

# Flexible and practical modeling of animal telemetry data: hidden Markov models and extensions

ROLAND LANGROCK,<sup>1,5</sup> RUTH KING,<sup>1</sup> JASON MATTHIOPOULOS,<sup>2</sup> LEN THOMAS,<sup>1</sup> DANIEL FORTIN,<sup>3</sup> AND  
JUAN M. MORALES<sup>4</sup>

<sup>1</sup>Centre for Research into Ecological and Environmental Modelling, School of Mathematics and Statistics, The Observatory, Buchanan Gardens, University of St Andrews, St Andrews, Fife KY16 PLZ Scotland, United Kingdom

<sup>2</sup>Scottish Oceans Institute, School of Biology, University of St Andrews, St Andrews, Fife KY16 8LB Scotland, United Kingdom

<sup>3</sup>Centre d'Étude de la Forêt and Département de Biologie, Université Laval, Québec City, Québec G1V 0A6 Canada

<sup>4</sup>Ecotono, INIBIOMA-CONICET, Universidad Nacional del Comahue, Quintral 1250, 8400 Bariloche, Argentina

**Abstract.** We discuss hidden Markov-type models for fitting a variety of multistate random walks to wildlife movement data. Discrete-time hidden Markov models (HMMs) achieve considerable computational gains by focusing on observations that are regularly spaced in time, and for which the measurement error is negligible. These conditions are often met, in particular for data related to terrestrial animals, so that a likelihood-based HMM approach is feasible. We describe a number of extensions of HMMs for animal movement modeling, including more flexible state transition models and individual random effects (fitted in a non-Bayesian framework). In particular we consider so-called hidden semi-Markov models, which may substantially improve the goodness of fit and provide important insights into the behavioral state switching dynamics. To showcase the expediency of these methods, we consider an application of a hierarchical hidden semi-Markov model to multiple bison movement paths.

**Key words:** behavioral state; Bison bison; maximum likelihood; random effects; random walk; semi-Markov model; state-space model; telemetry data.

## INTRODUCTION

Analyzing animal movement is essential for understanding processes such as inter- and intraspecific interactions, the dynamics of populations, and their distribution in space. Recent research has focused on dissecting movement patterns into different behavioral states, each with an associated set of parameters governing the movement process. This was accompanied by a focus on Bayesian methods, which allow inference for models with analytically intractable likelihoods (Morales et al. 2004, Jonsen et al. 2005, McClintock et al. 2012). Many movement models proposed in recent years in fact belong to the class of hidden Markov models (HMMs), with only a handful of implementations of likelihood-based (i.e., classical, non-Bayesian) HMM methodology for movement data (Franke et al. 2006, Holzmann et al. 2006, Patterson et al. 2009, Pedersen et al. 2011b). The principal aim of this manuscript is to outline how basic HMM-type movement models can be extended in different ecologically interesting and novel ways, thereby demonstrating that the full potential of HMMs for modeling animal movement has yet to be realized. HMMs cannot deal with all problems typically encountered with telemetry

data; see the sections *HMM basics* and *Discussion*. However, important advantages of using HMMs are computational tractability and mathematical simplicity. Thus, for many applications they constitute a competitive alternative to more flexible, yet more challenging, state-space models (SSMs) and associated Bayesian methods (Patterson et al. 2008).

Our presentation is as follows. The section *HMM Basics* introduces the relevant HMM theory. The section *HMMs for a single animal's movement* discusses how a single animal's movement path can be described using HMMs. Basic HMMs can be extended in various ways, and in particular by allowing for more complicated dependence structures and by including random effects. We outline such extensions and the associated statistical methodology in the section *Extensions*. In the section *Application: bison in Saskatchewan, Canada*, the implementation of two important extensions is demonstrated by applying a hierarchical hidden semi-Markov model to multiple bison movement paths.

## HMM BASICS

### Model formulation

An HMM is a time series model that comprises two components, an observable series and an underlying, non-observable state sequence. The observed variables, which we denote by  $Z_1, \dots, Z_T$ , in the context of this paper relate to some quantification of movement. When

Manuscript received 14 December 2011; revised 9 May 2012; accepted 14 June 2012. Corresponding Editor: B. D. Inouye.

