Check for updates

Journal name: *Ecology*
Manuscript type: *Article*
Handling Editor: Matthew J. Kauffman

Accepted Article

A multistate Langevin diffusion for inferring behavior-specific habitat selection and utilization distributions

Brett T. McClintock* and Michelle E. Lander

Marine Mammal Laboratory, Alaska Fisheries Science Center, NOAA, National Marine Fisheries Service, 7600 Sand Point Way NE, Seattle, WA 98115 USA

*Corresponding author. Email: brett.mcclintock@noaa.gov

Open Research: Supporting data and code (McClintock, 2023) are available in Zenodo at https://doi.org/10.5281/zenodo.8327405.

Key words: animal movement, continuous time, habitat selection, hidden Markov model, Langevin diffusion, resource selection, step selection, utilization distribution

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/ecy.4186

This article is protected by copyright. All rights reserved.

Accepted Article

Abstract. The identification of important habitat and the behavior(s) associated with it is critical to conservation and place-based management decisions. Behavior also links life-history requirements and habitat use, which is key to understanding *why* animals use certain habitats. Animal population studies often use tracking data to quantify space use and habitat selection, but they typically either ignore movement behavior (e.g., foraging, migrating, nesting) or adopt a two-stage approach that can induce bias and fail to propagate uncertainty. We develop a habitat-driven Langevin diffusion for animals that exhibit distinct movement behavior states, thereby providing a novel single-stage statistical method for inferring behavior-specific habitat selection and utilization distributions in continuous time. Practitioners can customize, fit, assess, and simulate our integrated model using the provided R package. Simulation experiments demonstrated that the model worked well under a range of sampling scenarios as long as observations were of sufficient temporal resolution. Our simulations also demonstrated the importance of accounting for different behaviors and the misleading inferences that can result when these are ignored. We provide case studies using plains zebra (*Equus quagga*) and Steller sea lion (*Eumetopias jubatus*) telemetry data. In the zebra example, our model identified distinct "encamped" and "exploratory" states, where the "encamped" state was characterized by strong selection for grassland and avoidance of other vegetation types, which may represent selection for foraging resources. In the sea lion example, our model identified distinct movement behavior modes typically associated with this marine central-place forager and, unlike previous analyses, found foraging-type movements to be associated with steeper offshore slopes characteristic of the continental shelf, submarine canyons, and seamounts that are believed to enhance prey concentrations. This is the first single-stage approach for inferring behavior-specific habitat selection and utilization distributions from tracking data that can be readily implemented with user-friendly software. As certain behaviors are often more relevant to specific conservation or management objectives, practitioners can use our model to help inform

This article is protected by copyright. All rights reserved.

19399170, ja, Downloaded from https://esajournals.onlinelibrary.wiley.com/doi/10.1002/ecy.4186 by University Of Washington, Wiley Online Library on [10/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

the identification and prioritization of important habitats. Moreover, by linking individual-level movement behaviors to population-level spatial processes, the multistate Langevin diffusion can advance inferences at the intersection of population, movement, and landscape ecology.

1 Introduction

The identification of important habitat is critical to wildlife population management and conservation, as well as our understanding of the links between animal behavior and their environment (e.g., Camaclang et al., 2015; Rice et al., 2017; Lennox et al., 2019; Dougherty et al., 2022). This is often accomplished through animal-borne telemetry studies that attempt to quantify habitat selection and/or utilization distributions (e.g., Hooten et al., 2017; Matthiopoulos et al., 2020). These statistical methods include kernel density estimators (e.g., Worton, 1989), Brownian bridges (e.g., Horne et al., 2007), resource selection functions (e.g., Manly et al., 2010), step selection functions (e.g., Thurfjell et al., 2014; Fieberg et al., 2021), discrete-space movement models (e.g., Wilson et al., 2018), and, more generally, spatio-temporal point process models (e.g., Johnson et al., 2013). There has also been recent work developing local (i.e., step-by-step) individual-level movement models that scale in time and space to the expected (population-level) utilization distribution in a coherent framework (e.g., Michelot et al., 2019*a*; Potts and Schlägel, 2020), including the habitat-driven Langevin diffusion of Michelot et al. (2019*b*). Each approach has its pros and cons (for brief reviews, see Michelot et al., 2019*b*; Northrup et al., 2022), but all suffer from one clear deficiency: they assume movement characteristics and/or habitat selection do not change as a function of behaviors (e.g., resting, foraging, nesting) arising from an animal's internal state (e.g., fatigue, hunger, reproductive instinct).

This would be less important if animals selected for the same habitat characteristics regardless

This article is protected by copyright. All rights reserved.

Accepted Article

19399170, ja, Downloaded from https://esajournals.onlinelibrary.wiley.com/doi/10.1002/ecy.4186 by University Of Washington, Wiley Online Library on [10/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

of behavior, but this is known to be untrue (e.g., Bouyer et al., 2015; Northrup et al., 2022). For example, migratory birds prefer different habitat types for winter, summer, and stop-over sites, and their movement properties will depend on whether they are resting, nesting, foraging, or migrating. Marine central-place foragers, such as pinnipeds and seabirds, typically use land (or ice) for resting, molting, and reproduction, but forage at sea. Other specific examples include behavior-dependent habitat selection in foraging and non-foraging movements of zebras (Dougherty et al., 2022), encamped and exploratory movements of African elephants (Roever et al., 2014), sage grouse (Picardi et al., 2022), and black bears (Karelus et al., 2017), resting, foraging, and travelling in wild pigs (Clontz et al., 2021) and African lions (Suraci et al., 2019), within-patch and between-patch movements in caribou (Johnson et al., 2002) and California pumas (Zeller et al., 2014), resting and active movements in Canada lynx (Squires et al., 2013), resting, predatory, and travelling movements in coyotes (Wilson et al., 2012), and resting, running, and travelling movements in African wild dogs (Abrahms et al., 2016). Clearly, attempting to infer habitat selection and utilization distributions across individuals exhibiting a range of different behaviors in response to internal and external drivers could result in erroneous inferences and potentially ineffective management decisions (e.g., Wilson et al., 2012; Roever et al., 2014; Abrahms et al., 2016; Dougherty et al., 2022; Picardi et al., 2022). As certain behaviors are often more relevant to specific conservation or management objectives, behavior-specific inferences could help improve our ability to identify and prioritize important habitats.

Linking habitat and the behavior(s) associated with it has benefits beyond informing behavior- and place-based management decisions that are critical to conservation (e.g., Lennox et al., 2019; Glennie et al., 2022; Northrup et al., 2022). For example, ignoring behavior limits our ability to understand habitat use in the context of fitness and natural selection. Behavior provides the fundamental link between life-history requirements and habitat use, and it is therefore key to un-

This article is protected by copyright. All rights reserved.

Accepted Article

19399170, ja, Downloaded from https://esajournals.onlinelibrary.wiley.com/doi/10.1002/ecy.4186 by University Of Washington, Wiley Online Library on [10/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

derstanding *why* animals use certain habitats (e.g., Nathan et al., 2008; Beyer et al., 2010; Owen-Smith et al., 2010). Linking such individual-level movement behaviors to population-level spatial processes are also key to advancing inferences at the intersection of population, movement, and landscape ecology (e.g., Morales et al., 2010; McClintock et al., 2022).

Despite a clear need and desire, current statistical approaches to inferring behavior-specific habitat selection and utilization distributions from tracking data are limited (e.g., Thurfjell et al., 2014; Northrup et al., 2022). Most adopt a two-stage approach (e.g., Johnson et al., 2002; Squires et al., 2013; Roever et al., 2014; Zeller et al., 2014; Picardi et al., 2022), where commonly a discrete-time hidden Markov model is first used for behavioral state classification based on movement step lengths and turn angles (e.g., "encamped" vs. "exploratory" *sensu* Morales et al., 2004), and then separate step selection analyses (e.g., Thurfjell et al., 2014) are performed using only the locations corresponding to the behavior(s) of interest. However, such approaches typically ignore habitat in the behavioral classification stage and fail to propagate state uncertainty into the habitat selection stage. This can conflate the movement and selection processes, particularly if the habitat-independent movement properties are not sufficiently distinct among the states, thereby inducing bias and underestimation of uncertainty (e.g., Prima et al., 2022; Pohle et al., 2023). Other approaches attempt to account for behavioral changes by performing independent analyses of subsets of the data (e.g., by season; Rice et al., 2017), but these require *a priori* determination of when these shifts occur and do not allow for the pooling of information among subsets or testing hypotheses related to when and why these changes occur. The dynamic Brownian bridge model of Kranstauber et al. (2012) does account for variability in the variance parameter of the Brownian motion using behavioral change point analysis, but the resulting utilization distribution is an amalgam of this behavioral variation that, as with other Brownian bridge approaches, is not a function of habitat covariates (i.e., linking space use to habitat thus still requires a two-stage approach). Ped-

This article is protected by copyright. All rights reserved.

Accepted Article

19399170, ja, Downloaded from https://esajournals.onlinelibrary.wiley.com/doi/10.1002/ecy.4186 by University Of Washington, Wiley Online Library on [10/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

ersen et al. (2011) proposed a Bayesian continuous-time, discrete-space hidden Markov model to simultaneously account for behavior state switching and heterogeneous space use, but the resulting "residency distributions" are not a function of habitat covariates and cannot be directly interpreted as utilization distributions.

The multistate step selection model of Nicosia et al. (2017) is, to our knowledge, the first single-stage approach that can account for different movement behavior states and propagate state uncertainty into the estimation of habitat selection parameters. Their model relaxes the standard step selection analysis assumptions of time-invariant movement and selection processes, as well as the assumption of serially uncorrelated movement (by modeling the latent movement behavior state sequence as a discrete-time Markov chain). However, unlike other commonly used hidden Markov models of animal movement (e.g., Morales et al., 2004; Jonsen et al., 2005; Langrock et al., 2012; McClintock et al., 2012), the utility of their approach is limited because it does not accommodate multiple individuals, missing or temporally-irregular data, or habitat covariates for the state transition probabilities (e.g., for testing hypotheses related to the drivers of movement behavior dynamics). Recent extensions have attempted to address several of these limitations (e.g., Prima et al., 2022; Klappstein et al., 2023), but, as a step selection model, the approach of Nicosia et al. (2017) describes selection at a local scale that strongly depends on the (temporally-regular) scale of the observed data and the specific choice of the movement (or "availability") kernel (e.g., Thurfjell et al., 2014). Although not addressed by Nicosia et al. (2017), estimation of the utilization distribution from step selection analyses typically requires simulation (e.g., Signer et al., 2017; Klappstein et al., 2023) or solving for the steady-state distribution of the underlying movement model (if it exists; e.g., Potts et al., 2014). With model fitting based on a custom expectation-maximization algorithm, the approach of Nicosia et al. (2017) is also computationally demanding and, in the absence of flexible user-friendly software, is unfortunately not readily accessible to

This article is protected by copyright. All rights reserved.

Accepted Article

19399170, ja, Downloaded from https://esajournals.onlinelibrary.wiley.com/doi/10.1002/ecy.4186 by University Of Washington, Wiley Online Library on [10/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

practitioners.

We address these shortcomings by developing a habitat-driven Langevin diffusion for animals that exhibit distinct movement behavior states, thereby providing a novel single-stage approach for inferring behavior-specific habitat selection and utilization distributions in continuous time. To facilitate implementation by non-statisticians, we provide a user-friendly R package for customizing, fitting, assessing, and simulating the model. Practitioners can therefore readily apply the multistate Langevin diffusion to better understand why animals use certain habitats and to help inform behavior- and place-based management decisions. After presenting the model in the next section, we perform simulation experiments assessing its performance under a range of sampling scenarios. We then provide two case studies using plains zebra (*Equus quagga*) and Steller sea lion (*Eumetopias jubatus*) data that have served as illustrative examples in previous (single-state) habitat selection analyses, and we highlight some of the potential advantages and challenges of the proposed approach.

2 Methods

2.1 Multistate Langevin diffusion

Michelot et al. (2019*b*) recently formulated a habitat-driven Langevin diffusion for inferring (individual-level) movement and (population-level) utilization distributions from animal tracking data. This continuous-time model can be written as the solution to the stochastic differential equation (SDE):

$$\mathrm{d}\boldsymbol{\mu}_t = \frac{1}{2}\boldsymbol{\Sigma}\nabla \log \pi\left(\boldsymbol{\mu}_t\right)\mathrm{d}t + \sqrt{\boldsymbol{\Sigma}}\mathrm{d}\mathbf{B}_t, \tag{1}$$

where $\boldsymbol{\mu}_t$ is location at time t, $\boldsymbol{\Sigma}$ is a variance-covariance matrix for the random Brownian motion

This article is protected by copyright. All rights reserved.

Accepted Article

19399170, ja, Downloaded from https://esajournals.onlinelibrary.wiley.com/doi/10.1002/ecy.4186 by University Of Washington, Wiley Online Library on [10/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

component, ∇ is the gradient operator, $\pi(\boldsymbol{\mu}_t)$ is the utilization distribution evaluated at $\boldsymbol{\mu}_t$, and $\mathbf{B}_t$ stands for standard Brownian motion (i.e., a continuous-time random walk). A key feature of the model is that the solution for Eq. 1 is a continuous-time Markov process with stationary (utilization) distribution π, and this is specified using the form of a resource selection function (e.g., Manly et al., 2010):

$$\pi(\boldsymbol{\mu} \mid \boldsymbol{\beta}) = \frac{\exp\left(\sum_{k=1}^{K} x_k(\boldsymbol{\mu})\,\beta_k\right)}{\int_{\mathcal{M}} \exp\left(\sum_{k=1}^{K} x_k(\mathbf{z})\,\beta_k\right) d\mathbf{z}}, \tag{2}$$

where $x_k(\boldsymbol{\mu})$ is the value of the kth habitat covariate (with continuous first-order partial derivatives) evaluated at location $\boldsymbol{\mu}$, β_k is the corresponding coefficient for covariate k, and $\mathcal{M}$ is the study area. In other words, the movement model is a continuous-time random walk with drift towards higher values of the utilization distribution. The potential function governing local (step-by-step) individual-level movements thus scales up in space and time to the expected (population-level) utilization distribution, which is expressed as a simple parametric function of the habitat covariates (e.g., elevation, slope, distance to forest, snow depth, relative prey abundance). The potential function steers movement toward (or away from) particular habitat features based on their gradients, much like a ball rolling along a hilly surface. The gradient is a vector field of partial derivatives pointing in the direction of the greatest rate of increase in the habitat covariate, with length corresponding to the rate of increase in that direction. Thus $\beta_k < 0$ indicates the animal generally moves to areas with lower values of covariate k, $\beta_k > 0$ indicates movement to areas with higher values, and $\beta_k = 0$ indicates no movement response.

