

# ANIMAL MOVEMENT MODELLING: INDEPENDENT OR DEPENDENT MODELS?

*V. L. Goodall*<sup>1</sup>

Nelson Mandela Metropolitan University & University of the Witwatersrand  
e-mail: [victoriagoodall@gmail.com](mailto:victoriagoodall@gmail.com)

*L. P. Fatti*

University of the Witwatersrand

*N. Owen-Smith*

University of the Witwatersrand

---

**Key words:** Activity states, African buffalo, Hidden Markov models, Independent Mixture models, Latent states.

---

**Abstract:** Hidden Markov models have become a popular time series method for the analysis of GPS tracked animals. Their advantage for identifying latent behavioural states compared with Independent Mixture models is that they take into account the time series dependency of successive displacement distances by the tracked animals. However, little is known about how the analysis results may differ depending on which of these approaches is used. We compared the results and interpretations obtained from fitting Hidden Markov and Independent Mixture models to simulated movement data as well as to field data recording the hourly movements of sable antelope and buffalo within the Kruger National Park, South Africa. Hidden Markov models consistently yielded narrower confidence intervals around parameters and smaller standard errors than simpler time independent mixture models, but for some data the improvement was marginal and the Independent Mixture model provided an adequate alternative for identifying the latent behavioural states of the animal. In general, it is expected Hidden Markov models will provide the better balance between model complexity and extensibility for animal movement modelling from a statistical perspective. However, in some cases, Independent Mixture models could provide an adequate alternative method and might be more faithful biologically.

---

## 1. Introduction

A vast array of different analysis types have been applied to animal tracking data collected using GPS devices (McClintock, King, Thomas, Matthiopoulos, McConnell and Morales, 2012). Thus far, no single method is universally accepted for inferring the behavioural states generating the recorded GPS locations for terrestrial or marine species. Various techniques for the analysis of these data

---

<sup>1</sup>Corresponding author.

have appeared in the ecological (McClintock et al., 2012; Schick, Loarie, Colchero, Best, Boustany, Conde, Halpin, Joppa, McClellan and Clark, 2008), rather than statistical literature. Few studies have compared the model output and interpretations drawn when different methods are applied to the same dataset. Latent behavioural states influencing movement patterns can be inferred by using metrics calculated from the successive locations of the animal, such as the distance between locations and the turning angle. These can be clustered into groups using a variety of different modelling techniques. These underlying groups can then be interpreted in terms of the associated activity of the animal at the time of the observation. Some papers review a few methods, highlighting advantages and disadvantages of each method and then customize or develop their own model, rather than testing their data with all of the reviewed models (Schick et al., 2008). It is important to have an understanding of how the decision about which analysis to use can influence the results and inferences obtained.

Hidden Markov models have become popular as they provide a flexible and practical modelling approach to segment the movement path into latent behavioural states. Independent Mixture models (IMMs) have a very similar framework to Hidden Markov models (HMMs) except that they do not include the time dependence between successive observations which HMMs do via the Markov process. HMMs are in fact also known as Dependent Mixture models. Both of these methods have been used in the literature to model movement metrics in order to segment a movement path into discrete clusters which are inferred to correspond to the various behavioural states of the animal (Langrock, Hopcraft, Blackwell, Goodall, King, Niu, Patterson, Pedersen, Skarin and Schick, 2014; Langrock, King, Matthiopoulos, Thomas, Fortin and Morales, 2012; Owen-Smith, Goodall and Fatti, 2012; Owen-Smith and Goodall, 2014; Franke, Caelli and Hudson, 2004; Franke, Caelli, Kuzyk and Hudson, 2006). The latent states are described by the model parameters and then interpreted in terms of the likely activity of the animal. The interpretation of the behavioural state is dependent on the time scale of the observations, as changes in behaviour that occur within the time steps are not visible in the data.

While both of these methods have been shown to provide easily interpretable results for various large mammals (Langrock et al., 2012; Owen-Smith et al., 2012; Patterson, Basson, Bravington and Gunn, 2009; Franke et al., 2004; Franke et al., 2006), it is not known how the results would differ if the two methods were applied to the same movement data. In this paper, we focus on the comparison of these two approaches used to model the movements of sable antelope (*Hippotragus niger*) and African buffalo (*Syncerus caffer*) in the Kruger National Park, South Africa. In addition, a simulation study was performed to examine the similarities of the output when the model generating the sequence of states and displacements was known and in the presence of extreme transition probabilities. Bayesian state-space models (Morales, Haydon, Friar, Holsinger and Fryxell, 2004, among others) may also be used to segment movement paths into discrete clusters assumed to correspond to behavioural states. However, we found that when fitted to large GPS datasets using noninformative priors, they gave very similar parameter values to the HMMs, but took much longer to fit (Goodall, 2014).

## 2. Models

### 2.1. Independent Mixture Models

Mixture models have been used in a wide variety of applications including astronomy, biology, genetics, medicine, economics and marketing among others (McLachlan and Peel, 2000). They are suitable for modelling multimodal distributions since their definition allows them to cope with heterogeneity within the population (Zucchini and MacDonald, 2009). A mixture distribution is a statistical distribution which can be expressed as a combination of simpler component distributions (Everitt and Hand, 1981). The mixture models assume that the observations are drawn from a variety of different categories or states, and that each state has its own distribution (Bolker, 2008; Zucchini and MacDonald, 2009) with its corresponding proportional contribution. Although these models do not take into account the time series nature of the data, they have been shown to provide simple and realistic results and state allocations for the movements of ungulates (Owen-Smith et al., 2012; Owen-Smith and Goodall, 2014).

The density function of a random variable  $X$  with finite mixture distribution can be defined as:  $f(x) = \sum_{i=1}^m \delta_i g(x, \theta_i)$  where  $\theta_i$  is the vector of parameters of the  $i$ -th of the  $m$  component distributions whose probability density functions are given by  $g(x, \theta_i)$  and  $\delta_1, \delta_2, \dots, \delta_m$  are the proportions of each component in the model, known as the mixing parameters. Once the model has been fitted to a dataset and the parameters estimated, a probabilistic clustering of each observation to one of the  $m$  'states' is done using the estimated posterior probabilities of the state membership (McLachlan and Basford, 1988). The observation ( $x_j$ ) is allocated to the state with the maximum estimated posterior probability where the  $\delta_i$  are taken as the prior probabilities for the  $i$  states and  $p_i(x_j, \theta_i)$  are the state dependent distributions with parameters  $\theta_i$ . The posterior probabilities,  $\tau_i$  are given by:

$$\tau_i(x_j) = \frac{\delta_i p_i(x_j, \theta_i)}{\sum_{k=1}^m \delta_k p_k(x_j, \theta_k)}$$

This method is known as the 'Mixture Likelihood' clustering method (McLachlan and Peel, 2000).

