Incorporating periodic variability in hidden Markov models for animal movement

Michael Li and Benjamin M. Bolker

February 29, 2016

Abstract

Clustering time-series data into discrete groups can improve prediction as well as providing insight into the nature of underlying, unobservable states of the system. However, temporal variation in the probability that individuals occupy or move between groups can obscure such signals. We use finite mixture and hidden Markov models, two standard clustering techniques, to model high-resolution hourly movement data from Florida panthers. Allowing for temporal heterogeneity in transition probabilities, a straightforward but rarely explored model extension, resolves some shortcomings of current modeling frameworks and clarifies the behavioural patterns of panthers. More generally, incorporating previously neglected structure in statistical models can allow researchers to more accurately identify models of appropriate complexity.

Keywords: hidden Markov model, animal movement, temporal autocorrelation, temporal heterogeneity, Florida panther

1 Introduction

Given a sequence of animal movements observed directly or via telemetry, movement models aim to find a parsimonious description that can be used to understand past movements as well as 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 (Patterson et al., 2008; McKenzie et al., 2009; Pal et al., 1998). More recently, modelers have developed *hidden Markov models* (HMMs) (Firle et al., 1998; Nathan et al., 2008; Langrock et al., 2012) — used in animal ecology under the rubric of the "multiphasic movement framework" (Fryxell et al., 2008) — that consider the effects of organisms' *internal* states; in particular, HMMs model animal movement as though individual animals' movement behaviour 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 (Okubo, 1980; Turchin, 1998).

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 remote-sensing 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 (Patterson et al., 2009), or by more *ad hoc* comparisons between inferred states and environmental conditions (Fryxell et al., 2008). Schliehe-Diecks et al. (2012) 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.

For movement behaviours that change on a fast time scale, such that movement behaviours recorded at successive observations are effectively independent, *finite mixture models* (FMM) — 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 (Tracey et al., 2012). 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.

do we want a figure showing this?

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. (2015), 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. (2015) found that the best-fitting HSMMs incorporated a surprisingly large number of hidden behavioural states (as many as 6 for individuals with a large amount of available data); for practical reasons, they restricted their detailed analysis to models with only 3 underlying states. In contrast, most studies using HMM have chosen the number of underlying states

a priori, typically using either two (Schliehe-Diecks et al., 2012; McKellar et al., 2014; Langrock et al., 2012; Fryxell et al., 2008), or three states (Dean et al., 2012; Morales et al., 2004; Franke et al., 2006; van de Kerk et al., 2015). As van de Kerk et al. (2015) comment, 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 might be driven at least in part by structural problems with the HMM, i.e. the assumption of temporally homogeneous behaviour. When large data sets are available, 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.

4

2 Methods

2.1 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. (2015). 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) (Turchin, 1998; van de Kerk et al., 2015).

not sure quite where we should use citeauthor vs. cite ...

van de Kerk et al. (2015) used hidden semi-Markov models (HSMM), an extension of HMM that permits explicit modelling of dwell times (Langrock et al., 2012), considering both Poisson and negative binomial distributions for dwell times. As shown by van de Kerk et al. (2015) (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 calculated), 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 makes precisely this assumption.

van de Kerk et al. (2015) 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

could say here (or maybe better in the conclusions) that revisiting HSMM in this context would be worthwhile future work. ML: conclusion is good, I still need to add viterbi paragraph

5

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 (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 1).(Note that most movement analyses, including van de Kerk et al. (2015), are only partially multivariate, treating step length and turning angle at a particular time as independent observations while neglecting possible correlations between them.)

have we run Gamma as well?

W/VM/HMM (blue/solid/left panel) should correspond to vdK's results, but that has ΔBIC = 0 for 5 states rather than 6. Is that right? Is that HMM/HSMM diff?

van de Kerk et al. (2015) 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 3 to 6 states, suggesting that the 6-state model was favoured statistically; however, the authors used 3-state models in most of their analyses for ease of biological interpretation. We follow van de Kerk et al. (2015) 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 (Richards, 2005). 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 (Burnham and Anderson, 1998); we suspect that the qualitative conclusions stated here for BIC-optimality will carry over to analyses using AIC instead.

2.2 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 $\boldsymbol{\theta}$ and covariates $\mathbf{X}_{1:T} = \mathbf{x}_1, ..., \mathbf{x}_T$, can be written as:

$$P(\mathbf{Y}_{1:T}, \mathbf{Z}_{1:T} | \boldsymbol{\theta}, \mathbf{X}_{1:T}) = P(z_1 \mid \mathbf{x}_1) P(\mathbf{y}_1 | z_1, \mathbf{x}_1) \times \\ \prod_{k=2}^{T} P(z_k | z_{k-1}, \mathbf{x}_k) P(\mathbf{y}_k | z_k, \mathbf{x}_k) \tag{1}$$

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 $\mathbf{x}_1$; the vector of observation densities $\mathbf{y}_1$ conditions 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 observation density $\mathbf{y}_k$ condition 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 (Zucchini and

7

MacDonald, 2009); we use those implemented in the `depmixS4` package for R (Visser and Speekenbrink, 2010; R Core Team, 2015).

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 $\pi_{ij} = \pi_{i*}$.

> BMB: is this right?

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 (Visser and Speekenbrink, 2010), 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 nevertheless index j from 2 to n for notational clarity):

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

(2)

We considered three different transition models for diurnal variation in behaviour, incorporating hour-of-day as a covariate following the general approach of Morales et al. (2004) of incorporating covariate dependence in the transition matrix.

Multiple block transition Here we assume piecewise-constant transition probabilities. The transition probability π_{ij} is a function of time (hour of day),

8

where it is assigned to one of M different time blocks:

$$\eta_{ij}(t) = \sum_{m=1}^{M} a_{ijm}\delta_{m=t}$$

where a_{ijm} are parameters, and $\delta_{m=t}$ is a Kronecker delta ($\delta_{m=t} = 1$ for the time block at the corresponding time t, and 0 otherwise).

