

REVIEW

Realising the Promise of Large Data and Complex Models

Hidden Markov models: Pitfalls and opportunities in ecology

Richard Glennie¹  | Timo Adam¹ | Vianey Leos-Barajas²  | Théo Michelot¹  |
Theoni Photopoulou¹  | Brett T. McClintock³ 

¹Centre for Research into Ecological and Environmental Modelling, University of St Andrews, St Andrews, UK

²Department of Statistical Sciences, University of Toronto, Toronto, ON, Canada

³Marine Mammal Laboratory, NOAA-NMFS Alaska Fisheries Science Center, Seattle, WA, USA

Correspondence

Richard Glennie

Email: rg374@st-andrews.ac.uk

Handling Editor: Ruth King

Abstract

1. Hidden Markov models (HMMs) and their extensions are attractive methods for analysing ecological data where noisy, multivariate measurements are made of a hidden, ecological process, and where this hidden process is represented by a sequence of discrete states. Yet, as these models become more complex and challenging to understand, it is important to consider what pitfalls these methods have and what opportunities there are for future research to address these pitfalls.
2. In this paper, we review five lesser known pitfalls one can encounter when using HMMs or their extensions to solve ecological problems: (a) violation of the snapshot property in continuous-time HMMs; (b) biased inference from hierarchical HMMs when applied to temporally misaligned processes; (c) sensitive inference from using random effects to partially pool across heterogeneous individuals; (d) computational burden when using HMMs to approximate models with continuous state spaces; and (e) difficulty linking the hidden process to space or environment.
3. This review is for ecologists and ecological statisticians familiar with HMMs, but who may be less aware of the problems that arise in more specialised applications. We demonstrate how each pitfall arises, by simulation or example, and discuss why this pitfall is important to consider. Along with identifying the problems, we highlight potential research opportunities and offer ideas that may help alleviate these pitfalls.
4. Each of the methods we review are solutions to current ecological research problems. We intend for this paper to heighten awareness of the pitfalls ecologists may encounter when applying these more advanced methods, but we also hope that by highlighting future research opportunities, we can inspire ecological statisticians to weaken these pitfalls and provide improved methods.

KEYWORDS

animal movement, continuous time, hidden Markov model, hierarchical model, population ecology, random effects, state space models, time series

1 | INTRODUCTION

The statistical methods used in ecology are becoming increasingly complex. What statistical and computational pitfalls do these methods have? What future research opportunities are there for these methods to be improved? In this paper, we discuss these two questions for a popular class of statistical models: hidden Markov models (HMMs; Zucchini et al., 2017).

Hidden Markov models are widely applied in ecology, from individual-level to ecosystem-level modelling [see McClintock et al. (2020) for a comprehensive overview of the application of HMMs in ecology]. Their intuitive structure often corresponds with our conceptual models for ecological systems: there is a hidden process unfolding over time from which we obtain noisy, multivariate observations. In their commonest form, HMMs consist of two time series (S_t, Y_t) , in discrete time, where S_t is a hidden (unobserved) state that can take one of a finite number of values and Y_t is a collection of observed variables whose distribution we assume depends on the hidden state S_t , termed its state-dependent distribution. Crucially, it is assumed the observations are independent given the states. Over time, states evolve as a Markov process which is described by a transition probability matrix Γ_t whose $(i, j)^{\text{th}}$ entry is the probability that $S_{t+1} = j$ given $S_t = i$. Key properties of HMMs are that they commonly deal with observations in regular, discrete time; the state, S_t , switches between a pre-specified finite number of possible values; and there are computationally efficient algorithms to quickly fit HMMs to large amounts of data.

The successful application of HMMs in ecology has inspired further use that goes beyond the standard application of HMMs, and it is these extended uses that this paper will discuss (Readers unfamiliar with HMMs are encouraged to consult McClintock et al. (2020) or Zucchini et al. (2017) for an introduction). We consider five extended uses of HMMs: (a) HMMs in continuous time; (b) hierarchical HMMs; (c) HMMs with random effects; (d) approximating state space models with HMMs; and (e) two-stage analyses of decoded states from HMMs. Each extended use adds complexity to what is already a complex statistical analysis and so we ask the two questions above: what pitfalls are revealed and what future research opportunities are there? We hope this discussion will serve two audiences. For ecologists familiar with HMMs, it provides a statistical overview of more advanced uses of HMMs, focusing on what possible problems

a practitioner may encounter. For statistical ecologists or ecological statisticians, we hope this paper serves as a resource to inspire future research, either to weaken the pitfalls we highlight in HMMs or to introduce alternative methods to solve these problems.

2 | CONTINUOUS-TIME HIDDEN MARKOV MODELS

2.1 | Introduction

A continuous-time HMM has the same dependence structure as a discrete-time HMM, but the underlying state S_t of the system is determined by a continuous-time Markov chain. An N -state continuous-time Markov chain is defined by an $N \times N$ transition rate matrix,

$$Q = \begin{pmatrix} -q_1 & q_{12} & \cdots & q_{1N} \\ q_{21} & -q_2 & \cdots & q_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ q_{N1} & q_{N2} & \cdots & -q_N \end{pmatrix}$$

where $q_i = \sum_{j \neq i} q_{ij}$ for all i . The time spent in each state i follows an exponential distribution with mean $1/q_i$. The non-diagonal elements of the matrix are strictly positive and proportional to the transition probabilities out of state i (within each row). For any given time interval of length $\Delta > 0$, the corresponding transition probability matrix over that time interval can be computed from the transition rate matrix with a matrix exponential:

$$\Gamma_\Delta = e^{Q\Delta}. \quad (1)$$

The key property of continuous-time HMMs is that they do not require observations made at regular time intervals, and the times of observations do not need to match the times of state transitions. A schematic representation of the dependence structure of a continuous-time HMM is shown in Figure 1.

Continuous-time HMMs have been popular in medical statistics, in particular to study disease progression in patients from consultations done at irregular time intervals (e.g. Bureau et al., 2003; Jackson et al., 2003; Liu et al., 2015). However, they remain rare

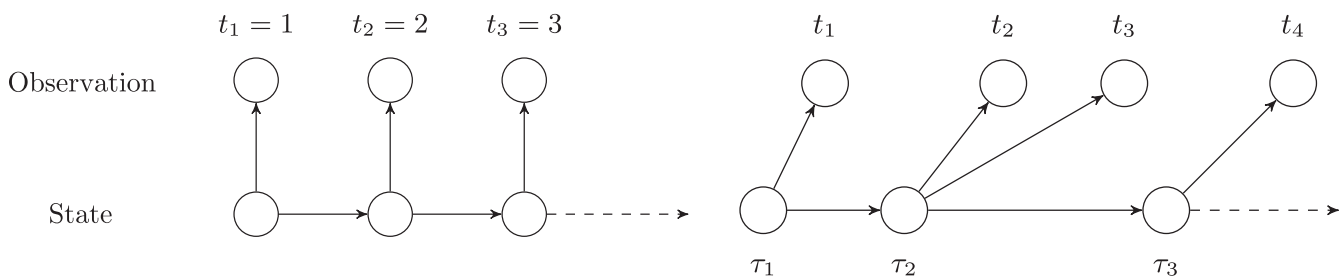


FIGURE 1 Comparison of dependence graphs for discrete-time (left) and continuous-time (right) HMMs. In both cases, the distribution of an observation only depends on the current value of the state. The key difference is that, in a continuous-time HMM, the times of state transitions (τ_1, τ_2, \dots) and the times of observation (t_1, t_2, \dots) do not need to match, and both may be irregularly spaced

in ecological studies, despite their broad applicability to irregular datasets. We posit that they have been underutilised because the mathematical theory behind continuous-time Markov processes is less intuitive than its counterpart in discrete time, and because of the apparent lack of accessible software to readily apply this method. However, much of the inferential framework developed for discrete-time HMMs can also be used in continuous time, based on the relationship between transition rates and transition probabilities in Equation (1). This includes the forward algorithm and the Viterbi algorithm (Zucchini et al., 2017). In many situations, the implementation of a continuous-time HMM is virtually identical to that of a discrete-time HMM, with the only difference being that the likelihood function is parametrised in terms of transition rates. Yet, it is once you move from discrete to continuous time that a central property of HMMs becomes less intuitive: the snapshot property.

2.2 | Pitfall: Snapshot property violation

The snapshot property is satisfied if the value of the observation process at time t only depends on the state at that time, rather than on past values of the state (Patterson et al., 2017). In discrete time, this is a well-understood assumption; however, in continuous time, state transitions could happen *at any time* (Figure 1). Continuous-time HMMs are therefore only suitable when the distribution of each observation does not depend on the intermediate sequence of switches that have occurred since the previous observation. When this is not the case, the snapshot property may still be a reasonable approximation if observations occur at a high temporal resolution relative to the scale of state switching, that is, if only a small proportion of observation intervals contain state switches. We illustrate this below using simulations. One context where the snapshot property is usually violated is state-switching models based on relatively infrequent observations of animal locations or velocities, because these variables depend on the behaviour of the animal over the whole time interval between observations rather than at the time of observation (Blackwell et al., 2016; Michelot

& Blackwell, 2019; Patterson et al., 2017). We therefore investigated the approximation error arising from using a continuous-time HMM in cases where the snapshot property is not satisfied. As a simple example of a movement model, we considered a two-state Brownian motion process with no drift, where the diffusion parameter was small in state $S_t = 1$ (slow movement) and large in $S_t = 2$ (fast movement). To facilitate interpretation of the simulations, we used parameters estimated from a real dataset of Antarctic petrels (*Thalassoica antarctica*) from the Movebank data repository (Descamps et al., 2016a, 2016b). The diffusion parameters estimated from the petrel data were $(\sigma_1\sigma_2) = (1.2, 14.9)$ (i.e. mean distance travelled over 1 h was 1.5 km in state 1 and 18.7 km in state 2), and the transition rates were $q_{12} = 0.29$, and $q_{21} = 0.33$ (i.e. expected dwelling time was 3.4 h in state 1 and 3 h in state 2). The details of the petrel analysis and the formulation of the Brownian motion are given in Section S1.1 of the Supplementary Material.