While mathematically and theoretically appealing, the Langevin diffusion unrealistically assumes that the movement properties do not change with behavior. The movement scale and habitat selection parameters remain constant regardless of whether an animal is, for example, resting,

This article is protected by copyright. All rights reserved.

Accepted Article

19399170, ja, Downloaded from https://esajournals.onlinelibrary.wiley.com/doi/10.1002/ecy.4186 by University Of Washington, Wiley Online Library on [10/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

nesting, foraging, dispersing, or migrating. Just as in step selection analysis (e.g., Roever et al., 2014), this can induce bias in habitat selection estimates when applied to tracks that exhibit multiple movement behavior modes. However, several challenges arise when attempting to formulate a Langevin diffusion for animals that can switch among S movement behavior states, each exhibiting its own scale (or speed), potential surface, and corresponding utilization distribution. We require a mixture of S Langevin diffusions, where the (microscopic) movement properties and the (macroscopic) utilization distribution are conditional on the state active at time t (s_t):

$$\mathrm{d}\boldsymbol{\mu}_t = \frac{1}{2}\boldsymbol{\Sigma}_{s_t}\nabla \log \pi_{s_t}\left(\boldsymbol{\mu}_t\right)\mathrm{d}t + \sqrt{\boldsymbol{\Sigma}_{s_t}}\mathrm{d}\mathbf{B}_t \tag{3}$$

and

$$\pi_s\left(\boldsymbol{\mu} \mid \boldsymbol{\beta}_s\right) = \frac{\exp\left(\sum_{k=1}^{K_s} x_k\left(\boldsymbol{\mu}\right)\beta_{s,k}\right)}{\int_{\mathcal{M}} \exp\left(\sum_{k=1}^{K_s} x_k\left(\mathbf{z}\right)\beta_{s,k}\right)\mathrm{d}\mathbf{z}} \tag{4}$$

for state $s \in \{1, 2, \ldots, S\}$. We therefore have both an observed (continuous) movement process $(\boldsymbol{\mu}_t)$ and an unobserved (discrete) state process (s_t) to contend with. While this task may seem intimidating, we will rely on several simplifications and approximations to make it more tractable.

Hidden Markov models (HMMs; e.g., Zucchini et al., 2016) provide a framework for simultaneously modeling the latent (or "hidden") state process and, conditional on the state process, observations arising from the movement process through time. Discrete-time HMMs are routinely used in ecology to infer hidden state dynamics from indirect observations (e.g., McClintock et al., 2020), where it is assumed that the observation and state-switching times are discrete and regular. They characterize the state process (s_t) as a first-order Markov chain with $S \times S$ state transition probability matrix $\boldsymbol{\Gamma} = \{\gamma_{i,j}\}$, where $\gamma_{i,j}$ is the state transition probability from state i at time t to state j at time $t+1$ and $\sum_{j=1}^{S} \gamma_{i,j} = 1$. Although still relatively rarely used for ecological applica-

This article is protected by copyright. All rights reserved.

Accepted Article

19399170, ja, Downloaded from https://esajournals.onlinelibrary.wiley.com/doi/10.1002/ecy.4186 by University Of Washington, Wiley Online Library on [10/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

tions, HMMs can also be formulated for state and observation processes occurring in continuous time. The hidden state process of a continuous-time HMM is characterized by its infinitesimal generator matrix:

$$\mathbf{Q} = \begin{array}{c} \begin{array}{cccc} s_{t+1}=1 & s_{t+1}=2 & \dots & s_{t+1}=S \end{array} \\ \begin{bmatrix} -q_{1,1} & q_{1,2} & \dots & q_{1,S} \\ q_{2,1} & -q_{2,2} & \dots & q_{2,S} \\ \vdots & \vdots & \ddots & \vdots \\ q_{S,1} & q_{S,2} & \dots & -q_{S,S} \end{bmatrix} \begin{array}{l} s_t = 1 \\ s_t = 2 \\ \\ s_t = S \end{array} \end{array}, \tag{5}$$

where $q_{i,j} \geq 0$ for $i \neq j$ and $q_{i,i} = \sum_{j \neq i} q_{i,j}$. If $\mathbf{Q}$ is constant from time t to time $t+1$, the state transition probabilities can be calculated using the matrix exponential, $\mathbf{\Gamma}_t = \exp(\mathbf{Q}\Delta_t)$, where Δ_t is the (irregular) time interval between observations t and $t+1$. For the continuous-time observation process $(\boldsymbol{\mu}_t)$, we can approximate the solution to Eq. 3 using a more intuitive discrete approximation (e.g., Kloeden and Platen, 1992). By using an Euler discretization scheme to approximate the stochastic differential equation, the multistate Langevin diffusion can be expressed as

$$[\boldsymbol{\mu}_{t+1} \mid \boldsymbol{\mu}_t, \boldsymbol{\Sigma}, \boldsymbol{\beta}, s_t] \equiv \mathcal{N}\left(\boldsymbol{\mu}_t + \frac{\Delta_t}{2}\boldsymbol{\Sigma}_{s_t} \nabla \log \pi_{s_t}(\boldsymbol{\mu}_t \mid \boldsymbol{\beta}_{s_t}), \Delta_t \boldsymbol{\Sigma}_{s_t}\right) \tag{6}$$

for $t \in \{1, \dots, T\}$, where $\nabla \log \pi_s(\boldsymbol{\mu} \mid \boldsymbol{\beta}_s) = \sum_{k=1}^{K_s} \beta_{s,k} \nabla x_k(\boldsymbol{\mu})$.

Crucially, continuous-time HMMs for animal movement behavior typically suffer from a violation of the so-called "snapshot" property (e.g., Glennie et al., 2022). The snapshot property is violated when observations depend on Δ_t as in Eq. 6, and, therefore, the observed locations $(\boldsymbol{\mu}_t)$ depend on the behavioral state sequence over the entire interval between times t and $t+1$. While exact solutions exist for reconstructing the entire sequence, they are currently prohibitively de-

This article is protected by copyright. All rights reserved.

Accepted Article

19399170, ja, Downloaded from https://esajournals.onlinelibrary.wiley.com/doi/10.1002/ecy.4186 by University Of Washington, Wiley Online Library on [10/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

manding in terms of computation (e.g., Blackwell, 2018). We will instead adopt a more pragmatic approach and assume the Δ_t are sufficiently small relative to the scale of state switching (as is also required in applications of discrete-time HMMs). For this to be reasonably satisfied, an intuitive rule-of-thumb is $\Delta_t \leq \frac{1}{\max_i(q_{i,i})}$ (Glennie et al., 2022). Michelot et al. (2019*b*) demonstrated that the Langevin diffusion requires Δ_t to be sufficiently small in order for the Euler approximation scheme to recover the habitat selection ($\boldsymbol{\beta}$) and scale ($\boldsymbol{\Sigma}$) parameters. If Δ_t is too large, the magnitude of $\boldsymbol{\beta}$ tends to be underestimated (although the sign of the effects, i.e., selection or avoidance, tends to be estimated correctly) and, to a lesser extent, $\boldsymbol{\Sigma}$ tends to be underestimated. Thus, given Δ_t must already be relatively small for the Euler approximation to be reliable, the snapshot property is also likely to be reasonably satisfied for hidden state processes with moderate to high levels of serial correlation relative to the temporal resolution of the observations (see *Simulation study*).

With the state and observation process models in place, the continuous-time HMM likelihood can be evaluated using the efficient forward algorithm to integrate over the latent states:

$$\mathcal{L}\left(\boldsymbol{\theta} \mid \boldsymbol{\mu}\right) = \boldsymbol{\delta}\mathbf{P}_1 \left[\prod_{t=2}^{T} \boldsymbol{\Gamma}_t \mathbf{P}_t\right] \mathbf{1}, \tag{7}$$

where $\boldsymbol{\delta} = (\delta_1, \delta_2, \ldots, \delta_S)$ are the initial state probabilities,

$$\mathbf{P}_t = \begin{bmatrix} [\boldsymbol{\mu}_{t+1} \mid \boldsymbol{\mu}_t, \boldsymbol{\theta}, s_t = 1] & 0 & \ldots & 0 \\ 0 & [\boldsymbol{\mu}_{t+1} \mid \boldsymbol{\mu}_t, \boldsymbol{\theta}, s_t = 2] & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & [\boldsymbol{\mu}_{t+1} \mid \boldsymbol{\mu}_t, \boldsymbol{\theta}, s_t = S] \end{bmatrix}, \tag{8}$$

$\boldsymbol{\theta}$ is a vector of the model parameters, and $\mathbf{1}$ is an S-vector of ones. Hypotheses about the drivers of

This article is protected by copyright. All rights reserved.

Accepted Article

19399170, ja, Downloaded from https://esajournals.onlinelibrary.wiley.com/doi/10.1002/ecy.4186 by University Of Washington, Wiley Online Library on [10/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

behavioral state dynamics can be incorporated by expressing $\boldsymbol{\delta}$ and/or $\boldsymbol{\Gamma}_t$ as a function of explanatory covariates (e.g., sex, age, habitat), but for $\boldsymbol{\Gamma}_t$ they are assumed to be either time-independent or piecewise-constant time-dependent covariates. A movement trajectory alone can often be insufficient for distinguishing certain behaviors (e.g., "resting" vs. "foraging"), and additional data streams (e.g., dive activity, accelerometer, heart rate) can be readily included in $\mathbf{P}_t$ to help distinguish states with similar movement properties (e.g., McClintock et al., 2013; DeRuiter et al., 2017, see *Illustration: Steller sea lions*).

While the most general model includes state-dependent coefficients $(\boldsymbol{\beta}_s)$ and scale $(\boldsymbol{\Sigma}_s)$ parameters, this could in some cases result in overfitting and loss of biological interpretation (e.g., Phillips et al., 2015; Pohle et al., 2017). A more parsimonious parameterization could maintain state-dependent coefficients and a constant scale either for all states or for a smaller subsets of states (see *Illustration: Steller sea lions*). Alternatively, some or all of the coefficients could be kept constant while maintaining state-dependent scale parameters. In the material that follows, we will assume the scales of movement in each direction are equal and independent, that is, $\boldsymbol{\Sigma}_s = \sigma_s^2\mathbf{I}$, where σ_s^2 is the scale (or speed) parameter for state $s \in \{1, \ldots, S\}$ and $\mathbf{I}$ is the identity matrix.

2.2 Model fitting and checking

The multistate Langevin diffusion benefits from standard HMM machinery for inference, including direct numerical maximization of the likelihood (Eq. 7), local and global state decoding based on the forward-backward and Viterbi algorithms, stationary state probability distribution calculation (by solving $\boldsymbol{\lambda}\mathbf{Q} = \mathbf{0}$ for $\sum_{s=1}^{S} \lambda_s = 1$), likelihood-based model selection procedures such as the Akaike or Bayesian Information Criterion (AIC or BIC), and assessment of goodness-of-fit based on pseudo-residuals (e.g., Zucchini et al., 2016). Pseudo-residuals rely on evaluating the cumulative distribution function (CDF) of the observations, and, while the the CDF of Eq. 6 could be used,

This article is protected by copyright. All rights reserved.

Accepted Article

19399170, ja, Downloaded from https://esajournals.onlinelibrary.wiley.com/doi/10.1002/ecy.4186 by University Of Washington, Wiley Online Library on [10/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

a more meaningful interpretation for animal movement is provided by the Mahalanobis distance (see Appendix S1: Section S1 of *Supporting Information*). All of these inferential procedures for the multistate Langevin model have been incorporated into R package `momentuHMM` (version 2.0.0; McClintock and Michelot, 2018) via the `fitCTHMM` function for fitting continuous-time HMMs using maximum likelihood methods. The data preparation function `prepData` can also now automatically calculate gradients, $\nabla x_k\left(\boldsymbol{\mu}\right)$, for regular discretized grids of habitat covariates using bilinear interpolation (e.g., Kirkland 2010, pp. 261–263; Gardner et al. 2022).

2.3 Model simulation

If the hidden state process does not depend on location or time (i.e., no spatio-temporal covariates are included on the state transition rates), then it is straightforward to simulate the waiting time from time index τ until the next state switch at time index $\tau+1$ as $\Delta_\tau \sim \text{Exponential}\left(-q_{s_\tau,s_\tau}\right)$, where Δ_τ is the time interval between time indices τ and $\tau+1$, $s_{\tau+1}=j \mid s_\tau \sim \text{Categorical}\left(\frac{q_{s_\tau,j}}{\sum_{l\neq s_\tau} q_{s_\tau,l}}\right)$ for $j \neq s_\tau$, and $q_{s_\tau,j}$ is the jth element in row s_τ of $\mathbf{Q}$. If the hidden state process depends on location or time, it can be simulated exactly by adapting the approach of Blackwell et al. (2016) as described in Appendix S1: Section S2. The state-dependent movement process can then be approximated using Eq. 6, where the quality of the approximation increases as Δ_t decreases. `momentuHMM` now includes the `simCTHMM` function for HMM data simulation in continuous time (either "from scratch" or from a fitted model) , thereby facilitating further exploration of the multistate Langevin model's properties and simulation-based assessments of goodness-of-fit.

This article is protected by copyright. All rights reserved.