<sup>5</sup> E-mail: roland@mcs.st-and.ac.uk

modeling movement data, one can either model the positions themselves or some derived quantities. In the models that we consider,  $Z_t$  is bivariate and comprises step lengths and turning angles (or directions). Each (bivariate) observation is assumed to be generated by one of  $N$  distributions, and the unobservable (hidden) state sequence,  $S_1, \dots, S_T$ , modeled as a Markov chain, determines which of the  $N$  distributions is selected at each time  $t$ . In the context of animal movement, the states can be interpreted as behavioral states of the observed animal, such as “foraging” or “encamped”; Morales et al. (2004) provide a more detailed elaboration of this concept.

HMMs have precisely the same dependence structure as SSMs (see Appendix A for an illustration). Some authors in fact regard HMMs and SSMs as the same type of objects (Cappé et al. 2005), although an HMM is usually regarded as a special case of an SSM where the number of hidden states is finite. In SSM approaches to animal movement modeling, it is commonly assumed that the movement metrics are part of the hidden component, i.e., not directly observed due to measurement error (Patterson et al. 2008). Here, we assume that measurement error is negligible and, thus, that we observe the movement metrics directly. This reduces flexibility in terms of incorporating measurement error, with the advantage that the resulting models are much more mathematically tractable than SSMs, because the hidden component takes only a finite number of values.

The Markov chain  $S_t$  is usually assumed to be of first order, which means that the probability of state occurrences at time  $t + 1$  depends only on which state the chain is in at time  $t$ . The state transition probabilities (Pr) at time  $t$  are summarized in the  $N \times N$  matrix  $\Gamma^{(t)} = (\gamma_{ij}^{(t)})$ , where

$$\gamma_{ij}^{(t)} = \Pr(S_{t+1} = j | S_t = i), i, j = 1, \dots, N.$$

*Estimation, decoding, model selection, and model checking*

In the homogeneous case, i.e., if the transition probabilities do not depend on the time index  $t$ , the likelihood ( $\mathcal{L}$ ) of an HMM is given by the following matrix product:

$$\mathcal{L} = \delta^{(1)} \mathbf{P}(z_1) \Gamma \mathbf{P}(z_2) \Gamma \times \dots \times \Gamma \mathbf{P}(z_{T-1}) \Gamma \mathbf{P}(z_T) \mathbf{1}^T \quad (1)$$

where  $\mathbf{P}(z) = \text{diag}(f_1(z), \dots, f_N(z))$  with  $f_n = f_{Z_t | S_t=n}$  denoting the conditional probability density function of  $Z_t$ , given  $S_t = n$ ;  $\mathbf{1}$  is a row vector of ones;  $z_1, \dots, z_T$  denote the observations;  $\delta^{(1)}$  is the initial distribution of the Markov chain; and  $\text{diag}$  refers to a diagonal matrix. The initial distribution is usually assumed to be the stationary distribution (Patterson et al. 2009). It is straightforward to deal with missing data: for a corresponding time index  $t$ , the matrix  $\mathbf{P}(z_t)$  is replaced by the  $N \times N$  unity matrix. Maximum likelihood estimates of the parameters can be found by numerical

likelihood maximization, subject to well-known technical issues arising in all optimization problems (Zucchini and MacDonald 2009). Alternatively, one may apply the expectation-maximization (EM) algorithm.

Via the Viterbi algorithm, one can obtain estimates of the underlying hidden behavioral states,  $s_1, \dots, s_T$ , given an observed sequence of movements (Zucchini and MacDonald 2009: Chapter 5).

Furthermore, because there is an explicit, closed-form likelihood, it is straightforward to employ information criteria for model selection. Such measures compare the relative goodness of fit, while for example an analysis of the residuals provides a measure of absolute goodness of fit. Pseudo-residuals offer a convenient way for model checking in HMMs (Patterson et al. 2009; see also Appendix B).

HMMs FOR A SINGLE ANIMAL’S MOVEMENT

We consider multistate random walks in the spirit of Morales et al. (2004). These can be regarded as HMMs where each state is associated with a distinct random walk behavior. The class of building blocks that we consider for the state-dependent process  $Z_t$  comprises correlated and biased random walks (CRWs and BRWs, respectively), as well as walks that are both correlated and biased (BCRWs); see Codling et al. (2008) and Appendix C for more details. A typical HMM for animal movement will involve some combination of CRWs, BRWs, and BCRWs, each of them being allocated to a different state of the underlying Markov chain. Four sample movement trajectories simulated from HMMs with different components are given in Appendix D, illustrating the flexibility of these models. More ecological realism may be obtained, e.g., by allowing for seasonality or modeling parameters as a function of environmental covariates (see, e.g., Morales et al. 2004). Including covariates is straightforward in HMMs (Patterson et al. 2009).