We generated data using those parameters and mimicked the real-life scenario where the behavioural switching times are unknown, to evaluate the error caused by violations of the snapshot property. We then assessed the performance of the approximation based on bias between true and estimated parameters, and on the proportion of correctly estimated states, for different time intervals of observation. The simulation procedure is described in Section S1.2 of the Supplementary Material, and the results are shown in Figure 2. As expected, the approximation error increases with the mean time interval because more switches can occur between observations (but each observation is assumed to only depend on the state active at that time).

The performance of the continuous-time HMM methodology in contexts where the snapshot property does not hold depends on the time-scale at which the hidden states occur (i.e. on the transition rates) relative to the time-scale of the observations. Our results support an intuitive rule of thumb for determining whether or not any given sampling interval may reasonably satisfy the snapshot property:

$$\Delta \leq \frac{1}{\max_i(q_i)}, \quad (2)$$

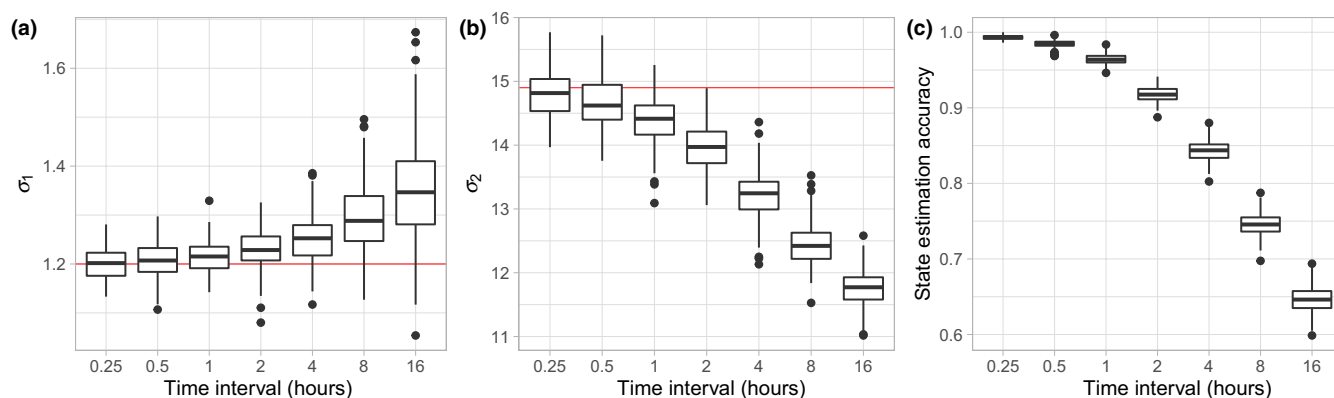


FIGURE 2 Results of continuous-time HMM simulation study. Estimated parameters σ_1 (a) and σ_2 (b), and state estimate accuracy (c) for simulated data with different mean time intervals, from 0.25 h to 16 h. Each box contains 200 replications, for each simulation scenario

that is, the mean time interval of observation should be shorter than the shortest expected dwelling time (3 h in our example). However, as with all HMMs, performance will also depend on the amount of state-dependent observation distribution overlap, serial correlation in the hidden state sequence, and other properties of the data (e.g. McClintock, 2021; Zucchini et al., 2017). These results suggest that continuous-time HMMs should not be applied naively in ecological studies where the snapshot property is known to be violated, in particular when the time intervals of observations are long relative to the time-scale of the state process.

2.3 | Opportunities

Our simulation results indicated that it is important to consider whether or not the snapshot property is reasonably satisfied when using continuous-time HMMs for ecological data. Continuous-time HMMs have recently been proposed for the analysis of capture–recapture data (Choquet et al., 2017; Mews, Langrock, King, et al., 2020), where the observation is a categorical variable which indicates whether an animal was captured or not at a given capture occasion, and the hidden state is the existential state of that animal (usually ‘alive’ or ‘dead’). The probability of capturing the animal at time t only depends on whether the animal is alive at time t , rather than on whether it was alive over the whole interval since the last capture occasion, and the system therefore has the snapshot property. Continuous-time HMMs are a natural framework for this type of data because capture occasions are often irregular in time. Similarly, HMMs developed for occupancy data could be extended to the continuous-time case, because the observation (‘detected’ or ‘not detected’) does not depend on the history of the state process, conditional on the current state (‘present’ or ‘absent’; see Royle & Kéry, 2007; MacKenzie et al., 2018).

For other types of data where the snapshot property cannot be satisfied by study design (e.g. by ensuring time intervals are sufficiently short relative to the transition rates), there are existing remedies. One option is to augment the data with missing observation times (thereby yielding sufficiently short intervals) and then integrate over the missing observations during model fitting. This integration is relatively straightforward in Bayesian analysis (e.g. Gelman et al., 2013) and maximum likelihood analyses using expectation–maximisation algorithms (e.g. McLachlan & Krishnan, 2007). When the forward algorithm is used to maximise the likelihood directly, multiple imputation methods can account for the missing observations (e.g. McClintock, 2017; Rubin, 2004). These approaches are approximate but relatively easy to implement. The challenge with exact methods is that the state transition times are unknown. An exact but relatively difficult method to implement is the integrated continuous-time HMM (Blackwell, 2018), which is a Bayesian approach that utilises the efficient forward algorithm to estimate the whole state process together with the model parameters. Further development of exact methods that efficiently account for violations of the snapshot property is a promising area of future research.

While discrete-time HMMs for ecological data have been widely applied and extended in recent decades (e.g. McClintock et al., 2020), continuous-time HMMs have received far less attention. This is unfortunate because many types of ecological data are collected in continuous time, but are then ‘shoehorned’ into a discrete-time framework for analysis (e.g. Borchers et al., 2014; McClintock et al., 2014). Although they have not yet seen wide use by ecologists, there are several R (R Core Team, 2020) packages that make continuous-time HMMs readily applicable to ecological datasets that satisfy the snapshot property. In particular, the package `msm` includes many common observation distributions and allows for covariate dependence in the model parameters (Jackson, 2011). More recently, the package `momentuHMM` has been extended to accommodate continuous-time HMMs, with an emphasis on observation distributions commonly used in animal movement behaviour models for biotelemetry data (McClintock & Michelot, 2018). Other options include `HMMCont`, which is limited to normally distributed observations (Beketov, 2014), and `JAGS` using the `msm` module (Plummer, 2017). In addition to refinements that can more efficiently account for violations of the snapshot property, there remain many opportunities to develop continuous-time analogues to recent extensions of discrete-time HMMs, including semi-Markov models (e.g. Langrock & Zucchini, 2011), hierarchical HMMs (e.g. Fine et al., 1998, see Section 3), random effects (e.g. Altman, 2007, see Section 4) and covariates that vary in continuous time (e.g. Mews, Langrock, King, et al., 2020). Such developments will help improve continuous-time HMMs for ecology and facilitate their application to a broad range of systems and taxa.

3 | HIERARCHICAL HIDDEN MARKOV MODELS

3.1 | Introduction

Hierarchical HMMs (HHMMs; Fine et al., 1998) extend basic HMMs by having multiple processes operate at different time-scales (Adam et al., 2019; Leos-Barajas et al., 2017). The key property of HHMMs is that they model multi-scale data jointly (e.g. hourly step lengths from GPS tags and accelerations recorded from accelerometers several times per second), that are driven by multiple hidden Markov chains that evolve at different time-scales (e.g. behavioural state each hour and within that behavioural substate each second). They are already successfully applied in ecology, for example, to the dive behaviour of harbour porpoises (Leos-Barajas et al., 2017; Sacchi & Swallow, 2021), to horizontal and vertical movements of Atlantic cod (Adam et al., 2019) and white sharks (Aquino-Baleytó et al., 2021), and to the kinematic movements of northern resident killer whales (Sidrow et al., 2021).

In their simplest form, HHMMs have two hierarchical hidden processes, each of which operates at a different time-scale: (a) a coarse-scale (e.g. hours) hidden process S_t for coarse-scale time t and (b) a fine-scale (e.g. seconds) hidden process S_{t^*} for fine-scale

time t^* within coarse-scale time t . Each of these hidden processes has its own transition probability matrix and each can have an associated observation process for observations that occur at either coarse-scale or fine-scale resolution. Essentially, if the coarse-scale state process has N states, the fine-scale observations are modelled by N different fine-scale HMMs. The coarse-scale state that is active determines which fine-scale HMM models the fine-scale observations. As the fine-scale observations do not only depend on the fine-scale state process but also on the coarse-scale state process, HHMMs can capture how switches between fine-scale behaviours depend on the coarse-scale behavioural mode. This added flexibility can then better capture how animals determine their behaviour in both the short and long term (Adam et al., 2019).

3.2 | Pitfalls

Not all multi-scale data follow the dependence structure assumed by HHMMs. As described in Section 3.1, a central assumption is that the coarse-scale observations and the fine-scale observations are ultimately driven by a coarse-scale state process; the dependence structure is thus determined by the resolution of the coarse-scale observations (see Figure 3 [a] for an illustration of two state processes that satisfy this assumption). Consider a scenario with coarse-scale step lengths observed once per hour and fine-scale accelerations observed once per second. HHMMs assume that given the coarse-scale state for that hour, the accelerations during that *entire* hour arise from a single fine-scale HMM; in the next hour, the coarse-scale state may change and so the accelerations may arise from a different HMM. The important insight is that the accelerations cannot arise from two different fine-scale HMMs within the same hour.

