefficient models can capture daily patterns of zebra movement speed. This approach can be generalised to other SSF covariates, where any time-varying coefficient could capture phases of movement or habitat selection corresponding to behavioural changes. This accomplishes a similar goal as state-switching SSFs (Klappstein et al., 2023; Nicosia et al., 2017), but rather than model discrete behavioural switches, movement or selection patterns are allowed to vary smoothly and continuously through time (similar to approaches in Hanks et al., 2015; Michelot et al., 2021). We could extend this approach to allow selection coefficients to vary over space if there is reason to believe that animals may modulate their behaviour in response to environmental features (e.g. in response to conspecifics; Smith et al., 2023). These time- or space-varying coefficients could be accomplished with specific covariates (e.g. time of day, season, habitat covariates), or with general temporal or spatial smooth interactions (e.g. a spatially varying coefficient model; Comber et al., 2022). Although we focused on 'varying coefficients', which refers to allowing a *linear* effect to vary smoothly with another covariate, smooth effects can also interact smoothly with other covariates. Interactions between covariates on different scales are possible with tensor products, which can include the marginal effects of one or both variables (Wood, 2017). For example, this could capture a non-parametric movement kernel with dependence between step lengths and turning angles (similar to copula-based kernels; Hodel & Fieberg, 2022).

It is often of interest to capture inter-group (e.g. inter-individual) variability in movement or habitat selection patterns. For this purpose, random slopes have been adopted for SSFs, but these are limited to linear habitat selection and parametric movement (Chatterjee et al., 2024; Duchesne et al., 2010; Muff et al., 2020). We showed how simple random slopes can be fitted as a smooth term, and how we can improve on this linear framework by incorporating hierarchical smooths into the SSF framework. Hierarchical smooths are more flexible and can capture inter-group variability in *non-linear* (i.e. smooth) terms. McCabe et al. (2021) explored the utility of hierarchical smooths in the context of large-scale species distributions across multiple individuals via RSFs, which do not explicitly model movement. By extending hierarchical smooths to SSFs, we can further investigate inter-individual differences in both animal habitat selection *and* movement patterns that vary through time and space (Chatterjee et al., 2024). This framework also makes it straightforward to account for inter-individual variability by including individual-specific covariates, via continuous or 'factor-smooth' interactions (e.g. animals may change their habitat selection throughout their life cycle or there may be inter-sex differences).

5.2 | Modelling challenges

Smooth effects afford modellers a vast range of options, and this can make model formulation and model selection increasingly difficult. Although AIC and shrinkage are useful tools, expert opinion and tenets of causal inference are necessary for selecting variables to include in models (Arif & MacNeil, 2022; Fieberg & Johnson, 2015). We also strongly encourage practitioners to carefully consider where additional model complexity is needed to answer their research question, while taking into account the limitations of their dataset (e.g. sample size, spatial and temporal resolution). Statistical confounding should also be considered; for example, although spatial smoothing is a powerful way to account for unexplained spatial variation, it does not necessarily resolve parameter bias (as identified in Dejeante, Lemaire-Patin, & Chamaillé-Jammes, 2024) if environmental features are spatially correlated (Hodges & Reich, 2010).

Further, the additional flexibility of smooth and random effect models comes with more computational and estimation challenges. Model fitting may be less numerically stable and more sensitive to the choice of random points (i.e. integration methods; Michelot et al., 2024). Depending on the basis dimension K, there can be many parameters to estimate, adding computational burden, and requiring trial and error to appropriately specify K. Unfortunately, comprehensive model checking is still challenging for SSFs. We suggest that users carefully consider model formulation, and inspect model outputs for signs of estimation problems such as very large standard errors, convergence issues, or high sensitivity of outputs to model formulation. Future work could explore simulation-based checks which have recently been proposed for simpler SSFs (Fieberg et al., 2018, 2024; Signer et al., 2024).

6 | CONCLUSIONS

The smooths framework provides a simple way to formulate and fit several recent, complex extensions of SSFs. Non-linear effects can be modelled without needing to specify model complexity a priori (e.g. using polynomials) and more flexible cyclic patterns can be captured without the use of trigonometic functions (Feldmann et al., 2023). Further, random slopes can be implemented with `mgcv`, similar to other approaches in the packages `glmmTMB` or `inlabru` (Muff et al., 2020). Arce Guillen et al. (2023) also used `inlabru` to implement spatial random effects for SSFs. In `mgcv`, it is possible to specify spatial smooths with several forms, including Gaussian processes (to match previous approaches; Arce Guillen et al., 2023; Miller et al., 2020). Our proposed implementation allows individual-specific and spatial random effects to be combined with other smooth terms described in this paper, using software for GAMs that many analysts are already familiar with (Pedersen et al., 2019; Wood, 2017). We hope this framework encourages practitioners to think about their habitat selection model formulations more flexibly, via the inclusion of interactions, random effects, and more realistic movement dynamics.