A more specific question concerns the choice of the state-dependent distributions. Step lengths are, by nature, positive and continuous, which renders gamma (with the exponential as a special case) and Weibull distributions plausible candidates for modeling this component. Plausible distributions for modeling the circular-valued statistics given by turning angles (or directions) include the von Mises and wrapped Cauchy. All of these distributions have been used previously in animal movement models (Morales et al. 2004, Holzmann et al. 2006) and are equally easy to fit in the HMM framework.

EXTENSIONS

*Other formulations of the state process*

*Semi-Markov state processes.*—One important limitation of basic HMMs is that the state dwell-time distributions (i.e., the times spent in a state before changing to a different state) are necessarily geometric. In some applications this may not be biologically reasonable, in which case models that provide more

flexibility are called for. So-called hidden *semi*-Markov models (HSMMs) are designed to relax this condition: the dwell time in states of an HSMM is explicitly modeled from some distribution on the positive integers (Guédon 2003). The state process of an HSMM is determined by the probability mass functions (pmfs) of the dwell-time distributions,  $p_1(r), \dots, p_N(r)$  (with  $p_n(r)$  the probability that the duration of a stay in state  $n$  is  $r$  time units), and the conditional state transition probabilities, given the current state is left:  $\gamma_{ij} = \Pr(S_{t+1} = j | S_t = i, S_{t+1} \neq i)$ ,  $i, j = 1, \dots, N$ . The difference to HMMs is that the probabilities of self-transitions,  $\gamma_{ii}$ , which determine the times spent in the states, are now modeled separately by the distributions  $p_i$ .

Langrock and Zucchini (2011) demonstrate how any given HSMM can be approximated with arbitrary accuracy by a suitably structured HMM, which means that HSMMs can be fitted in the well-developed HMM framework. We illustrate the basic idea via a simple example. Suppose that a two-state HSMM is to be fitted, with state-dependent distribution  $f_n$  for state  $n = 1, 2$ , pmf  $p_1(r)$  from some non-geometric distribution with infinite support (e.g., a Poisson distribution),  $p_2(r)$  a geometric distribution, and a uniform initial state distribution. We now construct an HMM that approximates this HSMM by expanding the state that has a non-geometric dwell time into a set of  $N^*$  states (where  $N^*$  is some large number) and suitably structuring the transition probabilities between those  $N^*$  states. Consider an HMM with  $N^* + 1$  states, initial distribution  $\delta^{(1)} = (0.5, 0, \dots, 0, 0.5)$ , and state-dependent distributions given by  $f_k^* = f_1$  for  $k = 1, \dots, N^*$  and  $f_{N^*+1}^* = f_2$ . Each state in the set  $I = \{1, \dots, N^*\}$  is associated with the same distribution of the state-dependent process ( $I$  is called a state aggregate). We further define the  $(N^* + 1) \times (N^* + 1)$  transition probability matrix of the HMM as

$$\Gamma = \begin{pmatrix} 0 & 1-c(1) & 0 & \dots & 0 & c(1) \\ & 0 & 1-c(2) & & & c(2) \\ \vdots & & & \ddots & & \vdots \\ & & & & 0 & 1-c(N^*-1) & c(N^*-1) \\ 0 & & \dots & & 1-c(N^*) & c(N^*) \\ p_2(1) & 0 & \dots & & 0 & 1-p_2(1) \end{pmatrix}.$$

Here

$$c(r) = p_1(r) / \left( 1 - \sum_{k=1}^{r-1} p_1(k) \right)$$

with the usual convention that the empty sum equals 0. It can easily be shown (see Appendix G) that the pmf  $p_1^*(r)$  of the dwell time in  $I$  matches  $p_1(r)$  for  $r = 1, \dots, N^*$ . For  $r > N^*$ ,  $p_1^*(r)$  decays geometrically, which, in general, is only an approximation of  $p_1(r)$ . The approximation, however, can be made arbitrarily accurate by increasing  $N^*$ . We have thus represented the dwell time in state 1 of the HSMM as the dwell time in the state aggregate  $I$  of the approximating HMM (and recall that at the observation level there is no difference between state 1 of the HSMM and states  $1, \dots, N^*$  of

the HMM). This formulation does not entail an increase in the number of parameters, compared to the HSMM that is to be represented. In order to get an accurate approximation, this representation typically will involve large matrices, which makes these models computationally more demanding to fit than basic HMMs. The same is true for alternative strategies for estimating HSMMs (Guédon 2003). Crucially, the structure of the transition probability matrix is the same irrespective of what dwell-time distribution is to be fitted. For a more comprehensive description of this method, we refer the reader to Langrock and Zucchini (2011).

