RESEARCH

Incorporating Periodic Variability in Hidden Markov Models for Animal Movement

Michael Li[1*] and Benjamin M. Bolker[1,2]

Sample of title note

*Correspondence:
lim88@mcmaster.ca
[1] Department of Biology,
McMaster University, 1280 Main
St. West, L8S 4K1, Hamilton,
Ontario, Canada
Full list of author information is
available at the end of the article
[†]Equal contributor

Abstract

Clustering time-series data into discrete groups can improve prediction and provide insight into the nature of underlying, unobservable states of the system. However, temporal variation in the rates at which individuals move between groups can obscure such signals. We use finite mixture and hidden Markov models (HMMs), two standard clustering techniques, to model high-resolution hourly movement data from Florida panthers (*Puma concolor coryi*). Allowing for temporal heterogeneity in transition probabilities, a straightforward but little-used extension of the standard HMM framework, resolves some shortcomings of current models and clarifies the behavioural patterns of panthers. More generally, we point out that model misspecification (omitting important sources of variation) can lead to overfitting, and as a corollary that incorporating previously neglected structure in statistical models can allow more accurate assessment of appropriate model complexity.

Background: Text for this section.

Results: Text for this section.

Conclusion: Text for this section.

Keywords: Hidden Markov Model; Animal Movement; Temporal Autocorrelation; Temporal Heterogeneity; Florida Panther

Background

Given a sequence of animal movements, movement models aim to find a parsimonious description that can be used to understand past movements and predict future movements. Ecologists have long considered the effects of individual-level covariates (sex, age, nutritional status) and environmental covariates (habitat type, location of predators or prey) on movement [1, 2, 3]. More recently, modelers have developed *hidden Markov models* (HMMs) [4, 5, 6] — used in animal ecology under the rubric of the "multiphasic movement framework" [7] — that consider the effects of organisms' *internal* states; in particular, HMMs model animal movement as though individual animals' movement behaviour at particular times is determined by which of a discrete set of unobserved movement states (e.g. "foraging", "traveling", "resting") they currently occupy. Conditional on the state occupied by an individual, HMMs typically assume that animals follow a standard correlated walk model [8, 9].

Ever-increasing capabilities of remote sensors are making movement data available over an ever-wider range of time scales, at both higher resolution (e.g. hourly data from GPS collars vs. daily or weekly fixes for radio or VHF collars) and longer

extent (e.g. from a few days to significant fractions of a year, or longer). When analyzing such long-term data, ecologists will more often have to account for temporal variability in movement behaviour at diurnal and seasonal scales that were previously not captured in the data.

HMMs have typically been used to model movements over short time scales, where the probability of transitioning between movement states is approximately constant. Changes in latent/hidden behavioural state/mode transition probabilities based on the local environment can be accounted for incorporating environmental covariates in the HMM [10], or by more direct comparisons between inferred states and environmental conditions [7]. Schliehe-Diecks et al. [11] consider temporal trends in behavioural transitions over the time scales of a six-hour observation period, but in general ecologists have turned to other tools to describe behavioural changes over longer (diurnal, seasonal, or ontogenetic) time scales [12].

For movement behaviours that change on a fast time scale, such that movement behaviours recorded at successive observations are effectively independent, *finite mixture models* (FMMs) — which can be considered a special case of HMMs where the probability of state occupancy is independent of the previous state — can adequately describe movement [13]. When movement varies over long time scales (relative to the time between observations) with little short-term persistence or correlation, movement could be well represented by FMMs where the occupancy probabilities change deterministically over time. Thus FMMs and HMMs, with or without temporal variation in the occupancy or transition probabilities, form a useful family of models for capturing changes in movement behaviour over a range of time scales.

Our primary goal in this paper is to introduce the use of HMMs with temporally varying transition probabilities – in particular, transition probabilities that follow a diurnal cycle – for modeling animal movement recorded over long time scales. In addition to simulation-based examples, we also re-analyze data from van de Kerk et al.[14], who used temporally homogeneous hidden semi-Markov models (HSMMs: an extension of HMMs that allow flexible modelling of the distribution of *dwell times*, the lengths of consecutive occupancy of a behavioural state) to describe the movement and putative underlying behavioural states of Florida panthers (*Puma concolor coryi*). van de Kerk et al.[14] found that the best-fitting HSMMs incorporated a surprisingly large number of hidden behavioural states (as many as six for individuals with a large amount of available data); for reasons of computational practicality and biological interpretability, they restricted their detailed analysis to models with only three underlying states. In contrast, most studies using HMM have chosen the number of underlying states *a priori*, typically using either two [11, 15, 6, 7], or three states [16, 17, 18]. In contrast, [19] evaluated models with up to 10 states, but like [14] they chose to consider only models with three states. As van de Kerk et al. [14] comment, and as we discuss further below, behavioural repertoires with more than three distinct states are difficult to interpret — one possible reason that other authors have not adopted van de Kerk et al.'s model-based approach to identifying the number of latent states.