AUTHOR CONTRIBUTIONS

Natasha J. Klappstein and Théo Michelot conceived the original idea for the paper. Natasha J. Klappstein conducted the analyses (with assistance from Théo Michelot and John Fieberg), and wrote the first draft of the manuscript. All authors contributed ideas throughout the project and provided feedback on the written drafts of the manuscript.

ACKNOWLEDGEMENTS

We are grateful to the associate editor and two anonymous reviewers for the constructive feedback. We thank Andrew Derocher for providing the polar bear data and Simon Chamaillé-Jammes for providing the zebra data used in the illustrative examples. Also, we thank Tal Avgar and Chris Field for taking the time to discuss the project in its early stages. Natasha J. Klappstein and Joanna Mills Flemming received support from the National Sciences and Engineering Research Council of Canada. John Fieberg was supported by National Aeronautics and Space Administration award 80NSSC21K1182 and received partial salary support from the Minnesota Agricultural Experimental Station.

CONFLICT OF INTEREST STATEMENT

The authors have no conflict of interest to declare.

PEER REVIEW

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/2041-210X.14367.

DATA AVAILABILITY STATEMENT

Code and data are archived on Zenodo https://doi.org/10.5281/zenodo.11390020 (Klappstein et al., 2024) and a working version of the code is available at GitHub (https://github.com/NJKlappstein/smoothSSF).

ORCID

Natasha J. Klappstein https://orcid.org/0000-0001-6527-9629
Théo Michelot https://orcid.org/0000-0002-3838-4113
John Fieberg https://orcid.org/0000-0002-3180-7021
Eric J. Pedersen https://orcid.org/0000-0003-1016-540X
Joanna Mills Flemming https://orcid.org/0000-0002-3517-3089

REFERENCES

Arce Guillen, R., Lindgren, F., Muff, S., Glass, T. W., Breed, G. A., & Schlägel, U. E. (2023). Accounting for unobserved spatial variation in step selection analyses of animal movement via spatial random effects. *Methods in Ecology and Evolution*, *14*(10), 2639–2653.

Arel-Bundock, V. (2023). *Marginaleffects: Predictions, comparisons, slopes, marginal means, and hypothesis tests*. R package version 0.16.0.

Arif, S., & MacNeil, M. A. (2022). Predictive models aren't for causal inference. *Ecology Letters*, *25*(8), 1741–1745.

Avgar, T., Lele, S. R., Keim, J. L., & Boyce, M. S. (2017). Relative selection strength: Quantifying effect size in habitat- and step-selection inference. *Ecology and Evolution*, *7*(14), 5322–5330.

Avgar, T., Potts, J. R., Lewis, M. A., & Boyce, M. S. (2016). Integrated step selection analysis: Bridging the gap between resource selection and animal movement. *Methods in Ecology and Evolution*, *7*, 619–630.

Chatterjee, N., Wolfson, D., Kim, D., Velez, J., Freeman, S., Bacheler, N. M., Shertzer, K., Taylor, J. C., & Fieberg, J. (2024). Modelling individual variability in habitat selection and movement using integrated step-selection analysis. *Methods in Ecology and Evolution*, *15*(6), 1034–1047.

Comber, A., Harris, P., & Brunsdon, C. (2022). Spatially varying coefficient regression with GAM gaussian process splines: GAM(e)-on. *AGILE: GIScience Series*, *3*, 1–6.

Dejeante, R., Lemaire-Patin, R., & Chamaillé-Jammes, S. (2024). Can overlooking 'invisible landscapes' bias habitat selection estimation and population distribution projections? *bioRxiv*.

Dejeante, R., Valeix, M., & Chamaillé-Jammes, S. (2024). Time-varying habitat selection analysis: A model and applications for studying diel, seasonal and post-release changes. *Ecology*, *105*, e4233.

DiRenzo, G. V., Hanks, E., & Miller, D. A. W. (2023). A practical guide to understanding and validating complex models using data simulations. *Methods in Ecology and Evolution*, *14*(1), 203–217.

Duchesne, T., Fortin, D., & Courbin, N. (2010). Mixed conditional logistic regression for habitat selection studies. *Journal of Animal Ecology*, *79*(3), 548–555.

Duchesne, T., Fortin, D., & Rivest, L.-P. (2015). Equivalence between step selection functions and biased correlated random walks for statistical inference on animal movement. *PLoS One*, *10*(4), e0122947.