Accepted Article

19399170, ja, Downloaded from https://esajournals.onlinelibrary.wiley.com/doi/10.1002/ecy.4186 by University Of Washington, Wiley Online Library on [10/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

2.4 Simulation study

We performed a simulation study to evaluate the performance of the multistate Langevin model relative to the single-state model when fitted to data generated from a $S = 2$ state model across a range of expected values for Δ_t. We constructed $K = 4$ spatial covariates over a discretized 201×201 grid of $G = 40401$ cells centered at the origin, where the first three covariates (x_1, x_2, x_3) were randomly generated with the `simSpatialCov` function in R package `Rhabit` (Gloaguen and Michelot, 2020) and the fourth covariate (x_4) was calculated as the squared Euclidean distance between the centroid of each cell and the origin. The habitat coefficients for state 1 were $\beta_{1,1} = 6$, $\beta_{1,2} = -4$, $\beta_{1,3} = -5$, and $\beta_{1,4} = -0.1$, while the coefficients for state 2 were $\beta_{2,1} = -4$, $\beta_{2,2} = 6$, $\beta_{2,3} = 5$, and $\beta_{2,4} = -0.1$. The movement scale was smaller in state 1 ($\sigma_1^2 = 5$) than in state 2 ($\sigma_2^2 = 7.5$). Michelot et al. (2019*b*) proposed a Metropolis-adjusted Langevin algorithm (MALA) for assessing the accuracy of the Euler approximation (Eq. 6), where higher acceptance rates indicate a better approximation of the true process. At $\Delta_t = 0.01$, the MALA acceptance rates for each state were 99.5% (state 1) and 99.3% (state 2), indicating that the simulated (discretized) process is a good approximation of the true SDE for both states. The MALA acceptance rates were slightly lower for the movement properties of state 2 (likely attributable to the faster speed of state 2), but mean acceptance rates for both states were $> 93\%$ for $\Delta_t \leq 0.1$ (i.e., at $\geq 10\%$ of the temporal resolution of the generating process; see Appendix S1: Section S3, Figure S1). Michelot et al. (2019*b*) demonstrated that MALA acceptance rates that were $\geq 95\%$ generally resulted in little bias, but the performance of the approximation is expected to decline as Δ_t increases (and MALA acceptance rates decrease). To demonstrate how the speed of movement (relative to Δ_t, $\boldsymbol{\beta}$, and the scale of the covariate grid) can affect model performance, we performed an additional simulation study using the same design points but with $\sigma_1^2 = 0.5$ and $\sigma_2^2 = 2.5$. At these slower

This article is protected by copyright. All rights reserved.

Accepted Article

19399170, ja, Downloaded from https://esajournals.onlinelibrary.wiley.com/doi/10.1002/ecy.4186 by University Of Washington, Wiley Online Library on [10/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

speeds, the mean MALA acceptance rates were $\geq 95\%$ for all $\Delta_t \leq 1$ (i.e., at $\geq 1\%$ of the temporal resolution of the generating process) for state 1 and $\geq 95\%$ for all $\Delta_t \leq 0.2$ for state 2 (see Appendix S1: Section S3, Figure S2).

Tracks of length $T = 50000$ for $N = 20$ individuals were simulated at a "high" temporal resolution using Eq. 6 with $\Delta_t \sim \text{Exponential}(100)$ (such that $\text{E}(\Delta_t) = 0.01$), with $\log(q_{1,2}) = \log(q_{2,1}) = -2$ (such that serial correlation in the hidden state process is relatively high). Initial locations were assigned for individual $n \in \{1, 2, \ldots, N\}$ by: 1) drawing the initial state $s_{n,1} \in \{1, 2\}$ from a categorical distribution with equal probability; 2) drawing the initial grid cell $g_n \mid s_{n,1} \sim \text{Categorical}\left(\tilde{\pi}_{s_{n,1}}\right)$, where $\tilde{\boldsymbol{\pi}}_{s_{n,1}} = \left(\tilde{\pi}_{s_{n,1},1}, \tilde{\pi}_{s_{n,1},2}, \ldots, \tilde{\pi}_{s_{n,1},G}\right)$, $\tilde{\pi}_{s_{n,1},g} = \frac{\exp\left(\sum_{k=1}^{K} x_k(\mathbf{c}_g)\beta_{s_{n,1},k}\right)}{\sum_{l=1}^{G}\exp\left(\sum_{k=1}^{K} x_k(\mathbf{c}_l)\beta_{s_{n,1},k}\right)}$, and $x_k(\mathbf{c}_g)$ is the value of the kth covariate evaluated at the centroid of cell g ($\mathbf{c}_g$); and 3) setting $\boldsymbol{\mu}_{n,1} = \mathbf{c}_{g_n}$. After the "high" resolution tracks were simulated, they were subsampled uniformly at random to form seven datasets with increasing Δ_t such that $\text{E}(\Delta_t) \in \{0.01, 0.02, 0.05, 0.1, 0.2, 0.5, 1\}$ (respectively corresponding to 100%, 50%, 20%, 10%, 5%, 2%, and 1% of the temporal resolution of the data-generating process). For each subsampled dataset, both the generating 2-state Langevin model (Eq. 7) and the single-state Langevin model were fitted to the pooled data including all N individuals using maximum likelihood methods. Thus the pooled datasets consisted of $NT = 1,000,000$ observations for the scenario with $\text{E}(\Delta_t) = 0.01$, but only $NT/100 = 10,000$ observations for the scenario with $\text{E}(\Delta_t) = 1$. This process was repeated 100 times for summarizing estimator performance under each Δ_t scenario. Data simulation and model fitting were performed using R package `momentuHMM` (version 2.0.0; McClintock and Michelot, 2018), with code available in *Supporting Information*.

This article is protected by copyright. All rights reserved.

Accepted Article

19399170, ja, Downloaded from https://esajournals.onlinelibrary.wiley.com/doi/10.1002/ecy.4186 by University Of Washington, Wiley Online Library on [10/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

2.5 Illustration: plains zebra

To illustrate a potential workflow for practitioners, we use a plains zebra GPS telemetry track from previous habitat selection analyses (Michelot et al., 2020; Klappstein et al., 2023). Locations were collected every 30 minutes from January to May 2014 in Hwange National Park (Zimbabwe). The track consists of $T = 7121$ locations, with 125 missing observations resulting in $\Delta_t \in \{30, 60, 90, 120, 150\}$ mins with respective frequencies of 7002, 114, 1, 1, and 2 observations. The habitat layer is a vegetation map consisting of 1510×1686 0.03 km^2 grid cells, with four categories (grassland, bushed grassland, bushland, and woodland; see Fig. 1). Because the Langevin model requires habitat covariates to have continuous first-order partial derivatives, we calculated the shortest Euclidean distance from each cell to each habitat category using the `raster::distance` function (Hijmans, 2023) and used these distances as the habitat covariates (x_1 = "d2grass", x_2 = "d2bushgrass", x_3 = "d2bush", and x_4 = "d2wood", respectively). Because the habitat covariates are the distance (in km) from each vegetation type, a negative coefficient indicates selection (i.e., movement toward the vegetation type) and a positive coefficient indicates avoidance (i.e., movement away from the vegetation type). We calculated the gradients for each distance covariate (∇x_k) via bilinear interpolation using the `momentuHMM::prepData` function. We then fit a $S = 2$ state model using the `momentuHMM::fitCTHMM` function, where such models typically identify a slow (i.e., "encamped") and a fast (i.e., "exploratory") movement behavior state (e.g., Morales et al., 2004). To facilitate comparisons with previous analyses, we extended Eq. 6 to include state-dependent terms for short-term persistence in velocity ($\beta_{s,0}$):

$$[\boldsymbol{\mu}_{t+1} \mid \boldsymbol{\mu}_t, \boldsymbol{\Sigma}, \boldsymbol{\beta}, s_t] \equiv \mathcal{N}\left(\boldsymbol{\mu}_t + (\boldsymbol{\mu}_t - \boldsymbol{\mu}_{t-1})\,\Delta_t \beta_{s_t,0} + \frac{\Delta_t}{2}\boldsymbol{\Sigma}_{s_t} \sum_{k=1}^{4} \beta_{s_t,k} \nabla x_k\left(\boldsymbol{\mu}_t\right), \Delta_t \boldsymbol{\Sigma}_{s_t}\right) \quad (9)$$

This article is protected by copyright. All rights reserved.

Accepted Article

19399170, ja, Downloaded from https://esajournals.onlinelibrary.wiley.com/doi/10.1002/ecy.4186 by University Of Washington, Wiley Online Library on [10/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

for $s_t \in \{1, 2\}$. Similar to Klappstein et al. (2023), we included a cyclical time of day covariate on the state transition rates:

$$\log\left(q_{i,j,t}\right) = \alpha_{i,j,0} + \alpha_{i,j,1}\cos\left(\frac{2\pi h_t}{24}\right) + \alpha_{i,j,2}\sin\left(\frac{2\pi h_t}{24}\right) \tag{10}$$

for $i \neq j$, where h_t is the time of day in hours (i.e., $h_t \in \{0, 1, \ldots, 23\}$). After fitting the model, we calculated the (state-dependent) utilization distributions over the discretized grid of covariates as:

$$\pi_s\left(\mathbf{c}_g \mid \boldsymbol{\beta}_s\right) = \frac{\exp\left(\sum_{k=1}^{4} x_k(\mathbf{c}_g)\beta_{s,k}\right)}{\sum_{l=1}^{G}\exp\left(\sum_{k=1}^{4} x_k(\mathbf{c}_l)\beta_{s,k}\right)} \tag{11}$$

for $s \in \{1, 2\}$, where $\mathbf{c}_g$ is the coordinates for the centroid of cell g, $x_k(\mathbf{c}_g)$ is the value of the kth habitat covariate evaluated at $\mathbf{c}_g$, and $G = 2,545,860$ cells. Post analysis included calculation of the Viterbi-decoded states (using `momentuHMM::viterbi`), stationary state probablities (using `momentuHMM::plotStationary`), pseudo-residuals (using `momentuHMM::plotPR`), and MALA acceptance rates (using a modified version of `Rhabit:::simMALA`). Although there is no theoretical "overall" utilization distribution for the multistate Langevin model, we calculated a weighted average of the state-dependent utilization distributions for comparison with previous analyses, where the weights were determined from the time spent in each state based on the Viterbi algorithm:

$$\tilde{\pi}\left(\mathbf{c}_g \mid \boldsymbol{\beta}\right) = \sum_{s=1}^{2} w_s \pi_s\left(\mathbf{c}_g \mid \boldsymbol{\beta}_s\right), \tag{12}$$

where $w_s = \frac{\sum_{t=1}^{T} I(\hat{s}_t = s)\Delta_t}{\sum_{t=1}^{T}\Delta_t}$, $\hat{s}_t$ is the Viterbi-decoded state at time t, and I is the indicator function. Data and code for reproducing the entire analysis are available in *Supporting Information*.

This article is protected by copyright. All rights reserved.

Accepted Article

19399170, ja, Downloaded from https://esajournals.onlinelibrary.wiley.com/doi/10.1002/ecy.4186 by University Of Washington, Wiley Online Library on [10/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

2.6 Illustration: Steller sea lions

In Appendix S1: Section S5, we provide a fully-worked example application using Steller sea lion (SSL) satellite telemetry data, including background information, data collection, model specification, model fitting, model checking, and biological interpretation. The dataset consists of location and conductivity (i.e., wet/dry) data collected from $N = 3$ adult females tagged over winter in the Aleutian Islands of Alaska, USA (Fig. 2). Both Wilson et al. (2018) and Michelot et al. (2019*b*) found distance to the nearest haul-out or rookery site to be by far the most important predictor for the movements of these SSLs, where movements became faster and more directed toward the sites as this distance increased. This is not surprising given adult female SSLs are central-place foragers that must make frequent, but relatively short, foraging trips from these sites (e.g., Raum-Suryan et al., 2004). When not foraging, they tend to rest or care for young at or around these sites on land (i.e., "haulout"). The movement behavior therefore changes based on whether the animal is on a foraging trip or at a site. Furthermore, foraging trips can consist of benthic near-shore foraging and/or off-shelf foraging farther away from land. We therefore fitted a suite of multistate Langevin models (ranging from $S = 2$ to $S = 5$ states) intended to distinguish these movement behaviors (e.g., "outbound", "foraging", "inbound", "haulout") and compared these state-specific inferences with those from single-state analyses. By explicitly disentangling the different movement behaviors SSLs exhibit as central-place foragers, multistate models could potentially provide additional insights about foraging habitat selection and activity budgets for these physiologically important behaviors. For each movement behavior state, habitat selection (and the corresponding utilization distribution) was modeled as a function of depth, sea floor slope, a depth $\times$ slope interaction ("depth:slope"), distance to the nearest haul-out or rookery site ("d2site"), and squared distance to site ("d2site2"; Appendix S1: Section S5, Figure S12). Distance to site was also included as a

This article is protected by copyright. All rights reserved.

Accepted Article

19399170, ja, Downloaded from https://esajournals.onlinelibrary.wiley.com/doi/10.1002/ecy.4186 by University Of Washington, Wiley Online Library on [10/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

state transition rate covariate.

With S states, $K = 5$ habitat covariates, $R = 1$ state transition rate covariate, and $N = 3$ individuals, a fully unconstrained joint likelihood with individual-level effects on all parameters would include $(S-1) + (S + KS) + S(S-1)(R+1)$ parameters (corresponding to $\boldsymbol{\delta}$, $\mathbf{P}$, and $\boldsymbol{\Gamma}$, respectfully) for each individual:

$$\mathcal{L}\left(\boldsymbol{\theta} \mid \boldsymbol{\mu}\right) = \prod_{n=1}^{N} \boldsymbol{\delta}_n \mathbf{P}_{n,1} \left[\prod_{t=2}^{T_n} \boldsymbol{\Gamma}_{n,t} \mathbf{P}_{n,t}\right] \mathbf{1}. \tag{13}$$

We therefore imposed some biologically reasonable constraints to reduce the number of estimated parameters, thereby facilitating both model fitting and interpretation. Complete technical details can be found in Appendix S1: Section S5, but we found that additional constraints were necessary for the $S = 4$ and $S = 5$ state models because the inclusion of more than two or three $\sigma_{n,s}$ terms, respectively, for each individual resulted in much more unstable optimization of the likelihood and a merging of "outbound", "inbound", and "foraging" movements into one or more states with no plausible explanation in terms of central-place foraging. For the most heavily-constrained 4-state model ($1 =$ "outbound", $2 =$ "foraging", $3 =$"inbound", $4 =$ "haulout"), the scale of movement was divided into two main categories while allowing the general movement patterns of a marine central-place forager to be preserved. The first scale category, "at sea", is intended to capture the "outbound", "foraging", and "inbound" movement states associated with foraging trips. The second category, "haulout", is intended to capture periods that tend to be dry with little horizontal movement. In this more heavily constrained model, the "at sea" states shared a scale parameter

This article is protected by copyright. All rights reserved.

