



Original Article

# Using mixed hidden Markov models to examine behavioral states in a cooperatively breeding bird

Ann E. McKellar,<sup>a</sup> Roland Langrock,<sup>b</sup> Jeffrey R. Walters,<sup>c</sup> and Dylan C. Kesler<sup>a</sup>

<sup>a</sup>Department of Fisheries and Wildlife Sciences, University of Missouri, 1105 E Rollins, Columbia, MO 65211, USA, <sup>b</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, UK, and <sup>c</sup>Department of Biological Sciences, Virginia Tech, 1405 Perry Street, Blacksburg, VA 24061, USA

Received 30 April 2014; revised 11 August 2014; accepted 20 August 2014; Advance Access publication 30 September 2014.

Movement has important consequences for individual and population-level processes, but methods are only starting to become available for quantifying fine-scale movement paths of smaller animals. New techniques for inferring behavioral states and their relation to social and environmental factors provide a powerful way to test the influence of such factors on individuals. One such technique that has recently gained popularity is the use of hidden Markov models, which link time series of movement variables and the underlying behavioral states of individuals. We used hidden Markov models to evaluate behavioral states and their relation to environmental, seasonal, and social factors in the cooperatively breeding red-cockaded woodpecker (*Picoides borealis*) while accounting for individual heterogeneity with discrete random effects. We identified 2 distinct behavioral states, resting and foraging, which were related to covariates in our models. Using this approach, we concluded that woodpecker step lengths tended to be longest in winter, larger groups of woodpeckers tended to spend less time foraging and more time resting when compared with smaller groups, and woodpeckers foraged more and rested less when in higher-quality habitat. Our results demonstrate the impact that social and environmental factors can have on movement in a social species and, thus, reinforce the importance of including these factors in animal movement studies. The extensions of basic hidden Markov models considered here may prove valuable in forthcoming studies that involve high-resolution tracking to understand behavior of birds and other small animals.

**Key words:** behavioral state, maximum likelihood, *Picoides borealis*, red-cockaded woodpecker, telemetry data.

## INTRODUCTION

Animal movement is a process involving internal cues and external factors that interact to influence decisions and behaviors (Nathan et al. 2008). Despite its importance in ecology and evolution, and considerable interest in the study of animal movement in recent decades, research is only presently beginning to characterize individual movements in a quantitative way (Holyoak et al. 2008). Recent work on animal movement has categorized movement paths into states that link to underlying individual motivations and behaviors; examples include encamped and exploratory states in elk (*Cervus elaphus*) (Morales et al. 2004); transient and resident states in bowhead

whales (*Balaena mysticetus*) (Pomerleau et al. 2011); bedding, feeding, and relocating states in woodland caribou (*Rangifer tarandus*) (Franke et al. 2004); and 3 types of diving states in American mink (*Neovison vison*) (Bagniewska et al. 2013). Analyzing the behavioral states of individuals can lead to insights regarding resource and space use (Forester et al. 2007; Fryxell et al. 2008) and can be scaled up to examine population processes (Morales et al. 2010). Understanding these features in endangered and sensitive species may be particularly useful for management and conservation planning (Lusseau 2003; Pomerleau et al. 2011; Anadón et al. 2012).

Advances in tracking techniques, such as radio and satellite telemetry, have allowed researchers to record animal movement paths at high spatial (e.g., <1 m) and temporal resolutions (e.g., 1 min). Linking these data to social and environmental variables is a powerful way to test the influence of exogenous factors on movement behaviors (e.g., Morales et al. 2004; Patterson et al. 2009;

Address correspondence to A.E. McKellar, who is now at Environment Canada, 115 Perimeter Road, Saskatoon, Saskatchewan S7N 0X4, Canada. E-mail: ann.mckellar@ec.gc.ca.

Langrock et al. 2014). Statistical models for dissecting movement patterns into different underlying states have been formulated and fitted primarily in a discrete-time framework, typically using either hidden Markov models (HMMs; Morales et al. 2004; Holzmann et al. 2006; Patterson et al. 2009; Langrock et al. 2012) or more general state-space models (SSMs; Jonsen et al. 2005; Patterson et al. 2008; Schick et al. 2008; Jonsen et al. 2013), although we note that a similar approach has also been considered in a continuous-time framework (Blackwell 2003). HMMs and SSMs are doubly stochastic processes and have precisely the same dependence structure, with an observed time series such that any observation depends only on the current value of an underlying Markov state (or system) process. Some authors in fact do not distinguish between HMMs and SSMs (e.g., Cappé et al. 2005). However, the label HMM is usually used to indicate a model with a finite number of possible states, whereas in SSMs, the underlying state process typically takes continuous values and hence involves an infinite number of states. In the literature on movement modeling via state-switching processes, SSM approaches typically include both the continuous movement metrics and the discrete behavioral states in the hidden component of the model, using the link to the observations to describe potential measurement error (e.g., Patterson et al. 2008). In contrast, in HMM approaches, the measurement error is typically assumed to be negligible, so that the hidden component of the model only involves the behavioral states, with the observed process giving the movement metrics, typically step lengths and turning angles (Langrock et al. 2012). SSMs of the described type usually require the use of the Bayesian inference paradigm—with all its advantages and disadvantages—whereas HMMs can easily be fitted also in a frequentist framework using the corresponding powerful machinery.

Assuming negligible measurement error and adopting a likelihood-based frequentist estimation approach, we applied HMMs to a large data set of movements of foraging red-cockaded woodpeckers at Eglin Air Force Base (AFB), Florida, USA, with the goal of identifying behavioral states and their relation to the social (group size) and environmental context (season, habitat quality) within which they occurred. Red-cockaded woodpeckers breed cooperatively and occupy year-round territories centered on a cluster of nest and roost trees (Lennartz et al. 1987). Much work has been devoted to understanding the habitat requirements of this species. In general, higher woodpecker fitness is associated with open, park-like mature pine woodlands and savannahs characterized by open understory with herbaceous growth, large pines, low densities of small pines, and a sparse hardwood midstory (Hardesty et al. 1997; James et al. 1997, 2001; Walters et al. 2002; McKellar et al. 2014). Red-cockaded woodpeckers at Eglin AFB have been found to select forest stands with more large pines and fewer small pines when compared with random stands, and the strength of resource selection varied with group size (McKellar et al. 2013). Thus, we predicted that the woodpeckers would demonstrate differential behavioral states in relation to these factors.

Our methods extend standard HMM-type movement models in several ways. First, we consider zero-inflated step length distributions, which was valuable for our analysis due to the high temporal resolution of our data set, with observations made every minute and woodpeckers often remaining in the same foraging location for several consecutive minutes. Second, we include seasonal effects in our models to account for and to examine potential differences in the step length distributions across seasons. Movements by foraging woodpeckers are expected to vary seasonally as invertebrate prey

are more scarce and/or more difficult to access in winter, whereas resources are more abundant in spring and summer (Skorupa and McFarlane 1976; Bradshaw 1990; McKellar et al. 2013). Third, we allow for the possibility that the movement dynamics of woodpecker groups may vary across territories of different types by including random effects in our models. Here, we avoid the computational problems that have plagued previous attempts to include random effects in HMMs by incorporating discrete-valued random effects, following Maruotti and Rydén (2009). Compared to conventional strategies for incorporating random effects in HMMs (e.g., Altman 2007), this leads to a substantially decreased computing time and also avoids potentially unrealistic assumptions about the shape of the random effects distribution (Aitkin 1996). Discrete-valued random effects have been implemented in capture–recapture studies (Pledger and Schwarz 2002), but to the best of our knowledge, they have not yet been applied in animal movement analyses. Finally, though previous work with HMMs has included multiple covariates (e.g., Schliehe-Diecks et al. 2012), our study is unique in that we explicitly assess both social (group size) and environmental factors that may influence woodpecker movement behaviors.