While this assumption is reasonable in many scenarios, it may be questionable in other applications. An animal can perform (moderately) fast movements during the first few minutes of an hour that is, overall, characterised by resting behaviour, or no (slow) movements

during the first few minutes of an hour that is, overall, characterised by transiting behaviour (see the red-shaded areas in Figure 3 [b] for an illustration of such a scenario). Similarly, the fine-scale observations are not necessarily driven by the *same* coarse-scale state process as the coarse-scale observations. Accelerations can be driven by other behaviours than resting and transiting, which drive step lengths and turning angles. In both examples, multi-scale data do not follow the dependence structure that is assumed by HHMMs.

To show the consequences of such a violation of the dependence structure, we conducted a simulation experiment (further details on the simulation procedure are provided in Section S2.1 of the Supplementary Material). Over 200 replications, we simulated a two-state coarse-scale process on the hourly scale with 1,000 observations of step length and turning angle, and a two-state fine-scale process with 100 observations of acceleration. We then progressively shifted the fine-scale process by 0, 5, 10, 15 and 20 observations and computed the percentage bias in parameter estimates (full descriptions of the parameters are provided in Section S2.1 of the Supplementary Material). While we used deterministic shifts, in practice, they can also be probabilistically. In that regard, the deterministic shifts used for the simulation experiment can be thought of as being exemplary of a probabilistic shifting process with mean equal to the deterministic shifts and small variance. For probabilistic shifting processes with large variance, or scenarios where the fine-scale HMMs change probabilistically within a coarse-scale state, we expect that HHMMs often fail to infer distinct behavioural modes at the fine scale, as all fine-scale behaviours occur within all coarse-scale states, that is, the two processes are less correlated. In such cases, separate HMMs for the two processes should be preferred over HHMMs. All models were fit using the R package *momentuHMM* (McClintock & Michelot, 2018).

Example results are displayed in Figure 4 (full results are displayed in Section S2.2 of the Supplementary Material). Clearly, when the observations were simulated as assumed by the HHMM (i.e. without shifting the fine-scale process), the estimates are unbiased; however, the question is how severe the bias increases as the

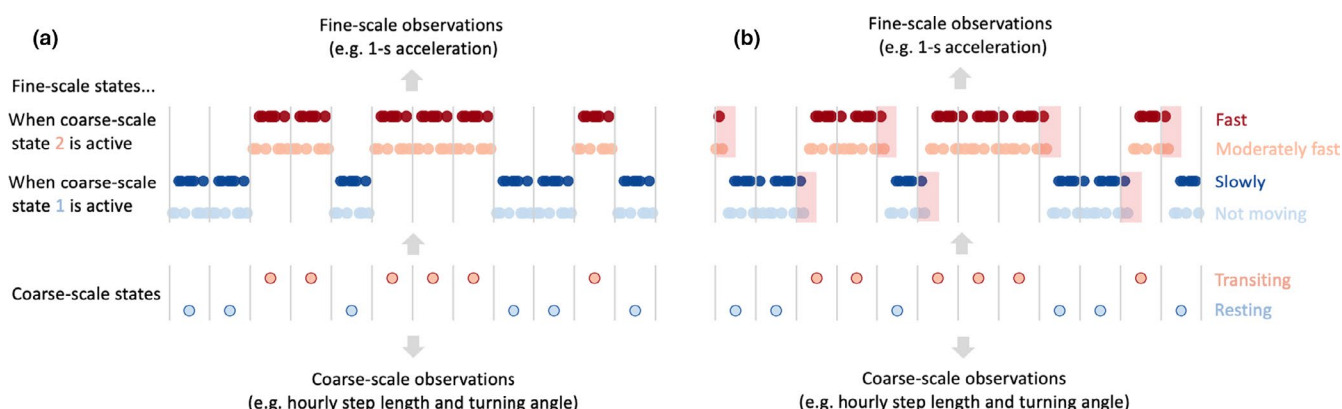


FIGURE 3 Illustration of the state processes of an Hierarchical HMM where the processes are aligned (a) and where this assumption is violated (b). A typical example for the latter scenario is an animal that performs (moderately) fast movements during the first few minutes of an hour that is, overall, characterised by resting behaviour, or no (slow) movements during an hour that is, overall, characterised by transiting behaviour (red-shaded areas). Hourly segments are indicated by vertical lines

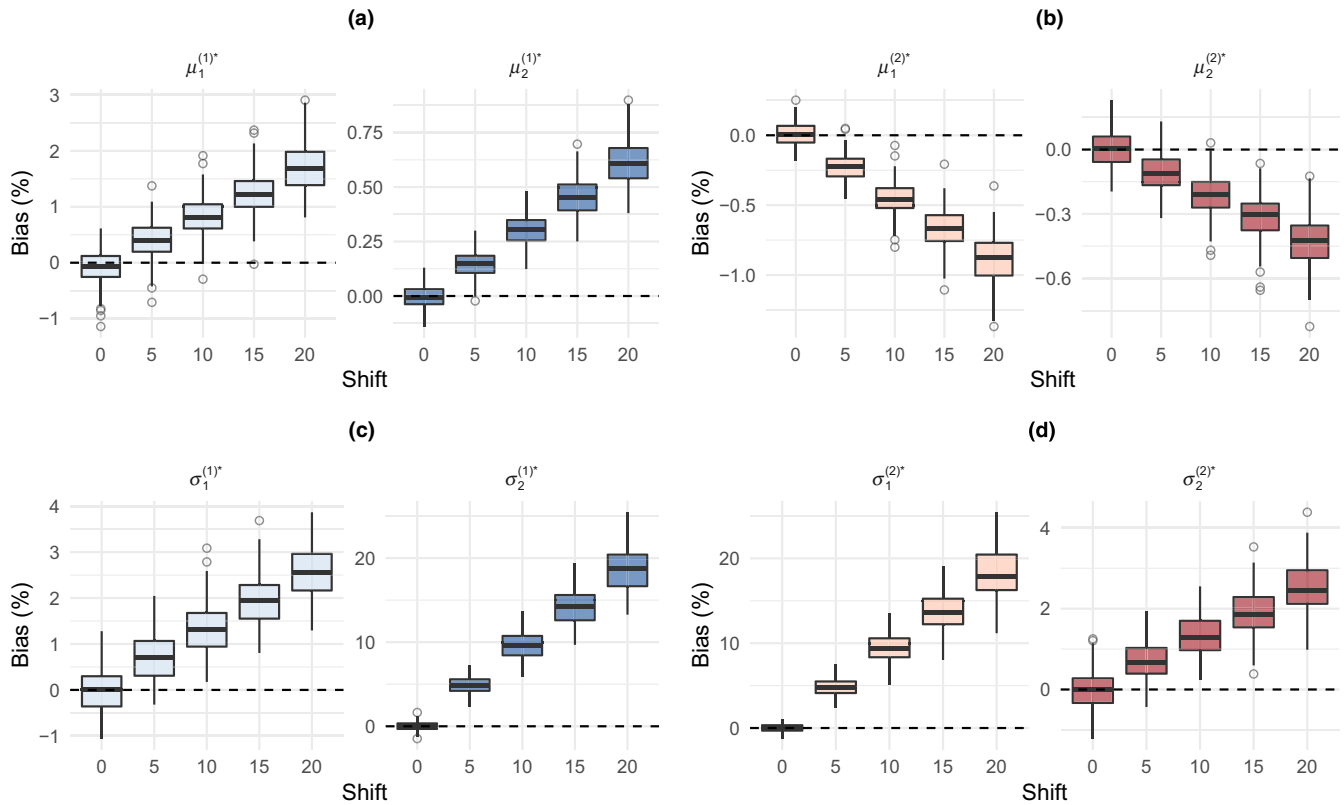


FIGURE 4 Sample of results from the simulation experiment. Displayed is the percentage bias obtained across all 200 replications. The means of the accelerations under state i associated with fine-scale HMM k are denoted by $\mu_i^{(k)*}$ (panels (a) and (b)); the corresponding variances are denoted by $\sigma_i^{(k)*}$ (panels (c) and (d)). Full results are displayed in section S2.2 of the supplementary material

fine-scale state process is shifted. While the bias remains relatively small for small shifts, it increases sharply with shifting of the fine-scale process. This severe bias is due to the fact that each of the two fine-scale HMMs must accommodate observations within each hour that truly belong to the alternate fine-scale HMM: a restriction imposed by having an hourly coarse-scale process.

As a consequence, ecological interpretations of the state-dependent distributions can be difficult to elicit (in the above example, for instance, the true within-state variance of the accelerations is much smaller than the estimated one, simply because of the temporal regularity of the observations and the dependence structure of the HHMM). Furthermore, pitfalls regarding model selection, which are often problematic even in basic HMMs (Pohle et al., 2017), can be exacerbated in HHMMs. In this example, it is likely a four-state fine-scale process will be preferred: two to represent the true process and two to represent the unintended overlap of the two processes. This can cause misleading ecological conclusions about patterns in animal behaviour.

3.3 | Opportunities and future prospects

Recent advances in biologging technology have led to the ability to track animals for increasingly long time periods at increasingly fine temporal resolutions (Kays et al., 2015; Lennox et al., 2017). As these

data are often collected by multiple sensors with different sampling frequencies (e.g. GPS tags, dive loggers or accelerometers), HHMMs provide a natural framework to jointly model these data and make inference on the multi-scale nature of animal behaviour depicted by these new types of data. However, as demonstrated in Section 3.2, the true data-generating processes underlying these multi-scale data do not necessarily follow the assumed dependence structure and thus the model can misconstrue inference on the underlying behavioural processes.