Our second goal, therefore, is to explore whether van de Kerk et al.'s results on optimal model complexity might be driven at least in part by structural problems with

their statistical model, i.e. the assumption of temporally homogeneous behaviour. For large data sets, information-theoretic model selection methods will typically choose complex, highly parameterized models; when there is only one way in which models can become more complex (e.g. by increasing the number of latent states), complexity that is present in the data but not accounted for in the model (e.g. spatial or temporal heterogeneity) can be misidentified as other forms of complexity. We predict that increasing volumes of data will increasingly lead researchers who are accustomed to fitting small models to sparse data into such traps. We examine whether allowing for diurnal variation in the Florida panther data leads to selection of models with smaller numbers of latent states; we also fit models to simulated data with varying numbers of latent states and degrees of temporal heterogeneity to test our conjecture that heterogeneity can be misidentified as behavioural complexity.

Methods

Data and previous analyses

GPS collars were fitted to 18 Florida panthers in 2005-2012 by Florida Fish and Wildlife and Conservation Commission staff using trained hounds and houndsmen. Of these animals, 13 had sufficient data to be used by van de Kerk et al.[14]. Here we focus on the four cats with the most data (all with approximately 10,000-15,000 observations: see Table 1 in Supplementary Material), in part because our goal is to understand the issues that arise when simple models are fitted to large data sets, and in part because the general trend in telemetry studies is toward larger data sets. As is typical in studies of animal movement, we took first differences of the data by decomposing contiguous sequences of hourly GPS coordinates into successive step lengths (in meters) and turning angles (in radians) [9, 14].

van de Kerk et al.[14] used hidden semi-Markov models (HSMM), an extension of HMM that permits explicit modelling of dwell times [6], considering both Poisson and negative binomial distributions for dwell times. As shown by van de Kerk et al.[14] (Figure S3b, top row, middle panel), the estimated shape parameter of the negative binomial dwell time distribution was typically close to 1 ($\approx 0.4 - 1.6$; confidence intervals were not given), implying that a geometric distribution (i.e., negative binomial with shape=1) might be adequate. In turn, this suggests that we might not lose much accuracy by reverting to a simpler HMM framework, which corresponds to making precisely this assumption.

van de Kerk et al.[14] considered time-homogeneous models with a variety of candidate distributions — log-Normal, Gamma, and Weibull distributions for step lengths and von Mises and wrapped Cauchy distributions for the turning angle — concluding on the basis of the Akaike information criterion (AIC) that Weibull step length and wrapped Cauchy turning angle distributions were best. Since our analysis aims for simplicity and qualitative conclusions rather than for picking the very best predictive model, we focus on models that treat each step as a univariate, log-Normally distributed observation, glossing over both the differences in shape between the three candidate step-length distributions and the effects of considering multivariate (i.e., step length plus turning angle) observations. However, we do briefly compare log-Normal and Weibull step-length distributions, with and without a von Mises-distributed turning angle included in the model (Figure **??**). (Note that

most movement analyses, including van de Kerk et al. [14], are only partially multi-variate, treating step length and turning angle at a particular time as multivariate observations for the purpose of HMM analysis but neglecting possible correlations between the two measures.)

van de Kerk et al.[14] used the Bayesian (Schwarz) information criterion (BIC) to test the relative penalized goodness of fit for models ranging from 2 to 6 latent states. In general, BIC values decreased as the number of states increased from three to six states, suggesting that the six-state model was favoured statistically; however, the authors used three-state models in most of their analyses for ease of biological interpretation. We follow van de Kerk et al.[14] in using BIC-optimality (i.e., minimum BIC across a family of models) as the criterion for identifying the best model, because we are interested in explaining the data generation process by identifying the "true" number of underlying movement states. Using BIC also simplifies evaluation of model selection procedures; it is easier to test whether our model selection procedure has selected the model used to simulate the data, rather than testing whether it has selected the model with the minimal Kullback-Leibler distance [20]. We recognize that ecologists will often be interested in maximizing predictive accuracy rather than selecting a true model, and that as usual in ecological systems the true model will be far more complicated than any candidate model [21]; we believe that the qualitative conclusions stated here for BIC-optimality will carry over to analyses using AIC instead.