## MATERIALS AND METHODS

### Field methods

Red-cockaded woodpeckers are nonmigratory and occupy year-round territories as solitary males, pairs, or cooperatively breeding groups (Walters et al. 1988, 1992; Jackson 1994). Pairs and groups typically forage together on their large (i.e., 50–100 ha) territories during the day, departing from their roosting cavities after sunrise and returning to them prior to sunset (Lennartz et al. 1987). After fledging, young woodpeckers that survive remain with the family group until they either disperse or become helpers. Dispersal is thought to peak several weeks after fledging (July or August), with a second peak occurring just prior to the following breeding season (March or April; Kesler et al. 2010; Kesler and Walters 2012). Breeding pairs can be assisted by up to 5 helpers, which are typically males fledged during previous breeding seasons (Walters et al. 1988).

We followed groups of red-cockaded woodpeckers at Eglin AFB (30°29'N, 86°32'W), Florida, USA, between February 2007 and November 2008. We performed observation sessions ( $n = 440$ ), defined as the set of location points for a given woodpecker group on a given day, at 97 territories (1–8 observation sessions per territory). Birds were followed during each session by a single observer outfitted with a global positioning system (GPS; Garmin eTrex Legend, Garmin Ltd, Olathe, KS), set to automatically record the geographic coordinates of its location once each minute. Red-cockaded woodpecker habitats on the study site were composed of open stands of trees with substantial sky exposure, so we anticipate that the accuracy of GPSs was  $<3$  m error (95% circular effort probability; Garmin Ltd 2005). This error was judged to be small in comparison to woodpecker step lengths (see Results). Observers arrived at territory centers at or before sunrise and before birds emerged from roost cavities. After woodpeckers emerged, observers recorded the number of birds in each group and traveled on foot with woodpecker groups for approximately 1 h (February 2007 to July 2007) or 2 h (September 2007 to November 2008) while birds foraged and moved throughout territories. Because observations were recorded during all months, except December, group sizes included single males, pairs, pairs with helpers, and groups with

fledglings not yet dispersed, and ranged from 1 to 9 individuals. We attempted to observe each group at least 3 times each year.

During observation sessions, woodpeckers moved from tree to tree, and although group members were not always on the same tree at the same time, birds remained in the same general areas and regularly made contact calls. Observers traveled with groups without difficulty and usually remained <30 m from a group member. Observers spent time near each woodpecker as they visually observed and recorded the combination of colored leg bands and recorded behavior observations, and thus, tracking locations were recorded very close to the birds and within the same forest stand in which woodpecker groups were located (mean area of forest stands at Eglin was 0.0427 ha). The population has been monitored using these methods for many years and there have been no indications that foraging behavior was affected by observer presence. Observers stood still while birds were at rest or when they spent periods of time foraging at a single site.

We removed locations recorded within 50 m of the territory center, defined as the geographic mean of coordinates for nesting and roosting trees, to reduce the influence of preforaging social interactions that occurred at sunrise and near roost cavities. We truncated observation sessions when birds made sudden long-distance movements and observers could not follow them, when inclement weather prevented the visual observations of foraging birds, or when GPS units failed to estimate a location and a period of >1 min elapsed between recorded locations. We removed data from 9 sessions that included <15 observations. In total, our data set consisted of 35 304 locations for the 440 observation sessions.

We used ArcMap version 10.0 (ESRI, Redlands, CA) to calculate step lengths and turning angles for each observed location. We used forest stand geodatabases (spatial and quantitative stand representations), provided by Eglin AFB, to determine the density of pine and hardwood trees at each observed GPS location in 2 size classes: 10.2–25.4 cm diameter at breast height (dbh) and >25.4 cm dbh. Previous work showed that hardwood densities did not affect resource selection of red-cockaded woodpeckers at Eglin AFB, but that selection was oriented to these 2 size classes of pines, with woodpeckers selecting stands with higher densities of large pines (>25.4 cm dbh) and lower densities of small pines (10.2–25.4 cm dbh) (McKellar et al. 2013).

## Behavior observations

Between January 2008 and November 2008, focal samples of behavior (Altmann 1974) were conducted for up to 4 woodpeckers in each group during sessions when observers were traveling with the birds. During focal sampling periods, observers recorded and classified behavioral interactions among individuals for use in a separate study. Focal sampling began immediately after colored leg bands had been inspected and all group members were identified. Each focal sample lasted 10 min (based on a stopwatch), and subsequent focal samples began 5 min after the conclusion of the previous sample. Initial and subsequent focal samples were initiated only when birds were actively foraging or moving. Thus, by pairing focal sampling time frames with location data from GPS units, we derived a set of locations during which birds were known to be foraging ( $n = 1048$  focal samples during 326 sessions).

## Movement analyses

We consider 2-state HMMs for analyzing the 440 observed multiple bivariate time series of step lengths and turning angles, associated

with the different territories and different sessions within territories. At any time, each observed group is assumed to move according to a correlated random walk, wherein turning angles are generated by a von Mises distribution—typically with mass centered around zero such that there is persistence in the movement direction (see Results; Figure 1)—and step lengths are drawn from a zero-inflated gamma distribution. Including a point mass (on zero) in the step length distribution is necessary in the given application because, due to the groups sometimes remaining stationary for some time, there is a considerable proportion of zero step lengths in the observations. In each session, an underlying, nonobservable Markov chain is assumed to determine the time-varying (behavioral) states of the group and the variation of the parameters of the von Mises and of the zero-inflated gamma distribution across these states. For example, the average step length will be longer in some states than in others. We further assume the parameters of the zero-inflated gamma distribution to vary across seasons. We consider the case of 2 states, which we subsequently reference as a resting state (state 1) and a foraging state (state 2), although such interpretations need to be made cautiously, because the nominal HMM states need not necessarily correspond to biologically meaningful states. In total, 2 sets of parameters are estimated for the turning angle distribution (one for each state), and 8 sets of parameters are estimated for the step length distribution (2 sets, corresponding to 2 states, in each of 4 seasons). The Markov chain generating the state sequence is assumed to be nonhomogeneous, with transition probability matrix at time  $t$  given by

$$\Gamma(c, s, t) = \begin{pmatrix} \gamma_{11}(c, s, t) & \gamma_{21}(c, s, t) \\ \gamma_{21}(c, s, t) & \gamma_{22}(c, s, t) \end{pmatrix}, \quad (1)$$