### 2.2. Hidden Markov Models

The underlying assumption of both the Mixture and Hidden Markov models is that the observed data come from a population with underlying groupings. For the HMM, these groups are associated with one another via a Markov process. The marginal distribution of an HMM is a mixture distribution. The probability distribution of an observation at any time is determined only by the current state of that Markov chain. A hidden Markov process  $\{X_t\}$  comprises two probabilistic mechanisms, firstly an unobserved Markov chain  $\{C_t\}$  on  $m$  states which Zucchini and MacDonald (2009) call the 'parameter process'. The 'state dependent process'  $\{X_t : t = 1, 2, \dots\}$  is defined such that the distribution of  $X_t$  depends only on the current state  $C_t$  and not on the previous sequence of observations or states, which satisfies the Markov process. This model can be defined simply as (Zucchini and MacDonald, 2009):

$$P(C_t | \mathbf{C}^{(t-1)}) = P(C_t | C_{t-1}), t = 2, 3, \dots$$

$$P\left(X_t | \mathbf{X}^{(t-1)}, \mathbf{C}^{(t)}\right) = P(X_t | C_t), t \in \mathbb{N}$$

where  $\mathbf{X}^{(t)}$  and  $\mathbf{C}^{(t)}$  represent the history from  $t = 1, 2, \dots, T$ . Such a Hidden Markov model can be characterised by the distribution of  $\{C_t\}$ , the transition probability matrix of the Markov chain ( $\mathbf{\Gamma}$ ), and the state dependent distributions defined by  $p_i(x) = P(X_t = x | C_t = i)$ . This equation applies to discrete and continuous distributions of any parametric family (Ephraim and Merhav, 2002) and it could comprise distributions of the same or different families.  $p_i(x)$  is the probability mass function of  $X_t$  if the Markov chain is in state  $i$  at time  $t$  for the discrete case and the probability density function in the continuous case.

The Viterbi algorithm is used to find the sequence of states which maximises the probability of the observation sequence given the model (Tucker and Anand, 2005). The algorithm determines the sequence of states  $c_1, c_2, \dots, c_T$  which maximises the conditional probability:  $P(\mathbf{C}(T) = \mathbf{c}(T) | \mathbf{X}(T) = \mathbf{x}(T))$  where  $\mathbf{X}(T)$  and  $\mathbf{C}(T)$  represent the full sequence of observations and states (Zucchini and MacDonald, 2009). The Viterbi algorithm is an example of a dynamic programming algorithm which makes it possible to determine the most likely sequence of states without maximising over all possible sequences of states, which would not be feasible except for a very small number of states since  $m^T$  calculations would be required (Zucchini and MacDonald, 2009; Cappe, Moulines and Ryden, 2005).

### 3. Methods

All models were fitted using R statistical software (R Core Team, 2016). The displacement between successive locations were calculated using the `adehabitatLT` (Calenge, 2006) and `rgdal` (Bivand, Keitt and Rowlingson, 2016) packages in R. Two methods of comparison were used in order to formally compare the results of applying these two different approaches to animal movement data. Two simulations were done using the log-normal distribution with parameters almost identical to those obtained from fitting hourly movement models for particular sable antelope and buffalo herds respectively (Goodall, 2014). The log-normal distribution was used since it provided a good fit to the data and did not require the use of offsets which had been needed when the movements were modelled using the Gamma distribution (Owen-Smith et al., 2012; Owen-Smith and Goodall, 2014). Exponential and Weibull distributions were also considered but did not fit the observed data, and also required the use of offsets. The displacement distances only were used as an input for the models, since for cohesive ungulate herds, the turning angle between successive displacements are less relevant (Owen-Smith et al., 2012). Simulated movement data similar to these species, indicated that the accuracy of the state allocation did not improve much by including the turning angle in the analysis, particularly for models with more than two latent states (Goodall, 2014). The Hidden Markov model was used as the ‘true’ model since the behavioural state of the animal will tend to persist, which will produce displacement data which are serially correlated. The parameters of the model were assumed to be known and the “true states” and associated “observed movements” simulated from the “true” model. Mixture and Hidden Markov models were fitted to the simulated displacements and the underlying states predicted using the Mixture Likelihood clustering method (McLachlan and Peel, 2000) or the Viterbi algorithm for the IMM and HMMs respectively. The two approaches were compared using two criteria, the state classification accuracy of the model and

the precision of the parameter estimates comparing the known and estimated states and parameters. The parameters were compared using 90% confidence intervals and bootstrap standard errors of the estimated parameters in order to investigate the precision of and uncertainty around the parameter estimates. This was done for the state dependent distribution parameters for both methods and for the transition probabilities for the HMM. In order to investigate the difference in the results obtained for field studies where the true parameters are not known, the magnitude of the standard error and the width of the confidence interval can be used as an indicator of the precision of the estimate. If the confidence intervals are very wide, there is greater uncertainty about the precision of the parameters. This comparison is potentially of more interest to the ecologists since it will provide an idea of the uncertainty around the parameters and hence the confidence with which inferences can be made about the movement states of the animal, and the probabilities of shifting from one state to another.

### 3.1. Simulated Data

Observations and their associated latent states were simulated from a 4-state log-normal HMM. The four states were interpreted biologically as corresponding to resting, foraging, mixed foraging plus movement and persistent travelling activity states. If  $\delta$  represents the stationary distribution parameters,  $\Gamma$  the transition probability matrix and  $\mathbf{1}'$  a vector of ones, the likelihood ( $L_T$ ) of the stationary true model, is given by:

$$L_T = \delta \Gamma \mathbf{P}(x_1) \Gamma \mathbf{P}(x_2) \Gamma \mathbf{P}(x_3) \dots \Gamma \mathbf{P}(x_T) \mathbf{1}'$$

where  $\mathbf{P}(x) = \text{diag}(p_i(x_t))$  is the diagonal matrix of state dependent log-normal probability density functions if the Markov chain is in state  $i$  at time  $t$  for  $i = 1, 2, \dots, 4$ . The parameters for the state-dependent distributions and the transition probabilities were selected to be almost identical to those obtained for models fitted to hourly movements of sable antelope and buffalo respectively (Goodall, 2014). The log-normal distribution has a density function given by:  $f(x) = \frac{1}{\sqrt{2\pi}\sigma x} \exp\left(-\frac{(\ln(x)-\mu)^2}{2\sigma^2}\right)$ . The log-normal location ( $\mu$ ) and scale ( $\sigma$ ) parameters, as well as the mixing probabilities ( $\delta$ ) and transition probability matrix ( $\Gamma$ ) for the sable data were defined as:

$$\delta = \begin{bmatrix} 0.50 \\ 0.26 \\ 0.18 \\ 0.06 \end{bmatrix}; \Gamma = \begin{bmatrix} 0.70 & 0.21 & 0.08 & 0.01 \\ 0.37 & 0.41 & 0.19 & 0.03 \\ 0.25 & 0.25 & 0.41 & 0.09 \\ 0.08 & 0.06 & 0.32 & 0.54 \end{bmatrix}; \mu = \begin{bmatrix} -3.71 \\ -1.68 \\ -0.40 \\ 0.69 \end{bmatrix}; \sigma = \begin{bmatrix} 1.13 \\ 0.67 \\ 0.50 \\ 0.32 \end{bmatrix}.$$

The simulation creates a “known” sequence of states which can be used to simulate observed displacements including a correlation between successive states via the Markov process. The parameters used to simulate data similar to the buffalo movements were defined as:

$$\delta = \begin{bmatrix} 0.37 \\ 0.23 \\ 0.32 \\ 0.08 \end{bmatrix}; \Gamma = \begin{bmatrix} 0.66 & 0.00 & 0.32 & 0.02 \\ 0.56 & 0.44 & 0.00 & 0.00 \\ 0.00 & 0.35 & 0.58 & 0.08 \\ 0.03 & 0.19 & 0.20 & 0.59 \end{bmatrix}; \mu = \begin{bmatrix} -4.06 \\ -1.98 \\ -1.01 \\ 0.06 \end{bmatrix}; \sigma = \begin{bmatrix} 1.11 \\ 0.81 \\ 0.58 \\ 0.48 \end{bmatrix}.$$

The distribution curves for simulated states are shown in the right hand plots of Figure 2, based on the fitted distributions from the HMMs for the field study. The transition probability matrix contains

some zero values based on the findings from the hourly buffalo HMM fitting for four states. For the fitted mixture models, the simulation comparison will indicate how well the simpler model is able to describe the underlying states within the data, even though the fitted model does not take into account the serial correlation between the observations. The simulation and model fitting were repeated 500 times for both the sable and the buffalo parameters, and the results were used to calculate the 90% confidence intervals. This will indicate whether this model is a useful approximation to the true model, under the assumption that in practice the animals behave according to an HMM (Langrock et al., 2012; Franke et al., 2004; Franke et al., 2006). The more complex HMM analysis takes into account the correlation and should provide more precise results.