Model description

In a HMM, the joint likelihood of *emissions* (i.e., direct observations) $\mathbf{Y} = \mathbf{y}_1, ..., \mathbf{y}_T$ and a hidden state sequence $\mathbf{Z}, z_t \in \{1, ..., n\}, t = 1, ..., T$, given model parameters θ and covariates $\mathbf{X}_{1:T} = \mathbf{x}_1, ..., \mathbf{x}_T$, can be written as:

$$
\begin{aligned}
P(\mathbf{Y}_{1:T}, \mathbf{Z}_{1:T}|\theta, \mathbf{X}_{1:T}) \;=\; & P(z_1 \mid \mathbf{x}_1)P(\mathbf{y}_1|z_1, \mathbf{x}_1) \cdot \\
& \prod_{k=2}^{T} P(z_k|z_{k-1}, \mathbf{x}_k)P(\mathbf{y}_k|z_k, \mathbf{x}_k) \quad (1)
\end{aligned}
$$

The emissions $\mathbf{y}_i$ are boldfaced to denote that we may have a vector of observations at each time point (e.g., step length and turning angle). The model contains three distinct components:

Initial probability $P(z_1 = i|\mathbf{x}_1)P(\mathbf{y}_1|z_1, \mathbf{x}_1)$: the probability of state i at time $t = 1$ where the covariate is $\mathbf{x}_1$, times the vector of observations $\mathbf{y}_1$ conditioned on covariates $\mathbf{x}_1$ and state z_1.

Transition probability $P(z_k = j|z_{k-1} = i, \mathbf{x}_k)$: the probability of a transition from state i at time $t = k - 1$ to state j with covariate $\mathbf{x}_k$ at time $t = k$.

Emission probability $P(\mathbf{y}_k|z_k, \mathbf{x}_k)$: a vector of observations $\mathbf{y}_k$ conditioned on covariates $\mathbf{x}_k$ at state z_k at time $t = k$.

Eq. 1 gives the likelihood of the observed sequence given (conditional on) a particular hidden sequence. In order to calculate the overall, unconditional (or marginal) likelihood of the observed sequence, we need to average over all possible hidden sequences. There are several efficient algorithms for computing the marginal likelihood and numerically estimating parameters [22]; we used those implemented in the `depmixS4` package for R [23, 24].

For any n-state HMM, we need to define a $n \times n$ matrix that specifies the probabilities π_{ij} of being in movement states j at time $t+1$ given that the individual is in state i. The FMM is a special case of HMM where the probabilities of *entering* a given state are identical across all states — i.e., the probability of occupying a state at the next time step is independent of the current state occupancy. It can be modelled in the HMM framework by setting the transition probabilities $\pi_{ij} = \pi_{i*}$.

In any case, the transition matrix π_{ij} must respect the constraints that (1) all probabilities are between 0 and 1 and (2) transition probabilities out of a given state sum to 1. As is standard for HMMs with covariates [23], we define this multinomial logistic model in terms of a linear predictor η_{ij}, where η_{i1} is set to 1 without loss of generality (i.e. we have only $n \times (n-1)$ distinct parameters; we index j from 2 to n for notational clarity):

$$
\begin{aligned}
\pi_{ij} &= \exp(\eta_{ij}(t))/\left(1 + \sum_{j=2}^{n} \exp(\eta_{ij}(t))\right), \text{ for } j = 2, ..., n \qquad (2)\\
\pi_{i1} &= 1 - \sum_{j=2}^{n} \pi_{ij}
\end{aligned}
$$

Sub-heading for section

Text for this sub-heading . . .

Sub-sub heading for section

Text for this sub-sub-heading . . .