where  $\gamma_{ij}(c, s, t)$  is the territory-specific (as indicated by  $c, c = 1, \dots, 97$ ) and session-specific (as indicated by  $s, s = 1, \dots, c_s$ , where  $c_s \in \{1, \dots, 8\}$ ) conditional probability of a group of woodpeckers being in state  $j$  in the time interval  $(t, t+1)$ , given it is in state  $i$  during the interval  $(t-1, t)$ . To investigate the influence of the group size and of environmental covariates, the state transition probabilities are assumed to be functions of 5 covariates, denoted by  $x_{1,c,s}$  (number of birds in group observed in territory  $c$  and session  $s$ ),  $x_{2,c,s,d}$  (PTPA\_4\_10—pine stems/ha 10.2–25.4 cm dbh),  $x_{3,c,s,d}$  (PTPA\_10—pine stems/ha > 25.4 cm dbh),  $x_{4,c,s,d}$  (HTPA\_4\_10—hardwood stems/ha 10.2–25.4 cm dbh), and  $x_{5,c,s,d}$  (HTPA\_10—hardwood stems/ha > 25.4 cm dbh), as follows:

$$\text{logit}(\gamma_{ij}(c, s, t)) = \beta_{i,0} + \beta_{i,1}x_{1,c,s} + \beta_{i,2}x_{2,c,s,d} + \beta_{i,3}x_{3,c,s,d} + \beta_{i,4}x_{4,c,s,d} + \beta_{i,5}x_{5,c,s,d}, \quad (2)$$

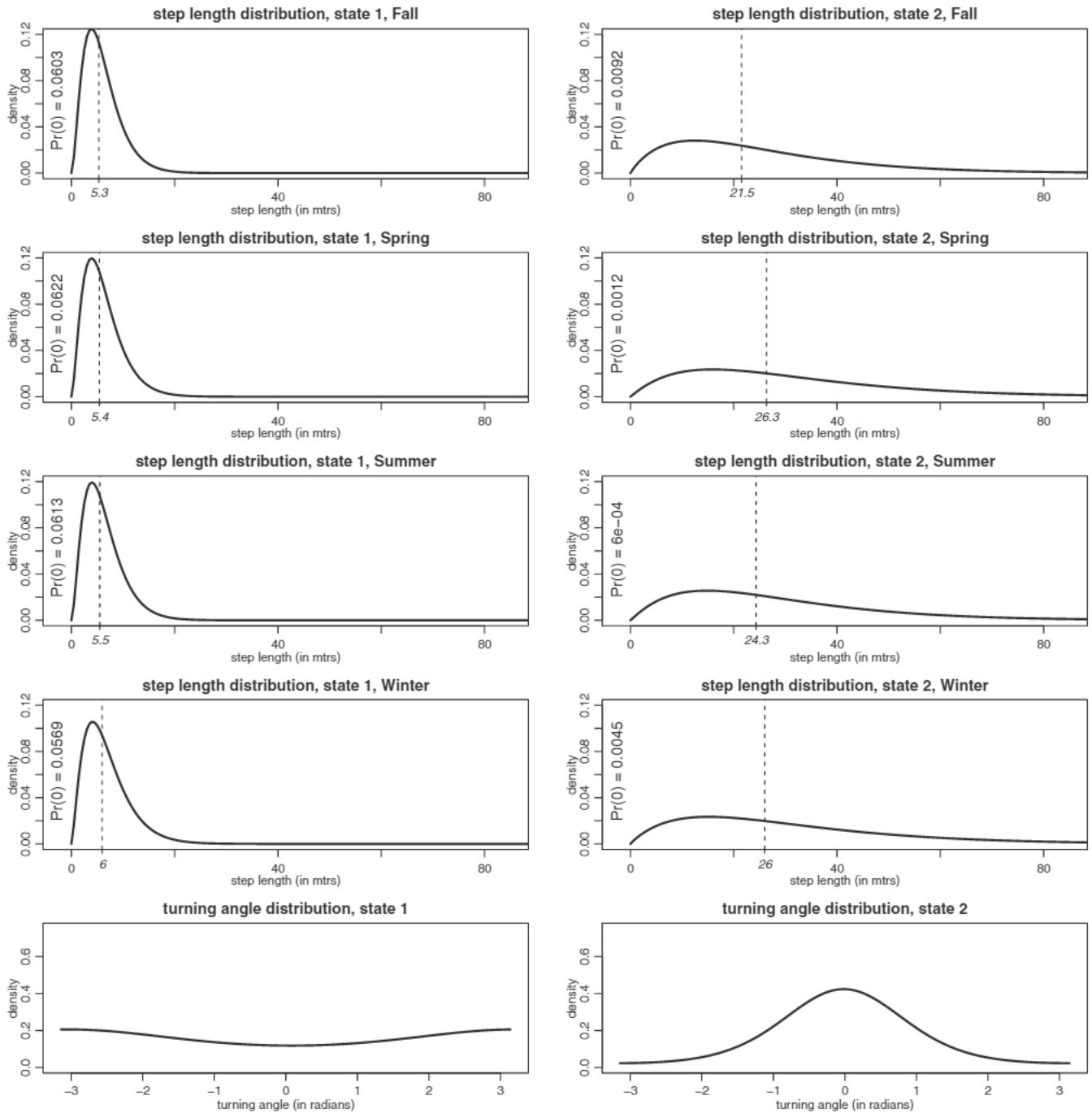
for  $i = 1, 2$ . For the single bivariate time series of step lengths and turning angles observed for the group in territory  $c$  and session  $s$ , given by  $\mathbf{z}_1^{(c,s)}, \dots, \mathbf{z}_{T_{c,s}}^{(c,s)}$ , the likelihood is a standard HMM likelihood:

$$L(c, s) = \delta \mathbf{P}_s(\mathbf{z}_1^{(c,s)}) \prod_{i=2}^{T_{c,s}} \Gamma(c, s, i) \mathbf{P}_s(\mathbf{z}_i^{(c,s)}) \mathbf{1}, \quad (3)$$

where the row vector  $\delta$  is the Markov chain initial state distribution, which we estimate,  $\mathbf{1} = (1, 1)^T$  and

$$\mathbf{P}(\mathbf{z}_i^{(c,s)}) = \begin{pmatrix} f_{1, \text{scas}}(\mathbf{z}_i^{(c,s)}) & 0 \\ 0 & f_{2, \text{scas}}(\mathbf{z}_i^{(c,s)}) \end{pmatrix}, \quad (4)$$

with  $f_{i, \text{scas}}(\mathbf{z}_i^{(c,s)})$  denoting the conditional density of the observation  $\mathbf{z}_i^{(c,s)}$ , given the current behavioral state,  $i$ , and the season in which



**Figure 1** State-specific, and in the case of step lengths also season-specific, conditional distributions of the step lengths (top 4 rows) and of the turning angles (bottom row). For the step lengths,  $Pr(0)$  is the estimated point mass on zero, the vertical dashed line indicates the median of the distribution, and the vertical dotted lines indicate the 5% and the 95% quantiles of the distribution. State 1 corresponds to a resting state, and state 2 corresponds to a foraging state.

the session took place ( $seas \in \{1,2,3,4\}$ ). Note that this conditional density simply is the product of the state-dependent density of the von Mises distribution and the state- and season-dependent density of the zero-inflated gamma distribution. The likelihood of interest is the product of the likelihoods corresponding to the different territories and sessions:

$$L = \prod_c \prod_s L(c,s). \tag{5}$$

Note that the matrix product expression for the likelihood of a single series is a consequence of applying a recursive scheme called the forward algorithm, which is an extremely powerful HMM tool and one of the main reasons for the popularity of these models (see Zucchini and MacDonald 2009 for a more detailed description). In particular, it allows for a numerical maximization of the likelihood above, which we conducted using R (R Core Team 2012). We considered several sets of initial values in the numerical maximization

in order to increase the chances of finding the global maximum of the likelihood in Equation 5. Further, we fitted all nested models that included only a subset of the 5 covariates and used the Akaike information criterion (AIC) to select among the competing models (Burnham and Anderson 2002). Note that we also considered a model that included seasonal effects in the state-switching process, but found this to be highly inferior, in terms of the AIC, to a model with seasonality only in the step length distributions.