A separate simulation exercise was done to investigate the effect of extreme values within the transition probability matrix. A sequence of known states and associated displacement distances were simulated from two-, three- and four-state Hidden Markov models. It is highly unlikely to have extreme parameters associated with the displacement states, since animals are restricted to how much they can move within a predefined time period by their maximum travel rate, biology and the extent of their home range. Therefore, the simulations were specified with “typical” parameters for the state dependent distributions and with “extreme” values selected for the transition probability parameters. Different models were simulated with either very high probabilities of remaining within a state, or very high probabilities of shifting from the current state to one of the other states. 5000 observations were simulated per model and the simulations were repeated ten times. Four scenarios were considered, ranging from moderately extreme to very extreme parameters. Scenario one had 0.75 as the highest probability of remaining within/transitioning from a state, 0.8, 0.9 and 0.95 were the highest probabilities for scenarios two, three and four respectively. The transition probability matrices used for the simulations are shown in Table 1. This meant that 12 different model specifications were used to simulate from for extreme values of remaining in a state, and 12 specifications for extreme values of transitioning out of the current state. These simulations were used to investigate if the IMMs are still able to obtain results similar to the HMM, even in the presence of extreme transition probabilities.

### 3.2. Real-world data

Data for this study were collected using GPS collars placed on an adult female sable antelope and an adult female buffalo moving with conspecifics in cohesive herds in the Punda Maria region of the Kruger National Park, South Africa. Data from these two collars were also used in (Owen-Smith, Fryxell and Merrill, 2010; Owen-Smith et al., 2012; Owen-Smith, 2013; Owen-Smith and Goodall, 2014). Hourly locations used for this comparison were recorded between 17th August 2007 and 9th January 2008 for the sable antelope ( $n = 3064$ ), and between 21st January 2008 and 29th June 2009 for the buffalo ( $n = 11241$ ). During these periods 12.46% and 10.77% of locations were recorded as missing for the sable and buffalo respectively. Missed locations are due to a failure to connect to sufficient satellites to obtain a location. Changes in the programmed frequency of location recordings means that the hourly displacement cannot be calculated when successive observations are longer than one hour apart and these instances are also included as missing data in the analysis. Missing data need to be included for the HMM since the models take into account the time series nature of the input. The HMMs are fitted using a modification of the observed log-likelihood which

allows for missing observed data (Goodall, 2014). Missing data are irrelevant in the mixture models since there is no dependency on the time or sequence of the observations.

## 4. Results

### 4.1. Simulated Data

#### Sable herd

2-, 3- 4-, and 5-state IMM and HMMs were fitted to the simulated data, and the “best” model for each method identified using the AIC and BIC criteria. The AIC and BIC both identified the 4-state models as the best-fitting models, with the HMM always providing the better fit compared to the IMM according to these criteria. Table 2 shows the true HMM parameters and the fitted parameters from the 4-state log-normal IMM and HMM. The fitted parameters are very close to the true parameters using both methods, and the 90% confidence intervals for the parameters from both models always include the true parameter value, and were calculated using 500 sequences of observations. 5000 observations were simulated per “sable” sequence and 10 000 observations per “buffalo” sequence. These were selected in order to be consistent with the analysis of the case study data. The confidence intervals are narrower for the HMM. The standard errors for every parameter are small for both methods, but are smaller for the fitted HMM parameters. The results suggest that the IMM is a very adequate method for approximating the distributions that describe the underlying states. However, as expected the HMM is the better method and provides results with greater certainty when there is serial correlation. The state allocations obtained using the fitted model from the IMM and HMM methods are shown in Table 3. The diagonal values in the table are by far the highest for each state, which represents observations correctly allocated to their true state. The incorrect state allocation is nearly always to the state closest to the correct state. There does not appear to be any allocation bias towards either the incorrect lower or upper state for either model, with similar numbers of incorrectly allocated observations going to the states on either side of the true state. The fitted IMM provides 80.54% state allocation accuracy with the HMM slightly better at 82.96%. On the basis of this, it seems reasonable to conclude that inferences drawn from the underlying states using these methods can be used in further ecological analyses. The HMM performs better than the IMM in terms of state allocation accuracy for all states except the fourth state, which is associated with the longest movements (92.1% for the IMM and 90.7% for the HMM). The allocation accuracy is lowest for the second state (65.3% for the IMM and 67.1% for the HMM), assumed to correspond to the foraging state. The errors in state allocation for the second state are evenly spread between the first and third states for both the IMM and the HMMs (Table 3). This lower allocation accuracy for this state is not necessarily as a result of a poor model, but rather due to the overlap in displacement distributions for the foraging state with the states on either side of it which are expected to represent resting and mixed movement behaviours. This overlap will result in an increased probability of an incorrect state allocation.

It is interesting to note that the improvement in state allocation accuracy from the independent to dependent model is only 2.42%, which shows that the IMM is performing very well in terms of predicting the true state despite ignoring the serial correlation within the input data. It seems that

the improvement in accuracy for the HMM comes from observations incorrectly allocated to the higher state (longer displacements) by the IMM, which are better classified by the Viterbi algorithm applied to the fitted HMM (see Table 3). Since the fitted distributions are almost identical for the two models, these differences are due to the transition probabilities and the Viterbi algorithm which takes into account the most likely sequence of states rather than just the most likely state for a particular displacement, whatever the previous state had been.

### **Buffalo herd**

However, the IMMs did not perform as well for the buffalo data. In Table 4, the fitted parameters with their 90% confidence intervals and standard errors are shown for the IMMs and HMMs fitted to the data simulated from a true HMM with parameters similar to those obtained from models fitted to buffalo movement data with certain transitions between states precluded. The standard errors are once again larger for the IMM parameters and the 90% confidence intervals much wider than for the fitted HMM parameters. The difference between the models is accentuated in Table 3, which shows the state allocation of the true and fitted models. 85.3% of the observations are correctly allocated using the HMM, but this drops to 73.6% for the IMM. The IMM incorrectly allocates many of the true State 2 observations to the incorrect state, with 499 observations being allocated to State 1 and 974 allocated to State 3. Note that it is transitions from state 1 into state 2 and from state 2 into states 3 or 4 that are precluded by the simulated HMM model. In this situation, the IMM still does reasonably but is clearly outperformed by the HMM in terms of state allocation accuracy. The IMM will have no restriction on which state can be allocated next, based on the prior activity state, leading to lower state allocation accuracy. Similar differences were found between the IMMs and the HMMs for a 3-state model.