Quadratic transition model We assume the elements of the linear predictor are quadratic functions of hour. The quadratic model is not diurnally continuous, i.e. there is no constraint that forces $\eta_{ij}(0) = \eta_{ij}(24)$; imposing a diurnal continuity constraint would collapse the model to a constant.

$$\eta_{ij}(t) = b_{ij1} + b_{ij2}\left(\frac{t}{24}\right) + b_{ij3}\left(\frac{t}{24}\right)^2$$

Sinusoidal transition model A sinusoidal model with a period of 24 hours is identical in complexity to the quadratic model, but automatically satisfies the diurnal continuity constraint.

$$\eta_{ij}(t) = b_{ij1} + b_{ij2}\cos\left(\frac{2\pi t}{24}\right) + b_{ij3}\sin\left(\frac{2\pi t}{24}\right)$$

Model complexity and number of parameters increase as the number of latent states increase, FMM to HMM, and lastly, FMM and HMM incorporating temporal heterogeneity.

2.3 Model evaluation

We used the `depmixS4` package to fit covariate-dependent transition HMMs, simulate states and step lengths using the estimated parameters, and estimate

Discuss model complexity/parameter-number computations here, or with each model. Also discuss the "hourly" model here. Make the table into a figure!

Give more details about parameters, structure of simulations, etc.

the most likely states with the Viterbi algorithm.

We used three approaches to assess the fit of both time-homogeneous and time-inhomogeneous HMMs with 3 to 6 states to step-length data from the four of the thirteen Florida panthers with the most data ($>$ observations). (1) BIC relative to the optimal-BIC model within each type of transition complexity(ΔBIC = BIC - min(BIC)). (2) Computed average step-length by hour of day for the observed data and for data simulated from the models. (3) Computed temporal autocorrelation function (ACF) for the observed data and for data simulated from the models. For the second and third metrics (hourly average and ACF), we simulated data from the models because the metrics are hard to compute from first principles for HMMs with temporally varying transition probabilities.

<div style="float:right; border:1px solid gray; padding:4px; font-size:small;">
describe the Viterbi algorithm here, and the autocorrelation function computation - everything that appears in the results should be described explicitly here....ML: at the bottom
</div>

<div style="float:right; border:1px solid gray; padding:4px; font-size:small;">
BMB: not sure what that means. Clarify?
</div>

<div style="float:right; border:1px solid gray; padding:4px; font-size:small;">
give at least a hint of why/what each of these three approaches (ΔBIC, hourly averages, ACF) will show. ML: What do you think?
</div>

3   Results

Figure 1 shows the BIC-optimal number of states for time homogenous models is consistent with van de Kerk et al.'s results. Furthermore, BIC-optimal number of states does not depend on the model for both time-homogenous and time-inhomogenous within their respective framework, but time-inhomogenous models have the BIC-optimal number of states overall.

As a complement, we also added FMM, FMM with sinusoidal temporal heterogeneities, and HMM with hourly transitions to compare the temporal effects in goodness of fit. Looking at Figure 2, models with temporal heterogenity is better than models without in both FMM and HMM frameworks, but time homogenous HMM models is better than FMM with sinusoidal temporal heterogeneity. Temporal heterogenity is good with good models such as the quadratic
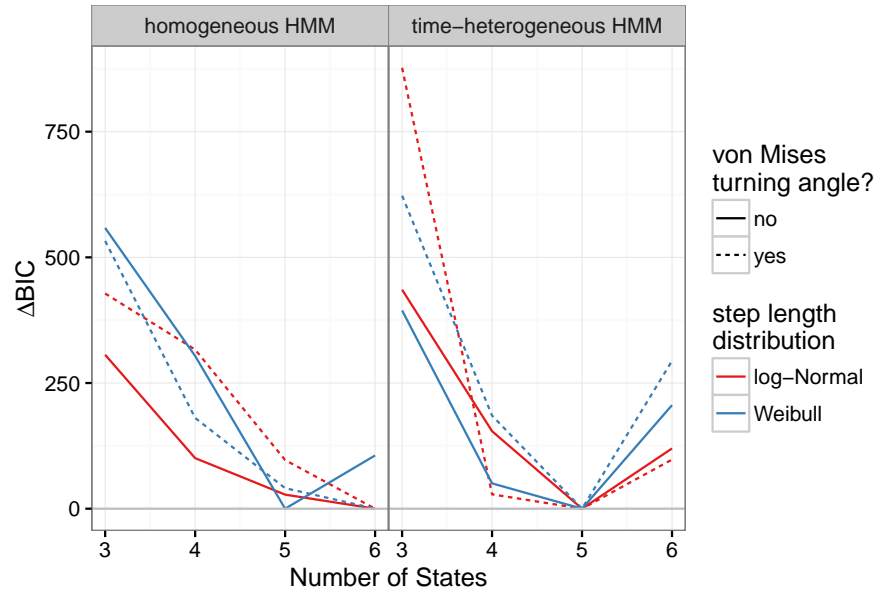
Figure 1: Change in BIC across models, cat #130. ΔBIC is scaled individually per model, so that ΔBIC = 0 represents the BIC-optimal number of states for a particular model.
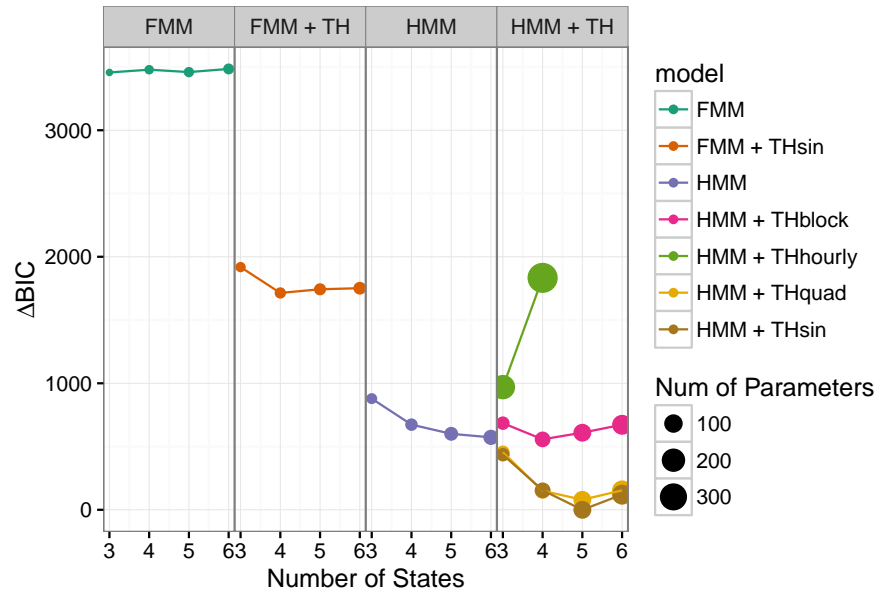