We further investigated potential individual heterogeneity not captured by the covariates, by including a territory-specific random effect in the model for the state transition probabilities, replacing  $\beta_{i,0}$  in Equation 2 by  $\epsilon_i^{(c)}$ . Here  $\epsilon^{(1)}, \dots, \epsilon^{(K)}$ , with  $\epsilon^{(c)} = (\epsilon_1^{(c)}, \epsilon_2^{(c)})$ , are bivariate random variables taking on different values for the different territories. Usually, such random effects are assumed to be Gaussian. However, in the frequentist HMM framework, this typically leads to computational problems because each continuous-valued random effect adds an integral to the likelihood, quickly rendering a numerical maximization infeasible, especially for data sets as large as the one we consider (cf. Langrock et al. 2012). Thus, we implemented the alternative approach suggested by Maruotti and Rydén (2009), assuming a discrete support of the  $\epsilon^{(c)}$ , so that

$$\epsilon^{(c)} = (u_{k,1}, u_{k,2}) \text{ with probability } \pi_k \text{ for } k = 1, \dots, K \text{ with } \sum_{k=1}^K \pi_k = 1.$$

The value of  $K$ , giving the number of possible values of  $\epsilon^{(c)}$ , is chosen based on the AIC. The likelihood of the corresponding model, which is a mixed HMM for multiple time series (in the terminology introduced by Altman 2007; see also Schliehe-Diecks et al. 2012), has the following form:

$$L = \prod_c \sum_{k=1}^K \prod_s L_k(c, s) \pi_k, \quad (6)$$

where  $L_k(c, s)$  is defined analogously as in Equation 5 but with the  $\beta_{i,0}$  in the predictor for  $\gamma_{i,c}(c, s, t)$  replaced by  $u_{k,i}$  (for  $i = 1, 2$ ).

Note that we model the (two) random effects using a bivariate distribution, rather than considering 2 independent univariate distributions. The motivation for this is the intuitive idea that each possible outcome of the bivariate distribution and associated transition probability matrix corresponds to a particular territory type (of which there are finitely many), and the estimated probabilities of the different outcomes correspond to the proportion of the associated type of territory in the sample. A random effects model with 2 independent univariate distributions would be much harder to interpret and would also be computationally more challenging to fit as it would involve 2 summations in the likelihood (rather than just one). Code for the HMM analysis is provided in the [Supplementary Material](#).

## Behavior and movement

Behavior data from focal samples were not recorded throughout each tracking session, so they could not be directly related to HMM-based analyses of movement. However, behavior and movement data did overlap for portions of tracking sessions, so we used the data to conduct a post hoc analysis of movement states identified in the HMM analysis to test whether the predicted movement states differed when focal sampling took place (i.e., birds were known to be foraging) and when focal sampling did not take place. We identified tracking locations recorded on GPS units during each of the 10-min focal samples, and for each, we identified the most likely state under the fitted model using the Viterbi algorithm

(Zucchini and MacDonald 2009). We also identified the most likely state for all other GPS locations recorded when focal sampling did not take place, but during the same sessions in which focal observations were recorded. We assigned a value of 1 to points classified by the fitted HMM in the resting state and a value of 2 to points classified in the foraging state. We used generalized linear mixed models in the R lme4 (Bates et al. 2014) library to compare the mean of state values predicted by the fitted model for all points observed during focal sampling and when birds were actively foraging, with the mean of state values for locations outside of focal sampling, and when birds were resting for at least some of the time. Mean state values for GPS locations were fitted as the response variable with focal sampling (yes or no) as the fixed effect and group identity as a random effect.

## RESULTS

Based on the AIC, the model including all 5 covariates (group size, pine stems/ha 10.2–25.4 cm dbh, pine stems/ha > 25.4 cm dbh, hardwood stems/ha 10.2–25.4 cm dbh, and hardwood stems/ha > 25.4 cm dbh) was selected over all 32 nested models that include only a subset of the covariates in the predictor. In terms of the likelihood-ratio statistics, HTPA\_10 was the most significant covariate affecting the state transition probabilities ( $P = 0.002$ ), and PTPA\_4\_10 was the least significant ( $P = 0.091$ ) among the 5 covariates considered. Assuming the step length distribution to depend not only on the state but also on the season led to a lower AIC compared with the more basic nonseasonal model ( $\Delta\text{AIC} = 117.5$ ), stressing the need to account for seasonal effects. However, in view of the very large sample size, the magnitude of this seasonal effect is small. Including territory-specific random effects also led to substantially increased values of the log likelihood, with the AIC indicating  $K = 4$  as the optimal number of possible values to be taken on by the random effect (Table 1).

Figure 1 displays the estimated state-specific, and in the case of step lengths also season-specific, conditional distributions estimated for the step lengths and turning angles. The typical pattern of one state is exemplified with generally smaller step lengths (median 5.3–6 m/min across seasons) and the other with larger step lengths (median 21.5–26.3 m/min across seasons). Previous research has referred to these states as encamped and exploratory, respectively (Morales et al. 2004; Fryxell et al. 2008) because of the generally greater movement activity of the exploratory state and the more

**Table 1**  
Log-likelihood and AIC values obtained for the mixed HMMs with different possible numbers ( $K$ ) of values of the territory-specific random effect variable and for basic models ( $K = 1$ ) without covariates and/or without seasonality

$K$	Log-likelihood	AIC	$\Delta\text{AIC}$
1 (no seasonality, no covariates)	-182252.5	364530.9	192.1
1 (seasonality, no covariates)	-182173.6	364409.1	70.3
1 (no seasonality, all 5 covariates)	-182230.4	364506.8	168
1 (seasonality, all 5 covariates)	-182153.6	364389.3	50.5
2	-182130.4	364348.7	9.9
3	-182126.1	364346.2	7.4
4	-182119.4	364338.8	0
5	-182119.4	364344.8	6

The models with  $K > 1$  include all 5 covariates and seasonal effects. Based on the AIC values, the best model is one with seasonality and all 5 covariates and with  $K = 4$  as the number of possible values of the random effect.

sedentary nature of the encamped state. The behaviors of foraging woodpeckers on territories are relatively well documented, however, so based on this understanding, we use the terms resting and foraging hereinafter in association with the movement states. The foraging state involves a turning angle distribution with mass centered on zero and a small variance, indicating a relatively high directional persistence. In contrast, in the resting state, there are much more frequent reversals. Though differences within states among seasons were small in comparison to differences between states, the unconditional median step lengths (averaged over the states according to the equilibrium probabilities—see below) were largest in winter, followed by spring, summer, and fall.