### **Extreme Value Simulation**

A summary of the results of the simulations are shown in Table 5. For the 2-state models, the HMMs perform very well with a very high state allocation accuracy. However, the IMMs also perform well, with 91.4% correct state allocation and over 93.5% of the predicted states were exactly the same as the predicted HMM states. As the transition probabilities become more extreme, the similarity between the state prediction of the IMM and the HMM does decrease, although not substantially. A similar pattern is observed for the 3-state models, in this case, 87.5% of state predictions were identical between the IMM and the HMM. For the 4-state model, the state allocation accuracy decreases, which is mainly due to the overlap between the latent states and the uncertainty associated with the state predictions. The state allocation accuracy decreases for both the IMM and HMM, although more substantially for the IMMs. For the 4-state models, slightly over 60% of observations are allocated to the same state by the IMM and HMMs, when the true model has extreme transition probabilities.

### **Simulation Summary**

The results for the simulated sable and buffalo data suggest that the additional fitting complexity for the HMM may be unnecessary in some cases when the only desired outcome from the model

is the behavioural prediction, rather than investigating the probability of moving from one state to another. Even though the mixture model ignores this serial dependence within the input data, the results provide an adequate representation of the underlying simulated groupings within the model. However, in the situation where certain transitions from one state to another are precluded, the IMM appears not to perform as well through ignoring the time sequence.

While this particular extreme value simulation provides interesting insight into how well the IMM is able to cope when the underlying model has very strong probabilities of transitioning from one state to another or remaining within a state, it is unlikely that such extreme probabilities would occur within animal movement data. An investigation into the autocorrelation between simulated displacements, with extreme probabilities of remaining within a state, shows highly significant correlations to over lag 20 and in some cases lag 30. This is not realistic for animal movement displacements, especially at the time scales that we are currently able to record GPS locations at. However, the results of this simulation indicate that even in the presence of very extreme probabilities of transitioning from one state to another, the IMM is able to perform reasonably well. For models with only two or three states, the IMMs perform very well, however for the more complex 4-state model with very extreme transition probabilities, the performance is not as good.

## 4.2. Real-world data

### Sable herd

The AIC for both IMMs and HMMs supported a four-state model as the best model, and the fitted model parameters are shown in Table 7. Although the HMM took longer to fit than the IMM, for datasets of this size and with a basic model specification, the time incurred is not prohibitive in fitting either model. It could become more problematic for very large datasets or models with many ecological covariates and more latent states. The fitted model parameters for the distributions and the mixing parameters/stationary distribution parameters are very similar between the two approaches, as shown in Table 7 with the fitted distributions illustrated in Figure 2. As with the models fitted to the simulated data, the 90% confidence intervals and the estimated parameter standard errors were always narrower for the HMMs. This confirms that there is more certainty around the parameter estimates for the HMM. The inclusion of the state transition probabilities matrix in the time dependent HMM improved the precision of the parameter estimates for the state dependent distributions compared to the IMM. The expected displacement distance in an hour for each state using the expected value for the state based on the fitted log-normal distribution parameters is shown in Table 6, along with the average displacement of the observations allocated to each state. The results are very similar for the two methods, and the interpretations of the latent states from these two methods would be the same. Figure 1 shows the frequency of state allocation by time of day for the observations using the two methods, with each bar in the plot representing an hour time interval. The pattern is almost identical for both methods with clear periods of resting, foraging and moving. This pattern is consistent with the expected behaviour for this species whose activities will vary through the day with resting dominating at night and during the heat of the day, foraging around sunrise and sunset and active states during the day rather than at night (Owen-Smith et al., 2010; Owen-Smith and Goodall, 2014). At a global level, the behavioural interpretations obtained from these two methods appear to be exactly the same and 86.1% of observations were allocated to the same state irrespective

of the modelling method used (Table 9). The observations not allocated to the same state are most likely as a result of the transition probabilities influencing the state allocation. These observations tend to be more likely to be allocated to the less active state for the IMMs and the more active state for the HMMs (Table 9).

### 4.3. Buffalo Herd

In contrast to the findings for the sable herd, those comparing the IMMs and HMMs differed for the buffalo herd (Table 8). Figure 1 shows the state allocation for the buffalo herd by time of day using the two different methods. State 1, state 2 and state 3 have a very similar pattern in both plots with clear peaks in resting before sunrise and during the heat of the day, and foraging peaks mid-morning, late afternoon and into the night. In particular, the state allocation pattern for the most active state is different. This can be seen from the fitted distributions in Figure 2, where the fourth state for the HMM has a lower expected value but wider distribution compared to the fourth state for the IMM. Furthermore, a greater proportion of observations was allocated to the fourth state for the HMM than for the IMM (Table 9). The interpretation of the fourth state is hence different for these two methods, with that provided by IMM more narrowly restricted to directed travel generating displacement distances sometimes exceeding 3km during an hour. Both models were re-run using different starting values but in all cases the models converged to the results presented here.

It is expected that the HMM will always provide the better statistical model for the analysis of animal movement, except in the situation of little serial correlation in the movement data. However, an alternative view is that the HMM will mostly persist in a state for more than one time step. In reality this might not be the case for the most active relocating state. For buffalo, this travelling state rarely persists for more than one hour since they do not need to undertake long journeys to water. By anticipating the persistence, the HMM might not perform as well for this particular state.

## 5. Discussion

Findings from the independent and the dependent mixture models are mostly similar, which is expected given the similarities of the model specifications. In the simulation study the HMM performed better than the IMM in terms of the standard errors and confidence intervals around the estimates. For the sable simulation, the HMM only showed a slight improvement in state allocation accuracy. The models for the buffalo showed differences between the two modelling techniques, leading to different interpretations of the activity states represented.

The zero transition probability between state 1 and state 2 implies that buffalo enter state 2 only from states 3 or 4, and never by a direct shift from state 1 to state 2. This was initially surprising, but is actually an outcome of the high probability of persisting in state 1, which overrides the assignment to the less probable state 2 that the IMM would make based on the displacement distance alone. Contributing to this is the lower expected value for state 2 assigned by HMM for buffalo compared to that for state 2 for the sable data. This suggests that state 2 may represent a restless mix of foraging and stationary behaviour for an individual buffalo within the large herds formed by this ungulate, rather than persistent foraging as for the sable moving in somewhat smaller herds. Notably, buffalo herds tend to remain close to water during the dry season, in contrast to the journeys undertaken by

sable herds to remaining water sources over distances of 3–5km (Cain, Owen-Smith and Macandza, 2012). To check that the zero transition probability between states 1 and 2 for the buffalo was not an anomalous peculiarity of the particular data set that we had used, we also fitted a HMM to the movements of another buffalo herd in a different region of the Kruger National Park, and obtained this same pattern. The confidence intervals around the parameter estimates for the HMM were always narrower than for the IMM parameter confidence intervals, which illustrates that the inclusion of the time dependence via the Markov chain even improves the precision of the state dependent distribution parameters.

In the case of the extreme value simulation, it is interesting to note that the IMM performs less well as the number of states increase and when the transition probabilities are more extreme. Reducing the time step would tend to increase the probability of remaining within the same state from one time to the next and would correspond to the “remain within a state” extreme scenarios. In this case, the IMMs perform poorly in comparison to the HMMs, which take into account the time dependence. In a corresponding situation if the time step was increased, the persistence in a state will decrease and the IMM is likely to perform better.

The HMMs provide information about the probability of shifting from one state to another, and the state allocation is not only based on the probability of observing a displacement of a particular distance. The transition matrix in effect determines how long an animal is likely to remain within a state. Nevertheless, the assumption of a first order Markov process may be misleading, because the longer an animal has been resting, the more likely that it is to shift towards foraging as it becomes more hungry, while persistent foraging eventually leads to satiation and hence resting. It is expected furthermore that the performance of the IMM relative to the HMM will depend on the temporal resolution of the data, with HMMs increasingly better supported when states commonly persist through multiple time steps.