*Higher-order state processes.*—On only a slightly different note, another crucial assumption in the basic HMM is that the state process has a memory of one time lag only. Theoretically, it is easy to increase the memory of the Markov chain, and there are simple means for implementing the associated models. For example, if the Markov chain,  $S_t$ , is of second order, then the Markov chain  $U_t = (S_{t-1}, S_t)$  is of first order. Although higher-order state processes can also imply non-geometric state dwell-time distributions, there is a conceptual difference to the semi-Markov approach: here one increases the flexibility in modeling the animal's memory, including the memory of particular state switches, whereas in the semi-Markov approach, the time of staying in a state is modeled explicitly, but is unaffected by which behavioral state the animal occupied previously.

*Feedback models.*—Another strategy for modifying the dependence structure is to incorporate feedback from the observed process on the subsequent generation of states. Zucchini et al. (2008) modeled feeding behavior using an HMM-type feedback model. In their application, the observed behavior is either feeding or non-feeding, and the particular behavior observed influences the probabilities of transitioning to subsequent states (i.e., there is additional dependence of  $S_{t+1}$  on  $Z_t$  as determined by a functional relation between the observation at time  $t$  and the probabilities of state transitions in the time interval  $[t, t + 1]$ ).

*Continuous-valued behavioral states.*—Another critical issue concerns the state space (i.e., the set of possible behavioral states), because in some applications the division into a finite number of states may seem counterintuitive. The likelihood of a model with a continuous state space (i.e., that of an SSM) is given by a high-order multiple integral that usually cannot be evaluated analytically. However, by finely discretizing the state space, one is back in the HMM framework, such that estimation is not considerably more challenging than for models with a finite number of states (Langrock 2011, Pedersen et al. 2011a).

#### Multiple animals

Scaling individual movement models up to the population level is an issue of great ecological relevance, but it first requires us to quantify individual variation. The class of HMMs provides several different strategies

for dealing with a set of time series of individual animal movement paths. The simplest approach is to assume that the model parameters are identical across individuals (complete pooling). This approach can be unrealistic because it does not permit any individual variability. Conversely, it could be assumed that each parameter is individual-specific, i.e., that each individual has its own independent set of parameters (no pooling). This approach involves a larger number of parameters and generally post analysis ad hoc comparisons among individuals. Patterson et al. (2009) compare these approaches in an HMM application to tagging data from southern bluefin tuna (*Thunnus maccoyii*). An important alternative to these extreme options is a hierarchical model (partial pooling), where one assumes a common distribution for (some) individual-level parameters; these are then called *random effects*. For implementations of such movement models in a Bayesian framework, see e.g., Jonsen et al. (2006) and Eckert et al. (2008). For a likelihood-based implementation of a hierarchical HMM in a different ecological context, see Schliehe-Diecks et al. (2012). This approach substantially reduces the number of parameters to be estimated, and allows joint inference about population-level parameters. The implementation of random effects can be computationally intensive because the likelihood needs to be integrated over the random effects' distribution (Altman 2007):

$$\mathcal{L} = \prod_{m=1}^M \int_{\theta} \delta_{\theta}^{(1)} \mathbf{P}_{\theta}(z_{1,m}) \Gamma_{\theta} \mathbf{P}_{\theta}(z_{2,m}) \Gamma_{\theta} \dots \Gamma_{\theta} \mathbf{P}_{\theta}(z_{T,m}) \mathbf{1}^t f(\theta) d(\theta).$$

Here  $\theta \in \mathbb{R}^d$  denotes the vector of random effects (from the space of all  $d$ -dimensional vectors with real-valued components);  $M$  is the number of individuals (animals);  $z_{t,m}$  denotes the observation made at time  $t$  for individual  $m$ ; and  $f$  is the joint density of the random effects, with its choice depending on which parameters are to be modeled as random effects. It is further assumed that individuals act independently of each other. Numerical likelihood maximization, which we conducted in the application to bison in the next section, is only feasible for models involving few random effects; the other parameters are assumed to be common to all individuals. Monte Carlo EM (expectation-maximization) has been suggested when many random effects are included (Altman 2007). A computationally less demanding way of accounting for heterogeneity across individuals is to incorporate individual-specific covariates in the model. This strategy requires that suitable covariate data are available.