Figure 2 displays the equilibrium probabilities of occupying the resting and the foraging state, respectively, as functions of the standardized covariates (separately for each of the covariates, in each case fixing the value of the other covariates at their respective means). These equilibrium probabilities are indicators for how likely it is to find a group of woodpeckers within a particular state under different conditions (Patterson et al. 2009). The figure also gives the expected state dwell times, that is, the expected durations of stays in the 2 different states, in equilibrium. Note that values for the resting state were always higher than those for the foraging state, indicating that for all the considered covariate combinations, a higher proportion of time was spent in the resting state and the expected dwell times for the resting state thus were longer, according to the fitted model. The following trends can be seen: 1) as group size increased, woodpeckers were more likely to be resting and less likely to be foraging, with the expected dwell time in the resting state increasing and the expected dwell time in the foraging state decreasing; 2) as the number of small pine stems/ha (10.2–25.4 cm dbh) increased, woodpeckers remained in both states for longer periods of time (i.e., fewer transitions between states); 3) as the number of large pine stems/ha (>25.4 cm dbh) increased, woodpeckers were less likely to be resting and more likely to be foraging, due to an increase in the mean duration of stays in the foraging state; 4) as the number of small hardwood stems/ha (10.2–25.4 cm dbh) increased, woodpeckers remained in both states for shorter periods of time (i.e., more transitions between states); and 5) as the number of large hardwood stems/ha (>25.4 cm dbh) increased, woodpeckers were more likely to be resting and less likely to be foraging, primarily due to an increase in the duration of stays in the resting state.

As an example, Figure 3 displays 3 sample tracks of observed locations for 3 different groups of woodpeckers as they travel across the landscape. The most likely state for each observation based on the fitted model is shown, as is the relative PTPA\_10 level (number of pine stems/ha > 25.4 cm dbh) of each forest stand used by the woodpeckers. Resting states were found more often in low PTPA\_10 stands.

The 4 different levels of the discrete random effects correspond to different state-switching dynamics. Again, fixing the value of all covariates at their respective means, the vectors of diagonal entries of the 4 estimated transition probability matrices, associated with the 4 different levels of the random effects, are (0.82, 0.82) (for  $k=1$  with proportion  $\pi_1=0.49$  in the mixture; i.e., about half of the observed groups have this state-switching pattern), (0.86, 0.74) ( $k=2$ , with  $\pi_2=0.38$ ), (0.90, 0.82) ( $k=3$ , with  $\pi_3=0.10$ ), and (0.73, 0.48) ( $k=4$ , with  $\pi_4=0.03$ ). These patterns indicate different amounts of time spent in the different states (e.g., with  $k=1$  leading to about 51% of the time spent in the resting state, in contrast to about 64% of the time spent in the resting state for

$k=2$ ) and also different levels of persistence in the states (with  $k=4$  involving the highest and  $k=3$  involving the smallest rate of switches). The considerable variation observed across the 4 different levels illustrates the need to adequately account for heterogeneity in the data analyzed here. Not only does this reduce the risk of introducing bias in the other parameters of the model but also gives a means to quantify variability across woodpecker territories.

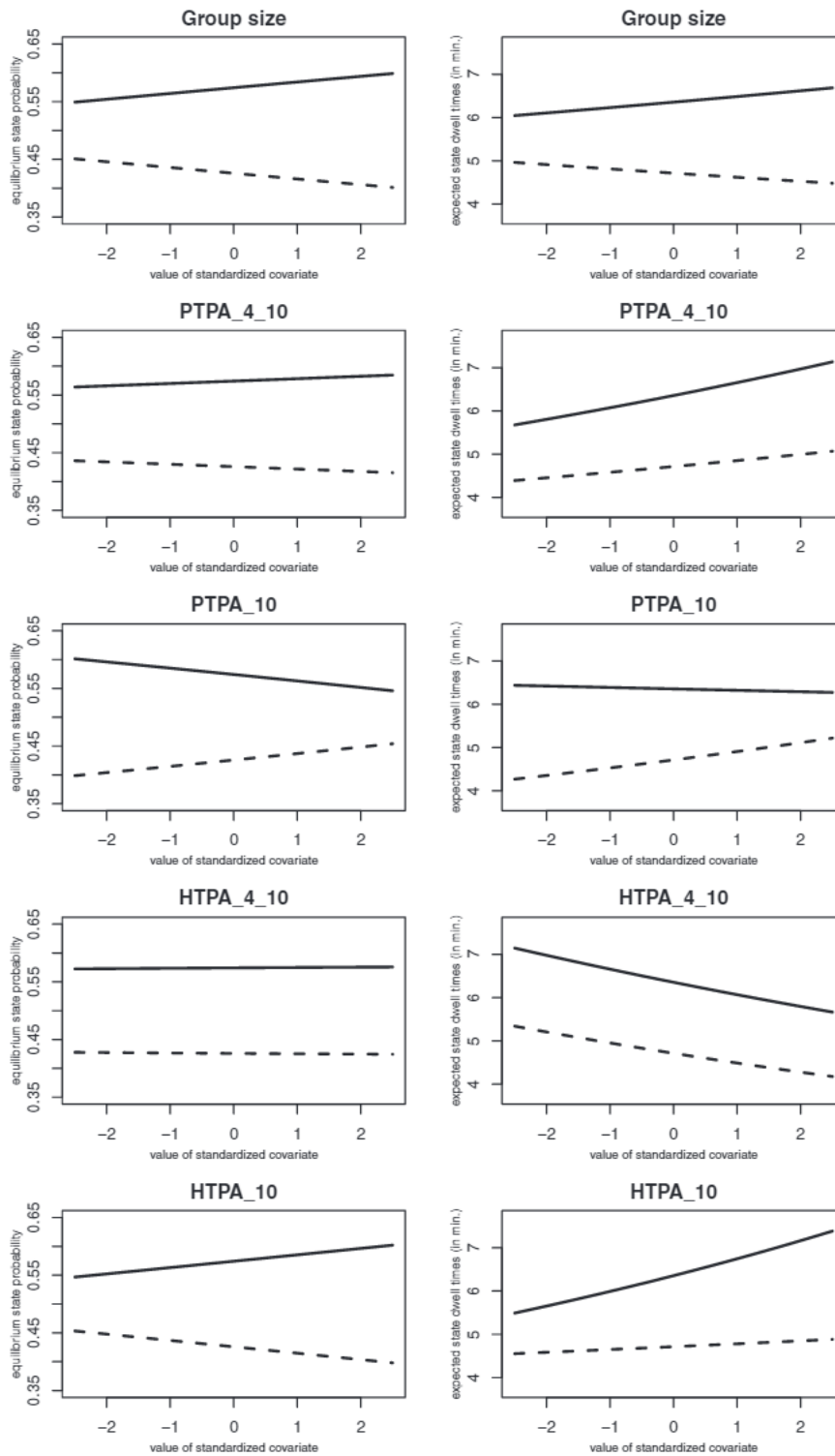
On a technical note, the computing times ranged from 0.2 h (model without random effects, seasonal effects, and covariates; 13 parameters) to 5 h (model without random effects, but including seasonal effects and all 5 covariates; 41 parameters) to 31 h (model with random effects,  $K=4$ , and including seasonal effects and all 5 covariates; 50 parameters), on an i7 CPU, at 3.4 GHz and with 8 GB RAM. For the very high number of observations considered ( $n=35\ 304$ ) and the relative complexity of the models, the computational effort is relatively low.