Feldmann, C. C., Mews, S., Coculla, A., Stanewsky, R., & Langrock, R. (2023). Flexible modelling of diel and other periodic variation in hidden Markov models. *Journal of Statistical Theory and Practice*, *17*(3), 45.

Fieberg, J., Forester, J. D., Street, G. M., Johnson, D. H., ArchMiller, A. A., & Matthiopoulos, J. (2018). Used-habitat calibration plots: A new procedure for validating species distribution, resource selection, and step-selection models. *Ecography*, *41*, 737–752.

Fieberg, J., Freeman, S., & Signer, J. (2024). Using lineups to evaluate goodness of fit of animal movement models. *Methods in Ecology and Evolution*, *15*(6), 1048–1059.

Fieberg, J., & Johnson, D. H. (2015). MMI: Multimodel inference or models with management implications? *The Journal of Wildlife Management*, *79*(5), 708–718.

Fieberg, J., Signer, J., Smith, B., & Avgar, T. (2021). A 'how to' guide for interpreting parameters in habitat-selection analyses. *Journal of Animal Ecology*, *90*(5), 1027–1043.

Forester, J., Kyung Im, H., & Rathouz, P. (2009). Accounting for animal movement in estimation of resource selection functions: Sampling and data analysis. *Ecology*, *90*(12), 3554–3565.

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

Hanks, E. M., Hooten, M. B., & Alldredge, M. W. (2015). Continuous-time discrete-space models for animal movement. *The Annals of Applied Statistics*, *9*(1), 145–165.

Hefley, T. J., Broms, K. M., Brost, B. M., Buderman, F. E., Kay, S. L., Scharf, H. R., Tipton, J. R., Williams, P. J., & Hooten, M. B. (2017). The basis function approach for modeling autocorrelation in ecological data. *Ecology*, *98*(3), 632–646.

Hodel, F., & Fieberg, J. (2022). Circular-linear copulae for animal movement data. *Methods in Ecology and Evolution*, *13*, 1001–1013.

Hodges, J. S. (2014). *Richly parameterized linear models: Additive, time series, and spatial models using random effects*. Texts in statistical sciences. CRC Press.

Hodges, J. S., & Reich, B. J. (2010). Adding spatially-correlated errors can mess up the fixed effect you love. *The American Statistician*, *64*(4), 325–334.

Kimeldorf, G. S., & Wahba, G. (1970). A correspondence between bayesian estimation on stochastic processes and smoothing by splines. *The Annals of Mathematical Statistics*, *41*(2), 495–502.

Klappstein, N. J., Michelot, T., Fieberg, J. R., Pedersen, E. J., & Mills Flemming, J. (2024). Reproducing the analyses of 'Step selection functions with non-linear and random effects'. *Zenodo*, https://doi.org/10.5281/zenodo.11390020

Klappstein, N. J., Potts, J. R., Michelot, T., Börger, L., Pilfold, N. W., Lewis, M. A., & Derocher, A. E. (2022). Energy-based step selection analysis: Modelling the energetic drivers of animal movement and habitat use. *Journal of Animal Ecology*, *91*, 946–957.

Klappstein, N. J., Thomas, L., & Michelot, T. (2023). Flexible hidden Markov models for behaviour-dependent habitat selection. *Movement Ecology*, *11*, 30.

Klappstein, N. J., Togunov, R. R., Reimer, J. R., Lunn, N. J., & Derocher, A. E. (2020). Patterns of sea ice drift and polar bear (*Ursus maritimus*) movement in Hudson Bay. *Marine Ecology Progress Series*, *641*, 227–240.

Manly, B. F., McDonald, L. L., Thomas, D. L., McDonald, T. L., & Erickson, W. P. (2002). *Resource selection by animals:Statistical design and analysis for field studies* (Vol. 34). Kluwer Academic Publishers.

Marra, G., & Wood, S. N. (2011). Practical variable selection for generalized additive models. *Computational Statistics and Data Analysis*, *55*, 2372–2387.

McCabe, J. D., Clare, J. D., Miller, T. A., Katzner, T. E., Cooper, J., Somershoe, S., Hanni, D., Kelly, C. A., Sargent, R., Soehren, E. C., Threadgill, C., Maddox, M., Stober, J., Martell, M., Salo, T., Berry, A., Lanzone, M. J., Braham, M. A., & McClure, C. J. W. (2021). Resource selection functions based on hierarchical generalized additive models provide new insights into individual animal variation and species distributions. *Ecography*, *44*(12), 1756–1768.

Michelot, T. (2023). hmmTMB: Hidden Markov models with flexible covariate effects in R. *arXiv*. http://arxiv.org/abs/2211.14139

Michelot, T., Blackwell, P. G., Chamaillé-Jammes, S., & Matthiopoulos, J. (2020). Inference in MCMC step selection models. *Biometrics*, *76*, 438–447.