Accepted Article

19399170, ja, Downloaded from https://esajournals.onlinelibrary.wiley.com/doi/10.1002/ecy.4186 by University Of Washington, Wiley Online Library on [10/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

$(\sigma^2_{n,AS})$ but had their own habitat selection coefficients ($\boldsymbol{\beta}_s$ for $s \in \{1,2,3\}$):

$$[\boldsymbol{\mu}_{n,t+1} \mid \boldsymbol{\mu}_{n,t}, \boldsymbol{\theta}, s_{n,t}] \equiv \begin{cases} \mathcal{N}\left(\boldsymbol{\mu}_{n,t} + \frac{\sigma^2_{n,AS}\Delta_t}{2} \nabla \log \pi_{s_{n,t}}\left(\boldsymbol{\mu}_{n,t} \mid \boldsymbol{\beta}_{s_{n,t}}\right), \sigma^2_{n,AS}\Delta_t \mathbf{I}\right) & \text{if } s_{n,t} \in \{1,2,3\} \\ \mathcal{N}\left(\boldsymbol{\mu}_{n,t}, \sigma^2_{HO}\Delta_t \mathbf{I}\right) & \text{if } s_{n,t} = 4 \end{cases}, \tag{14}$$

Thus the "haulout" state ($s_{n,t} = 4$) is a continuous-time simple random walk (with scale parameter σ_{HO}), and, while "at sea", each individual is assumed to have the same underlying state-dependent utilization distribution, $\pi_{s_{n,t}}\left(\boldsymbol{\mu}_{n,t} \mid \boldsymbol{\beta}_{s_{n,t}}\right)$ for $s_{n,t} \in \{1,2,3\}$, but each animal is allowed to traverse the different surfaces at their own speed ($\sigma^2_{n,AS}$). For the most heavily-constrained 5-state model including an additional "foraging" state ($1 =$ "outbound", $2 =$ "foraging1", $3 =$" foraging2", $4 =$ "inbound", $5 =$ "haulout"), both "foraging" states shared the same utilization surface (i.e., $\boldsymbol{\beta}_2 = \boldsymbol{\beta}_3$), but the additional "foraging1" state was allowed to have its own speed (thereby potentially capturing slower horizontal movements that can occur near shore or during benthic foraging trips):

$$[\boldsymbol{\mu}_{n,t+1} \mid \boldsymbol{\mu}_{n,t}, \boldsymbol{\theta}, s_{n,t}] \equiv \begin{cases} \mathcal{N}\left(\boldsymbol{\mu}_{n,t} + \frac{\sigma^2_{n,AS}\Delta_{n,t}}{2} \nabla \log \pi_{s_{n,t}}\left(\boldsymbol{\mu}_{n,t} \mid \boldsymbol{\beta}_{s_{n,t}}\right), \sigma^2_{n,AS}\Delta_{n,t} \mathbf{I}\right) & \text{if } s_{n,t} \in \{1,3,4\} \\ \mathcal{N}\left(\boldsymbol{\mu}_{n,t} + \frac{\sigma^2_{n,2}\Delta_{n,t}}{2} \nabla \log \pi_2\left(\boldsymbol{\mu}_{n,t} \mid \boldsymbol{\beta}_2\right), \sigma^2_{n,2}\Delta_{n,t} \mathbf{I}\right) & \text{if } s_{n,t} = 2 \\ \mathcal{N}\left(\boldsymbol{\mu}_{n,t}, \sigma^2_{HO}\Delta_{n,t} \mathbf{I}\right) & \text{if } s_{n,t} = 5 \end{cases}, \tag{15}$$

The models were compared and assessed using a combination of AIC, pseudo-residual plots, and biological interpretation (e.g., Pohle et al., 2017). We also examined state transition rate ($\mathbf{Q}$) estimates for evidence of potential violations of the snapshot property, as well as MALA acceptance rates for evaluating the quality of the Euler approximation (see Appendix S1: Section S5). Data and code for reproducing this analysis using `momentuHMM` are available in *Supporting Information*.

This article is protected by copyright. All rights reserved.

Accepted Article

19399170, ja, Downloaded from https://esajournals.onlinelibrary.wiley.com/doi/10.1002/ecy.4186 by University Of Washington, Wiley Online Library on [10/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

3 Results

3.1 Simulation study

The multistate model performed nominally under the data-generating scenarios with $\mathrm{E}(\Delta_t) = 0.01$. The magnitude of the habitat coefficients and (to a lesser extent) the scale parameters tended to be underestimated as Δ_t increased, but the sign of the habitat effect (i.e., selection or avoidance) was always preserved (Figs 3 and Appendix S1: Section S3, Figure S3). This nevertheless resulted in a gradual degradation of the estimated state-dependent utilization distributions as Δ_t increased (Figs 4 and Appendix S1: Section S3, Figure S4). The state transition rates were generally well estimated, but state decoding became less reliable as Δ_t increased (Appendix S1: Section S3, Figures S5–S6). This is likely attributable to the (fine-scale) movement properties of the states becoming more difficult to distinguish at coarser observation times, as well as violations of the snapshot property. With $\log(q_{1,2}) = \log(q_{2,1}) = -2$, we would expect violation of the snapshot property to become more problematic for scenarios with $\mathrm{E}(\Delta_t) \geq 0.1$, which is where bias in the scale parameters becomes more noticeable and state decoding accuracy begins to decline more rapidly. However, the degradation in model performance as Δ_t increased was mitigated in the simulation scenarios with slower speeds (i.e., with $\sigma_1^2 = 0.5$ and $\sigma_2^2 = 2.5$). This is likely attributable to the Euler discretization scheme being a better approximation at slower speeds (relative to Δ_t, $\boldsymbol{\beta}$, and the scale of the covariate grid) and more distinct state-dependent observation distributions.

The single-state Langevin model produced spurious inferences about all of the habitat effects with the exception of x_4, where the effect was the same for both states (Appendix S1: Section S3, Figures S7–S8). With $\sigma_1^2 = 5$ and $\sigma_2^2 = 7.5$, the single-state model tended to "split the difference" between the two states, with no estimated effect for x_1, a moderate positive effect for x_2, and a smaller positive effect for x_3 (Appendix S1: Section S3, Figure S7). However, as evident from the

This article is protected by copyright. All rights reserved.

Accepted Article

19399170, ja, Downloaded from https://esajournals.onlinelibrary.wiley.com/doi/10.1002/ecy.4186 by University Of Washington, Wiley Online Library on [10/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

simulation scenarios with $\sigma_1^2 = 0.5$ and $\sigma_2^2 = 2.5$ (Appendix S1: Section S3, Figure S8), these patterns are likely the result of a complicated combination of factors, including the spatial autocorrelation structure of the underlying state-dependent utilization distributions, the scale parameters of the two states, and the amount of time spent in each state. The estimated utilization distribution is therefore a spurious amalgam that reflects space use but cannot be reliably interpreted in terms of the habitat covariates (Appendix S1: Section S3, Figures S9–S10). This relatively simple simulation study demonstrates the importance of accounting for different behaviors when attempting to infer utilization distributions, as well as the misleading inferences that can result when these behaviors are ignored.

3.2 Illustration: plains zebra

The $S = 2$ state model identified an "encamped" state with less short-term directional persistence ($\beta_{1,0} = 0.0018$, 95% CI: 0.0014, 0.0021) and slower speed ($\sigma_1 = 0.0083$, 95% CI: 0.0081, 0.0086) and an "exploratory" state with more short-term directional persistence ($\beta_{1,0} = 0.020$, 95% CI: 0.019, 0.021) and faster speed ($\sigma_2 = 0.075$, 95% CI: 0.073, 0.077). The "encamped" state was characterized by strong selection for grassland and avoidance of the other vegetation types (Fig. 1), which is consistent with previous findings (Michelot et al., 2020) and may represent selection for foraging resources. Conversely, the "exploratory" state was characterized by avoidance of bushed grassland and woodland. We found strong evidence of an effect of time of day on the state transition rate from the "encamped" state to the "exploratory" state ($\alpha_{1,2,0} = -4.73$, 95% CI: -4.82, -4.64; $\alpha_{1,2,1} = -0.31$, 95% CI: -0.44, -0.19; $\alpha_{1,2,2} = 0.24$, 95% CI: 0.12, 0.35), but not vice versa ($\alpha_{2,1,0} = -4.14$, 95% CI: -4.22, -4.05; $\alpha_{2,1,1} = -0.01$, 95% CI: -0.12, 0.10; $\alpha_{2,1,2} = -0.07$, 95% CI: -0.20, 0.06). The resulting stationary state probabilities suggest the zebra wast more likely to be in the less active "encamped" state in the afternoon and evening, and

This article is protected by copyright. All rights reserved.

Accepted Article

19399170, ja, Downloaded from https://esajournals.onlinelibrary.wiley.com/doi/10.1002/ecy.4186 by University Of Washington, Wiley Online Library on [10/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

it was most likely to be in the more active "exploratory" state in the morning (Fig. 1). Based on the Viterbi-decoded states, the majority (63%) of the time series was spent in the "encamped" state. Plots of the Viterbi-decoded states relative to the state-dependent utilization distributions confirm the "encamped" state tended to be more clustered in the vicinity of grassland, whereas the "exploratory" state tended to be more dispersed (Fig. 1). Unlike the multistate step selection function of Klappstein et al. (2023), where selection for grassland or bushed grassland was found in both states, the multistate Langevin model appears to have more clearly distinguished the biological behaviours in this example. Nevertheless, the weighted average of the state-dependent utilization distributions (Eq. 12) is quite similar to the estimated utilization distributions from previous studies (Michelot et al., 2020; Klappstein et al., 2023, Fig. 1). While we emphasize that Eq. 12 has no theoretical underpinnings, it is nevertheless reassuring that this crude approximation is consistent with simulation-based approaches.

MALA acceptance rates were 99.96% overall (99.98% for the "encamped" state and 99.91% for the "exploratory" state), indicating the Euler discretization scheme is a good approximation for this dataset. Pseudo-residual plots indicate the model captures the overall patterns in the data fairly well, but some of the shorter and longer displacements were not well characterized by the model (see Appendix S1: Section S4, Figure S11). Additional model structure (or states) would likely be required to improve the fit, but the need for this will generally be case-dependent based on the objectives of the study (e.g., Pohle et al., 2017).

3.3 Illustration: Steller sea lions

As detailed in Appendix S1: Section S5, AIC values for the 1–3 state models and the more heavily constrained 4- and 5-state models (Eqs 14 and 15) overwhelmingly supported the 5-state model (Appendix S1: Section S5, Table S1). In contrast to the best-supported single-state model (Table

This article is protected by copyright. All rights reserved.

Accepted Article

19399170, ja, Downloaded from https://esajournals.onlinelibrary.wiley.com/doi/10.1002/ecy.4186 by University Of Washington, Wiley Online Library on [10/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

1), the 4- and 5-state models were able to identify movement properties (and corresponding utilization distributions) characteristic of a marine central-place forager, with movement behaviors that could be reasonably characterized as "outbound", "foraging", "inbound", and "haulout". State-decoded activity budgets indicate these individuals devoted the majority of their time (>65%) to "foraging" behavior (Appendix S1: Section S5, Table S3). While distance to site tended to dominate the "outbound" and "inbound" states, the "foraging" state of the 4-state model suggests selection of shallower depths ($\beta_{\text{depth},2} = 1.47$, 95% CI: $1.02, 1.92$) and steeper slopes ($\beta_{\text{slope},2} = 3.77$, 95% CI: $3.08, 4.46$), but not too shallow or sloped ($\beta_{\text{depth:slope},2} = -3.69$, 95% CI: $-4.36, -3.02$), and a relatively weak attraction to the sites ($\beta_{\text{d2site},2} = -2.30$, 95% CI: $-3.35, -1.25$; Table 2). The resulting "foraging" utilization distribution reflects these relationships between depth, slope, and distance from site, where the shallower and sloped areas within ≈250 km of the sites may be serving as a proxy for aggregations of available prey species (Fig. 5). Patterns for the "outbound", "inbound", and "foraging2" states of the 5-state model were similar to those of the 4-state model, but the "foraging1" state exhibited much slower speeds and, with a relatively strong attraction to the sites ($\beta_{\text{d2site},2} = -38.47$, 95% CI: $-49.81, -27.13$; Table 3), was almost entirely located near shore. This additional "foraging" state could be describing benthic near-shore foraging, but it could also simply be capturing periods in the water near sites in response to environmental conditions (e.g., tides, weather), perceived threats, and social dynamics. The multistate models were much more useful for elucidating state-specific relationships with slope, supporting the idea that only presumed foraging behaviors occurred in areas with steeper slopes. Seemingly, the 5-state model distinguished slower movements associated with nearshore gradual slopes ("foraging1") that characterize areas around preferred SSL sites (Ban, 2005) from faster movements associated with offshore steeper slopes ("foraging2") characteristic of the continental shelf, submarine canyons, and seamounts. These features coupled with oceanographic conditions are believed to enhance prey

This article is protected by copyright. All rights reserved.

Accepted Article

19399170, ja, Downloaded from https://esajournals.onlinelibrary.wiley.com/doi/10.1002/ecy.4186 by University Of Washington, Wiley Online Library on [10/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

concentrations and appear to be targeted by SSLs (Merrick and Loughlin, 1997; Lander et al., 2020).

MALA simulations for the 1- and 5-state models yielded overall acceptance rates that were $\geq 91\%$ (Appendix S1: Section S5, Table S4), but the 4-state model had overall acceptance rates ranging from 80% (track 2) to 90% (track 3). MALA acceptance rates during intervals assigned to the "foraging" states were generally highest, while those assigned to the "outbound" and "inbound" states were generally lowest (likely attributable to the faster speeds, steeper potential surfaces, and larger Δ_t observed for these states). Estimated sojourn times for the multistate models are more difficult to assess because transition rates were modeled as a function of distance to site, but, for the 4- and 5-state models excluding the "d2site" covariate, the minimum sojourn times $\left(\frac{1}{\max_i(q_{i,i})}\right)$ were estimated as 2.2 and 2.4 hrs, respectively. Approximately 20% of the observed Δ_t were greater than these minimum expected sojourn times, indicating there were likely some time steps that were too long for the snapshot property to be satisfied. When examining the estimated sojourn times as a function of distance to site, most of these potential violations of the snapshot property occurred in close proximity (< 5 km) of the sites and were absorbed by the "foraging" states (Appendix S1: Section S5, Figure S14). This suggests finer temporal resolutions may be required to accurately characterize Steller sea lion movement behavior in the vicinity of rookeries, where behaviors appear to frequently change in response to changing environmental conditions and complex social dynamics.

Similar to Michelot et al. (2019*b*), we found the single-state Langevin model to be a poor fit for these data. Pseudo-residual plots show that the 4- and 5-state models are much better fits, with the 5-state model best explaining slower movements and most of the residual autocorrelation (Appendix S1: Section S5, Figure S15). However, none of the models were able to capture some of the fastest movements, which becomes more apparent when comparing the observed and predicted

This article is protected by copyright. All rights reserved.