The mean value for movement states associated with tracking locations collected during focal sampling was higher than the mean value for movement states associated with points collected outside focal sampling periods ( $F_{1,531}=67.9$ ,  $P<0.0001$ ), where higher values are associated with a foraging state based on the fitted model. Focal sampling was only conducted when birds were actively foraging, whereas at least some of the other locations were associated with resting behaviors. Thus, post hoc results indicated that our designations of foraging and resting aligned with the behaviors of the same name.

## DISCUSSION

Considering HMMs incorporating multiple covariates, seasonal effects, and random effects, we described red-cockaded woodpecker behavioral states and their relation to environmental and social factors. We found that a model including seasonality, woodpecker group size, and 4 environmental covariates, in addition to a discrete random effect for individual territories, had substantially more support than the nested simpler models (Table 1). Our results revealed insights into the potential underlying motivation and behavior of a social species under different contexts while at the same time demonstrating the flexibility and utility of HMMs for the analysis of high-resolution animal tracking data.

We identified 2 distinct behavioral states of woodpecker groups, which we termed resting and foraging. The resting state was generally associated with shorter step lengths and higher variance in turning angles in comparison to the foraging state. Previous studies have often referred to these states as “encamped” and “exploratory” (Morales et al. 2004; Fryxell et al. 2008), respectively, but we feel that this terminology is inappropriate for describing the behavior of red-cockaded woodpeckers. Reasons for this difference can be explained by differing spatial and temporal scales of the data sets and by the ecology of the species at hand. Much previous work has focused on wide-ranging movements of large animals outfitted with GPSs, which provide hourly or daily observations on a movement scale of hundreds of meters to kilometers (Franke et al. 2004; Morales et al. 2004; Forester et al. 2007; Fryxell et al. 2008). Our data set had a very high temporal resolution (1 min), movements were on the scale of meters, and very long-distance movements were excluded (see Materials and Methods). Red-cockaded woodpecker habitat at Eglin AFB is characterized by open pine savannah, and birds can travel meters or tens of meters from tree to tree while foraging and often move quickly as a group (Hardesty et al. 1997). Thus, longer and more directed paths in our data set likely

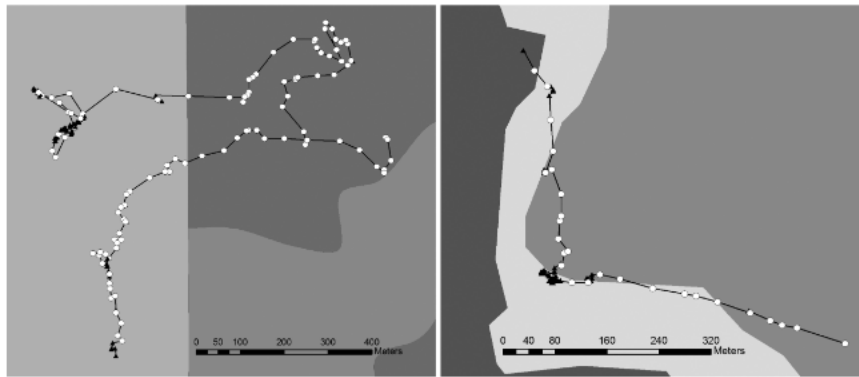


**Figure 2**

Left column: Equilibrium probabilities of occupying the different behavioral states (solid line: resting; dashed line: foraging), as functions of the covariates. Right column: expected state dwell times in minutes in equilibrium. For each covariate, the corresponding function was obtained by fixing the other covariates and also the random effect at their mean values.

correspond to foraging movements in this species. Red-cockaded woodpeckers also display long periods of resting or quiescence, which would be reflected by the smaller step lengths in the resting state. It should be noted, however, that long periods of time

can also be spent foraging at a single site, for instance, a dying tree with abundant prey, which would not accurately be categorized as foraging according to our models. Nonetheless, the foraging behavior state predicted by the fitted model was more represented during



**Figure 3**

Sample tracks taken by 3 groups of woodpeckers. Left panel shows the movements of a group of 5 woodpeckers in winter above and the movements of a group of 3 woodpeckers in fall below. Right panel shows the movements of a group of 7 woodpeckers in fall. Triangles denote the resting state, and circles denote the foraging state. Shading indicates the relative values of large pine stem density (>25.4 cm dbh) of the forest stands (darker = higher values).

focal sampling when observers noted that birds were actively foraging. We, therefore, feel that our interpretation of the 2 states was accurate in describing the general behaviors of red-cockaded woodpecker groups, even though the resting state likely includes a small amount of foraging behavior. Further, the classifications are intuitively appealing after considering their associations with habitat features known to be important in habitat selection, further discussed below.

Irrespective of potential minor caveats in the interpretation of the HMM states, the analysis led to intriguing insights into how movement metrics and patterns vary across seasons, habitat types, and territories. In particular, we found that step lengths tended to be largest in winter, though overall the differences among seasons were small (Figure 1). Larger step lengths in winter could be due to higher energetic demands and lower food availability in the season, requiring woodpeckers to travel farther and faster to obtain food (McKellar et al. 2013). An alternative is that larger step lengths in winter reflect the movement of the mixed species foraging flocks with which red-cockaded woodpeckers travel at this time of year (Schaefer et al. 2004). Behavioral state dynamics were also influenced by woodpecker group size (Figure 2). Larger groups were more likely to be in a resting state versus a foraging state. Larger groups may require more time to rest, especially because juveniles are often part of large groups. This result might also represent protection of juveniles, as the birds preferred habitat with better cover while in the resting state (see below). Alternatively, larger groups might occur on better territories or be able to locate food more efficiently, such that they required less time in the foraging state. We are aware of only one previous study that linked social factors to behavioral states in an HMM framework, in which groups of reindeer (*R. tarandus*) exhibited attraction to the group's center when in an "encamped" state (Langrock et al. 2014). Together, the results highlight the importance of considering the group environment as it relates to movement in social species.

The inclusion of 4 environmental covariates was found to substantially increase the likelihood of the fitted model. Interestingly, the influences of high densities of large pine trees and large hardwood trees (i.e., those in the >25.4 cm dbh size class) were opposite in direction. Woodpeckers were more likely to be in a foraging state when in forest stands with more large pine stems/ha and to be in a resting state when in forest stands with more large hardwood stems/ha (Figure 2). Similarly, dwell times for the foraging state were longer in forest stands with more large pine trees, and

dwell times for the resting state were longer in forest stands with more large hardwood trees. Forest stands with large pine trees are generally considered appropriate foraging habitat for red-cockaded woodpeckers, whereas stands with canopy hardwoods are considered unsuitable (Walters et al. 2002). Thus, our results suggest woodpecker groups made use of high-quality habitat for foraging movements and may have chosen low-quality habitat for resting or loafing. Lower quality habitats, like those in which birds were in a resting state, have greater structural complexity and thus may provide both thermal cover and visual cover from aerial predators (e.g., sharp-shinned hawks, *Accipiter striatus*; Bohall and Collopy 1984). Our findings with respect to foraging in high-quality habitat supported the fact that focal sampling periods were associated with the foraging state, though we lack direct behavioral observation data for resting birds.