Sub-sub-sub heading for section   Text for this sub-sub-sub-heading . . . In this section we examine the growth rate of the mean of Z_0, Z_1 and Z_2. In addition, we examine a common modeling assumption and note the importance of considering the tails of the extinction time T_x in studies of escape dynamics. We will first consider the expected resistant population at vT_x for some $v > 0$, (and temporarily assume $\alpha = 0$)

$$
E\big[Z_1(vT_x)\big] = E\bigg[\mu T_x \int_0^{v \wedge 1} Z_0(uT_x) \exp\big(\lambda_1 T_x (v - u)\big)\, du\bigg].
$$

If we assume that sensitive cells follow a deterministic decay $Z_0(t) = x e^{\lambda_0 t}$ and approximate their extinction time as $T_x \approx -\frac{1}{\lambda_0} \log x$, then we can heuristically estimate the expected value as

$$
\begin{aligned}
E\big[Z_1(vT_x)\big] &= \frac{\mu}{r} \log x \int_0^{v \wedge 1} x^{1-u} x^{(\lambda_1/r)(v-u)}\, du\\
&= \frac{\mu}{r} x^{1 - \lambda_1/\lambda_0 v} \log x \int_0^{v \wedge 1} x^{-u(1 + \lambda_1/r)}\, du\\
&= \frac{\mu}{\lambda_1 - \lambda_0} x^{1 + \lambda_1/rv}\left(1 - \exp\bigg[-(v \wedge 1)\bigg(1 + \frac{\lambda_1}{r}\bigg) \log x\bigg]\right) \quad (3)
\end{aligned}
$$

Thus we observe that this expected value is finite for all $v > 0$ (also see [25, 26, 27, 28, 29]).

Competing interests

The authors declare that they have no competing interests.

Author's contributions

Text for this section . . .

Author details

[1] Department of Biology, McMaster University, 1280 Main St. West, L8S 4K1, Hamilton, Ontario, Canada. [2] Department of Mathematics and Statistics, McMaster University, 1280 Main St. West, L8S 4K1, Hamilton, Ontario, Canada.

References

1. Patterson, T.A., Thomas, L., Wilcox, C., Ovaskainen, O., Matthiopoulos, J.: State–space models of individual animal movement. Trends in Ecology & Evolution **23**(2), 87–94 (2008)
2. McKenzie, H.W., Lewis, M.A., Merrill, E.H.: First passage time analysis of animal movement and insights into the functional response. Bulletin of Mathematical Biology **71**(1), 107–129 (2009)
3. Pal, S., Ghosh, B., Roy, S.: Dispersal behaviour of free-ranging dogs (*Canis familiaris*) in relation to age, sex, season and dispersal distance. Applied Animal Behaviour Science **61**(2), 123–132 (1998)
4. Firle, S., Bommarco, R., Ekbom, B., Natiello, M.: The influence of movement and resting behavior on the range of three carabid beetles. Ecology **79**(6), 2113–2122 (1998). doi:10.1890/0012-9658(1998)079[2113:TIOMAR]2.0.CO;2. Accessed 2015-04-14
5. Nathan, R., Getz, W.M., Revilla, E., Holyoak, M., Kadmon, R., Saltz, D., Smouse, P.E.: A movement ecology paradigm for unifying organismal movement research. Proceedings of the National Academy of Sciences **105**(49), 19052–19059 (2008). doi:10.1073/pnas.0800375105. Accessed 2015-04-29
6. Langrock, R., King, R., Matthiopoulos, J., Thomas, L., Fortin, D., Morales, J.M.: Flexible and practical modeling of animal telemetry data: hidden Markov models and extensions. Ecology **93**(11), 2336–2342 (2012). doi:10.1890/11-2241.1. Accessed 2013-10-24
7. Fryxell, J.M., Hazell, M., Börger, L., Dalziel, B.D., Haydon, D.T., Morales, J.M., McIntosh, T., Rosatte, R.C.: Multiple movement modes by large herbivores at multiple spatiotemporal scales. Proceedings of the National Academy of Sciences **105**(49), 19114–19119 (2008). doi:10.1073/pnas.0801737105. Accessed 2013-04-09
8. Okubo, A.: Diffusion and Ecological Problems: Mathematical Models (1980)
9. Turchin, P.: Quantitative Analysis of Movement: Measuring and Modeling Population Redistribution in Animals and Plants. Sinauer Associates, Sunderland, MA, USA (1998)
10. Patterson, T.A., Basson, M., Bravington, M.V., Gunn, J.S.: Classifying movement behaviour in relation to environmental conditions using hidden Markov models. Journal of Animal Ecology **78**(6), 1113–1123 (2009)
11. Schliehe-Diecks, S., Kappeler, P.M., Langrock, R.: On the application of mixed hidden Markov models to multiple behavioural time series. Interface Focus **2**(2), 180–189 (2012). doi:10.1098/rsfs.2011.0077. Accessed 2014-05-02
12. Gurarie, E., Andrews, R.D., Laidre, K.L.: A novel method for identifying behavioural changes in animal movement data. Ecology Letters **12**(5), 395–408 (2009)
13. Tracey, J.A., Zhu, J., Boydston, E., Lyren, L., Fisher, R.N., Crooks, K.R.: Mapping behavioral landscapes for animal movement: a finite mixture modeling approach. Ecological Applications **23**(3), 654–669 (2012). doi:10.1890/12-0687.1. Accessed 2015-04-20
14. van de Kerk, M., Onorato, D.P., Criffield, M.A., Bolker, B.M., Augustine, B.C., McKinley, S.A., Oli, M.K.: Hidden semi-Markov models reveal multiphasic movement of the endangered Florida panther. Journal of Animal Ecology **84**(2), 576–585 (2015)
15. McKellar, A.E., Langrock, R., Walters, J.R., Kesler, D.C.: Using mixed hidden Markov models to examine behavioral states in a cooperatively breeding bird. Behavioral Ecology, 171 (2014). doi:10.1093/beheco/aru171. Accessed 2015-04-21
16. Dean, B., Freeman, R., Kirk, H., Leonard, K., Phillips, R.A., Perrins, C.M., Guilford, T.: Behavioural mapping of a pelagic seabird: combining multiple sensors and a hidden Markov model reveals the distribution of at-sea behaviour. Journal of the Royal Society Interface, 20120570 (2012)
17. Morales, J.M., Haydon, D.T., Frair, J., Holsinger, K.E., Fryxell, J.M.: Extracting more out of relocation data: building movement models as mixtures of random walks. Ecology **85**(9), 2436–2445 (2004)
18. Franke, A., Caelli, T., Kuzyk, G., Hudson, R.J.: Prediction of wolf (*Canis lupus*) kill-sites using hidden Markov models. Ecological Modelling **197**(1–2), 237–246 (2006). doi:10.1016/j.ecolmodel.2006.02.043. Accessed 2015-04-29
19. Dean, B., Freeman, R., Kirk, H., Leonard, K., Phillips, R.A., Perrins, C.M., Guilford, T.: Behavioural mapping of a pelagic seabird: combining multiple sensors and a hidden Markov model reveals the distribution of at-sea behaviour. Journal of The Royal Society Interface **10**(78), 20120570 (2013). doi:10.1098/rsif.2012.0570. Accessed 2016-06-07
20. Richards, S.A.: Testing ecological theory using the information-theoretic approach: examples and cautionary results. Ecology **86**(10), 2805–2814 (2005)
21. Burnham, K.P., Anderson, D.R.: Model Selection and Inference: A Practical Information-Theoretic Approach. Springer, New York (1998)
22. Zucchini, W., MacDonald, I.L.: Hidden Markov Models for Time Series: An Introduction Using R. CRC Press, ??? (2009)
23. Visser, I., Speekenbrink, M.: depmixS4: An R package for hidden Markov models. Journal of Statistical Software **36**(7), 1–21 (2010)