#### APPLICATION: BISON IN SASKATCHEWAN, CANADA

##### Data

We consider locations of nine American bison (*Bison bison*), recorded between October 2005 and April 2006 with GPS radio collars in Prince Albert National Park,

Saskatchewan, Canada. The observations are spaced at regular time intervals every three hours. More data are available (Babin et al. 2011), but for simplicity we restricted the present analysis to the winter of 2005/2006. The bison behavior varies with the season, but within one winter, homogeneity over time can reasonably be assumed. Between 1427 and 1660 locations were observed for each of the nine bison (there were some missing data). Plots of the nine movement paths are given in Appendix E. From the locations, we calculated turning angles (in radians) between subsequent movement directions, and the associated Euclidean step lengths (in km).

##### Models

We describe the results of fitting two different types of models at the individual level (model 1 HMMs; model 2 HSMMs) and of fitting three different types of models simultaneously to all nine bison paths: HMMs (model 3) and HSMMs (model 4) with complete pooling, and finally a hierarchical HSMM (model 5). All models were estimated via numerical maximum likelihood; R code is given in the Supplement. In all cases, we assumed two states, stationarity of the (semi-)Markov chain and that each state is associated with a distinct CRW. We assumed Weibull distributions for the step lengths, and wrapped Cauchy distributions for the turning angles. (Gamma distributions and von Mises distributions, respectively, were outperformed in terms of the AIC; see Appendix F.) In the HSMMs, models 2 and 4, we assumed negative binomial distributions for the state dwell times, with probability mass function in state  $n$  ( $n = 1, 2$ ) given by

$$p_n(r) = \frac{\Gamma(r + k_n - 1)}{(r - 1)! \Gamma(k_n)} \pi_n^{k_n} (1 - \pi_n)^{r-1} \quad r = 1, 2, 3, \dots$$

with parameters  $k_n > 0$  and  $\pi_n \in [0, 1]$ . Note that the HMMs are nested ( $k_1 = k_2 = 1$ ). The HSMMs were fitted using the HMM approximation method, as described in *Extensions: Other formulations of the state process*. Each of the two state aggregates that we used, corresponding to  $p_1$  and  $p_2$ , respectively, consisted of 30 states, thus ensuring a very accurate approximation (with this size of the state aggregates, the HMM approximation is essentially exact).

Model 5, a hierarchical two-state HSMM, assumes the Weibull scale parameter  $\lambda_{n,i}$  ( $n = 1, 2$  denoting the behavioral state) for bison  $i$  to be drawn from a log-normal distribution, i.e.,  $\log(\lambda_{n,i}) \sim \mathbf{N}(\mu_n, \sigma_n)$ . All other parameters were assumed to be common across individuals in this model. Biologically this parameterization means that the expected state-dependent step lengths vary across animals (because the mean of the Weibull distribution is proportional to the scale parameter).

##### RESULTS

Details of the model-fitting results are given in Appendix F. According to the fitted individual-specific HMMs and HSMMs (models 1 and 2), the movement