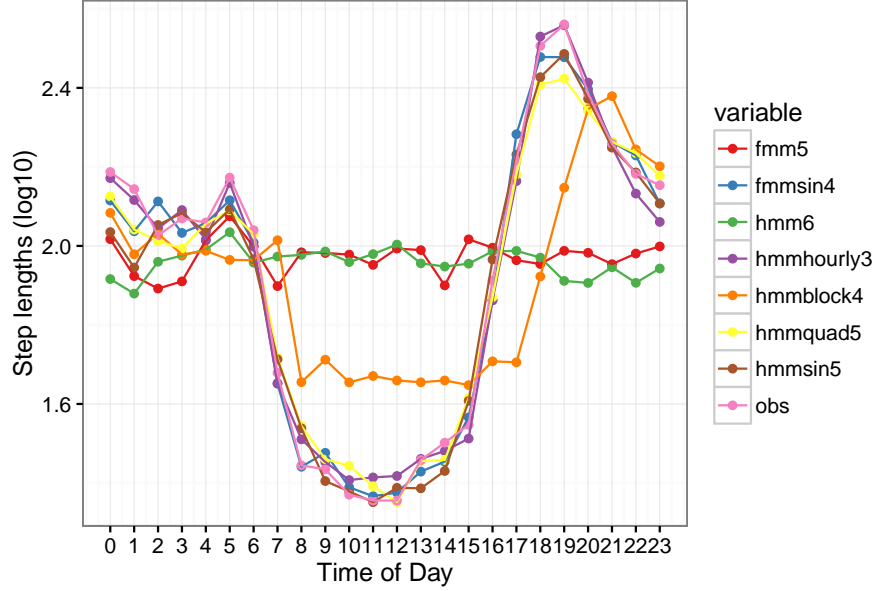Figure 2: Overall Adjusted BIC Comparisons between different HMM models

Figure 3: Average step length by time of day

and sinuosidal. Multiple block models is good for cats with distinguishable diurnal patterns and hourly model is too complicated.

Figure 3 shows a clear diurnal pattern from the average hourly step lengths from the observed data. Models that assume temporal homogeneity (FMM and HMM) fail to capture the diurnal activity cycle. Models incorporating temporal heterogeneity can capture the observed patterns without the Viterbi algorithm.

The observed data has strong 24 hour lag autocorrelation (correlation with itself) which suggest temporal heterogeneity. HMM without temporal heterogeneity can capture short-term autocorrelation (without Viterbi algorithm) but is unable to capture lags beyond 7. FMM with temporal heterogeneity can capture lags beyond 12, but not short-term autocorrelations. HMM with temporal heterogeneity can fully capture the autocorrelation pattern of the observed data whereas the FMM completely missed the autocorrelations.

The biggest issue with multiple states is identifying them biologically. We
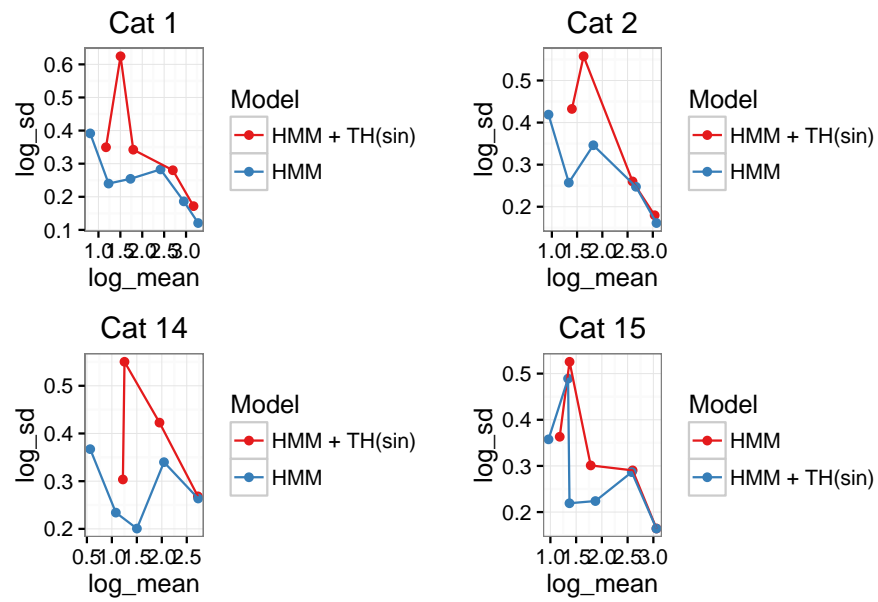
12
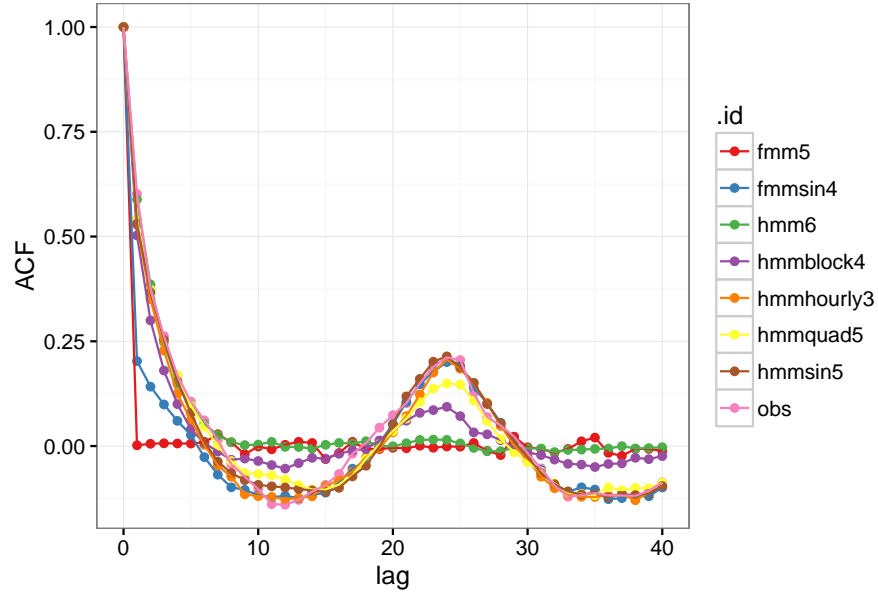
Figure 4: ACF

Figure 5: States

13

really do not know if there Florida panthers really moves according to hidden states, and even if they do, we don't know the true number nor able to observe them. Three distinct movement states seems biologically interpretable for Florida panthers: Short step length suggests resting states, intermediate step length might suggest foraging state and long step length might suggest traveling state. Long step lengths are easy to capture in all models, where as short step lengths are hard to capture given the GPS errors and unstable movements patterns suggesting multiple states associate with short step lengths. Thus, lower number of short step length states with higher variance can be equivalent or more realistic than multiple distinguishable short step length states with low variance.