## 6. Conclusion

Hidden Markov models provide a balance between computational requirements and model simplicity which allows for reasonably simple interpretation of the latent states in terms of the behaviour of an animal. A number of authors have identified them as promising for modelling animal movements in heterogeneous environments (Dragon, Bar-Hen, Monestiez and Guinet, 2012; Barraquand and Benhamou, 2008; Patterson, Thomas, Wilcox, Ovaskainen and Matthiopoulos, 2008; Langrock et al., 2012). In the end, it is expected that the decision about which analysis to use will be based on the questions being addressed about the ecology of the animal including the scale and regularity of the movements and the distances travelled, and the desired outcomes of the model (Dragon et al., 2012). If a simple state allocation is required from which the behaviour of the animal will be inferred, then the IMM may be sufficient. Gutenkunst, Newlands, Lutcavage and Edelstein-Keshet (2007) suggested that the major uncertainty in the analysis of animal movement data comes from an ignorance of the “perfect movement model”. This is a fundamental aspect of statistics, that the underlying true model is seldom known. However, with improvements in remote sensing technology, it is likely that the available data will become richer and have the potential to better characterize the true behavioural states, which can then be used to validate and refine the movement models.

## Acknowledgements

We acknowledge the contribution of two anonymous reviewers for providing helpful suggestions for further improvements.

## 7. Tables and Figures

**Table 1:** Transition probability matrices used for the extreme value simulation.

		2-state		3-state			4-state						2-state		3-state			4-state											
Remain within a state	Scenario 1	0.75	0.25	0.75	0.15	0.10	0.75	0.15	0.05	0.05	Transition to another state	Scenario 1	0.25	0.75	0.15	0.75	0.10	0.15	0.75	0.05	0.05	0.15	0.75	0.05	0.05				
		0.25	0.75	0.15	0.75	0.10	0.05	0.75	0.15	0.05			0.75	0.25	0.15	0.10	0.75	0.05	0.15	0.75	0.05	0.05	0.05	0.15	0.75				
				0.10	0.15	0.75	0.050	0.05	0.75	0.15					0.75	0.10	0.15	0.05	0.05	0.15	0.75	0.75	0.05	0.05	0.15				
							0.15	0.05	0.05	0.75								0.75	0.05	0.05	0.15								
	Scenario 2	0.80	0.20	0.80	0.15	0.05	0.80	0.10	0.05	0.05		Scenario 2	0.20	0.80	0.15	0.80	0.05	0.10	0.80	0.05	0.05	0.10	0.80	0.05	0.05	0.10	0.80	0.05	0.05
		0.20	0.80	0.05	0.80	0.15	0.05	0.80	0.10	0.05			0.80	0.20	0.05	0.15	0.80	0.05	0.10	0.80	0.05	0.05	0.05	0.10	0.80				
				0.15	0.05	0.80	0.05	0.05	0.80	0.10					0.80	0.05	0.15	0.05	0.05	0.10	0.80	0.80	0.05	0.05	0.10				
							0.10	0.05	0.05	0.80								0.80	0.05	0.05	0.10								
	Scenario 3	0.90	0.10	0.90	0.05	0.05	0.90	0.04	0.03	0.03		Scenario 3	0.10	0.90	0.05	0.90	0.05	0.04	0.90	0.03	0.03	0.04	0.90	0.03	0.03	0.04	0.90	0.03	0.03
		0.10	0.90	0.05	0.90	0.05	0.03	0.90	0.04	0.03			0.90	0.10	0.05	0.05	0.90	0.03	0.04	0.90	0.03	0.03	0.03	0.04	0.90				
				0.05	0.05	0.90	0.03	0.03	0.90	0.04					0.90	0.05	0.05	0.03	0.03	0.04	0.90	0.90	0.03	0.03	0.04				
							0.04	0.03	0.03	0.90								0.90	0.03	0.03	0.04								
	Scenario 4	0.95	0.05	0.95	0.03	0.02	0.95	0.02	0.02	0.01		Scenario 4	0.05	0.95	0.03	0.95	0.02	0.02	0.95	0.02	0.01	0.02	0.95	0.02	0.01	0.02	0.95	0.02	0.01
		0.05	0.95	0.02	0.95	0.03	0.01	0.95	0.02	0.02			0.95	0.05	0.02	0.03	0.95	0.01	0.02	0.95	0.02	0.02	0.01	0.02	0.95				
				0.02	0.03	0.95	0.02	0.01	0.95	0.02					0.95	0.03	0.02	0.02	0.01	0.02	0.95	0.02	0.01	0.02	0.95				
							0.02	0.02	0.01	0.95								0.95	0.02	0.01	0.02								

**Table 2:** Simulated Sable 4-state HMM parameters and fitted values for IMM and HMM.  $\delta_i$  denotes the stationary distribution or mixing parameters for the HMM or IMM respectively; and  $\mu_i$  and  $\sigma_i$  denote the location and scale parameters respectively of the log-normal distributions for the  $i$ -th state of either model.  $\gamma_{ij}$  is the transition probability from the  $i$ -th to  $j$ -th state for the HMM.

True Model		Fitted Models					
HMM		IMM	IMM 90% CI	Std Err - IMM	HMM	HMM 90% CI	Std Err - HMM
$\delta_1$	0.50	0.50	(0.41, 0.56)	0.05	0.51	(0.47, 0.55)	0.02
$\delta_2$	0.26	0.24	(0.10, 0.41)	0.09	0.23	(0.19, 0.29)	0.03
$\delta_3$	0.18	0.19	(0.09, 0.31)	0.07	0.19	(0.16, 0.23)	0.02
$\delta_4$	0.06	0.07	(0.05, 0.09)	0.01	0.06	(0.05, 0.08)	0.01
$\mu_1$	-3.71	-3.66	(-3.92, -3.47)	0.14	-3.63	(-3.75, -3.53)	0.07
$\mu_2$	-1.68	-1.79	(-2.07, -1.54)	0.17	-1.75	(-1.86, -1.61)	0.08
$\mu_3$	-0.40	-0.55	(-0.77, -0.42)	0.11	-0.52	(-0.60, -0.44)	0.05
$\mu_4$	0.69	0.62	(0.51, 0.71)	0.06	0.64	(0.59, 0.69)	0.03
$\sigma_1$	1.13	1.16	(1.05, 1.25)	0.06	1.17	(1.10, 1.23)	0.04
$\sigma_2$	0.67	0.64	(0.40, 0.92)	0.16	0.65	(0.56, 0.75)	0.06
$\sigma_3$	0.50	0.48	(0.34, 0.71)	0.11	0.52	(0.46, 0.58)	0.04
$\sigma_4$	0.32	0.33	(0.27, 0.38)	0.03	0.32	(0.29, 0.36)	0.02
$\gamma_{11}$	0.70				0.71	(0.67, 0.75)	0.03
$\gamma_{12}$	0.21				0.19	(0.15, 0.25)	0.03
$\gamma_{13}$	0.08				0.08	(0.05, 0.10)	0.02
$\gamma_{14}$	0.01				0.01	(0.01, 0.02)	0.00
$\gamma_{21}$	0.37				0.34	(0.29, 0.39)	0.03
$\gamma_{22}$	0.41				0.39	(0.33, 0.47)	0.04
$\gamma_{23}$	0.19				0.23	(0.17, 0.29)	0.04
$\gamma_{24}$	0.03				0.03	(0.02, 0.05)	0.01
$\gamma_{31}$	0.25				0.32	(0.27, 0.37)	0.03
$\gamma_{32}$	0.25				0.19	(0.14, 0.25)	0.03
$\gamma_{33}$	0.41				0.40	(0.35, 0.45)	0.03
$\gamma_{34}$	0.09				0.09	(0.07, 0.12)	0.01
$\gamma_{41}$	0.08				0.08	(0.04, 0.12)	0.03
$\gamma_{42}$	0.06				0.06	(0.00, 0.12)	0.03
$\gamma_{43}$	0.32				0.34	(0.28, 0.41)	0.04
$\gamma_{44}$	0.54				0.51	(0.46, 0.57)	0.03