Accepted Article

19399170, ja, Downloaded from https://esajournals.onlinelibrary.wiley.com/doi/10.1002/ecy.4186 by University Of Washington, Wiley Online Library on [10/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

steps under each model (Appendix S1: Section S5, Figure S16). The failure of the models to capture these features could be attributable to many factors, including additional unmodeled individual or temporal heterogeneity, missing habitat covariates, or location measurement error. When simulating from the fitted models, the single-state model does not produce foraging trips characteristic of central-place foragers, whereas the 4- and 5-state models yield multiple foraging trips both nearshore and off-shelf (Fig. 6). The simulated tracks also illustrate features that are missing from the models, including short-term directional persistence (i.e., the observed tracks appear smoother), fidelity to particular sites, and prohibition of movements across land. Despite these shortcomings and room for improvement, the multistate models nevertheless provide much better fits to the data, additional biological inferences, and more realistic simulations of central-place foraging relative to the single-state model.

4 Discussion

We have developed a habitat-driven Langevin diffusion for animals that switch among distinct movement behavior states and their corresponding utilization distributions. It combines desirable properties of the Langevin diffusion (e.g., continuous time, closed-form steady state distribution) with the inferential machinery of hidden Markov models. Unlike discrete-time HMMs or step selection functions, the multistate Langevin diffusion does not require that location data be collected (or interpolated) at regular time intervals. Because it is formulated in continuous space, the resulting inferences are also not sensitive to a particular space discretization (unlike the discrete-space models of Pedersen et al., 2011; Wilson et al., 2018). Unlike two-stage approaches, the multistate Langevin model accounts for state uncertainty and helps avoid confounding of the movement and habitat selection processes. An additional advantage is that, under the Langevin model, the mi-

This article is protected by copyright. All rights reserved.

Accepted Article

19399170, ja, Downloaded from https://esajournals.onlinelibrary.wiley.com/doi/10.1002/ecy.4186 by University Of Washington, Wiley Online Library on [10/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

croscopic rules of individual animal movement scale up in time and space to yield a closed form for the utilization distribution that is a simple parametric function of the spatial covariates (unlike other approaches that require simulation; e.g. Potts et al., 2014; Signer et al., 2017). To our knowledge, this is the first single-stage approach for estimating behavior-specific habitat selection and utilization distributions from tracking data that can be readily implemented with user-friendly software. The multistate Langevin model can be combined with auxiliary (biotelemetry) data streams to help distinguish movement behavior states with similar trajectories. In the absence of auxiliary data, the model can still disentangle such states as long as the underlying behavior-specific utilization distributions are sufficiently distinct (unlike standard "step and turn" movement HMMs; e.g. Langrock et al., 2012; McClintock et al., 2013). By facilitating the estimation of state-specific utilization distributions, such inferences can inform behavior- and place-based management decisions critical to conservation (e.g., Lennox et al., 2019; Glennie et al., 2022) and help improve our understanding of habitat use at both the individual (e.g., fitness) and population (e.g., density and dynamics) level (e.g., Beyer et al., 2010; Morales et al., 2010).

Compared with existing approaches for quantifying behavior-specific habitat selection, the multistate Langevin model has statistical advantages, theoretical appeal, and benefits from all of the inferential tools of other animal movement HMMs, such as the derivation of activity budgets, calculation of stationary state probabilities, and straightforward testing of hypotheses related to drivers of behavioral dynamics. However, it can be more limited in application owing to both its additional complexity and its simplifying assumptions. As we demonstrated, the model is not well suited for location data that are infrequently observed relative to the spatial scale of the movement behaviors of interest (owing to the Euler approximation) or for behaviors that are short-lived relative to the temporal resolution of the data (owing to violations of the snapshot property). For example, we found the temporal resolution of the SSL data to be inadequate for characterizing

This article is protected by copyright. All rights reserved.

Accepted Article

19399170, ja, Downloaded from https://esajournals.onlinelibrary.wiley.com/doi/10.1002/ecy.4186 by University Of Washington, Wiley Online Library on [10/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

movement behavior within close proximity of rookeries and haul-out sites, likely a result of rapid short-term responses to environmental conditions and complex social dynamics around the sites. Our model can also be more difficult and slower to fit, where, for more complex models, we found optimization of the likelihood to be more unstable and sensitive to initial parameter values.

The habitat-driven Langevin diffusion is still relatively new, and there remains more to learn about its properties, including performance under a broader range of simulated and applied scenarios, the scale of local movement relative to the scale of the spatial covariates, and the effects of smoothing the spatial scale of habitat covariates (e.g., as a proxy for local decision making based on perception or memory). The Langevin model also requires that habitat covariates have continuous first-order partial derivatives, which is not the case for categorical covariates (e.g., habitat type), although such features can often be incorporated using distance-based metrics (e.g., distance to water, distance to forest). These issues are also relevant to our multistate generalization, but with additional considerations for behavior and assumptions such as the snapshot property.

We have skirted the snapshot property by assuming Δ_t is reasonably small relative to the serial correlation in the state process. While we expect this to be less of a concern as tracking technology continues to improve, it will not always be reasonable or obtainable in practice. Obtaining reasonably small Δ_t could be accomplished through study design by programming tracking devices to record locations at a high temporal resolution relative to the behaviors of interest. For lower resolution observations, interpolation methods could potentially be used to augment the data with additional observations at a higher temporal resolution. The latter is routinely done in marine studies (e.g., DeRuiter et al., 2017; Wilson et al., 2018) to account for location measurement error and temporally-irregular observations. Alternatively, the model could be fitted in a Bayesian framework using Monte Carlo methods to integrate over the true (and/or missing) locations as well as the (unknown) times of state switches (e.g., Jonsen et al., 2005; Blackwell, 2018; Michelot et al.,

This article is protected by copyright. All rights reserved.

Accepted Article

19399170, ja, Downloaded from https://esajournals.onlinelibrary.wiley.com/doi/10.1002/ecy.4186 by University Of Washington, Wiley Online Library on [10/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

2019*a*) or, in a Frequentist framework, possibly using a two-stage multiple imputation approach based on interpolation at a high temporal resolution (e.g., McClintock, 2017). However, the multistate Langevin model is more complicated than those typically fitted when using interpolation approaches, and two-stage methods based on basic continuous-time correlated random walks may not be sufficient in this case. Bayesian or two-stage Frequentist approaches could also be employed for handling missing covariate values, or, if time-varying covariates are measured at a finer resolution than the observed locations, the approximation of Mews et al. (2022) could be used. These remain important topics warranting further investigation.

Practitioners wishing to customize, fit, assess, and simulate the multistate Langevin model can now do so using R package `momentuHMM` (version 2.0.0 and higher; McClintock and Michelot, 2018). Users already familiar with `momentuHMM` should find using the new functions associated with continuous-time HMMs (e.g., `fitCTHMM`, `simCTHMM`) relatively straightforward, but additional details can be found in the package help files and vignette. All utility functions that come with the package, such as those for AIC model selection, calculating pseudo-residuals, local and global state decoding, and visualizing results have been extended to continuous-time models. An option for fitting models using Template Model Builder (TMB; Kristensen et al., 2016; Michelot, 2023) has also been added to `momentuHMM`, where its use of automatic differentiation to calculate gradients can often result in faster and more stable optimization (relative to alternative gradient-based routines that rely on numerical differentiation).

Acknowledgments

Alaska Ecosystems Program at the Marine Mammal Laboratory, notably T. Gelatt and B. Fadely. Alaska Department of Fish and Game for assistance with SSL data collection. SSL data were

This article is protected by copyright. All rights reserved.

Accepted Article

19399170, ja, Downloaded from https://esajournals.onlinelibrary.wiley.com/doi/10.1002/ecy.4186 by University Of Washington, Wiley Online Library on [10/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

collected under MMPA/ESA Permit 14326-02 and IACUC Protocol No. A/NW-2010-4. Théo Michelot for helpful discussions and Michelot et al. (2020) for making the zebra data publicly available. Views expressed are those of the authors and do not necessarily represent findings or policy of any government agency. Use of trade or brand names does not indicate endorsement by the U.S. government.

Conflict of interest statement

The authors declare no conflicts of interest.

Literature Cited

Abrahms, B., N. Jordan, K. Golabek, J. McNutt, A. Wilson, and J. Brashares. 2016. Lessons from integrating behaviour and resource selection: activity-specific responses of African wild dogs to roads. Animal Conservation, **19**:247–255.

Ban, S. S. 2005. Modelling and characterization of Steller sea lion haulouts and rookeries using oceanographic and shoreline type data. Master's thesis, University of British Columbia, Vancouver, Canada.

Beyer, H. L., D. T. Haydon, J. M. Morales, J. L. Frair, M. Hebblewhite, M. Mitchell, and J. Matthiopoulos. 2010. The interpretation of habitat preference metrics under use–availability designs. Philosophical Transactions of the Royal Society B: Biological Sciences, **365**:2245–2254.

Blackwell, P. G. 2018. Integrated continuous-time hidden Markov models. arXiv preprint arXiv:1807.11907.

This article is protected by copyright. All rights reserved.

Accepted Article

19399170, ja, Downloaded from https://esajournals.onlinelibrary.wiley.com/doi/10.1002/ecy.4186 by University Of Washington, Wiley Online Library on [10/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

628 Blackwell, P. G., M. Niu, M. S. Lambert, and S. D. LaPoint. 2016. Exact Bayesian inference for
629 animal movement in continuous time. Methods in Ecology and Evolution, **7**:184–195.

630 Bouyer, Y., G. San Martin, P. Poncin, R. C. Beudels-Jamar, J. Odden, and J. D. Linnell. 2015.
631 Eurasian lynx habitat selection in human-modified landscape in Norway: effects of different
632 human habitat modifications and behavioral states. Biological Conservation, **191**:291–299.

633 Camaclang, A. E., M. Maron, T. G. Martin, and H. P. Possingham. 2015. Current practices in the
634 identification of critical habitat for threatened species. Conservation Biology, **29**:482–492.

635 Clontz, L. M., K. M. Pepin, K. C. VerCauteren, and J. C. Beasley. 2021. Behavioral state resource
636 selection in invasive wild pigs in the southeastern United States. Scientific Reports, **11**:1–14.

637 DeRuiter, S. L., R. Langrock, T. Skirbutas, J. A. Goldbogen, J. Calambokidis, A. S. Friedlaender,
638 and B. L. Southall. 2017. A multivariate mixed hidden Markov model for blue whale behaviour
639 and responses to sound exposure. The Annals of Applied Statistics, **11**:362 – 392.

640 Dougherty, E. R., D. P. Seidel, J. K. Blackburn, W. C. Turner, and W. M. Getz. 2022. A framework
641 for integrating inferred movement behavior into disease risk models. Movement Ecology, **10**:31.

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

644 Gardner, B., B. T. McClintock, S. J. Converse, and N. J. Hostetter. 2022. Integrated animal move-
645 ment and spatial capture-recapture models: simulation, implementation, and inference. Ecology,
646 **103**:e3771.

647 Glennie, R., T. Adam, V. Leos-Barajas, T. Michelot, T. Photopoulou, and B. T. McClintock. 2022.

This article is protected by copyright. All rights reserved.

Accepted Article

19399170, ja, Downloaded from https://esajournals.onlinelibrary.wiley.com/doi/10.1002/ecy.4186 by University Of Washington, Wiley Online Library on [10/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Hidden Markov models: Pitfalls and opportunities in ecology. Methods in Ecology and Evolution.

Gloaguen, P. and T. Michelot. 2020. Rhabit: Estimation of animal habitat selection using the Langevin movement model. URL https://github.com/papayoun/Rhabit. R package version 0.2.0.

Hijmans, R. J. 2023. raster: Geographic Data Analysis and Modeling. URL https://CRAN.R-project.org/package=raster. R package version 3.6-20.

Hooten, M. B., D. S. Johnson, B. T. McClintock, and J. M. Morales. 2017. Animal Movement: Statistical Models for Telemetry Data. CRC press.

Horne, J. S., E. O. Garton, S. M. Krone, and J. S. Lewis. 2007. Analyzing animal movements using Brownian bridges. Ecology, **88**:2354–2363.

Johnson, C. J., K. L. Parker, D. C. Heard, and M. P. Gillingham. 2002. Movement parameters of ungulates and scale-specific responses to the environment. Journal of Animal Ecology, **71**:225–235.

Johnson, D. S., M. B. Hooten, and C. E. Kuhn. 2013. Estimating animal resource selection from telemetry data using point process models. Journal of Animal Ecology, **82**:1155–1164.

Jonsen, I. D., J. M. Flemming, and R. A. Myers. 2005. Robust state–space modeling of animal movement data. Ecology, **86**:2874–2880.

Karelus, D. L., J. W. McCown, B. K. Scheick, M. van de Kerk, B. M. Bolker, and M. K. Oli. 2017. Effects of environmental factors and landscape features on movement patterns of florida black bears. Journal of Mammalogy, **98**:1463–1478.

This article is protected by copyright. All rights reserved.

Accepted Article

19399170, ja, Downloaded from https://esajournals.onlinelibrary.wiley.com/doi/10.1002/ecy.4186 by University Of Washington, Wiley Online Library on [10/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

669 Kirkland, E. J. 2010. Advanced Computing in Electron Microscopy. Springer, Boston.

670 Klappstein, N. J., L. Thomas, and T. Michelot. 2023. Flexible hidden Markov models for
671 behaviour-dependent habitat selection. Movement Ecology, **11**:30.

672 Kloeden, P. E. and E. Platen. 1992. Numerical Solution of Stochastic Differential Equations.
673 Springer.

674 Kranstauber, B., R. Kays, S. D. LaPoint, M. Wikelski, and K. Safi. 2012. A dynamic Brownian
675 bridge movement model to estimate utilization distributions for heterogeneous animal move-
676 ment. Journal of Animal Ecology, **81**:738–746.

677 Kristensen, K., A. Nielsen, C. W. Berg, H. Skaug, and B. M. Bell. 2016. TMB: Automatic differ-
678 entiation and Laplace approximation. Journal of Statistical Software, **70**:1–21.

679 Lander, M. E., B. S. Fadely, T. S. Gelatt, J. T. Sterling, D. S. Johnson, and N. A. Pelland. 2020.
680 Mixing it up in Alaska: Habitat use of adult female Steller sea lions reveals a variety of foraging
681 strategies. Ecosphere, **11**:e03021.

682 Langrock, R., R. King, J. Matthiopoulos, L. Thomas, D. Fortin, and J. M. Morales. 2012. Flex-
683 ible and practical modeling of animal telemetry data: hidden Markov models and extensions.
684 Ecology, **93**:2336–2342.