4   Discussion

Identifying behavioural states based on some set of observations is a common methodological problem in behavioural ecology; finding solutions that are both biologically and statistically interpretable and defensible is challenging. Models that suggest a higher number of states fit the data well but are hard to interpret biologically; models that are biologically interpretable are not statistically supported. While the HMM framework has been applied in many research areas, very few studies have considered modeling with more than 3 states. HMM is a simple straight forward framework, but HMM extensions of adding a small amount of complexity can complicate the model greatly and reveal deeper understanding. For example, van de Kerk et al. (2015) used 2 to 6 movement states HSMM to model Florida panthers, Morales et al. (2004) used habitat distance dependent transition HMM to model elk, and outside of ecology, Raffa and Dubin (2015) used mixed effect transition and emission HMM in a smoking study.

We have presented a relatively simple concept yet little-studied extension by covariate dependent HMM can overfit the number of true states in covariate dependent HMM (in our case, time-dependent transitions). HMM with more than three states are not biologically interpretable and often discarded, but we cannot ignore that increasing complexity (number of states) states can increase goodness of fit statistically. Our analyses revealed modeling temporal heterogeneity in HMM transition can reduce the number of states estimated in time homogeneous transitions HMM movement of Florida panthers and reduce the overall model complexity.

For some kinds of model predictions, the time homogeneous transition model can perform as good as the time inhomogeneous transition models in terms of predicting the states by the Viterbi algorithm (the tradition state prediction method) and can easily overlook more complicated models. The Viterbi algorithm solves/predicts for the most likely/probable sequence of movement states based on the observations (Zucchini and MacDonald, 2009; Langrock et al., 2012). This approach allows the time homogeneous transition model's prediction to capture the temporal heterogeneity from the observations. It is useful to predict missing data in the observation sequence but fail to future movements beyond the observations sequence. In general, the time homogeneous transition model gives the broader scope of movement pattern where as the time inhomogeneous transitions models gives a more realistic movement structure and future predictions.

Although the time inhomogeneous transition offer an objective approach to (1) reduce the number of states to a biologically interpretable level, (2) captures the observational autocorrelation and temporal pattern, (3) reduce the overall complexity as illustrated by our study, several challenges must be overcome in order to maximize its potential. The current model is based on the

Table 1: Cat-ID

(van de Kerk 2015)	(IR@UF)	Number of Observations
130	1	10286
131	2	9458
48	14	14645
94	15	10250

reduced univariate step length response emission without fitting turning angles. Although turning angles are strongly correlated with step length, it does not effect the overall prediction on number of states. But movement trajectories predictions, and spatial models still require both step length and turning angles. Furthermore, modelling temporal heterogeneity only reduce the number of predicted states by 1 resulting in 4/5 states might suggest (1) other factors such as spatial components that might also contribute to over predicting number of states, or (2) there really are 1 or 2 more different movement characteristics that are biologically unknown. Despite the challenges, our study provides an important extension in HMM modeling that is not in the seen before and a real life application that it exist and applicable in animal movement ecology.

A Supplementary material

```
## Error in data.frame(viterbi = vitdist, time = catt$Time, obs = catt$LogDist):
arguments imply differing number of rows:  10250, 10286
```

Table 2: Number of Parameters for each model type.