**Table 3:** Simulated state allocation for fitted IMM (left) and HMM (right).

Sable											
		State Allocation IMM				State Allocation HMM					
		State 1	State 2	State 3	State 4			State 1	State 2	State 3	State 4
True State	State 1	2158	294	26	0	True State	State 1	2233	222	23	0
	State 2	204	876	257	5		State 2	223	901	216	2
	State 3	0	78	725	86		State 3	0	78	750	61
	State 4	0	0	23	268		State 4	0	0	27	264
Buffalo											
		State Allocation IMM				State Allocation HMM					
		State 1	State 2	State 3	State 4			State 1	State 2	State 3	State 4
True State	State 1	3374	323	118	0	True State	State 1	3441	287	87	0
	State 2	499	736	974	37		State 2	157	1712	372	5
	State 3	10	201	2681	298		State 3	36	234	2813	107
	State 4	0	0	183	566		State 4	0	4	181	564

**Table 4:** Simulated Buffalo 4-state HMM parameters and fitted values for IMM and HMM.  $\delta_i$  denotes the stationary distribution or mixing parameters for the HMM or IMM respectively; and  $\mu_i$  and  $\sigma_i$  denote the location and scale parameters respectively of the log-normal distributions for the  $i$ -th state of either model.  $\gamma_{ij}$  is the transition probability from the  $i$ -th to  $j$ -th state for the HMM.

True Model		Fitted Models					
HMM		IMM	IMM 90% CI	Std Err - IMM	HMM	HMM 90% CI	Std Err - HMM
$\delta_1$	0.37	0.41	(0.36, 0.45)	0.03	0.37	(0.36, 0.38)	0.01
$\delta_2$	0.23	0.13	(0.04, 0.36)	0.10	0.24	(0.23, 0.25)	0.01
$\delta_3$	0.32	0.36	(0.11, 0.50)	0.13	0.31	(0.30, 0.33)	0.01
$\delta_4$	0.08	0.11	(0.02, 0.24)	0.07	0.08	(0.07, 0.09)	0.01
$\mu_1$	-4.06	-3.91	(-4.10, -3.79)	0.10	-4.06	(-4.11, -4.02)	0.03
$\mu_2$	-1.98	-2.24	(-2.49, -1.71)	0.24	-2.03	(-2.08, -1.98)	0.03
$\mu_3$	-1.01	-1.09	(-1.20, -0.95)	0.08	-1.01	(-1.04, -0.98)	0.02
$\mu_4$	0.06	-0.07	(-0.47, 0.33)	0.25	0.06	(0.00, 0.12)	0.03
$\sigma_1$	1.11	1.15	(1.07, 1.21)	0.04	1.10	(1.07, 1.13)	0.02
$\sigma_2$	0.81	0.49	(0.29, 0.87)	0.17	0.82	(0.79, 0.86)	0.02
$\sigma_3$	0.58	0.53	(0.34, 0.73)	0.12	0.57	(0.55, 0.60)	0.01
$\sigma_4$	0.48	0.52	(0.36, 0.63)	0.09	0.48	(0.45, 0.51)	0.02
$\gamma_{11}$	0.66				0.65	(0.64, 0.67)	0.01
$\gamma_{12}$	0.00				0.00	(0.00, 0.00)	0.00
$\gamma_{13}$	0.32				0.33	(0.31, 0.34)	0.01
$\gamma_{14}$	0.02				0.02	(0.01, 0.03)	0.01
$\gamma_{21}$	0.56				0.53	(0.50, 0.56)	0.02
$\gamma_{22}$	0.44				0.46	(0.43, 0.49)	0.02
$\gamma_{23}$	0.00				0.00	(0.00, 0.00)	0.00
$\gamma_{24}$	0.00				0.01	(0.00, 0.02)	0.01
$\gamma_{31}$	0.00				0.00	(0.00, 0.00)	0.00
$\gamma_{32}$	0.35				0.37	(0.35, 0.39)	0.01
$\gamma_{33}$	0.58				0.56	(0.54, 0.59)	0.02
$\gamma_{34}$	0.08				0.07	(0.06, 0.08)	0.01
$\gamma_{41}$	0.03				0.01	(0.00, 0.02)	0.01
$\gamma_{42}$	0.19				0.18	(0.14, 0.23)	0.03
$\gamma_{43}$	0.20				0.21	(0.17, 0.27)	0.03
$\gamma_{44}$	0.59				0.59	(0.55, 0.64)	0.03

**Table 5:** Mean results of extreme value simulations for IMM and HMM.

	2-state model				3-state model				4-state model			
	% IMM states allocated correctly	% HMM states allocated correctly	% state allocation identical IMM and HMM	% IMM states allocated correctly	% HMM states allocated correctly	% state allocation identical IMM and HMM	% IMM states allocated correctly	% HMM states allocated correctly	% state allocation identical IMM and HMM	% IMM states allocated correctly	% HMM states allocated correctly	% state allocation identical IMM and HMM
Remain within a state	Scenario 1	91.4	93.3	95.1	86.3	91.8	88.7	58.9	79.0	61.3		
	Scenario 2	91.3	94.3	94.2	87.0	94.7	88.6	64.9	82.7	65.8		
	Scenario 3	91.4	96.9	92.5	85.3	96.9	85.7	57.1	90.4	57.2		
	Scenario 4	91.5	98.4	92.1	86.7	98.5	87.1	65.1	95.3	65.5		
	Mean	91.4	95.7	93.5	86.3	95.5	87.5	61.5	86.8	62.4		
Shift to another state	Scenario 1	91.3	93.0	94.5	86.3	93.5	88.3	60.6	83.6	61.9		
	Scenario 2	91.6	94.0	93.7	86.8	95.4	88.6	63.2	88.7	64.0		
	Scenario 3	91.1	96.7	92.0	86.8	97.7	87.4	61.0	94.6	61.3		
	Scenario 4	91.4	98.4	91.8	86.7	99.0	87.0	59.9	97.6	60.4		
	Mean	91.4	95.5	93.0	86.7	96.4	87.8	61.2	91.1	61.9		

**Table 6:** Expected movement distances from the state dependent distributions and mean displacement of observations allocated to each state for sable and buffalo.