To best exploit the opportunities offered by multi-scale data using HHMMs, more flexible dependence structures are needed. There are three possible approaches that future research could pursue: (a) estimate the optimal resolution of the coarse-scale state process and compute coarse-scale observations on this time interval, rather than have this be determined by the sampling protocol; (b) develop models where switches between coarse-scale states and fine-scale HMMs are smoother and not a binary change; or (c) allow the time spent in each coarse-scale state to depend on the fine-scale state-switching dynamics and the fine-scale observations. If such extensions can overcome the problems outlined in this section, HHMMs have the opportunity to become a promising tool to draw a complete picture of animal behaviour, where an animal's movement decisions made at various time-scales, ranging from seasonal migration over diurnal activity to movements of individual body parts, can all be modelled in a joint modelling framework.

4 | RANDOM EFFECTS

4.1 | Introduction

HMMs are often used to make inference on multiple time series of data where each arises from a different sampling unit (e.g. movement data from different animals or abundance indices from different populations). One aim in such studies is to elicit common patterns that are exhibited across sampling units, for example, the relationship between an animal's movement and an environmental covariate.

It is common in this case to assume *complete pooling* of the parameters of the HMM for ease of interpretability, that is, one assumes the state-dependent distributions, the state-switching dynamics, and all parameters are the *same* across all sampling units. Nonetheless, ecological sampling units are often heterogeneous and so models that do not account for this variation can be too inflexible and lead to invalid inference. There are two alternatives to complete pooling: *no pooling*, where essentially one fits a separate HMM to each sampling unit and forgoes the ability to infer common patterns statistically, or *partial pooling*. Partial pooling refers to including random effects, either continuous valued or discrete valued (McClintock, 2021; Zucchini et al., 2017), in the observation or state processes to account for the heterogeneity among sampling units. McClintock et al. (2021) conducted an in-depth simulation to demonstrate when inclusion of random effects on the state-switching process would affect inference and state prediction. For this reason, we focus on when random effects are included on the state-dependent distributions instead.

As an example, consider the garter snake movement dataset in Leos-Barajas et al. (2017) where a time series of observed distances moved was recorded for multiple snakes. Suppose a priori we expect three underlying states, but know the movements under these states may vary by snake. We can model these data as a three-state HMM where the state-dependent distributions are gamma distributions with mean $\mu_{i,n}$ and standard deviation $\sigma_{i,n}$ for states $n = 1, 2, 3$ and snakes $i = 1, \dots, K$. For complete pooling, one would assume $\mu_{1,n} = \mu_{2,n} = \dots = \mu_{K,n}$ for all n (and similarly for σ). For no pooling, one would estimate each mean and standard deviation separately for each snake and state. For partial pooling, one could assume a hierarchical model where there is a population-level parameter, for example, μ_n for state n , and each snake's individual-level parameter varies around this population-level mean with some standard deviation τ_n for state n , that is, $\mu_{k,n} \sim N(\mu_n, \tau_n)$. This is the simplest way to partially pool across individuals, but it can lead to an unexpected pitfall: the added flexibility can overfit to certain individuals and make interpretation of the underlying states at the population-level difficult.

4.2 | Pitfall

At present, when including random effects in HMMs, two key assumptions are implicitly made (i) the K individual time series exhibit

the same number of distinct states N and (ii) the ecological interpretation of the states is consistent across sampling units. That is, state n across all individuals should be comparable and serve as a proxy for the same latent ecological behaviour. However, in practice, there is seldom a way to know a priori if all individuals exhibited the same number of states during the period of their observation or if the states are estimable given the data collected. Random effect modelling, as currently developed for HMMs, can account for individual heterogeneity, but does not have any theoretical grounding to prevent their accommodation of this heterogeneity from deforming interpretation of the population-level states. Even when the model assumed is correct, one may be unable to recover the true parameter values (see Section S3 of the Supplementary Material).

This is a key pitfall in HMMs. The difficulties that arise when accounting for individual heterogeneity in the state-dependent distributions stem from a lack of estimability and interpretability of the states across individuals. Thus, as it stands, partial pooling may not allow researchers to understand individual heterogeneity exhibited in the population of interest by simply incorporating random effects as this can lead to inaccurate inference due to a lack of estimability of the state-dependent parameters.

For the garter snake movement example, Figure 5 displays the estimated state-dependent distributions of four snakes, along with 95% pointwise credible intervals, for the model with complete pooling and partial pooling of the means. Estimated state-dependent distributions for all snakes are provided in Section S3 of the Supplementary Material. Across both models, the population-level estimates of the state-dependent distributions demonstrate three distinct states (more details in Leos-Barajas et al. (2017)), yet there are clear differences (especially in state 2) indicating the possible importance of accounting for individual variation. However, the individual-specific state-dependent distributions in the partial pooling model demonstrate crucial differences which are challenging to address. For example, for many individuals (e.g. snakes 1 and 18) there is a large overlap between state 1 and state 2 distributions, for example, $\mu_{k,2}$, for $k \in \{1 \dots K\}$, have lower bounds as small as 0.04, compared to a 95% credible interval for μ_2 of (0.11, 0.20). This raises a question: are we capturing true differences across states in individuals, or are we overfitting, or are the individual-specific parameters not estimable? At present, there is no mechanism in random effect modelling for HMMs to address this question without simply making more assumptions.

4.3 | Opportunities

Inclusion of random effects in a HMM provides opportunities to learn about individual heterogeneity, personality and preference within a given population of interest. However, we demonstrate that even under correct model specification, a HMM with random effects in the observation process may not be able to recover the true individual-specific state-dependent distributions and can lead to biased inferences for the state-switching dynamics (see Section

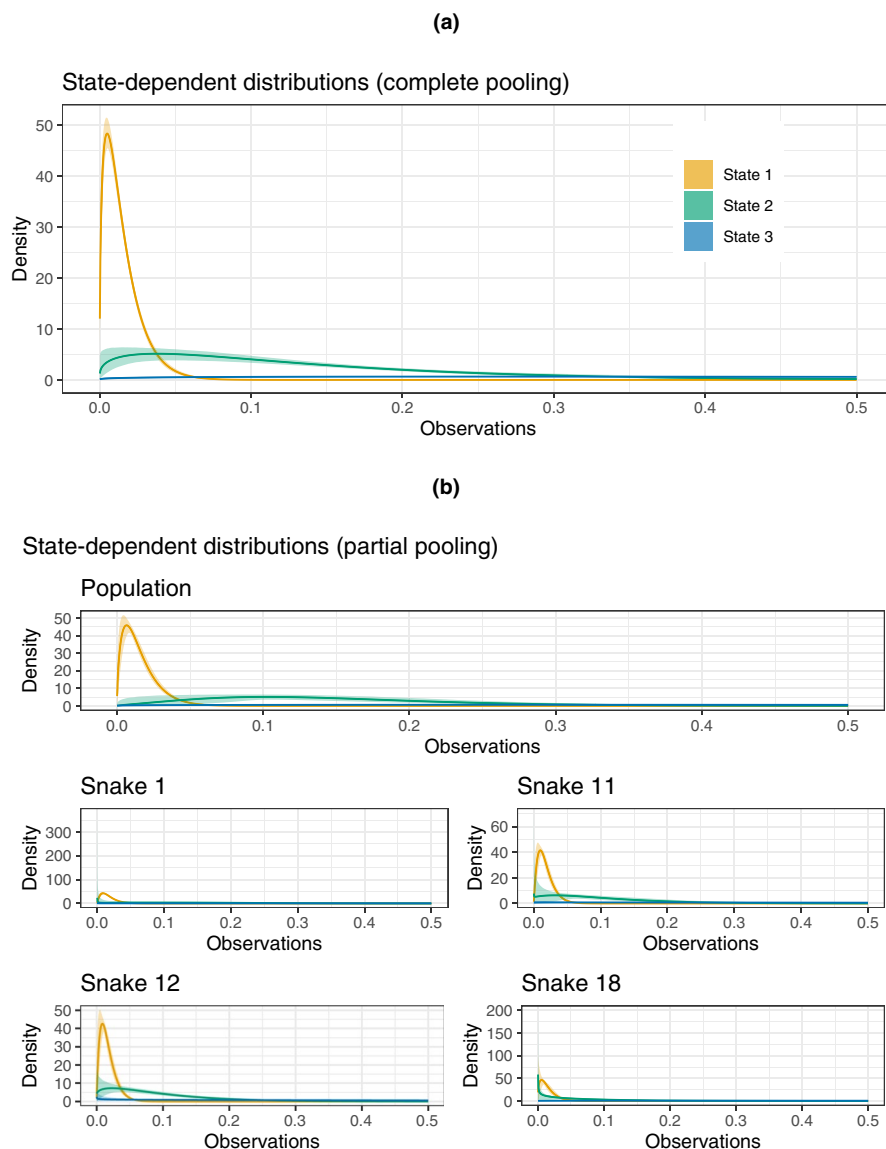


FIGURE 5 Estimated state-dependent distributions, unweighted, for models with complete pooling (a) and partial pooling (b) of the state-dependent means, along with 95% pointwise credible intervals

S3 of the Supplementary Material). As such, one opportunity is to understand under which experimental designs a HMM with random effects is estimable.