The density of small pine and hardwood trees (i.e., trees in the 10.2–25.4 cm dbh size class) also appeared to have similar, but opposing, effects on woodpecker behavioral states. Specifically, woodpeckers showed longer dwell times for both behavioral states in forest stands with more small pine stems/ha and shorter dwell times for both behavioral states in forest stands with more small hardwood stems/ha (Figure 2). These results imply fewer state switches in the former and more state switches in the latter. Previous work identifying state switches depending on covariates include a study of gray mouse lemurs (*Microcebus murinus*) in which frequency of state switches was related to sex, body mass, and time of night (Schliehe-Diecks et al. 2012) and a study of loggerhead turtles (*Caretta caretta*) in the western Mediterranean in which state dwell time depended on body size (Eckert et al. 2008). Forest stands containing high densities of small trees are generally considered low-quality foraging habitat for red-cockaded woodpeckers (Walters et al. 2002), and so frequent state switches could indicate foraging inefficiency or avoidance of such habitat. This interpretation can explain our findings for small hardwood tree densities but leaves us with a counterintuitive result for small pines. In any case, it is clear that woodpecker movement dynamics can be influenced by environmental resources critical to their fitness.

We found that a model including 4 distinct values for the random effect associated with individual woodpecker territories was preferred over models not incorporating variation among territories. This result indicates that differences among woodpecker territories can produce differences in foraging behavior, in addition to variation in other factors such as group size. The finding that 4 values



were preferred over greater or fewer distinct values is an interesting result, which suggests that woodpecker groups displayed about 4 different “types” of movement behavior according to their territories. More research would be needed to determine which characteristics of the territory are associated with movement dynamics, as red-cockaded woodpecker territories can vary substantially in size and in the relative proportion of good-quality foraging habitat available (Hooper et al. 1982; Conner et al. 2001; McKellar et al. 2014), factors that would no doubt influence movement behavior. Importantly, we emphasize that the use of discrete random effects, rather than more common continuous random effects, constitutes an important and broadly applicable tool for accounting for heterogeneity in longitudinal data, which to date has been widely neglected in the literature on animal movement modeling. In contrast, the same idea is routinely used in the capture–recapture literature (cf. the mixture models of Pledger and Schwarz 2002 and Pledger et al. 2003).

Advances in animal-mounted tracking technology will soon provide additional complex, high-resolution, individual behavioral time series, and development of new analytical tools is crucial to fully exploit these data sets. We have shown that HMMs can be useful for describing the influence of multiple covariates on movement while accounting for heterogeneity in a social bird, an approach that can potentially be applied to a range of species. For example, the increased deployment of transmitting devices on birds, both resident and migratory, in combination with improvements in the quality of remotely sensed habitat data will provide exceptional opportunities to better understand avian habitat selection, resource use, and migratory behavior, and our study indicates that HMMs are a powerful analytical tool for this purpose.

## SUPPLEMENTARY MATERIAL

Supplementary material can be found at <http://www.behco.oxfordjournals.org/>

## FUNDING

Funding for data collection was provided by the Department of Defense, Eglin Air Force Base (430022 to J.R.W.), and analysis and manuscript development was supported by a cooperative agreement between the Engineer Research and Development Center – Construction and Engineering Research Laboratory (ERDC – CERL) and the University of Missouri (W9132T-12-2-0035). This work was sponsored by the Department of Defense; the content of the information does not necessarily reflect the position or the policy of the government, and no official endorsement should be inferred.

We very much appreciate data collection by V. Genovese, S. Goodman, K. Jones, J. Kowalsky, J. Parker, and N. Swick. D. Delaney aided with data transfer and R. Campbell, K. Gault, B. Hagedorn, S. Hassell, B. Williams, and K. Hiers provided data from Eglin AFB.

**Handling editor:** Shinichi Nakagawa

## REFERENCES

- Aitkin M. 1996. A generalized maximum likelihood analysis of overdispersion in generalized linear models. *Stat Comput.* 6:251–262.
- Altman R. 2007. Mixed hidden Markov models: an extension of the hidden Markov model to the longitudinal data setting. *J Am Stat Assoc.* 102:201–210.
- Altmann J. 1974. Observational study of behavior: sampling methods. *Behaviour.* 49:227–266.
- Anadón JD, Wiegand T, Giménez A. 2012. Individual-based movement models reveals sex-biased effects of landscape fragmentation on animal movement. *Ecosphere.* 3:art64.
- Bagniewska JM, Hart T, Harrington LA, Macdonald DW. 2013. Hidden Markov analysis describes dive patterns in semiaquatic animals. *Behav Ecol.* 24:659–667.
- Bates D, Maechler M, Bolker B, Walker S, Christensen RHB, Singmann H. 2014. lme4: linear mixed-effects models using Eigen and S4. R package version 2.15.1 [cited 2014 Jun 19]. Available from: <http://cran.r-project.org/web/packages/lme4/index.html>.
- Blackwell PG. 2003. Bayesian inference for Markov processes with diffusion and discrete components. *Biometrika.* 90:613–627.
- Bohall PG, Collopy MW. 1984. Seasonal abundance, habitat use, and perch sites of four raptor species in north-central Florida. *J Field Ornithol.* 55:181–189.
- Bradshaw DS. 1990. Habitat quality and seasonal foraging patterns of the red-cockaded woodpecker (*Picoides borealis*) in southeastern Virginia [Master's thesis]. [Williamsburg (VA)]: College of William and Mary.
- Burnham KP, Anderson DR. 2002. Model selection and multimodel inference: a practical information-theoretic approach. 2nd ed. London: Springer.
- Cappé O, Moulines E, Rydén T. 2005. Inference in hidden Markov models. New York: Springer.
- Conner RN, Rudolph DC, Walters JR. 2001. The red-cockaded woodpecker: surviving in a fire-maintained ecosystem. Austin (TX): University of Texas Press.
- Eckert SA, Moore JE, Dunn DC, van Buiten RS, Eckert KI, Halpin PN. 2008. Modeling loggerhead turtle movement in the Mediterranean: importance of body size and oceanography. *Ecol Appl.* 18:290–308.
- Forester JD, Ives AR, Turner MG, Anderson DP, Fortin D, Beyer HL, Smith DW, Boyce MS. 2007. State-space models link elk movement patterns to landscape characteristics in Yellowstone National Park. *Ecol Monogr.* 77:285–299.
- Franke A, Caelli T, Hudson RJ. 2004. Analysis of movements and behavior of caribou (*Rangifer tarandus*) using hidden Markov models. *Ecol Model.* 173:259–270.
- Fryxell JM, Hazell M, Börger L, Dalziel BD, Haydon DT, Morales JM, McIntosh T, Rosatte RC. 2008. Multiple movement modes by large herbivores at multiple spatiotemporal scales. *Proc R Soc Lond B.* 105:19114–19119.
- Garmin Ltd. 2005. Garmin owner's manual eTrex Legend personal navigator. Olathe (KS): Garmin Ltd.
- Hardesty JL, Gault KE, Percival FP. 1997. Ecological correlates of red-cockaded woodpecker (*Picoides borealis*) foraging preference, habitat use, and home range size in northwest Florida (Eglin Air Force Base) [Final Report Research Work Order 99]. Gainesville (FL): Florida Cooperative Fish and Wildlife Research Unit, University of Florida.
- Holyoak M, Casagrandi R, Nathan R, Revilla E, Spiegel O. 2008. Trends and missing parts in the study of movement ecology. *Proc Natl Acad Sci USA.* 105:19060–19065.
- Holzmann H, Munk A, Suster M, Zucchini W. 2006. Hidden Markov models for circular and linear-circular time series. *Environ Ecol Stat.* 13:325–347.
- Hooper RG, Niles IJ, Harlow RF, Wood GW. 1982. Home ranges of red-cockaded woodpeckers in coastal South Carolina. *Auk.* 99:675–682.
- Jackson JA. 1994. Red-cockaded woodpecker (*Picoides borealis*). In: Poole A, editor. The birds of North America online. Ithaca (NY): Cornell Lab of Ornithology [cited 2014 Feb 1]. Available from: <http://bna.birds.cornell.edu/bna/species/085>.
- James FC, Hess CA, Kicklighter BC, Thum RA. 2001. Ecosystem management and the niche gestalt of the red-cockaded woodpecker in longleaf pine forests. *Ecol Appl.* 11:854–870.
- James FC, Hess CA, Kufirin D. 1997. Species-centered environmental analysis: indirect effects of fire history on red-cockaded woodpeckers. *Ecol Appl.* 7:118–129.
- Jonsen ID, Basson M, Bestley S, Bravington MV, Patterson TA, Pedersen MW, Thomson R, Thygesen UH, Wotherspoon SJ. 2013. State-space models for bio-loggers: a methodological road map. *Deep Sea Res Pt II.* 88–89:34–46.
- Jonsen ID, Flemming JM, Myers RA. 2005. Robust state-space modeling of animal movement data. *Ecology.* 86:2874–2880.