	Hidden Markov Model				Mixture Model			
	Stationary probabilities	Expected movement (km)	Average displacement using state allocation (km)	Prior probabilities	Expected movement (km)	Average displacement using state allocation (km)		
Sable	State 1	0.51	0.05	0.03	0.59	0.07	0.05	
	State 2	0.25	0.24	0.22	0.20	0.31	0.27	
	State 3	0.18	0.75	0.75	0.17	0.83	0.81	
	State 4	0.06	2.10	2.15	0.01	2.17	2.26	
Buffalo	State 1	0.37	0.03	0.03	0.36	0.03	0.02	
	State 2	0.23	0.19	0.16	0.33	0.24	0.16	
	State 3	0.32	0.43	0.44	0.30	0.58	0.62	
	State 4	0.08	1.19	1.32	0.01	2.00	2.29	

**Table 7:** Punda Maria sable - Fitted model parameters for 4-state IMM and HMMs.

	Parameter		90% CI	Std Error		Parameter		90% CI	Std Error
<b>4-state HMM</b>	-LL	-1477.34			<b>4-state IMM</b>	-LL	-1102.32		
	AIC	-2914.69				AIC	-2182.64		
	BIC	-2794.14				BIC	-2116.33		
	$\delta_1$	0.51	(0.46,0.55)	0.03		$\delta_1$	0.59	(0.45,0.66)	0.08
	$\delta_2$	0.25	(0.20,0.31)	0.03		$\delta_2$	0.20	(0.03,0.34)	0.10
	$\delta_3$	0.18	(0.14,0.22)	0.02		$\delta_3$	0.17	(0.05,0.31)	0.08
	$\delta_4$	0.06	(0.05,0.08)	0.01		$\delta_4$	0.05	(0.03,0.08)	0.02
	$\mu_1$	-3.68	(-3.81,-3.56)	0.08		$\mu_1$	-3.42	(-3.76,-3.20)	0.24
	$\mu_2$	-1.66	(-1.79,-1.51)	0.09		$\mu_2$	-1.42	(-2.15,-0.97)	0.38
	$\mu_3$	-0.41	(-0.50,-0.33)	0.05		$\mu_3$	-0.34	(-0.71,-0.04)	0.25
	$\mu_4$	0.69	(0.63,0.75)	0.04		$\mu_4$	0.73	(0.59,0.84)	0.08
	$\sigma_1$	1.13	(1.06,1.19)	0.04		$\sigma_1$	1.26	(1.12,1.36)	0.09
	$\sigma_2$	0.68	(0.57,0.81)	0.07		$\sigma_2$	0.71	(0.23,0.93)	0.22
	$\sigma_3$	0.50	(0.43,0.56)	0.04		$\sigma_3$	0.55	(0.28,0.73)	0.16
	$\sigma_4$	0.31	(0.27,0.35)	0.02		$\sigma_4$	0.29	(0.20,0.36)	0.05
	$\gamma_{11}$	0.71	(0.66,0.75)	0.03					
	$\gamma_{12}$	0.20	(0.15,0.26)	0.03					
	$\gamma_{13}$	0.08	(0.05,0.11)	0.02					
	$\gamma_{14}$	0.01	(0.00,0.02)	0.00					
	$\gamma_{21}$	0.36	(0.32,0.41)	0.03					
	$\gamma_{22}$	0.42	(0.34,0.50)	0.05					
	$\gamma_{23}$	0.18	(0.12,0.24)	0.04					
	$\gamma_{24}$	0.03	(0.01,0.05)	0.01					
	$\gamma_{31}$	0.27	(0.22,0.33)	0.04					
	$\gamma_{32}$	0.22	(0.16,0.29)	0.04					
	$\gamma_{33}$	0.41	(0.36,0.48)	0.04					
	$\gamma_{34}$	0.09	(0.06,0.12)	0.02					
	$\gamma_{41}$	0.09	(0.04,0.13)	0.03					
$\gamma_{42}$	0.06	(0.00,0.13)	0.04						
$\gamma_{43}$	0.32	(0.25,0.40)	0.05						
$\gamma_{44}$	0.53	(0.46,0.59)	0.04						

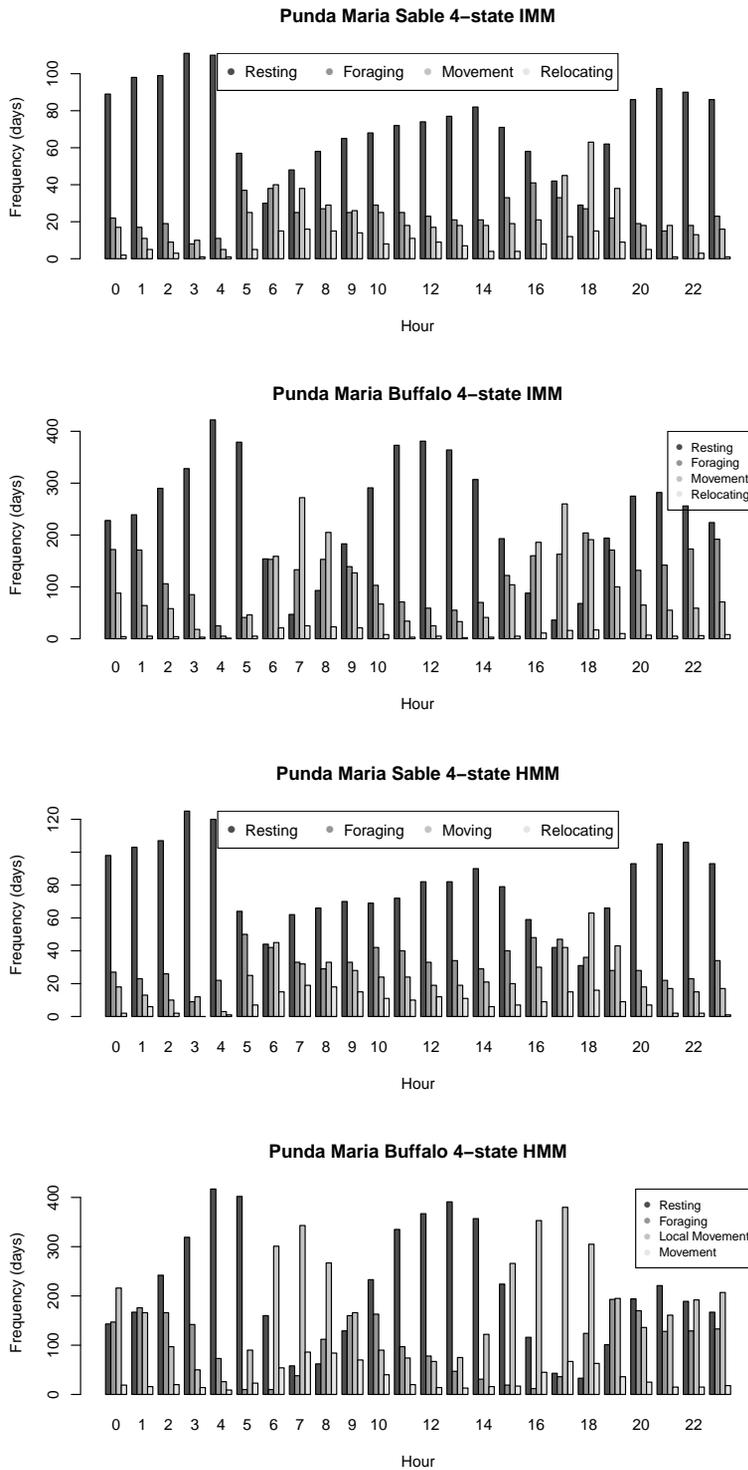
**Table 8:** Punda Maria buffalo - Fitted model parameters for 4-state IMMs and HMMs.