24. R Core Team: R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria (2015). R Foundation for Statistical Computing. https://www.R-project.org/
25. Koonin, E.V., Altschul, S.F., Bork, P.: Brca1 protein products: functional motifs. Nat Genet **13**, 266–267 (1996)
26. Kharitonov, S.A., Barnes, P.J.: Clinical Aspects of Exhaled Nitric Oxide. in press
27. Zvaifler, N.J., Burger, J.A., Marinova-Mutafchieva, L., Taylor, P., Maini, R.N.: Mesenchymal cells, stromal derived factor-1 and rheumatoid arthritis [abstract]. Arthritis Rheum **42**, 250 (1999)
28. Jones, X.: Zeolites and synthetic mechanisms. In: Smith, Y. (ed.) Proceedings of the First National Conference on Porous Sieves: 27-30 June 1996; Baltimore, pp. 16–27 (1996). Stoneham: Butterworth-Heinemann
29. Margulis, L.: Origin of Eukaryotic Cells. Yale University Press, New Haven (1970)

Figures

Figure 1 Sample figure title. A short description of the figure content should go here.

Figure 2 Sample figure title. Figure legend text.

Tables

Table 1 Sample table title. This is where the description of the table should go.

	B1	B2	B3
A1	0.1	0.2	0.3
A2	...	..	.
A3	..	.	.

Additional Files
Additional file 1 — Sample additional file title
Additional file descriptions text (including details of how to view the file, if it is in a non-standard format or the file extension). This might refer to a multi-page table or a figure.

Additional file 2 — Sample additional file title
Additional file descriptions text.