In addition to tackling the issue of estimability, another opportunity is to develop methodology that better respects the usual aim of using a HMM: to discern individual-specific state-dependent distributions that are *different enough* from one another across states and are coherent with the population-level inference. In Figure 5, the densities for state 1 and 2 have more overlap than is present in the population-level estimates, even when the aim is to capture three *distinct* states. Enforcing these criteria may be possible via selection of informative prior distributions, if inference is conducted in a Bayesian framework, or appropriate regularisation, constructed through the elicitation of domain expertise. For instance, one possibility is to assign an informative prior on the difference between $\mu_{n+1} - \mu_n$, for $n \in \{1 \dots N-1\}$, in order to enforce separation between densities.

If further research into the methodological development of HMMs with random effects proves to provide robust results for a

variety of real-world data collection scenarios, it opens up the opportunity for HMMs to provide insights into individual-specific movement dynamics and how distinct animal behaviours manifest across a population.

5 | CONTINUOUS STATE SPACES

In a HMM, the underlying state S_t takes one of a finite number of values; however, HMMs can be used for *approximate* inference when S_t varies over an infinite number of states (Zucchini et al., 2017). Models where S_t varies in a continuous state space are called 'state space models' (SSMs; Auger-Méthé et al., 2021). The connection between state space models, hidden Markov models, and discretisation is well known (Anderson-Sprecher & Ledolter, 1991; Kitagawa, 1987). In ecology, the approximate HMMs are used to model animal movement (Pedersen et al., 2011), population dynamics (Besbeas & Morgan, 2019), distance sampling sightings (Glennie et al., 2021), missing continuous covariates in capture-recapture

(Langrock & King, 2013) and for moving activity centres in spatial capture–recapture (Glennie et al., 2019). Yet, the method remains an obscurity for many ecological statisticians, for example, Patterson et al. (2017), in a review of animal movement modelling, described it as ‘underutilised’.

In this section, we describe the ideas involved when approximating a SSM with a HMM, the current pitfalls of this method, and the opportunities improving this method can bring to ecological applications.

5.1 | Example

In this section, we provide an example of constructing an approximate HMM for a continuous-time, state-switching animal movement SSM. This is to introduce the reader to the important ideas behind this method. The idea is to break the continuous hidden state in the SSM into a discrete, binned state for a HMM. We use continuous-time, state-switching animal movement with Brownian motion as an example (Pedersen et al., 2011), where the data consist of observations of the animal's location over (possibly irregular) time. We assume an animal has two latent behavioural states and the diffusion rate depends on its behaviour. The hidden variable in this SSM is the animal's location in 2D space coupled with its behavioural state. Figure 6 (Panel b) shows an example where 2D space has been split into a 5×2 grid. The approximate HMM will have 20 hidden states in this case as each hidden state represents where the animal is on the 2D grid and what behaviour the animal is in.

Once the states are defined for the HMM, the transition probability matrix or, if working in continuous time, transition rate matrix is defined. The transition probabilities must be derived from the parameters and state process in the SSM. If one can compute from the SSM the transition probabilities in continuous space (as in Mews, Langrock, Ötting, et al., 2020) then one can compute the transition probabilities directly in the discrete space. However, this is not always possible, for example, with spatially varying SSMs or state-switching SSMs. Following Pedersen et al. (2011), we can derive a transition rate matrix for the HMM that mimics the state process of the SSM using a partial differential equation (PDE; e.g. see Okubo & Levin, 2001). There are several methods (e.g. finite differencing, finite volume or finite element) to convert this PDE into a transition matrix (Quarteroni & Valli, 2008). The parameters of the SSM determine the switching rates for the approximating HMM. Figure 6 (Panel a) shows an example of a derived transition rate matrix for a behaviour-switching Brownian motion model on the 5×2 grid. Note that for most PDE methods the derived transition matrix is *sparse* (i.e. most of its entries are zero).

Once the transition matrix over the grid is specified, one can fit the approximate HMM in the usual way and estimate the parameters of the SSM. Figure 6 (Panel c) shows two examples of using a derived transition rate matrix to update the probability distribution over the 2D-behaviour space. Section S4 of the Supplementary

Material provides a full example of building HMM approximations from PDEs, including all code for constructing the necessary matrices and computing the likelihood.

5.2 | Pitfalls

The biggest pitfall to using approximate HMMs to fit SSMs is the curse of dimensionality. For higher dimensional hidden processes (e.g. where an animal's location and velocity are both hidden, a four-dimensional space), the number of states in the HMM quickly becomes computationally infeasible. For animal movement models, this inhibits the maximum number of behavioural states one can consider and precludes the possibility of incorporating directional persistence. For population dynamics, this limits the number of groups in age-structured models. Although this curse is unavoidable, future research could alleviate it: irregular gridding of the latent space can focus inference on important parts (Pedersen & Weng, 2013), only parts of the space can be updated when necessary, parallelised sparse matrix–vector products and sparse matrix exponential methods can be more fully exploited (Sherlock, 2021; Sidje, 1998), or sparse grids can reduce the number of states without compromising accuracy (Garcke, 2012). Many of these proposed ideas for future research have direct analogy with extensions of simulation-based methods.

The second pitfall is encountered when deciding whether to compute the transition probabilities directly or by PDE. Both have limitations. The direct approach requires a known solution to the SSM and, for large state spaces, a threshold be set such that transition probabilities below that threshold are taken to be zero as computing all pairwise transition probabilities is infeasible. The PDE approach has the advantage that the transition rate matrix's sparsity is fixed no matter the time interval between observations and no thresholding is necessary; however, the PDE approach is, at present, limited to SSM with Gaussian state processes and requires one to compute a sparse matrix *exponential*. Future research is needed to expand one or both of these approaches to a wider range of SSMs.

The final pitfall concerns the PDE approach in particular. The transition rate matrix is derived using PDE techniques from applied mathematics. These techniques are designed to compute an accurate solution to the PDE given the parameters are known. In ecological applications, however, the aim is different: to compute the solution many times and to determine the optimal parameters for the PDE given the data. Some techniques suitable for the former purpose may be subtly unsuitable for the latter, for example, when dealing with artificial diffusion or cross-diffusion (Quarteroni & Valli, 2008). In Section S4.4 of the Supplementary Material, we show an example of this problem for directed animal movement. Future research should consider the effect this may have on inference: most applications of the PDE method in ecology have investigated the efficacy of state prediction and not parameter recovery (Gatti et al., 2021).

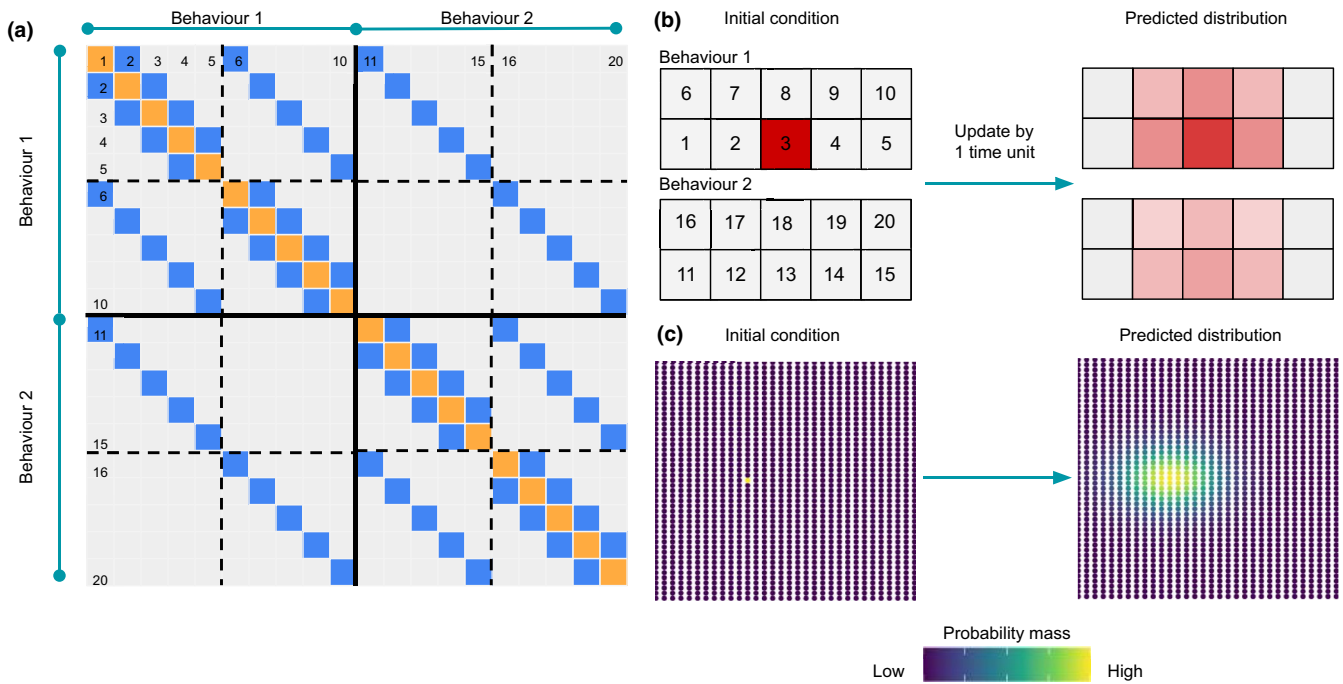


FIGURE 6 Example of hidden Markov model for state-switching animal movement over 2D (x, y) space: (a) the transition rate matrix derived from the continuous-space, continuous-time partial differential equation model (Pedersen et al., 2011) where non-zero entries are coloured and some rows and columns are numbered to indicate what grid cell in 2D-behaviour space (panel [b]) they refer to, solid lines demarcate blocks of the matrix corresponding to different behavioural states and dotted lines demarcate blocks corresponding to grid cells with different y values; (b) the 2D-behaviour space the transition rate matrix corresponds to, with two behavioural states, each with five grid cells in the x direction and two in the y direction, each grid cell is numbered and has colouring to depict the initial distribution of an animal over this space (here the animal is in grid cell 3 in behaviour 1 with probability 1) and the predicted distribution of the animal's location after one time unit (greater transparency indicates less probability mass); (c) a numerical example with one behavioural state using a higher grid resolution with an initial distribution of the animal in a single grid cell with probability 1 and the predicted distribution one time unit later under 2D Brownian motion, see section S4 of the supplementary material for the code to compute this approximation

5.3 | Opportunities

This method could have wide application in ecology if future research into its pitfalls led to improved computational and statistical implementation. To show this, we consider three areas where this method has shown some promise, but could be further exploited: animal movement, population dynamics and encounter modelling.