Type	Model	Number of States	Number of Parameters
FMM	FMM	3	9
FMM	FMM	4	12
FMM	FMM	5	15
FMM	FMM	6	18
FMM + TH	FMM + TH (Sin)	3	15
FMM + TH	FMM + TH (Sin)	4	20
FMM + TH	FMM + TH (Sin)	5	25
FMM + TH	FMM + TH (Sin)	6	30
HMM	HMM	3	18
HMM	HMM	4	28
HMM	HMM	5	40
HMM	HMM	6	54
HMM + TH	HMM + TH (Block)	3	36
HMM + TH	HMM + TH (Block)	4	60
HMM + TH	HMM + TH (Block)	5	90
HMM + TH	HMM + TH (Block)	6	126
HMM + TH	HMM + TH (Quad)	3	36
HMM + TH	HMM + TH (Quad)	4	60
HMM + TH	HMM + TH (Quad)	5	90
HMM + TH	HMM + TH (Quad)	6	126
HMM + TH	HMM + TH (Sin)	3	36
HMM + TH	HMM + TH (Sin)	4	60
HMM + TH	HMM + TH (Sin)	5	90
HMM + TH	HMM + TH (Sin)	6	126
HMM + TH	HMM + TH (Hourly)	3	225
HMM + TH	HMM + TH (Hourly)	4	396
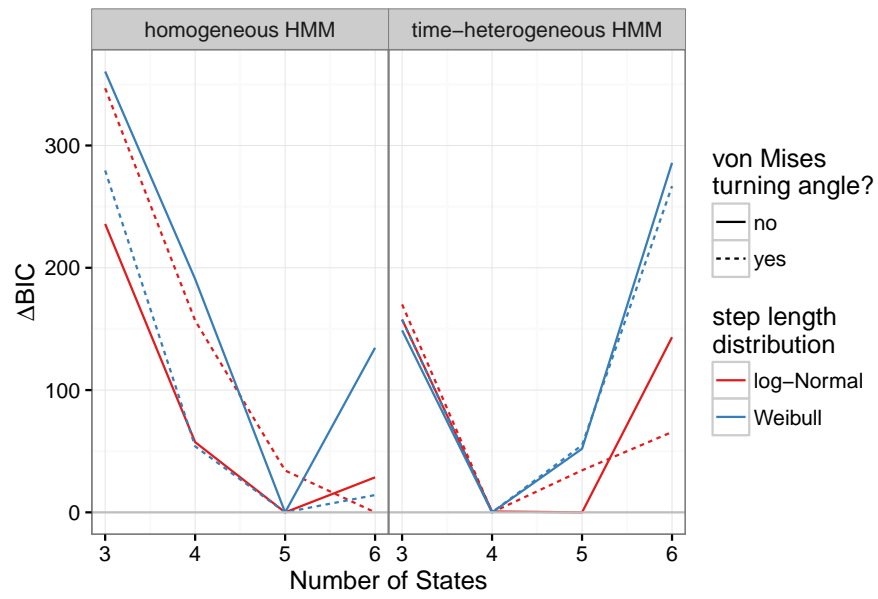
Figure 6: Change in BIC across models, cat #131. ΔBIC is scaled individually per model, so that ΔBIC $= 0$ represents the BIC-optimal number of states for a particular model.
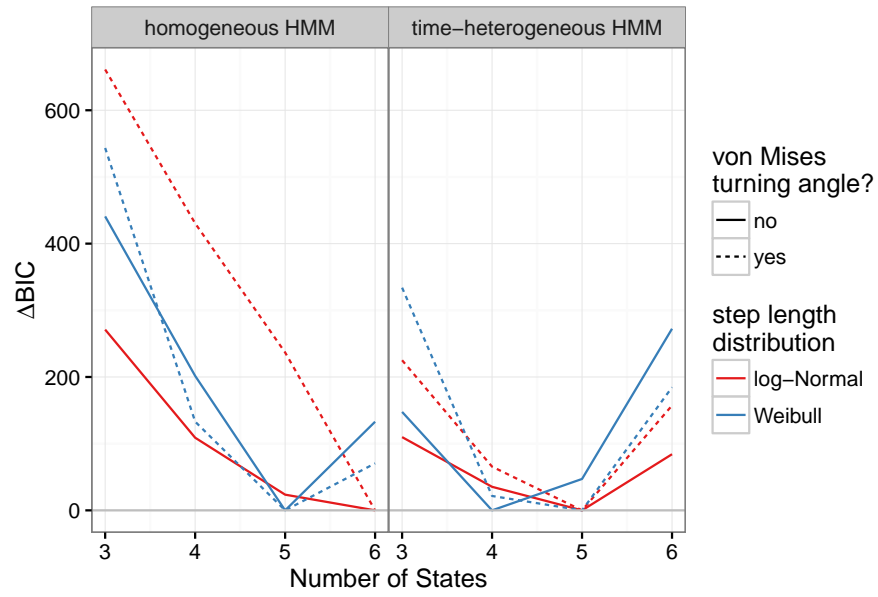
Figure 7: Change in BIC across models, cat #94. ΔBIC is scaled individually per model, so that ΔBIC $= 0$ represents the BIC-optimal number of states for a particular model.
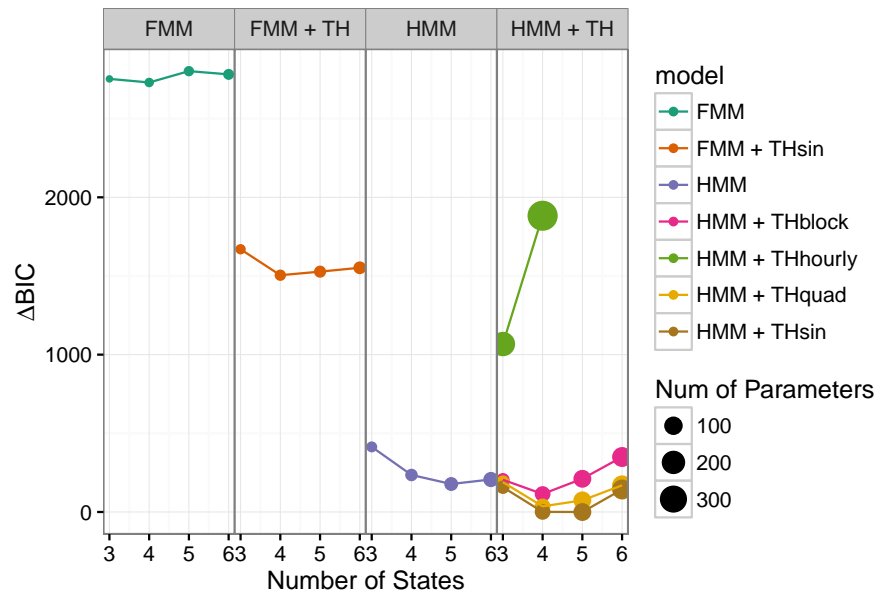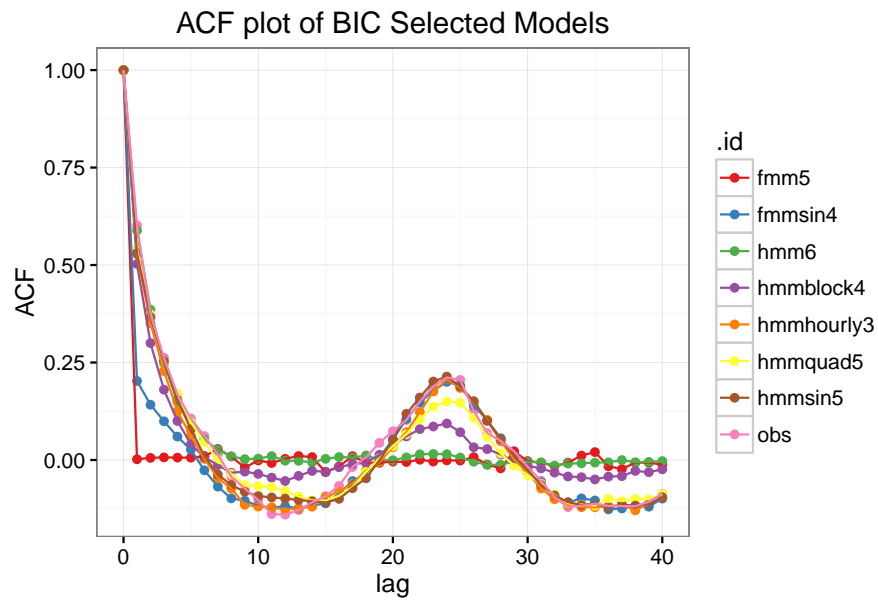
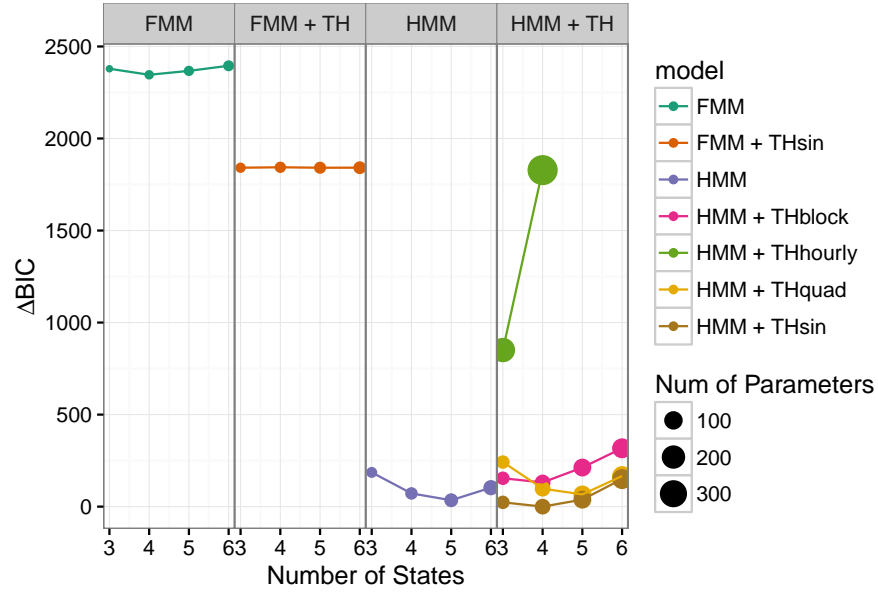Figure 8: Overall Adjusted BIC Comparisons between different HMM models