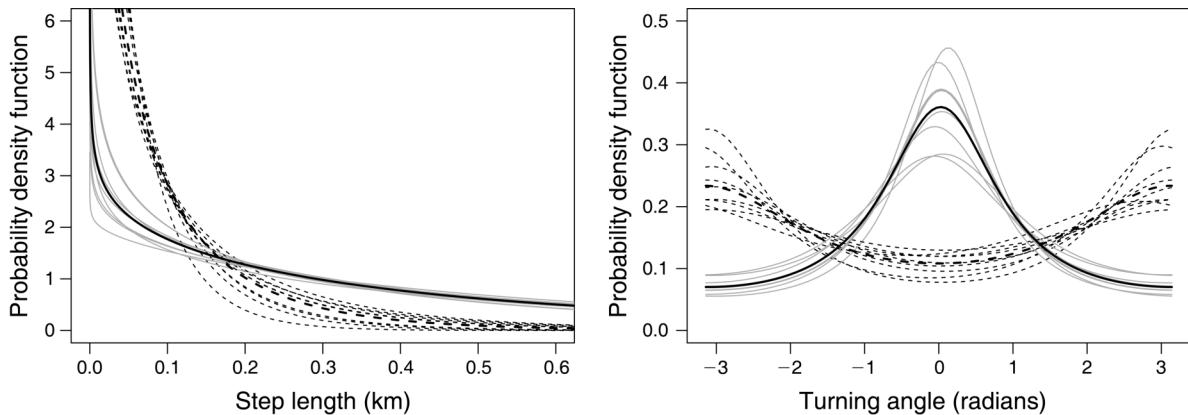


FIG. 1. Models 2 and 4 for the movement paths of nine American bison: fitted state-dependent Weibull distributions for step lengths (left) and wrapped Cauchy distributions for turning angles (right). Solid lines show “exploratory” movement; dashed lines show “encamped” movement. Gray lines are results for the individual-specific HSMs, and black lines are results for HSM with parameters common to all individuals.

paths of the different bison show a similar (stochastic) pattern. In particular, each bison switches between an “exploratory” (or “commuting”) state, with many long steps and few turnings, and an “encamped” state, with short steps and more frequent reversals (Fig. 1). The results are similar to those obtained for elk by Morales et al. (2004). Similarity in the patterns of the nine movement paths is stressed by the fact that when all nine bison are modeled simultaneously (models 3 and 4), the resulting AICs are lower than those of their respective counterparts with individual-specific parameters (i.e., the AICs resulting from the joint likelihood of the individual-specific models 1 and 2, respectively).

The HSM with complete pooling, model 4, has lower AIC than the HMM with complete pooling, model 3. At the individual level, HSMs are, in terms of the AIC, superior to their HMM competitors in seven out of nine cases. In the “exploratory” state, for seven out of nine bison the mode of the HSM-derived dwell-time distribution is greater than one. In the “encamped” state, the most apparent difference between the fitted dwell-time distributions of the HMMs and those of the HSMs is that in the latter case there is more mass on one in seven out of nine cases. A typical movement pattern thus consists of several successive commuting-type steps that are disrupted by pauses that mostly last

only one time unit (here: three hours). Residual analyses conducted for the fitted individual-based HSMs show that these models provide adequate fits (see Appendix B). However, the fitted models predict more steps of extremely short length (<10 m) than were observed (possibly related to GPS measurement error).

In terms of the AIC, the hierarchical HSM, model 5, performs better than any other model that we considered ( $\Delta\text{AIC} = 5.0$ , compared to the HSM with complete pooling). The parameter estimates and 95% confidence intervals for the hierarchical HSM are given in Table 1. Note that neither of the two confidence intervals associated with the size parameters of the state dwell-time distributions contains the value 1 (which corresponds to the special case of a geometric distribution, i.e., the HMM case). The relatively high uncertainty in the estimates associated with the random effects distributions is most likely due to the small number of individuals. According to the fitted random effects distributions, 95% of the individual-specific mean step lengths are in the intervals [0.073, 0.112] in the “encamped” state and [0.464, 0.564] in the “exploratory” state.

#### DISCUSSION

We outlined various extensions of basic HMM methods that may prove useful in animal movement

TABLE 1. Estimated parameters with 95% confidence limits for model 5 (hierarchical hidden semi-Markov model) where state  $i = 1$  is the “encamped” state and state  $i = 2$  is the “exploratory” state.

Term	Definition	State $i = 1$		State $i = 2$	
		Estimate	95% CL	Estimate	95% CL
$\beta_i$	Weibull shape	0.778	(0.753, 0.804)	0.870	(0.841, 0.900)
$\mu_i$	mean, logged Weibull scale	-2.492	(-2.628, -2.356)	-0.723	(-0.802, -0.645)
$\sigma_i$	SD, logged Weibull scale	0.097	(0.043, 0.217)	0.050	(0.020, 0.124)
$\omega_i$	wrapped Cauchy mean	3.131	(3.008, 3.255)	0.031	(-0.012, 0.073)
$\rho_i$	wrapped Cauchy scale	0.155	(0.125, 0.190)	0.394	(0.316, 0.491)
$k_i$	negative binomial size	0.394	(0.316, 0.491)	3.732	(1.917, 7.268)
$\pi_i$	negative binomial success probability	0.144	(0.114, 0.180)	0.596	(0.440, 0.735)

modeling. Many of these extensions have been developed only recently in the statistical literature, implying that the full scope of HMMs for modeling animal movement data has not yet been recognized in the ecological literature.