For animal movement, our example described one possible approach (Pedersen et al., 2008; Pedersen et al., 2011; Thygesen et al., 2009). It continues to be used (Braun et al., 2018; Haase et al., 2021; Jonsen et al., 2013), but only in limited contexts (focusing on light-based geolocation). Despite this, the flexibility of the HMM approach offers solutions to many common modelling needs that arise with animal telemetry: non-Gaussian, multivariate observations are easily accommodated, unlike common alternatives based on Kalman filtering (Johnson et al., 2008); continuous-time behaviour-switching can be accounted for and inference on where, not just when, behaviours are exhibited is possible (Pedersen et al., 2011), while alternatives make it more difficult to make this connection (see Section 6); barriers and obstacles to movement (e.g. land for marine animals) are trivially accounted for, while this remains difficult for simulation-based techniques; and environmental covariates can drive step-selection in continuous time, similar

to continuous-time discrete-space models (Hanks et al., 2015), but with flexibility to include behaviour-switching and link movement to a continuous-space model.

For population dynamics, De Valpine and Hastings (2002) introduced a discrete-time HMM approximation to general state-space population dynamics models. Besbeas and Morgan (2019) and Besbeas and Morgan (2020) generalise the approach and make the connection to HMMs explicit. These approaches differ from the example in Section 5.1: they do not consider continuous time and compute transition probabilities directly rather than via a PDE. The advantage of the HMM approach is the ability to model nonlinear dynamics that the standard Kalman filter cannot accommodate. It also allows for multivariate observations on population dynamics to be incorporated into a single model. Note that the state-switching models, continuous-time modelling or sparse matrix algorithms, all present in the animal movement applications of this method, are yet to be fully explored for population dynamics models.

For encounter models, we refer to applications where animal movement is inferred from encounters animals have with detectors, for example, human observers, cameras or acoustic devices. The approximate HMM is easily extended to allow for custom detection models. As with telemetry data, these detections can be used to infer individual animal movement (Dorazio & Price, 2019; Pedersen

& Weng, 2013). Yet, more powerfully, these individual animal movement models can be incorporated into population-level models such as distance sampling or spatial capture–recapture (e.g. McClintock et al., 2021). In each of these methods, a critical quantity is the probability *any* given animal in the population could be detected. When animals can move, this detection probability is calculated by averaging over all the possible (yet unobserved) paths an individual animal *could* have travelled. This is precisely the quantity the forward algorithm for HMMs can be used to compute. This approach has been taken to incorporate animal movement into both distance sampling (Glennie et al., 2021) and spatial capture–recapture (Glennie et al., 2019). Furthermore, this method provides the opportunity to build a general statistical approach to encounter modelling (Gurarie & Ovaskainen, 2013).

6 | SPATIAL ANALYSIS AND SPATIAL INTERPRETATION OF DECODED STATES

Inferring the behavioural state of an animal based on telemetry data is an increasingly popular application of HMMs in ecology. Telemetry data are primarily made up of geographical locations, which are inherently spatial, and HMMs for such data often characterise animal movement as a time series of steps and turns (e.g. Langrock et al., 2012), either on the plane (2D) or in a volume (3D). Inferences from movement HMMs can therefore assign behavioural states to spatial locations, which can be visualised in space by plotting the predicted state assignments on a map. Under these basic HMMs, the model is blind to the spatial mechanisms that give rise to certain behaviours. The spatial locations of the decoded states are therefore irrelevant to the model, but highly relevant to the ecological application, and the ecological interpretation of decoded states. A common extension to bring spatial mechanism into the model is to incorporate spatial covariate effects on the transition probabilities or parameters of the state-dependent observation distributions (e.g. Langrock et al., 2012; McClintock et al., 2012; Morales et al., 2004; Mul et al., 2020; Rivest et al., 2016). This creates an explicit link

between space and movement behaviour. The pitfall, however, with this approach is that the appropriate structure for a given research question can quickly lead to a model that is very complex, has a large number of parameters, and whose results are difficult to interpret. For these reasons, an attractive alternative is a two-stage approach, where a relatively simple HMM is first used to get state assignments for spatial locations, and post-hoc analyses are conducted to infer the relationship between the states and the spatial context in which they occur (e.g. Breed et al., 2009; Nickel et al., 2021). Although this is a simpler approach, there has been little research into how best to build a statistically robust and rigorous two-stage model.

To illustrate the pitfall in using HMMs to make spatial inferences and spatially interpreting the decoded states, we present a case study where HMMs were used to make inferences about animal movement from tracking data, and secondary spatial analyses were carried out to address specific spatial, ecological and behavioural questions.

6.1 | Pitfall example: Do hummingbirds use landmarks to remember spatial locations?

The movement trajectories of 14 rufous hummingbirds *Selasphorus rufus* were recorded in three-dimensions during a field experiment (Westcastle Valley, Canadian Rockies, Alberta, Canada, May–July 2014) to investigate spatial memory and learning. The data (step length, pitch angle, yaw angle) were analysed in (Pritchard et al., 2021) using a HMM, with distance to the location where a flower-shaped feeder was previously situated, and the presence or absence of landmarks, as covariates on the transition probability matrix.

The objective of the study was to gain a better understanding of how hummingbirds learn rewarding spatial locations, and as such there was an interest in *where*, with respect to the feeder and landmarks, certain movement behaviours occurred, as illustrated in Figure 7. Knowing if birds were more likely to display targeted searching behaviour in particular parts of space would allow inferences about what specific features were learned and remembered. Hypothesis tests

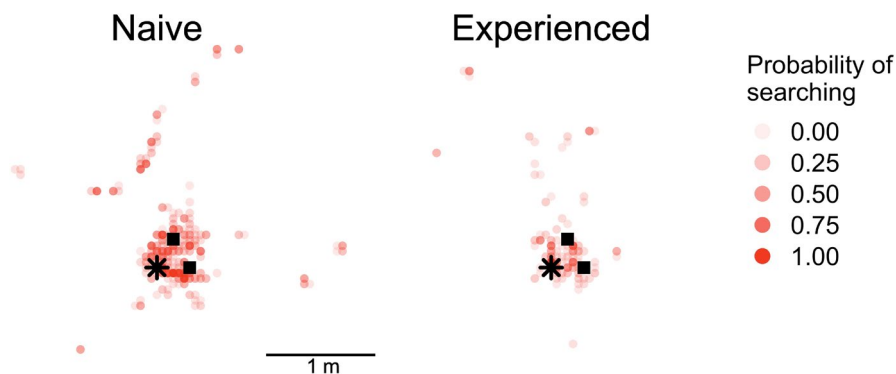


FIGURE 7 The spatial distribution of locations assigned the searching state for naive birds, who only had one prior visit to the feeder (asterisk) before it was removed, and experienced birds who had several prior visits. The black squares mark the location of two artificial landmarks that were present during the experiment, while the feeder had been removed. the colour brightness gives the state probability, to convey uncertainty in state assignments

(t-tests) were carried out to determine whether inferred searching state locations and stops made by experienced birds (with several prior visits to the feeder in the presence of landmarks) were closer to the feeder location than would have been expected by chance, compared to naive birds with a single prior visit. While this is a reasonable approach and provides some insight into where behaviours occur, it does not fully utilise the information contained in the spatial distribution of states, or account for uncertainty in the state assignments.

6.2 | Opportunities

This pitfall is at odds with what many, if not most, ecological and conservation science studies want to achieve when analysing telemetry data: an understanding of why behaviours occur where they do. In addition, the most easily implemented conservation and management interventions are often place based, for example, marine protected areas. To inform place-based management decisions with inferences from movement HMMs, we must be able to identify both important habitat (e.g. stop-over sites) and the behaviour(s) associated with it (e.g. migration). This is key for identifying what parts of space have disproportionate functional value to an animal population and deserve further protection or management (Lennox et al., 2019). Extending the use of HMMs to better answer these questions or promoting alternative methods will have widespread impact on conservation practice, feeding into policy and management decisions about place- and time-based conservation of wildlife, as well as their spatial and temporal exposure to risk.

There are two questions any such proposed extension must address: (a) How can spatial inference on behavioural state be made rigorously and robustly? (b) How can uncertainty in state be accounted for? For HMMs, future research could consider a more refined two-stage approach (e.g. taking predicted states, or, repeatedly drawing from the predicted state probabilities from a HMM and then performing spatial analyses), implementing carefully thought-through spatio-temporal transition probability matrices and/or observation distributions where interpretability permits, or exploiting the existing ability to make spatial inference with continuous-space HMMs (as described in Section 5). In the hummingbird example, the spatial research question could have instead been addressed within the HMM by including an interaction term between level of experience and distance to the flower on the state transitions. It is possible, however, that alternative, explicitly spatial methods, such as continuous-time discrete-space models (Hanks et al., 2015) or Langevin diffusion models (Michelot et al., 2019), could be extended to multiple behavioural states, thereby yielding state-specific utilisation distributions that could inform behaviour- and place-based management decisions. Such extensions could be embedded within the continuous-time HMM framework by formulating the observation distribution accordingly, possibly using a discrete-time approximation [e.g. Equation (S1) in Section S1.1 of the Supplementary Material], as has been recently done (e.g. *momentuHMM* version 2.0.0; McClintock & Michelot, 2018).