Figure 9: Overall Adjusted BIC Comparisons between different HMM models
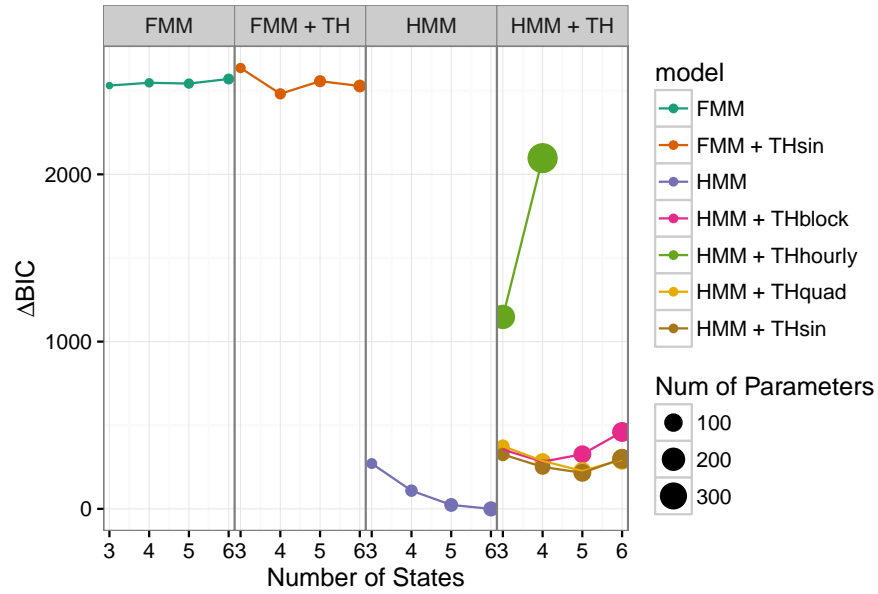
Figure 10: Overall Adjusted BIC Comparisons between different HMM models
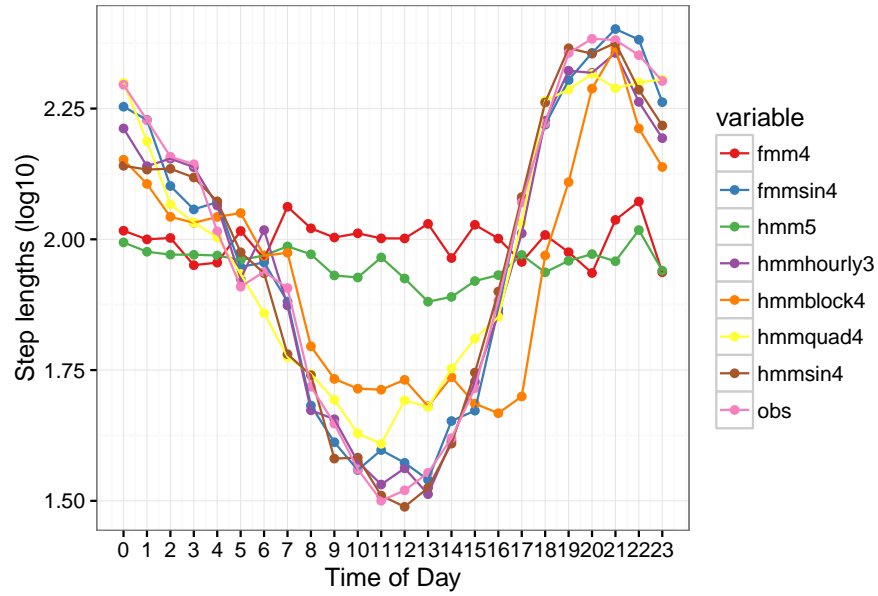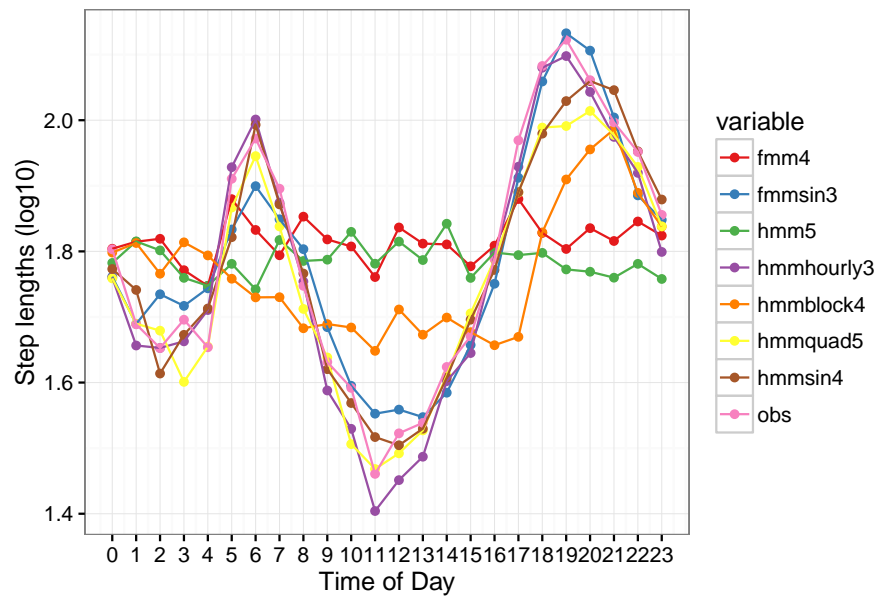
Figure 11: Average step length by time of day