In particular, we have demonstrated how considerable flexibility can be gained with modest effort by allowing for semi-Markov state processes. In an HSMM, one explicitly models the time an animal stays in a behavioral state, rather than simply accepting the geometric decay of the dwell-time distribution imposed by basic HMMs. We believe this extension to be potentially very important for ecological time series, not only because it may be more realistic and thus improve the fit, but also because it may lead to important insights into the dynamics of behavioral changes that basic Markov models cannot provide.

Another important issue is hierarchical modeling of animal movement paths. Hierarchical models, in principle, allow us to quantify the difference between movement patterns of different animals. In the ecological literature, hierarchical movement models have been fitted several times using Bayesian methods (e.g., Jonsen et al. 2006, Eckert et al. 2008). We have demonstrated that such models can also be fitted in a likelihood-based HMM framework, provided that the number of random effects is small. We expect that in many applications the simultaneous incorporation of both individual-specific covariates and (few) random effects in the model will adequately account for possible heterogeneity, while being feasible and directly interpretable.

Two limitations of the HMM approach to animal movement modeling need to be mentioned. First, the approach discussed here is suited only to regularly timed observations. If time intervals between successive observations vary, it makes no sense to assume state transition probabilities and state-dependent distributions to be homogeneous. Missing observations do not represent a problem, but for data that involve observations that are completely irregularly spaced in time, continuous-time models such as the Ornstein-Uhlenbeck process (Blackwell 2003) or a continuous-time CRW (Johnson et al. 2008) seem better suited. Such models allow for a better independence between the scales of behavior and those of observations, but are considerably more challenging.

Second, we suggest that the type of models presented here should only be used for time series with negligible observation error. In principle, one may assume that the observations are subject to some random error, but this quickly leads to technical problems (e.g., state-dependent distributions become convolutions that, in most cases, cannot be evaluated analytically). In order to account for measurement error, Pedersen et al. (2011b) present an alternative continuous-time HMM approach, based on a multistate advection–diffusion model that involves a grid discretization of space.

Finally, as with other movement modeling techniques, care should be taken in considering the temporal scale of analysis. The analyzed bison data have a resolution of 3

hours and our HMM formulations imply that animals are allowed to make a “behavioral switch” at this temporal scale, which is unrealistic. Thus, the classification of “encamped” vs. “exploring” should not be taken literally. Apart from this caveat, the models that we implemented were capable of capturing important features of the data.

Despite these limitations, the scope of HMMs in animal movement modeling is very wide, particularly as the relevance of the measurement error diminishes with technological advancements, and because in many studies locations are indeed observed at regular time intervals. We believe that their computational tractability (fitting an HMM to the movement path of one bison takes about 1 minute, or 10 minutes for an HSMM; and fitting the hierarchical HSMM took about 2 days) and mathematical simplicity will make this class of models attractive to ecologists.

#### ACKNOWLEDGMENTS

The authors thank Tiago Marques, Martin Wæver Pedersen, and Ian Jonsen for very helpful comments on earlier versions of the manuscript. R. Langrock was funded by the Engineering and Physical Sciences Research Council (ESPRC reference EP/F069766/1).

#### LITERATURE CITED

- Altman, R. 2007. Mixed hidden Markov models: An extension of the hidden Markov model to the longitudinal data setting. *Journal of the American Statistical Association* 102:201–210.
- Babin, J.-S., D. Fortin, J. F. Wilmshurst, and M.-E. Fortin. 2011. Energy gains predict the distribution of plains bison across populations and ecosystems. *Ecology* 92:240–252.
- Blackwell, P. G. 2003. Bayesian inference for Markov processes with diffusion and discrete components. *Biometrika* 90:613–627.
- Cappé, O., E. Moulines, and T. Rydén. 2005. *Inference in hidden Markov models*. Springer, New York, New York, USA.
- Codling, E. A., M. J. Plank, and S. Benhamou. 2008. Random walk models in biology. *Journal of the Royal Society Interface* 5:813–834.
- Eckert, S. A., J. E. Moore, D. C. Dunn, R. S. van Buiten, K. L. Eckert, and P. N. Halpin. 2008. Modeling loggerhead turtle movement in the Mediterranean: importance of body size and oceanography. *Ecological Applications* 18:290–308.
- Franke, A., T. Caelli, G. Kuzyk, and R. J. Hudson. 2006. Prediction of wolf (*Canis lupus*) kill-sites using hidden Markov models. *Ecological Modelling* 197:237–246.
- Guédon, Y. 2003. Estimating hidden semi-Markov chains from discrete sequences. *Journal of Computational and Graphical Statistics* 12:604–639.
- Holzmann, H., A. Munk, M. Suster, and W. Zucchini. 2006. Hidden Markov models for circular and linear-circular time series. *Environmental and Ecological Statistics* 13:325–347.
- Johnson, D. S., J. M. London, M.-A. Lea, and J. W. Durban. 2008. Continuous-time correlated random walk model for animal telemetry data. *Ecology* 89:1208–1215.
- Jonsen, I. D., J. M. Flemming, and R. A. Myers. 2005. Robust state-space modeling of animal movement data. *Ecology* 86:2874–2880.
- Jonsen, I. D., R. A. Myers, and M. C. James. 2006. Robust hierarchical state-space models reveal diel variation in travel rates of migrating leatherback turtles. *Journal of Animal Ecology* 75:1046–1057.