685 Lennox, R. J., C. Engler-Palma, K. Kowarski, A. Filous, R. Whitlock, S. J. Cooke, and M. Auger-
686 Méthé. 2019. Optimizing marine spatial plans with animal tracking data. Canadian Journal of
687 Fisheries and Aquatic Sciences, **76**:497–509.

688 Manly, B., L. McDonald, D. Thomas, T. McDonald, and W. Erickson. 2010. Resource Selection
689 by Animals, 2nd Edition. Kluwer Academic Publishers, Dordrecht, The Netherlands.

This article is protected by copyright. All rights reserved.

Accepted Article

19399170, ja, Downloaded from https://esajournals.onlinelibrary.wiley.com/doi/10.1002/ecy.4186 by University Of Washington, Wiley Online Library on [10/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

690 Matthiopoulos, J., J. R. Fieberg, and G. Aarts. 2020. Species-Habitat Associations: Spatial Data,
691 Predictive Models, and Ecological Insights. University of Minnesota Libraries Publishing. URL
692 https://hdl.handle.net/11299/217469.

693 McClintock, B. T. 2017. Incorporating telemetry error into hidden Markov models of animal
694 movement using multiple imputation. Journal of Agricultural, Biological and Environmental
695 Statistics, **22**:249–269.

696 McClintock, B. T. 2023. Supporting data and R code for 'A multistate Langevin diffusion
697 for inferring behavior-specific habitat selection and utilization distributions'. Zenodo. URL
698 https://doi.org/10.5281/zenodo.8327405.

699 McClintock, B. T., B. Abrahms, R. B. Chandler, P. B. Conn, S. J. Converse, R. Emmet, B. Gardner,
700 N. J. Hostetter, and D. S. Johnson. 2022. An integrated path for spatial capture-recapture and
701 animal movement modeling. Ecology, **103**:e03473.

702 McClintock, B. T., R. King, L. Thomas, J. Matthiopoulos, B. J. McConnell, and J. M. Morales.
703 2012. A general discrete-time modeling framework for animal movement using multistate ran-
704 dom walks. Ecological Monographs, **82**:335–349.

705 McClintock, B. T., R. Langrock, O. Gimenez, E. Cam, D. L. Borchers, R. Glennie, and T. A.
706 Patterson. 2020. Uncovering ecological state dynamics with hidden Markov models. Ecology
707 Letters, **23**:1878–1903.

708 McClintock, B. T. and T. Michelot. 2018. momentuHMM: R package for generalized hidden
709 Markov models of animal movement. Methods in Ecology and Evolution, **9**:1518–1530.

710 McClintock, B. T., D. J. Russell, J. Matthiopoulos, and R. King. 2013. Combining individual an-

This article is protected by copyright. All rights reserved.

Accepted Article

19399170, ja, Downloaded from https://esajournals.onlinelibrary.wiley.com/doi/10.1002/ecy.4186 by University Of Washington, Wiley Online Library on [10/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

imal movement and ancillary biotelemetry data to investigate population-level activity budgets. Ecology, **94**:838–849.

Merrick, R. L. and T. R. Loughlin. 1997. Foraging behavior of adult female and young-of-the-year Steller sea lions in Alaskan waters. Canadian Journal of Zoology, **75**:776–786.

Mews, S., R. Langrock, R. King, and N. Quick. 2022. Multistate capture–recapture models for irregularly sampled data. The Annals of Applied Statistics, **16**:982 – 998.

Michelot, T. 2023. hmmTMB: Hidden Markov models with flexible covariate effects in R. arXiv preprint arXiv:2211.14139.

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

Michelot, T., P. G. Blackwell, and J. Matthiopoulos. 2019*a*. Linking resource selection and step selection models for habitat preferences in animals. Ecology, **100**:e02452.

Michelot, T., P. Gloaguen, P. G. Blackwell, and M.-P. Étienne. 2019*b*. The Langevin diffusion as a continuous-time model of animal movement and habitat selection. Methods in Ecology and Evolution, **10**:1894–1907.

Morales, J. M., D. T. Haydon, J. Frair, K. E. Holsinger, and J. M. Fryxell. 2004. Extracting more out of relocation data: building movement models as mixtures of random walks. Ecology, **85**:2436–2445.

Morales, J. M., P. R. Moorcroft, J. Matthiopoulos, J. L. Frair, J. G. Kie, R. A. Powell, E. H. Merrill, and D. T. Haydon. 2010. Building the bridge between animal movement and population

This article is protected by copyright. All rights reserved.

Accepted Article

19399170, ja, Downloaded from https://esajournals.onlinelibrary.wiley.com/doi/10.1002/ecy.4186 by University Of Washington, Wiley Online Library on [10/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

731 dynamics. Philosophical Transactions of the Royal Society B: Biological Sciences, **365**:2289–
732 2301.

733 Nathan, R., W. M. Getz, E. Revilla, M. Holyoak, R. Kadmon, D. Saltz, and P. E. Smouse. 2008. A
734 movement ecology paradigm for unifying organismal movement research. Proceedings of the
735 National Academy of Sciences, **105**:19052–19059.

736 Nicosia, A., T. Duchesne, L.-P. Rivest, and D. Fortin. 2017. A multi-state conditional logistic re-
737 gression model for the analysis of animal movement. The Annals of Applied Statistics, **11**:1537–
738 1560.

739 Northrup, J. M., E. Vander Wal, M. Bonar, J. Fieberg, M. P. Laforge, M. Leclerc, C. M.
740 Prokopenko, and B. D. Gerber. 2022. Conceptual and methodological advances in habitat-
741 selection modeling: guidelines for ecology and evolution. Ecological Applications, **32**:e02470.

742 Owen-Smith, N., J. Fryxell, and E. Merrill. 2010. Foraging theory upscaled: the behavioural
743 ecology of herbivore movement. Philosophical Transactions of the Royal Society B: Biological
744 Sciences, **365**:2267–2278.

745 Pedersen, M. W., T. A. Patterson, U. H. Thygesen, and H. Madsen. 2011. Estimating animal
746 behavior and residency from movement data. Oikos, **120**:1281–1290.

747 Phillips, J. S., T. A. Patterson, B. Leroy, G. M. Pilling, and S. J. Nicol. 2015. Objective classifi-
748 cation of latent behavioral states in bio-logging data using multivariate-normal hidden markov
749 models. Ecological Applications, **25**:1244–1258.

750 Picardi, S., P. Coates, J. Kolar, S. O'Neil, S. Mathews, and D. Dahlgren. 2022. Behavioural
751 state-dependent habitat selection and implications for animal translocations. Journal of Applied
752 Ecology, **59**:624–635.

This article is protected by copyright. All rights reserved.

Accepted Article

19399170, ja, Downloaded from https://esajournals.onlinelibrary.wiley.com/doi/10.1002/ecy.4186 by University Of Washington, Wiley Online Library on [10/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

753 Pohle, J., R. Langrock, F. M. van Beest, and N. M. Schmidt. 2017. Selecting the number of states
754 in hidden Markov models: pragmatic solutions illustrated using animal movement. Journal of
755 Agricultural, Biological and Environmental Statistics, **22**:270–293.

756 Pohle, J., J. Signer, J. A. Eccard, M. Dammhahn, and U. E. Schlägel. 2023. How to ac-
757 count for behavioural states in step-selection analysis: a model comparison. arXiv preprint
758 arXiv:2304.12964.

759 Potts, J. R., G. Bastille-Rousseau, D. L. Murray, J. A. Schaefer, and M. A. Lewis. 2014. Predicting
760 local and non-local effects of resources on animal space use using a mechanistic step selection
761 model. Methods in Ecology and Evolution, **5**:253–262.

762 Potts, J. R. and U. E. Schlägel. 2020. Parametrizing diffusion-taxis equations from animal move-
763 ment trajectories using step selection analysis. Methods in Ecology and Evolution, **11**:1092–
764 1105.

765 Prima, M.-C., T. Duchesne, J. A. Merkle, S. Chamaillé-Jammes, and D. Fortin. 2022. Multi-mode
766 movement decisions across widely ranging behavioral processes. Plos one, **17**:e0272538.

767 Raum-Suryan, K. L., M. J. Rehberg, G. W. Pendleton, K. W. Pitcher, and T. S. Gelatt. 2004.
768 Development of dispersal, movement patterns, and haul-out use by pup and juvenile steller sea
769 lions (eumetopias jubatus) in alaska. Marine Mammal Science, **20**:823–850.

770 Rice, M., A. Apa, and L. Wiechman. 2017. The importance of seasonal resource selection when
771 managing a threatened species: targeting conservation actions within critical habitat designa-
772 tions for the gunnison sage-grouse. Wildlife Research, **44**:407–417.

773 Roever, C. L., H. Beyer, M. J. Chase, and R. J. Van Aarde. 2014. The pitfalls of ignoring behaviour
774 when quantifying habitat selection. Diversity and Distributions, **20**:322–333.

This article is protected by copyright. All rights reserved.

Accepted Article

19399170, ja, Downloaded from https://esajournals.onlinelibrary.wiley.com/doi/10.1002/ecy.4186 by University Of Washington, Wiley Online Library on [10/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

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

Squires, J. R., N. J. DeCesare, L. E. Olson, J. A. Kolbe, M. Hebblewhite, and S. A. Parks. 2013. Combining resource selection and movement behavior to predict corridors for Canada lynx at their southern range periphery. Biological Conservation, **157**:187–195.

Suraci, J. P., L. G. Frank, A. Oriol-Cotterill, S. Ekwanga, T. M. Williams, and C. C. Wilmers. 2019. Behavior-specific habitat selection by african lions may promote their persistence in a human-dominated landscape. Ecology, **100**:e02644.

Thurfjell, H., S. Ciuti, and M. S. Boyce. 2014. Applications of step-selection functions in ecology and conservation. Movement Ecology, **2**:1–12.

Wilson, K., E. Hanks, and D. Johnson. 2018. Estimating animal utilization densities using continuous-time Markov chain models. Methods in Ecology and Evolution, **9**:1232–1240.

Wilson, R. R., L. Gilbert-Norton, and E. M. Gese. 2012. Beyond use versus availability: behaviour-explicit resource selection. Wildlife Biology, **18**:424–430.

Worton, B. J. 1989. Kernel methods for estimating the utilization distribution in home-range studies. Ecology, **70**:164–168.

Zeller, K. A., K. McGarigal, P. Beier, S. A. Cushman, T. W. Vickers, and W. M. Boyce. 2014. Sensitivity of landscape resistance estimates based on point selection functions to scale and behavioral state: pumas as a case study. Landscape Ecology, **29**:541–557.

Zucchini, W., I. L. MacDonald, and R. Langrock. 2016. Hidden Markov Models for Time Series: An Introduction Using R. CRC Press.

This article is protected by copyright. All rights reserved.

Accepted Article

19399170, ja, Downloaded from https://esajournals.onlinelibrary.wiley.com/doi/10.1002/ecy.4186 by University Of Washington, Wiley Online Library on [10/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Accepted Article

Table 1: Estimated habitat selection coefficient (β) and scale (σ^2) parameters from the best-supported 1-state Langevin movement model for three Steller sea lion tracks in the Aleutian Islands of Alaska, USA. Standard errors (SE), lower 95% confidence intervals (LCI), and upper 95% confidence intervals (UCI) are also reported. Note that positive habitat coefficients for "depth", "slope", and "d2site" indicate selection of (i.e., movement toward) shallower waters, steeper slopes, and greater distances, respectively. Habitat selection coefficient confidence intervals that do not overlap zero are highlighted in bold.

Parm.	Est.	SE	LCI	UCI
β_{depth}	-0.05	0.12	-0.30	0.19
$\boldsymbol{\beta_{\text{slope}}}$	**0.25**	**0.08**	**0.10**	**0.40**
$\boldsymbol{\beta_{\text{depth:slope}}}$	**-0.28**	**0.08**	**-0.42**	**-0.13**
$\boldsymbol{\beta_{\text{d2site}}}$	**-1.14**	**0.38**	**-1.88**	**-0.40**
$\boldsymbol{\beta_{\text{d2site2}}}$	**-2.75**	**0.58**	**-3.89**	**-1.62**
σ_1	2.48	0.02	2.43	2.52
σ_2	3.09	0.03	3.03	3.15
σ_3	2.55	0.04	2.47	2.63

This article is protected by copyright. All rights reserved.

19399170, ja, Downloaded from https://esajournals.onlinelibrary.wiley.com/doi/10.1002/ecy.4186 by University Of Washington, Wiley Online Library on [10/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Accepted Article

Table 2: Estimated habitat selection coefficient (β) and scale (σ^2) parameters from the best-supported 4-state Langevin movement model for three Steller sea lion tracks in the Aleutian Islands of Alaska, USA. Standard errors (SE), lower 95% confidence intervals (LCI), and upper 95% confidence intervals (UCI) are also reported. Note that positive habitat coefficients for "depth", "slope", and "d2site" indicate selection of (i.e., movement toward) shallower waters, steeper slopes, and greater distances, respectively. Habitat selection coefficient confidence intervals that do not overlap zero are highlighted in bold.

Parm.	Est.	SE	LCI	UCI
$\beta_{\text{depth},1}$	0.41	0.81	-1.19	2.00
$\boldsymbol{\beta_{\text{slope},1}}$	**-0.44**	**0.21**	**-0.87**	**-0.02**
$\boldsymbol{\beta_{\text{depth:slope},1}}$	**0.69**	**0.35**	**0.01**	**1.38**
$\boldsymbol{\beta_{\text{d2site},1}}$	**95.55**	**2.81**	**90.04**	**101.07**
$\boldsymbol{\beta_{\text{d2site2},1}}$	**-30.04**	**2.41**	**-34.76**	**-25.32**
$\boldsymbol{\beta_{\text{depth},2}}$	**1.47**	**0.23**	**1.02**	**1.92**
$\boldsymbol{\beta_{\text{slope},2}}$	**3.77**	**0.35**	**3.08**	**4.46**
$\boldsymbol{\beta_{\text{depth:slope},2}}$	**-3.69**	**0.34**	**-4.36**	**-3.02**
$\boldsymbol{\beta_{\text{d2site},2}}$	**-2.30**	**0.53**	**-3.35**	**-1.25**
$\boldsymbol{\beta_{\text{depth},3}}$	**-4.05**	**0.74**	**-5.50**	**-2.61**
$\beta_{\text{slope},3}$	-0.30	0.16	-0.62	0.02
$\beta_{\text{depth:slope},3}$	0.17	0.25	-0.33	0.67
$\boldsymbol{\beta_{\text{d2site},3}}$	**-82.05**	**2.45**	**-86.86**	**-77.25**
$\boldsymbol{\beta_{\text{d2site2},3}}$	**14.68**	**1.86**	**11.03**	**18.33**
$\sigma_{1,AS}$	2.06	0.02	2.02	2.09
$\sigma_{2,AS}$	2.43	0.02	2.38	2.48
$\sigma_{3,AS}$	2.17	0.03	2.11	2.23
σ_{HO}	0.00	0.00	0.00	0.00

This article is protected by copyright. All rights reserved.