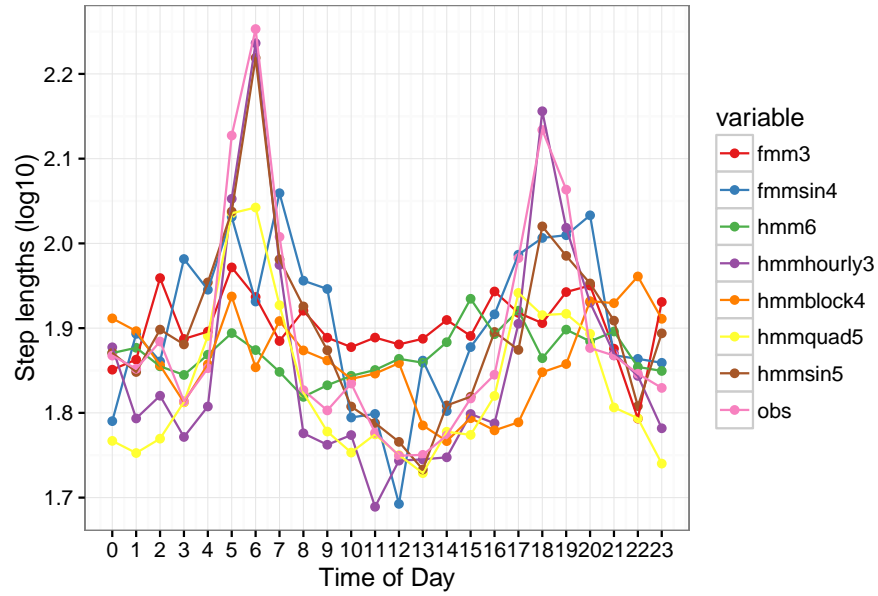Figure 12: Average step length by time of day

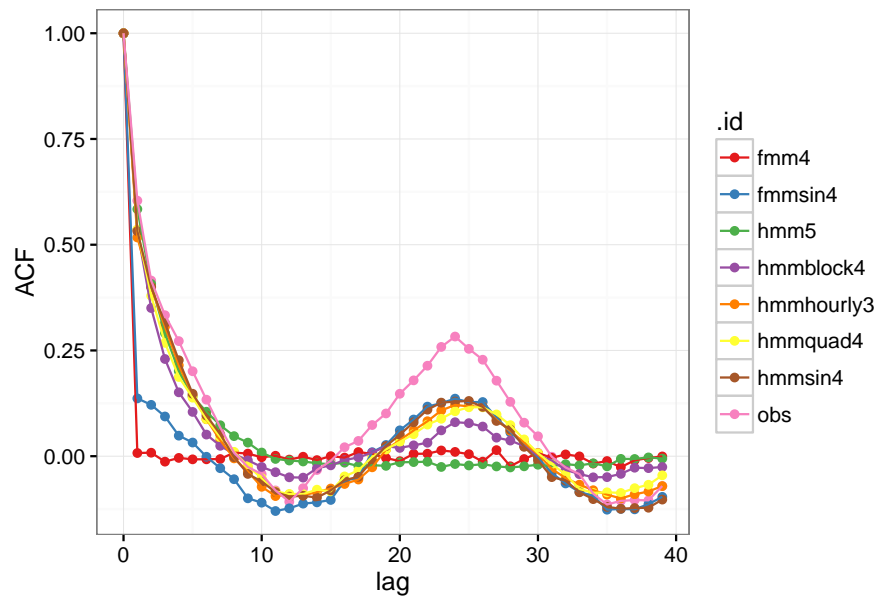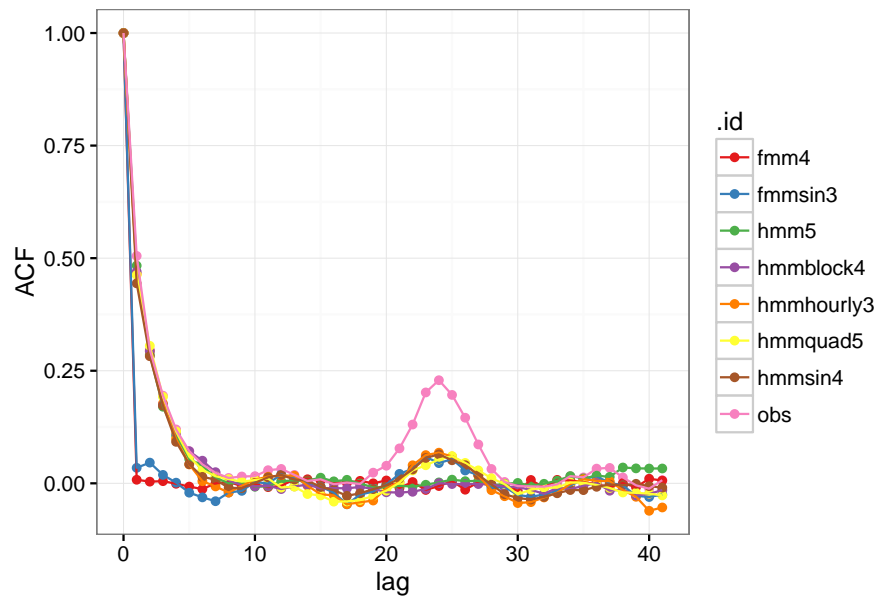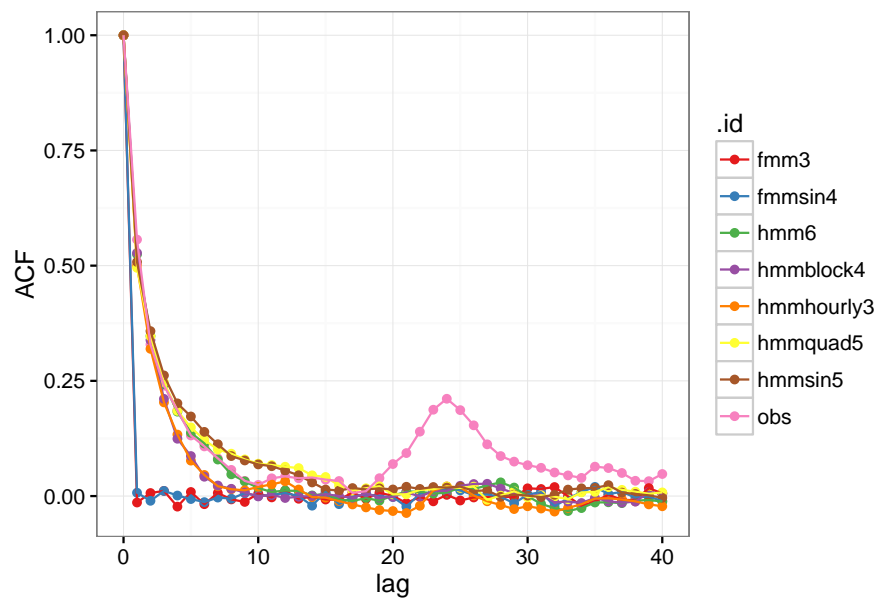Figure 13: Average step length by time of day