7 | CONCLUSION

Hidden Markov models are a versatile class of models that will continue to be developed for and applied to ecological problems. In this review, we have highlighted capabilities of HMMs that could be fruitful for ecological applications, but where one must be aware of current statistical and computational pitfalls. HMMs are becoming more complex in structure to better realise the potential of the more heterogeneous, multi-scale, multi-dimensional data being collected and the more detailed research questions being asked. This can make HMMs a valuable tool to capture a wide variety of observations, to model multiple processes acting at difference scales, to describe individual and population-level effects, and to link state-switching to space as well as time. With these developments, however, come greater difficulties in interpreting these models, justifying their assumptions, and fitting these models with current computational capabilities. We hope this paper heightens awareness of the trouble more complex models can bring and what opportunities for future solutions these problems can inspire.

ACKNOWLEDGEMENTS

The authors have no conflict of interest. The findings and conclusions in the manuscript are those of the author(s) and do not necessarily represent the views of the National Marine Fisheries Service, NOAA. Any use of trade, product or firm names does not imply an endorsement by the US Government. R.G.'s contribution was funded by the Biometrika Trust.

CONFLICT OF INTEREST

None of the authors have a conflict of interest.

AUTHORS' CONTRIBUTIONS

All authors contributed equally to the conception and implementation of the review; R.G. coordinated the writing of the manuscript and all authors contributed substantially and critically to the manuscript. All authors gave final approval for submission.

PEER REVIEW

The peer review history for this article is available at <https://publons.com/publon/10.1111/2041-210X.13801>.

DATA AVAILABILITY STATEMENT

Garter snake data (used in Section 4) are available from this link <https://link.springer.com/article/10.1007/s13253-017-0282-9#Sec19>.

ORCID

Richard Glennie  <https://orcid.org/0000-0003-3806-4280>

Vianey Leos-Barajas  <https://orcid.org/0000-0001-8016-773X>

Théo Michelot  <https://orcid.org/0000-0002-3838-4113>

Theoni Photopoulou  <https://orcid.org/0000-0001-9616-9940>

Brett T. McClintock  <https://orcid.org/0000-0001-6154-4376>

REFERENCES

- Adam, T., Griffiths, C. A., Leos-Barajas, V., Meese, E. N., Lowe, C. G., Blackwell, P. G., Righton, D., & Langrock, R. (2019). Joint modelling of multi-scale animal movement data using hierarchical hidden Markov models. *Methods in Ecology and Evolution*, 10(9), 1536–1550.
- Altman, R. M. (2007). Mixed hidden markov models: An extension of the hidden markov model to the longitudinal data setting. *Journal of the American Statistical Association*, 102(477), 201–210.
- Anderson-Sprecher, R., & Ledolter, J. (1991). State-space analysis of wildlife telemetry data. *Journal of the American Statistical Association*, 86(415), 596–602.
- Aquino-Baleytó, M., Leos-Barajas, V., Adam, T., Hoyos-Padilla, M., Santana-Morales, O., Galván-Magaña, F., González-Armas, R., Lowe, C., Ketchum, J., & Villalobos-Ortiz, H. (2021). Diving deeper into the underlying white shark behaviours at Guadalupe Island, Mexico. *Ecology and Evolution*, 11(21), 14932–14949.
- Auger-Méthé, M., Newman, K., Cole, D., Empacher, F., Gryba, R., King, A. A., Leos-Barajas, V., Mills Flemming, J., Nielsen, A., Petris, G., & Thomas, L. (2021). A guide to state-space modeling of ecological time series. *Ecological Monographs*, 91(4), e01470. <https://doi.org/10.1002/ecm.1470>
- Beketov, M. A. (2014). *HMMCont: Hidden Markov model for continuous observations processes*. R Package Version 1.0. Retrieved from <https://cran.r-project.org/web/packages/HMMCont/index.html>
- Besbeas, P., & Morgan, B. J. T. (2019). Exact inference for integrated population modelling. *Biometrics*, 75(2), 475–484.
- Besbeas, P., & Morgan, B. J. T. (2020). A general framework for modeling population abundance data. *Biometrics*, 76(1), 281–292.
- Blackwell, P. G. (2018). Integrated continuous-time hidden Markov models. *arXiv preprint arXiv:1807.11907*.
- Blackwell, P. G., Niu, M., Lambert, M. S., & LaPoint, S. D. (2016). Exact Bayesian inference for animal movement in continuous time. *Methods in Ecology and Evolution*, 7(2), 184–195.
- Borchers, D., Distiller, G., Foster, R., Harmsen, B., & Milazzo, L. (2014). Continuous-time spatially explicit capture-recapture models, with an application to a jaguar camera-trap survey. *Methods in Ecology and Evolution*, 5(7), 656–665.
- Braun, C. D., Galuardi, B., & Thorrold, S. R. (2018). HMMoce: An R package for improved geolocation of archival-tagged fishes using a hidden Markov method. *Methods in Ecology and Evolution*, 9(5), 1212–1220.
- Breed, G. A., Jonsen, I. D., Myers, R. A., Bowen, W. D., & Leonard, M. L. (2009). Sex-specific, seasonal foraging tactics of adult grey seals (*halichoerus grypus*) revealed by state-space analysis. *Ecology*, 90(11), 3209–3221.
- Bureau, A., Shiboski, S., & Hughes, J. P. (2003). Applications of continuous time hidden Markov models to the study of misclassified disease outcomes. *Statistics in Medicine*, 22(3), 441–462.
- Choquet, R., Garnier, A., Awuve, E., & Besnard, A. (2017). Transient state estimation using continuous-time processes applied to opportunistic capture-recapture data. *Ecological Modelling*, 361, 157–163.
- De Valpine, P., & Hastings, A. (2002). Fitting population models incorporating process noise and observation error. *Ecological Monographs*, 72(1), 57–76.
- Descamps, S., Tarroux, A., Cherel, Y., Delord, K., Godø, O. R., Kato, A., Krafft, B. A., Lorentsen, S., Ropert-Coudert, Y., Skaret, G., & Varpe, Ø. (2016a). Data from: At-sea distribution and prey selection of Antarctic petrels and commercial krill fisheries. Movebank Data Repository. <https://doi.org/10.5441/001/1.q4gn4q56>
- Descamps, S., Tarroux, A., Cherel, Y., Delord, K., Godø, O. R., Kato, A., Krafft, B. A., Lorentsen, S., Ropert-Coudert, Y., Skaret, G., & Varpe, Ø. (2016b). At-sea distribution and prey selection of Antarctic petrels and commercial krill fisheries. *PLoS One*, 11(8), e0156968.
- Dorazio, R. M., & Price, M. (2019). State-space models to infer movements and behavior of fish detected in a spatial array of acoustic receivers. *Canadian Journal of Fisheries and Aquatic Sciences*, 76(4), 543–550.
- Fine, S., Singer, Y., & Tishby, N. (1998). The hierarchical hidden Markov model: Analysis and applications. *Machine Learning*, 32(1), 41–62.
- Garcke, J. (2012). Sparse grids in a nutshell. In *Sparse grids and applications* (pp. 57–80). Springer.
- Gatti, P., Fisher, J. A., Cyr, F., Galbraith, P. S., Robert, D., & Le Bris, A. (2021). A review and tests of validation and sensitivity of geolocation models for marine fish tracking. *Fish and Fisheries*, 22(5), 1041–1066.
- Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis*. CRC Press.
- Glennie, R., Borchers, D. L., Murchie, M., Harmsen, B. J., & Foster, R. J. (2019). Open population maximum likelihood spatial capture-recapture. *Biometrics*, 75(4), 1345–1355.
- Glennie, R., Buckland, S. T., Langrock, R., Gerrodette, T., Ballance, L., Chivers, S., & Scott, M. (2021). Incorporating animal movement into distance sampling. *Journal of the American Statistical Association*, 116(533), 107–115.
- Gurarie, E., & Ovaskainen, O. (2013). Towards a general formalization of encounter rates in ecology. *Theoretical Ecology*, 6(2), 189–202.
- Haase, S., Krumme, U., Gräwe, U., Braun, C. D., & Temming, A. (2021). Validation approaches of a geolocation framework to reconstruct movements of demersal fish equipped with data storage tags in a stratified environment. *Fisheries Research*, 237, 105884.
- Hanks, E. M., Hooten, M. B., & Alldredge, M. W. (2015). Continuous-time discrete-space models for animal movement. *The Annals of Applied Statistics*, 9(1), 145–165.
- Jackson, C. H. (2011). Multi-state models for panel data: The msm package for R. *Journal of Statistical Software*, 38(8), 1–29.
- Jackson, C. H., Sharples, L. D., Thompson, S. G., Duffy, S. W., & Couto, E. (2003). Multistate Markov models for disease progression with classification error. *Journal of the Royal Statistical Society: Series D (the Statistician)*, 52(2), 193–209.
- Johnson, D. S., London, J. M., Lea, M. A., & Durban, J. W. (2008). Continuous-time correlated random walk model for animal telemetry data. *Ecology*, 89(5), 1208–1215.
- Jonsen, I., Basson, M., Bestley, S., Bravington, M., Patterson, T., Pedersen, M. W., Thomson, R., Thygesen, U. H., & Wotherspoon, S. (2013). State-space models for bio-loggers: A methodological road map. *Deep Sea Research Part II: Topical Studies in Oceanography*, 88, 34–46.
- Kays, R., Crofoot, M. C., Jetz, W., & Wikelski, M. (2015). Terrestrial animal tracking as an eye on life and planet. *Science*, 348(6240), aaa2478–1–aaa2478–9.
- Kitagawa, G. (1987). Non-gaussian state-space modeling of nonstationary time series. *Journal of the American Statistical Association*, 82(400), 1032–1041.
- Langrock, R., & King, R. (2013). Maximum likelihood estimation of mark-recapture-recovery models in the presence of continuous covariates. *The Annals of Applied Statistics*, 7(3), 1709–1732.
- Langrock, R., King, R., Matthiopoulos, J., Thomas, L., Fortin, D., & Morales, J. M. (2012). Flexible and practical modeling of animal telemetry data: Hidden Markov models and extensions. *Ecology*, 93(11), 2336–2342.
- Langrock, R., & Zucchini, W. (2011). Hidden markov models with arbitrary state dwell-time distributions. *Computational Statistics & Data Analysis*, 55(1), 715–724.
- Lennox, R. J., Aarestrup, K., Cooke, S. J., Cowley, P. D., Deng, Z. D., Fisk, A. T., Harcourt, R. G., Heupel, M., Hinch, S. G., Holland, K. N., et al. (2017). Envisioning the future of aquatic animal tracking: Technology, science, and application. *Bioscience*, 67(10), 884–896.
- Lennox, R. J., Engler-Palma, C., Kowarski, K., Filous, A., Whitlock, R., Cooke, S. J., & Auger-Méthé, M. (2019). Optimizing marine spatial plans with animal tracking data. *Canadian Journal of Fisheries and Aquatic Sciences*, 76(3), 497–509.
- Leos-Barajas, V., Gangloff, E. J., Adam, T., Langrock, R., Van Beest, F. M., Nabe-Nielsen, J., & Morales, J. M. (2017). Multi-scale modeling of