Michelot, T., Glennie, R., Harris, C., & Thomas, L. (2021). Varying-coefficient stochastic differential equations with applications in ecology. *Journal of Agricultural, Biological and Environmental Statistics*, *26*(3), 446–463.

Michelot, T., Klappstein, N. J., Potts, J. R., & Fieberg, J. (2024). Understanding step selection analysis through numerical integration. *Methods in Ecology and Evolution*, *15*(1), 24–35.

Miller, D. L. (2021). Bayesian views of generalized additive modelling. *arXiv*. http://arxiv.org/abs/1902.01330

Miller, D. L., Glennie, R., & Seaton, A. E. (2020). Understanding the stochastic partial differential equation approach to smoothing. *Journal of Agricultural, Biological and Environmental Statistics*, *25*(1), 1–16.

Muff, S., Signer, J., & Fieberg, J. (2020). Accounting for individual-specific variation in habitat-selection studies: Efficient estimation of mixed-effects models using Bayesian or frequentist computation. *Journal of Animal Ecology*, *89*(1), 80–92.

Nicosia, A., Duchesne, T., Rivest, L.-P., & Fortin, D. (2017). A multi-state conditional logistic regression model for the analysis of animal movement. *The Annals of Applied Statistics*, *11*(3), 1537–1560.

Pedersen, E. J., Miller, D. L., Simpson, G. L., & Ross, N. (2019). Hierarchical generalized additive models in ecology: An introduction with mgcv. *PeerJ*, *7*, e6876.

Pilfold, N. W., Derocher, A. E., & Richardson, E. S. (2014). Influence of intraspecific competition on the distribution of a wide-ranging, non-territorial carnivore. *Global Ecology and Biogeography*, *23*(4), 425–435.

Pohle, J., Signer, J., Eccard, J. A., Dammhahn, M., & Schlägel, U. E. (2023). How to account for behavioural states in step-selection analysis: A model comparison. *PeerJ*, *12*, e16509.

Potts, J. R., & Börger, L. (2022). How to scale up from animal movement decisions to spatiotemporal patterns: An approach via step selection. *Journal of Animal Ecology*, *00*, 1–14.

Rhodes, J. R., Mcalpine, C. A., Lunney, D., & Possingham, H. P. (2005). A spatially explicit habitat selection model incorporating home range behavior. *Ecology*, *86*(5), 1199–1205.

Richter, L., Balkenhol, N., Raab, C., Reinecke, H., Meißner, M., Herzog, S., Isselstein, J., & Signer, J. (2020). So close and yet so different: The importance of considering temporal dynamics to understand habitat selection. *Basic and Applied Ecology*, *43*, 99–109.

Signer, J., Fieberg, J., & Avgar, T. (2017). Estimating utilization distributions from fitted step-selection functions. *Ecosphere*, *8*(4), e01771.

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

Signer, J., Fieberg, J., Reineking, B., Schlägel, U., Smith, B., Balkenhol, N., & Avgar, T. (2024). Simulating animal space use from fitted integrated step-selection functions (iSSF). *Methods in Ecology and Evolution*, *15*(1), 43–50.

Simpson, G. L. (2023). *Gratia: Graceful ggplot-based graphics and other functions for GAMs fitted using mgcv*. R package version 0.8.1.38.

Smith, B. J., MacNulty, D. R., Stahler, D. R., Smith, D. W., & Avgar, T. (2023). Density-dependent habitat selection alters drivers of population distribution in northern Yellowstone elk. *Ecology Letters*, *26*(2), 245–256.

Therneau, T. (2022). *A package for survival analysis in R*. R package version 3.3-1.

Wood, S. N. (2017). *Generalized additive models: An introduction with R*. Chapman and Hall/CRC.

Wood, S. N., Pya, N., & Säfken, B. (2016). Smoothing parameter and model selection for general smooth models. *Journal of the American Statistical Association*, *111*(516), 1548–1563.

Wood, S. N., Scheipl, F., & Faraway, J. J. (2013). Straightforward intermediate rank tensor product smoothing in mixed models. *Statistics and Computing*, *23*(3), 341–360.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Appendix A: Relationship to a Cox proportional hazards model.

Appendix B: Simulations.

How to cite this article: Klappstein, N. J., Michelot, T., Fieberg, J., Pedersen, E. J., & Mills Flemming, J. (2024). Step selection functions with non-linear and random effects. *Methods in Ecology and Evolution*, *15*, 1332–1346. https://doi.org/10.1111/2041-210X.14367