19399170, ja, Downloaded from https://esajournals.onlinelibrary.wiley.com/doi/10.1002/ecy.4186 by University Of Washington, Wiley Online Library on [10/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Accepted Article

Table 3: Estimated habitat selection coefficient (β) and scale (σ^2) parameters from the best-supported 5-state Langevin movement model for three Steller sea lion tracks in the Aleutian Islands of Alaska, USA. Standard errors (SE), lower 95% confidence intervals (LCI), and upper 95% confidence intervals (UCI) are also reported. Note that positive habitat coefficients for "depth", "slope", and "d2site" indicate selection of (i.e., movement toward) shallower waters, steeper slopes, and greater distances, respectively. Habitat selection coefficient confidence intervals that do not overlap zero are highlighted in bold.

Parm.	Est.	SE	LCI	UCI
$\beta_{\text{depth},1}$	0.89	0.72	-0.52	2.31
$\beta_{\text{slope},1}$	-0.13	0.19	-0.50	0.24
$\beta_{\text{depth:slope},1}$	0.32	0.30	-0.27	0.91
$\boldsymbol{\beta_{\text{d2site},1}}$	**54.52**	**2.35**	**49.92**	**59.13**
$\boldsymbol{\beta_{\text{d2site2},1}}$	**-18.14**	**1.91**	**-21.89**	**-14.40**
$\beta_{\text{depth},2}$	8.32	5.84	-3.12	19.76
$\beta_{\text{slope},2}$	11.48	12.92	-13.84	36.81
$\beta_{\text{depth:slope},2}$	-10.77	12.45	-35.18	13.64
$\boldsymbol{\beta_{\text{d2site},2}}$	**-38.47**	**5.79**	**-49.81**	**-27.13**
$\beta_{\text{depth},3}$	0.24	0.25	-0.24	0.73
$\boldsymbol{\beta_{\text{slope},3}}$	**1.65**	**0.34**	**0.99**	**2.32**
$\boldsymbol{\beta_{\text{depth:slope},3}}$	**-1.71**	**0.33**	**-2.37**	**-1.06**
$\boldsymbol{\beta_{\text{d2site},3}}$	**-3.20**	**0.72**	**-4.62**	**-1.79**
$\boldsymbol{\beta_{\text{depth},4}}$	**-2.20**	**0.65**	**-3.48**	**-0.92**
$\beta_{\text{slope},4}$	-0.22	0.13	-0.48	0.04
$\beta_{\text{depth:slope},4}$	0.20	0.20	-0.20	0.60
$\boldsymbol{\beta_{\text{d2site},4}}$	**-48.85**	**2.17**	**-53.10**	**-44.61**
$\boldsymbol{\beta_{\text{d2site2},4}}$	**9.62**	**1.72**	**6.26**	**12.99**
$\sigma_{1,AS}$	2.58	0.03	2.51	2.64
$\sigma_{2,AS}$	3.28	0.05	3.18	3.37
$\sigma_{3,AS}$	2.80	0.05	2.70	2.91
$\sigma_{1,2}$	0.34	0.01	0.32	0.36
$\sigma_{2,2}$	0.23	0.01	0.22	0.25
$\sigma_{3,2}$	0.47	0.02	0.43	0.51
σ_{HO}	0.00	0.00	0.00	0.00

This article is protected by copyright. All rights reserved.

19399170, ja, Downloaded from https://esajournals.onlinelibrary.wiley.com/doi/10.1002/ecy.4186 by University Of Washington, Wiley Online Library on [10/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Figure 1. Data and results from the 2-state plains zebra analysis, including: (a) the track relative to the four vegetation types (grassland = "grass", bushed grassland = "bushgrass", bushland = "bush", woodland = "wood"); (b) the estimated habitat selection coefficients and 95% confidence intervals for the two states (blue = "encamped", orange = "exploratory"); (c) the estimated stationary state probabilities and 95% confidence intervals as a function of time of day; and the estimated (log-scale) utilization distributions for (d) the "encamped" state (π_1), (e) the "exploratory" state (π_2), and (f) the weighted average of the state-dependent utilization distributions ($\tilde{\pi}$) based on the estimated amount of time spent in each state.

Figure 2. Three Steller sea lion tracks in the Aleutian Islands of Alaska, USA, relative to sea floor depth (km from maximum depth).

Figure 3. Boxplots of 100 estimates of the habitat selection parameters for state 1 ($\beta_{1,1}, \beta_{1,2}, \beta_{1,3}, \beta_{1,4}$; top row), habitat selection parameters for state 2 ($\beta_{2,1}, \beta_{2,2}, \beta_{2,3}, \beta_{2,4}$; middle row), the state transition rates ($\log(q_{1,2})$ and $\log(q_{2,1})$; bottom left), and the scale parameters (σ_1^2, σ_2^2; bottom right) for the 2-state Langevin model under different time intervals of observation. Dashed colored lines indicate the true values of the parameters for state 1 (red) and state 2 (blue).

Figure 4. One realization of the true (top row, first two columns) and estimated state-dependent utilization distributions for the 2-state Langevin model under different time intervals of observation, $\Delta_t \in \{0.01, 0.02, 0.05, 0.1, 0.2, 0.5, 1\}$, where "UD1" indicates the utilization distribution for state 1 and "UD2" indicates the utilization distribution for state 2.

Figure 5. Estimated (state-dependent) utilization distributions and Viterbi-decoded state assignments for the best supported 1- (left column), 4- (middle column), and 5-state (right column) Langevin movement models for three Steller sea lion tracks in the Aleutian Islands of Alaska,

This article is protected by copyright. All rights reserved.

Accepted Article

19399170, ja, Downloaded from https://esajournals.onlinelibrary.wiley.com/doi/10.1002/ecy.4186 by University Of Washington, Wiley Online Library on [10/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

USA. Utilization distributions are plotted on the log scale. Enlarged versions of each panel can be found in *Supporting Information*.

Figure 6. Simulated tracks for a) three Steller sea lions in the Aleutian Islands of Alaska, USA, generated from the parameter estimates of the best-supported b) 1-state, c) 4-state, and d) 5-state Langevin movement models. Simulated tracks are plotted on top of the estimated (state-dependent) utilization distributions ("foraging" and "foraging2" for the 4- and 5-state models, respectively), which are presented on the log scale. Enlarged versions of each panel can be found in *Supporting Information*.

This article is protected by copyright. All rights reserved.

Accepted Article

19399170, ja, Downloaded from https://esajournals.onlinelibrary.wiley.com/doi/10.1002/ecy.4186 by University Of Washington, Wiley Online Library on [10/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Accepted Article

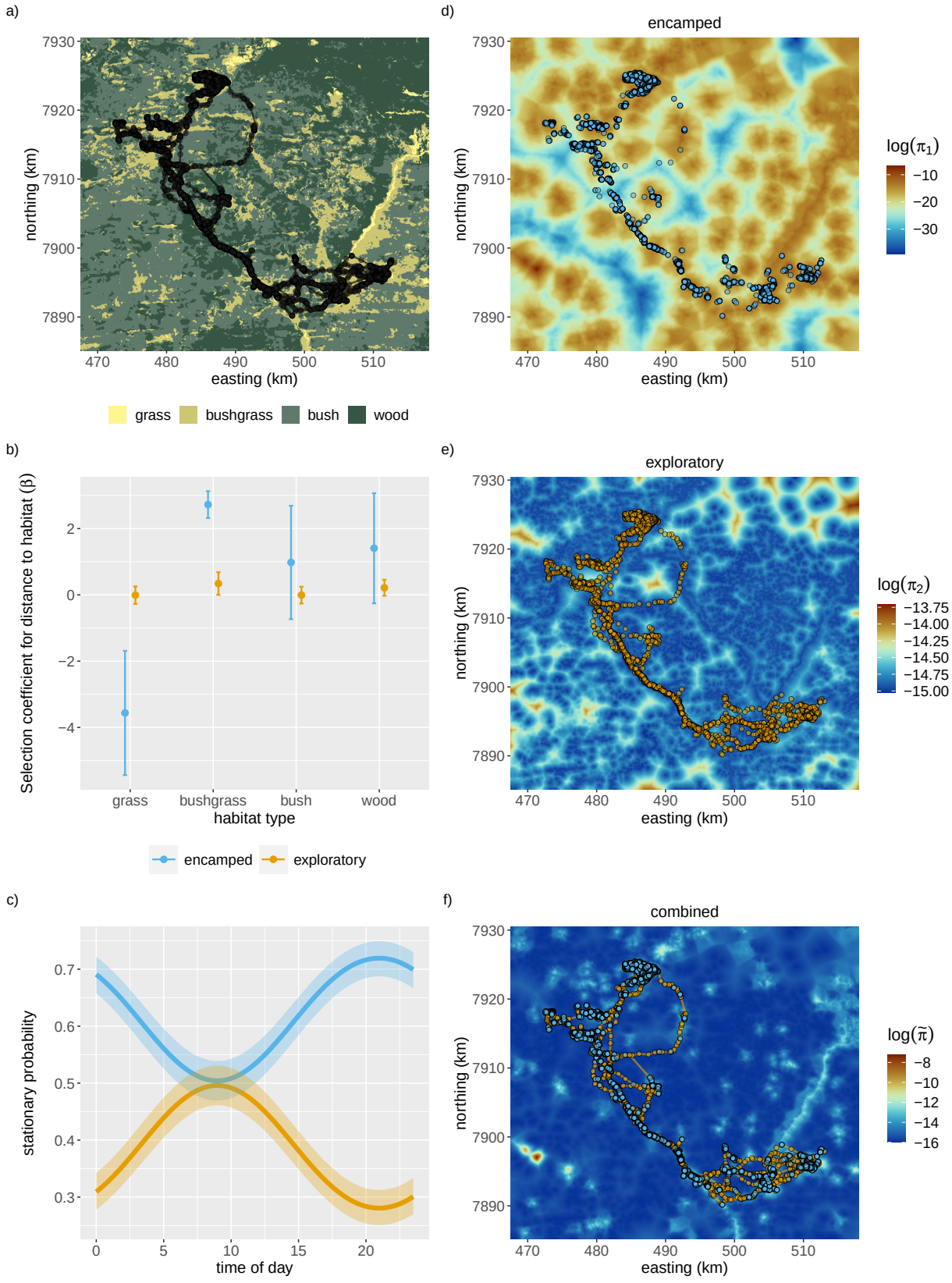

Figure 1

This article is protected by copyright. All rights reserved.

19399170, ja, Downloaded from https://esajournals.onlinelibrary.wiley.com/doi/10.1002/ecy.4186 by University Of Washington, Wiley Online Library on [10/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Accepted Article

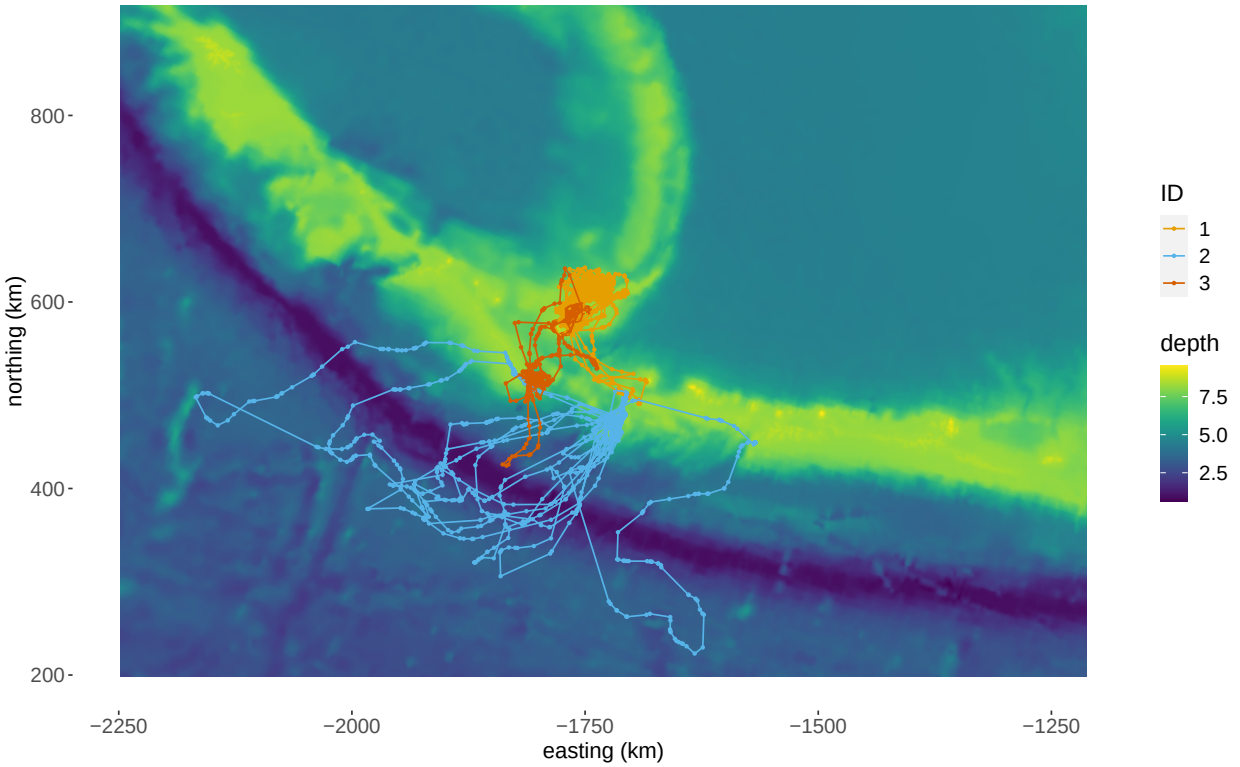

Figure 2

This article is protected by copyright. All rights reserved.