- animal movement and general behavior data using hidden Markov models with hierarchical structures. *Journal of Agricultural, Biological and Environmental Statistics*, 22(3), 232–248.
- Liu, Y.-Y., Li, S., Li, F., Song, L., & Rehg, J. M. (2015). Efficient learning of continuous-time hidden Markov models for disease progression. *Advances in Neural Information Processing Systems*, 28, 3500–3607.
- MacKenzie, D. I., Nichols, J. D., Royle, J. A., Pollock, K. H., Bailey, L., & Hines, J. E. (2018). *Occupancy estimation and modeling: Inferring patterns and dynamics of species occurrence* (2nd ed.). Elsevier.
- McClintock, B. T. (2017). Incorporating telemetry error into hidden Markov models of animal movement using multiple imputation. *Journal of Agricultural, Biological and Environmental Statistics*, 22(3), 249–269.
- McClintock, B. T. (2021). Worth the effort? A practical examination of random effects in hidden Markov models for animal telemetry data. *Methods in Ecology and Evolution*, 12(8), 1475–1497.
- McClintock, B. T., Abrahms, B., Chandler, R. B., Conn, P. B., Converse, S. J., Emmet, R., Gardner, B., Hostetter, N. J. & Johnson, D. S. (2021). An integrated path for spatial capture-recapture and animal movement modeling. *Ecology*, e03473. <https://doi.org/10.1002/ecy.3473>.
- McClintock, B. T., Johnson, D. S., Hooten, M. B., Ver Hoef, J. M., & Morales, J. M. (2014). When to be discrete: The importance of time formulation in understanding animal movement. *Movement Ecology*, 2(1), 1–14.
- McClintock, B. T., King, R., Thomas, L., Matthiopoulos, J., McConnell, B. J., & Morales, J. M. (2012). A general discrete-time modeling framework for animal movement using multistate random walks. *Ecological Monographs*, 82(3), 335–349.
- McClintock, B. T., Langrock, R., Gimenez, O., Cam, E., Borchers, D. L., Glennie, R., & Patterson, T. A. (2020). Uncovering ecological state dynamics with hidden Markov models. *Ecology Letters*, 23(12), 1878–1903.
- McClintock, B. T., & Michelot, T. (2018). momentuHMM: R package for generalized hidden Markov models of animal movement. *Methods in Ecology and Evolution*, 9(6), 1518–1530.
- McLachlan, G. J., & Krishnan, T. (2007). *The EM algorithm and extensions*. John Wiley & Sons.
- Mews, S., Langrock, R., King, R., and Quick, N. (2020). Continuous-time multi-state capture-recapture models. *arXiv preprint arXiv:2002.10997*.
- Mews, S., Langrock, R., Ötting, M., Yaqine, H., and Reinecke, J. (2020). Maximum approximate likelihood estimation of general continuous-time state-space models. *arXiv preprint arXiv:2010.14883*.
- Michelot, T., & Blackwell, P. G. (2019). State-switching continuous-time correlated random walks. *Methods in Ecology and Evolution*, 10(5), 637–649.
- Michelot, T., Gloaguen, P., Blackwell, P. G., & Étienne, M.-P. (2019). The Langevin diffusion as a continuous-time model of animal movement and habitat selection. *Methods in Ecology and Evolution*, 10(11), 1894–1907.
- Morales, J. M., Haydon, D. T., Frair, J., Holsinger, K. E., & Fryxell, J. M. (2004). Extracting more out of relocation data: Building movement models as mixtures of random walks. *Ecology*, 85(9), 2436–2445.
- Mul, E., Blanchet, M.-A., McClintock, B. T., Grecian, W. J., Biuw, M., & Rikardsen, A. (2020). Killer whales are attracted to herring fishing vessels. *Marine Ecology Progress Series*, 652, 1–13.
- Nickel, B. A., Suraci, J. P., Nisi, A. C., & Wilmsers, C. C. (2021). Energetics and fear of humans constrain the spatial ecology of pumas. *Proceedings of the National Academy of Sciences of the United States of America*, 118(5), e2004592118.
- Okubo, A., & Levin, S. A. (2001). *Diffusion and ecological problems: Modern perspectives*. Springer.
- Patterson, T. A., Parton, A., Langrock, R., Blackwell, P. G., Thomas, L., & King, R. (2017). Statistical modelling of individual animal movement: An overview of key methods and a discussion of practical challenges. *AStA Advances in Statistical Analysis*, 101(4), 399–438.
- Pedersen, M. W., Patterson, T. A., Thygesen, U. H., & Madsen, H. (2011). Estimating animal behavior and residency from movement data. *Oikos*, 120(9), 1281–1290.
- Pedersen, M. W., Righton, D., Thygesen, U. H., Andersen, K. H., & Madsen, H. (2008). Geolocation of North Sea cod (*Gadus morhua*) using hidden markov models and behavioural switching. *Canadian Journal of Fisheries and Aquatic Sciences*, 65(11), 2367–2377.
- Pedersen, M. W., & Weng, K. C. (2013). Estimating individual animal movement from observation networks. *Methods in Ecology and Evolution*, 4(10), 920–929.
- Plummer, M. (2017). JAGS: Just another Gibbs sampler. Version 4.0.3 user manual. Retrieved from https://people.stat.sc.edu/hansont/stat740/jags_user_manual.pdf
- Pohle, J., Langrock, R., van Beest, F. M., & Schmidt, N. M. (2017). Selecting the number of states in hidden Markov models: Pragmatic solutions illustrated using animal movement. *Journal of Agricultural, Biological and Environmental Statistics*, 22(3), 270–293.
- Pritchard, D. J., Hurly, T. A., Photopoulou, T., and Healy, S. D. (2021). 3d flightpaths reveal the development of spatial memory in wild hummingbirds. In prep.
- Quarteroni, A., & Valli, A. (2008). *Numerical approximation of partial differential equations*. Springer Science & Business Media.
- R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.
- Rivest, L.-P., Duchesne, T., Nicosia, A., & Fortin, D. (2016). A general angular regression model for the analysis of data on animal movement in ecology. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 65(3), 445–463.
- Royle, J. A., & Kéry, M. (2007). A Bayesian state-space formulation of dynamic occupancy models. *Ecology*, 88(7), 1813–1823.
- Rubin, D. B. (2004). *Multiple imputation for nonresponse in surveys*. John Wiley & Sons.
- Sacchi, G., & Swallow, B. (2021). Toward efficient Bayesian approaches to inference in hierarchical hidden markov models for inferring animal behavior. *Frontiers in Ecology and Evolution*, 9, 249.
- Sherlock, C. (2021). Direct statistical inference for finite markov jump processes via the matrix exponential. *Computational Statistics*, 36, 2863–2887.
- Sidje, R. B. (1998). Expokit: A software package for computing matrix exponentials. *ACM Transactions on Mathematical Software (TOMS)*, 24(1), 130–156.
- Sidrow, E., Heckman, N., Fortune, S. M., Trites, A. W., Murphy, I., and Auger-Méthé, M. (2021). Modelling multi-scale state-switching functional data with hidden markov models. *arXiv preprint arXiv:2101.03268*.
- Thygesen, U. H., Pedersen, M. W., & Madsen, H. (2009). Geolocating fish using hidden Markov models and data storage tags. In *Tagging and tracking of marine animals with electronic devices* (pp. 277–293). Springer.
- Zucchini, W., MacDonald, I. L., & Langrock, R. (2017). *Hidden Markov models for time series: An introduction using R* (2nd ed.). CRC Press.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Glennie, R., Adam, T., Leos-Barajas, V., Michelot, T., Photopoulou, T., & McClintock, B. T. (2022). Hidden Markov models: Pitfalls and opportunities in ecology. *Methods in Ecology and Evolution*, 00, 1–14. <https://doi.org/10.1111/2041-210X.13801>