Figure 14: ACF

Figure 15: ACF

Figure 16: ACF

References

Burnham, K. P. and D. R. Anderson (1998). *Model Selection and Inference: A Practical Information-Theoretic Approach*. New York: Springer.

Dean, B., R. Freeman, H. Kirk, K. Leonard, R. A. Phillips, C. M. Perrins, and T. Guilford (2012). 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*, rsif20120570.

Firle, S., R. Bommarco, B. Ekbom, and M. Natiello (1998, September). The influence of movement and resting behavior on the range of three carabid beetles. *Ecology 79*(6), 2113–2122.

Franke, A., T. Caelli, G. Kuzyk, and R. J. Hudson (2006, August). Prediction of wolf (*Canis lupus*) kill-sites using hidden Markov models. *Ecological Modelling 197*(1–2), 237–246.

Fryxell, J. M., M. Hazell, L. Börger, B. D. Dalziel, D. T. Haydon, J. M. Morales, T. McIntosh, and R. C. Rosatte (2008, December). Multiple movement modes by large herbivores at multiple spatiotemporal scales. *Proceedings of the National Academy of Sciences 105*(49), 19114–19119.

Langrock, R., R. King, J. Matthiopoulos, L. Thomas, D. Fortin, and J. M. Morales (2012, June). Flexible and practical modeling of animal telemetry data: hidden Markov models and extensions. *Ecology 93*(11), 2336–2342.

McKellar, A. E., R. Langrock, J. R. Walters, and D. C. Kesler (2014, September). Using mixed hidden Markov models to examine behavioral states in a cooperatively breeding bird. *Behavioral Ecology*, aru171.

McKenzie, H. W., M. A. Lewis, and E. H. Merrill (2009). First passage time

analysis of animal movement and insights into the functional response. *Bulletin of Mathematical Biology 71*(1), 107–129.

Morales, J. M., D. T. Haydon, J. Frair, K. E. Holsinger, and J. M. Fryxell (2004). Extracting more out of relocation data: building movement models as mixtures of random walks. *Ecology 85*(9), 2436–2445.

Nathan, R., W. M. Getz, E. Revilla, M. Holyoak, R. Kadmon, D. Saltz, and P. E. Smouse (2008, December). A movement ecology paradigm for unifying organismal movement research. *Proceedings of the National Academy of Sciences 105*(49), 19052–19059.

Okubo, A. (1980). Diffusion and Ecological Problems: Mathematical Models.

Pal, S., B. Ghosh, and S. Roy (1998). 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.

Patterson, T. A., M. Basson, M. V. Bravington, and J. S. Gunn (2009, November). Classifying movement behaviour in relation to environmental conditions using hidden Markov models. *Journal of Animal Ecology 78*(6), 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 & evolution 23*(2), 87–94.

R Core Team (2015). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Raffa, J. D. and J. A. Dubin (2015, March). Multivariate longitudinal data analysis with mixed effects hidden Markov models. *Biometrics*, n/a–n/a.

Richards, S. A. (2005). Testing ecological theory using the information-theoretic approach: examples and cautionary results. *Ecology 86*(10), 2805–2814.

Schliehe-Diecks, S., P. M. Kappeler, and R. Langrock (2012, April). On the application of mixed hidden Markov models to multiple behavioural time series. *Interface Focus 2*(2), 180–189.

Tracey, J. A., J. Zhu, E. Boydston, L. Lyren, R. N. Fisher, and K. R. Crooks (2012, October). Mapping behavioral landscapes for animal movement: a finite mixture modeling approach. *Ecological Applications 23*(3), 654–669.

Turchin, P. (1998). *Quantitative analysis of movement: measuring and modeling population redistribution in animals and plants*, Volume 1. Sinauer Associates Sunderland.

van de Kerk, M., D. P. Onorato, M. A. Criffield, B. M. Bolker, B. C. Augustine, S. A. McKinley, and M. K. Oli (2015). Hidden semi-Markov models reveal multiphasic movement of the endangered Florida panther. *Journal of Animal Ecology 84*(2), 576–585.

Visser, I. and M. Speekenbrink (2010). `depmixS4`: An R package for hidden Markov models. *Journal of Statistical Software 36*(7), 1–21.

Zucchini, W. and I. L. MacDonald (2009, April). *Hidden Markov Models for Time Series: An Introduction Using R*. CRC Press.