19399170, ja, Downloaded from https://esajournals.onlinelibrary.wiley.com/doi/10.1002/ecy.4186 by University Of Washington, Wiley Online Library on [10/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Accepted Article

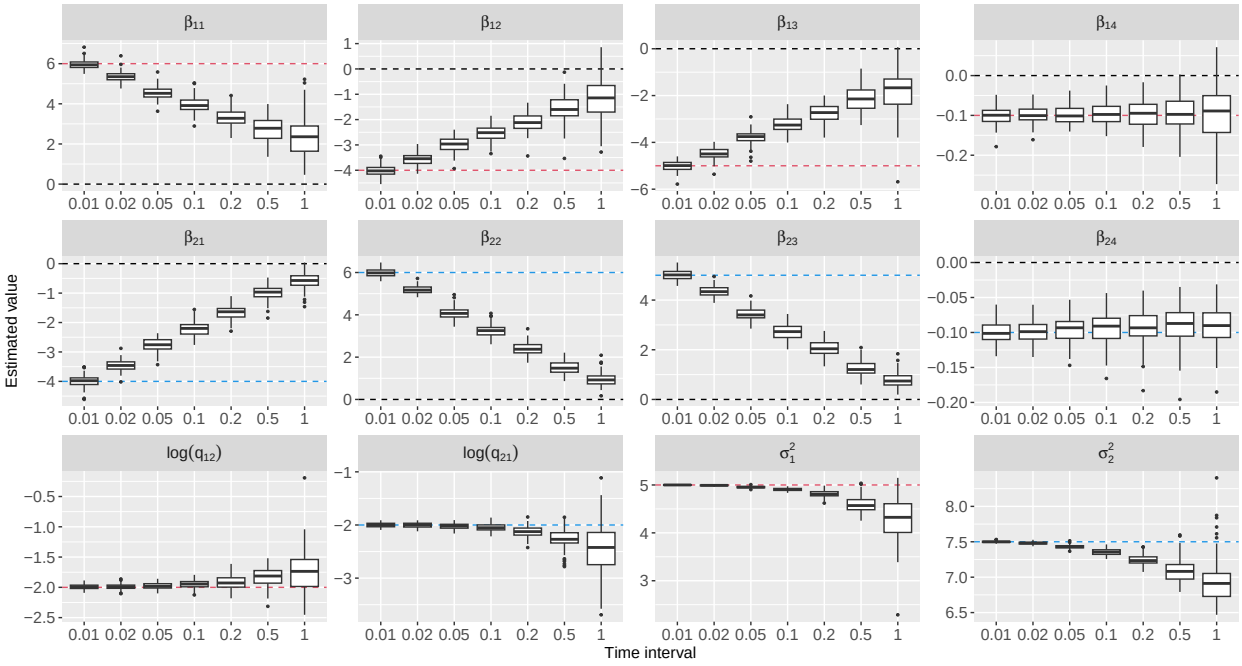

Figure 3

This article is protected by copyright. All rights reserved.

19399170, ja, Downloaded from https://esajournals.onlinelibrary.wiley.com/doi/10.1002/ecy.4186 by University Of Washington, Wiley Online Library on [10/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Accepted Article

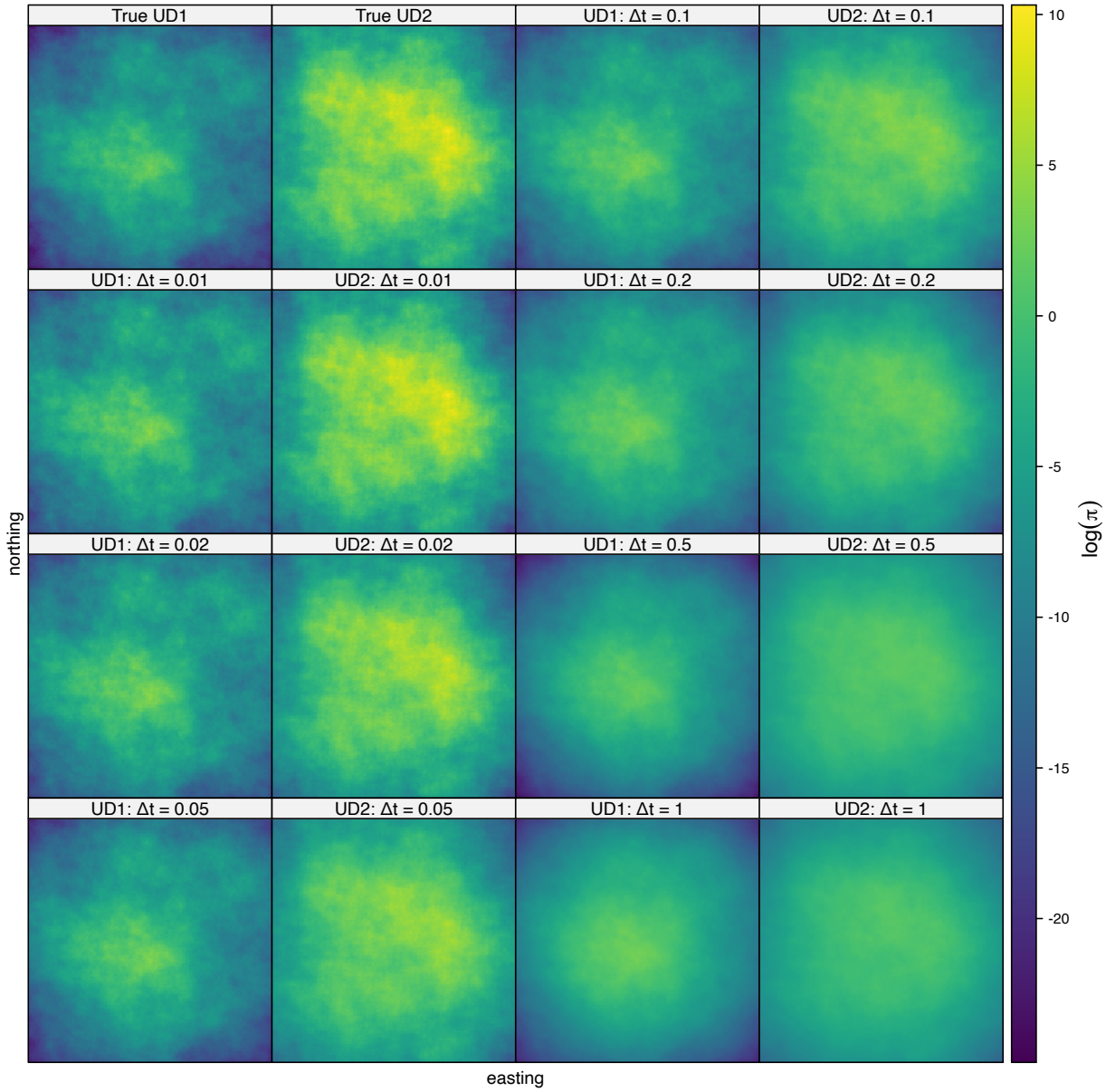

Figure 4

This article is protected by copyright. All rights reserved.

19399170, ja, Downloaded from https://esajournals.onlinelibrary.wiley.com/doi/10.1002/ecy.4186 by University Of Washington, Wiley Online Library on [10/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Accepted Article

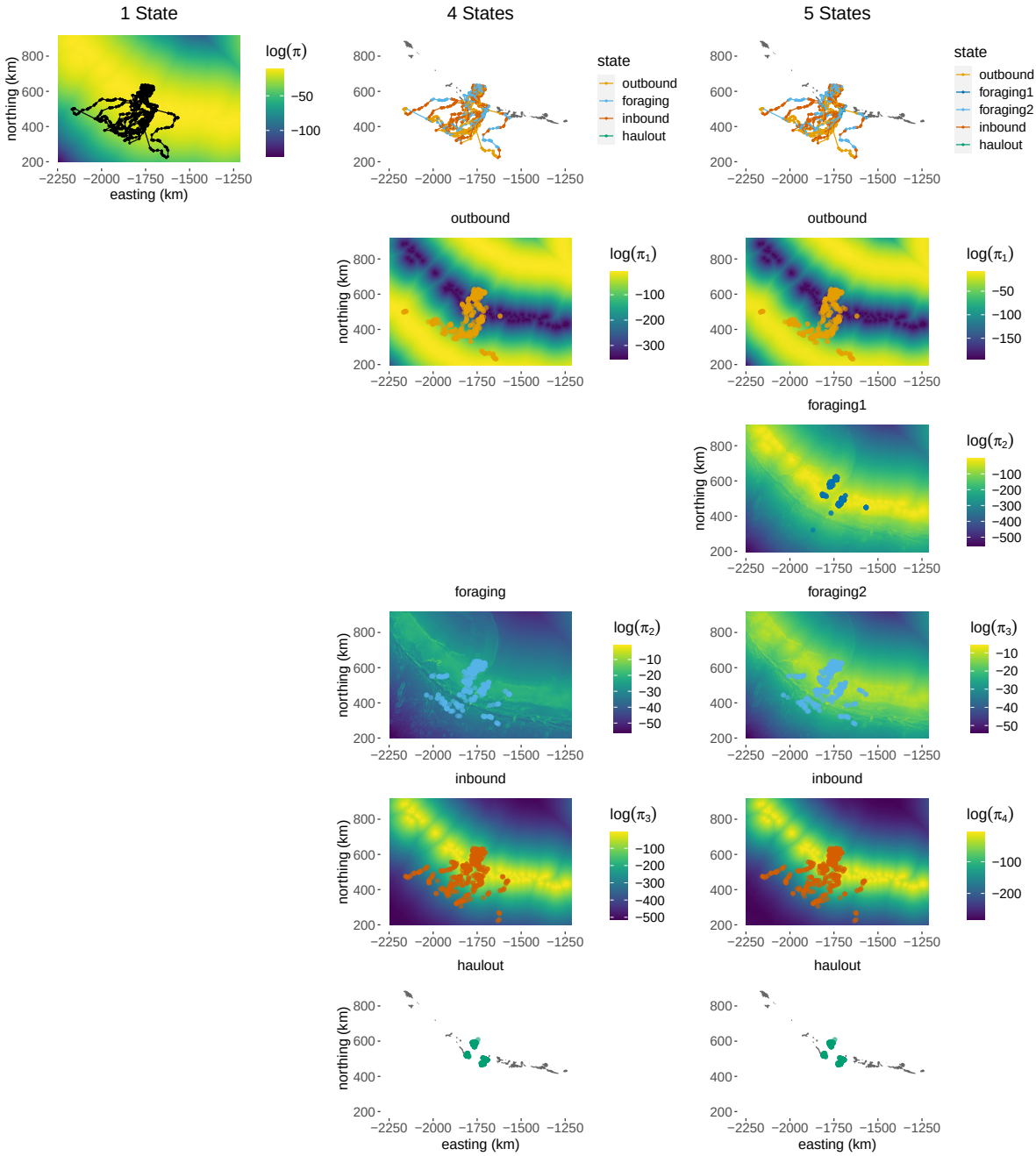

Figure 5

This article is protected by copyright. All rights reserved.

19399170, ja, Downloaded from https://esajournals.onlinelibrary.wiley.com/doi/10.1002/ecy.4186 by University Of Washington, Wiley Online Library on [10/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Accepted Article

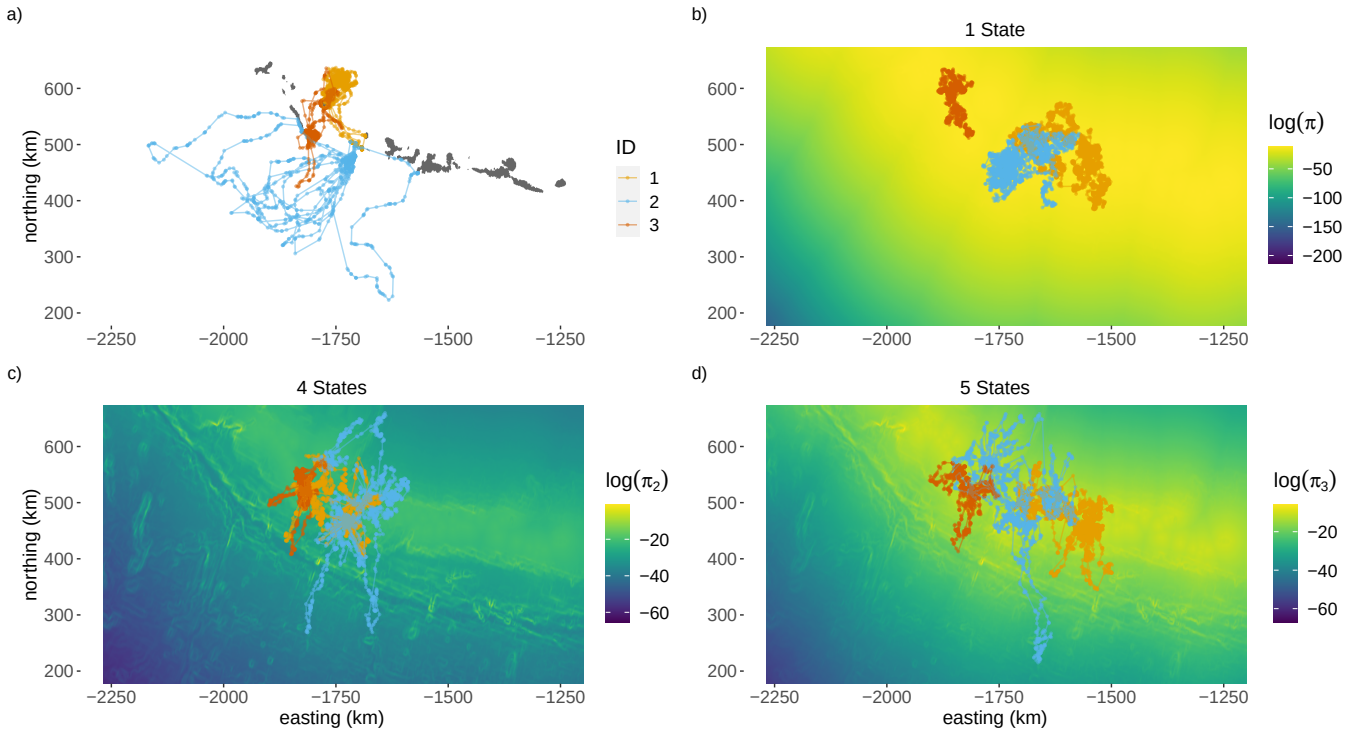

Figure 6

This article is protected by copyright. All rights reserved.

19399170, ja, Downloaded from https://esajournals.onlinelibrary.wiley.com/doi/10.1002/ecy.4186 by University Of Washington, Wiley Online Library on [10/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License