- Langrock, R. 2011. Some applications of nonlinear and non-Gaussian state-space modelling by means of hidden Markov models. *Journal of Applied Statistics* 38:2955–2970.
- Langrock, R., and W. Zucchini. 2011. Hidden Markov models with arbitrary dwell-time distributions. *Computational Statistics and Data Analysis* 55:715–724.
- McClintock, B. T., R. King, L. Thomas, J. Matthiopoulos, B. J. McConnell, and J. M. Morales. 2012. A general discrete-time modeling framework for animal movement using multi-state random walks. *Ecological Monographs*: forthcoming.
- Morales, J. M., D. T. Haydon, J. L. 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.
- Patterson, T. A., M. Basson, M. V. Bravington, and J. S. Gunn. 2009. Classifying movement behaviour in relation to environmental conditions using hidden Markov models. *Journal of Animal Ecology* 78:1113–1123.
- Patterson, T. A., L. Thomas, C. Wilcox, O. Ovaskainen, and J. Matthiopoulos. 2008. State-space models of individual animal movement. *Trends in Ecology and Evolution* 23:87–94.
- Pedersen, M. W., C. W. Berg, U. H. Thygesen, and H. Madsen. 2011a. Estimation methods for nonlinear state-space models in ecology. *Ecological Modelling* 222:1394–1400.
- Pedersen, M. W., T. A. Patterson, U. H. Thygesen, and H. Madsen. 2011b. Estimating animal behaviour and residency from movement data. *Oikos* 120:1281–1290.
- Schliehe-Diecks, S., P. M. Kappeler, and R. Langrock. 2012. On the application of mixed hidden Markov models to multiple behavioural time series. *Interface Focus* 2:180–189.
- Zucchini, W., and I. L. MacDonald. 2009. Hidden Markov models for time series: an introduction using R. Chapman and Hall, London, UK.
- Zucchini, W., D. Raubenheimer, and I. L. MacDonald. 2008. Modeling time series of animal behavior by means of a latent-state model with feedback. *Biometrics* 64:807–815.

#### SUPPLEMENTAL MATERIAL

##### Appendix A

Graphical model of an HMM for animal movement (*Ecological Archives* E093-220-A1).

##### Appendix B

Description of HMM model checking via pseudo-residuals, and plots of pseudo-residuals obtained in the bison application (*Ecological Archives* E093-220-A2).

##### Appendix C

Outline of the different types of random walks, and description of how biased random walks can be fitted in the HMM framework (*Ecological Archives* E093-220-A3).

##### Appendix D

Plots of movement trajectories simulated from four different types of HMMs (*Ecological Archives* E093-220-A4).

##### Appendix E

Plots of the bison movement paths (*Ecological Archives* E093-220-A5).

##### Appendix F

Model fitting results, including parameter estimates for the individual-specific models, log-likelihood and AIC values for all considered models, and plots of the fitted state dwell-time distributions in the individual-specific models (*Ecological Archives* E093-220-A6).

##### Appendix G

Verification of the HMM representation being an approximate representation of the example HSMM given in the section *Extensions: Other formulations of the state process* (*Ecological Archives* E093-220-A7).

##### Supplement

R code for fitting the individual-specific HMMs and HSMMs, for computing the HSMM residuals, for fitting the hierarchical HSMM, and observations for one of the bison (*Ecological Archives* E093-220-S1).