- Kesler DC, Walters JR. 2012. Social composition of destination territories and matrix habitat affect red-cockaded woodpecker dispersal. *J Wildl Manage.* 76:1028–1035.
- Kesler DC, Walters JR, Kappes JJ. 2010. Social influences on dispersal and the fat-tailed dispersal distribution in red-cockaded woodpeckers. *Behav Ecol.* 21:1337–1343.
- Langrock R, Hopcraft G, Blackwell P, Goodall V, King R, Niu M, Patterson T, Pedersen M, Skarin A, Schick R. 2014. Modelling group dynamic animal movement. *Methods Ecol Evol.* 5:190–199.
- Langrock R, King R, Matthiopoulos J, Thomas L, Fortin D, Morales JM. 2012. Flexible and practical modeling of animal telemetry data: hidden Markov models and extensions. *Ecology.* 93:2336–2342.
- Lennartz MR, Hooper RG, Harlow RF. 1987. Sociality and cooperative breeding in red-cockaded woodpeckers (*Picoides borealis*). *Behav Ecol Sociobiol.* 20:77–78.
- Lusseau D. 2003. Effects of tour boats on the behavior of bottlenose dolphins: using Markov chains to model anthropogenic impacts. *Conserv Biol.* 17:1785–1793.
- Maruotti A, Rydén T. 2009. A semiparametric approach to hidden Markov models under longitudinal observations. *Stat Comput.* 19:381–393.
- McKellar AE, Kesler DC, Mitchell RJ, Delaney DK, Walters JR. 2013. Range-wide meta-analysis of red-cockaded woodpecker foraging habitat suitability. Final technical report to the Defense Technical Information Center, Fort Belvoir, Virginia. Columbia (MO): University of Missouri.
- McKellar AE, Kesler DC, Mitchell RJ, Delaney DK, Walters JR. 2014. Geographic variation in fitness and foraging habitat quality in an endangered bird. *Biol Conserv.* 175:52–64.
- Morales JM, Haydon DT, Frair J, Holsinger KE, Fryxell JM. 2004. Extracting more out of relocation data: building movement models as mixtures of random walks. *Ecology.* 85:2436–2445.
- Morales JM, Moorcroft PR, Matthiopoulos J, Frair JL, Kie JG, Powell RA, Merrill EH, Haydon DT. 2010. Building the bridge between animal movement and population dynamics. *Philos Trans R Soc Lond B Biol Sci.* 365:2289–2301.
- Nathan R, Getz WM, Revilla E, Holyoak M, Kadmon R, Saltz D, Smouse PE. 2008. A movement ecology paradigm for unifying organismal movement research. *Proc Natl Acad Sci USA.* 105:19052–19059.
- Patterson TA, Basson M, Bravington MV, Gunn JS. 2009. Classifying movement behaviour in relation to environmental conditions using hidden Markov models. *J Anim Ecol.* 78:1113–1123.
- Patterson TA, Thomas L, Wilcox C, Ovaskainen O, Matthiopoulos J. 2008. State-space models of individual animal movement. *Trends Ecol Evol.* 23:87–94.
- Pledger S, Pollock KH, Norris JL. 2003. Open capture-recapture models with heterogeneity: I. Cormack-Jolly-Seber model. *Biometrics.* 59:786–794.
- Pledger S, Schwarz CJ. 2002. Modelling heterogeneity of survival in band-recovery data using mixtures. *J Appl Stat.* 29:315–327.
- Pomerleau C, Patterson TA, Luque S, Lesage V, Heide-Jørgensen MP, Dueck LL, Ferguson SH. 2011. Bowhead whale *Balaena mysticetus* diving and movement patterns in the eastern Canadian Arctic: implications for foraging ecology. *Endanger Species Res.* 15:167–177.
- R Core Team. 2012. R: a language and environment for statistical computing. Vienna (Austria): R Foundation for Statistical Computing [cited 2014 Feb 1]. Available from: <http://www.R-project.org/>.
- Schaefer RR, Rudolph DC, Conner RN, Saenz D. 2004. Composition of mixed-species foraging flocks associated with red-cockaded woodpeckers. In: Costa R, Daniels SJ, editors. Red-cockaded woodpecker: road to recovery. Blaine (WA): Hancock House Publishers. p. 672–677.
- Schick RS, Loarie SR, Colchero E, Best BD, Boustany A, Conde DA, Halpin PN, Joppa LN, McClellan CM, Clark JS. 2008. Understanding movement data and movement processes: current and emerging directions. *Ecol Lett.* 11:1338–1350.
- Schliehe-Diecks S, Kappeler PM, Langrock R. 2012. On the application of mixed hidden Markov models to multiple behavioural time series. *Interface Focus.* 2:180–189.
- Skorupa JP, McFarlane RW. 1976. Seasonal variation in foraging territory of red-cockaded Woodpeckers. *Wilson Bull.* 88:662–665.
- Walters JR, Copeyon CK, Carter JH III. 1992. Test of the ecological basis of cooperative breeding in red-cockaded woodpeckers. *Auk.* 109:90–97.
- Walters JR, Daniels SJ, Carter JH III, Doerr PD. 2002. Defining quality of red-cockaded woodpecker foraging habitat based on habitat use and fitness. *J Wildl Manage.* 66:1064–1082.
- Walters JR, Doerr PD, Carter JH III. 1988. The cooperative breeding system of the red-cockaded woodpecker. *Ethology.* 78:275–305.
- Zucchini W, MacDonald IL. 2009. Hidden Markov models for time series: an introduction using R. London: Chapman and Hall.