	Parameter	90% CI	Std Error		Parameter	90% CI	Std Error	
4-state HMM	-LL	-6501.99			-LL	-4728.16		
	AIC	-12963.97			AIC	-9434.33		
	BIC	-12817.43			BIC	-9353.73		
	$\delta_1$	0.37	(0.36, 0.39)	0.01	$\delta_1$	0.36	(0.33, 0.40)	0.02
	$\delta_2$	0.23	(0.21, 0.24)	0.01	$\delta_2$	0.33	(0.12, 0.44)	0.10
	$\delta_3$	0.32	(0.31, 0.33)	0.01	$\delta_3$	0.30	(0.17, 0.48)	0.09
	$\delta_4$	0.08	(0.07, 0.09)	0.01	$\delta_4$	0.01	(0.01, 0.02)	0.01
	$\mu_1$	-4.06	(-4.09, -4.02)	0.02	$\mu_1$	-4.13	(-4.25, -3.98)	0.08
	$\mu_2$	-1.98	(-2.02, -1.94)	0.03	$\mu_2$	-1.75	(-2.26, -1.50)	0.24
	$\mu_3$	-1.01	(-1.04, -0.99)	0.02	$\mu_3$	-0.74	(-0.94, -0.62)	0.10
	$\mu_4$	0.06	(0.01, 0.10)	0.03	$\mu_4$	0.66	(0.52, 0.77)	0.08
	$\sigma_1$	1.11	(1.09, 1.13)	0.02	$\sigma_1$	1.06	(1.00, 1.13)	0.04
	$\sigma_2$	0.81	(0.78, 0.84)	0.02	$\sigma_2$	0.83	(0.53, 0.94)	0.13
	$\sigma_3$	0.58	(0.56, 0.61)	0.01	$\sigma_3$	0.62	(0.51, 0.70)	0.06
	$\sigma_4$	0.48	(0.45, 0.51)	0.02	$\sigma_4$	0.25	(0.15, 0.33)	0.06
	$\gamma_{11}$	0.66	(0.64, 0.67)	0.01				
	$\gamma_{12}$	0.00	(0.00, 0.00)	0.00				
	$\gamma_{13}$	0.32	(0.31, 0.34)	0.01				
	$\gamma_{14}$	0.02	(0.01, 0.03)	0.00				
	$\gamma_{21}$	0.56	(0.53, 0.59)	0.02				
	$\gamma_{22}$	0.44	(0.41, 0.47)	0.02				
	$\gamma_{23}$	0.00	(0.00, 0.00)	0.00				
	$\gamma_{24}$	0.00	(0.00, 0.00)	0.00				
	$\gamma_{31}$	0.00	(0.00, 0.00)	0.00				
	$\gamma_{32}$	0.35	(0.33, 0.36)	0.01				
	$\gamma_{33}$	0.57	(0.55, 0.59)	0.01				
	$\gamma_{34}$	0.08	(0.07, 0.09)	0.01				
	$\gamma_{41}$	0.03	(0.01, 0.04)	0.01				
$\gamma_{42}$	0.19	(0.16, 0.23)	0.02					
$\gamma_{43}$	0.20	(0.15, 0.24)	0.03					
$\gamma_{44}$	0.59	(0.55, 0.62)	0.02					

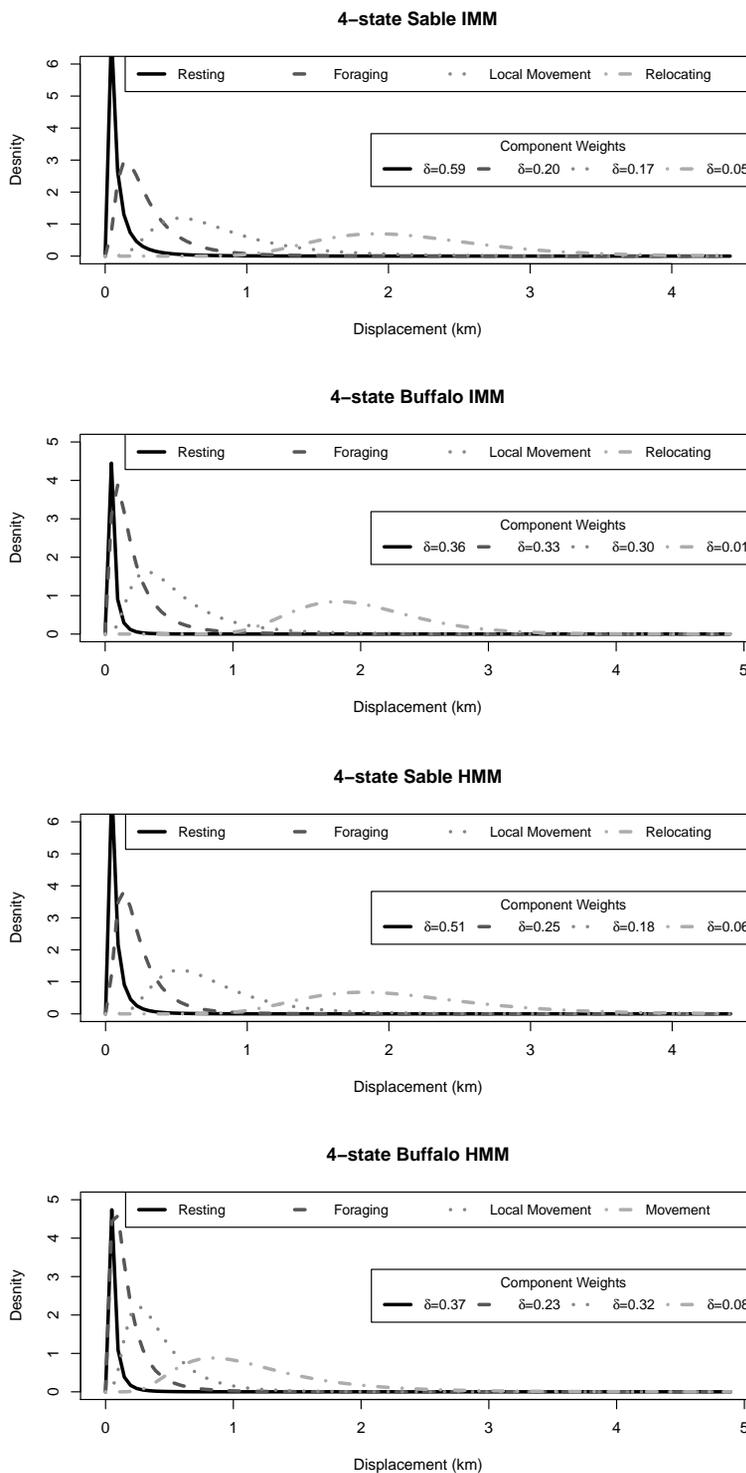
**Table 9:** State allocation for Sable and Buffalo using IMM and HMM.

		Sable			
		State allocation - IMM			
		State 1	State 2	State 3	State 4
State allocation - HMM	State 1	1495	10	0	0
	State 2	259	487	32	0
	State 3	0	82	494	13
	State 4	0	0	31	161

		Buffalo			
		State allocation - IMM			
		State 1	State 2	State 3	State 4
State allocation - HMM	State 1	3718	402	6	0
	State 2	260	1788	282	0
	State 3	0	1266	2787	1
	State 4	0	0	591	140



**Figure 1:** State allocation by time of day using 4-state Mixture Model and Hidden Markov model - Sable and Buffalo.



**Figure 2:** Fitted distributions for IMM and HMM models - Sable and Buffalo. Distributions are shown unweighted, and would be adjusted by the Component Weights shown in the Figure to form the fitted distribution.

## References

- BARRAQUAND, F. AND BENHAMOU, S. (2008). Animal movements in heterogeneous landscapes: Identifying profitable places and homogenous movement bouts. *Ecology*, **89** (12), 3336–3348.
- BIVAND, R., KEITT, T., AND ROWLINGSON, B. (2016). rgdal: Bindings for the Geospatial Data Abstraction Library.  
URL: <https://CRAN.R-project.org/package=rgdal>
- BOLKER, B. M. (2008). *Ecological Models and Data in R*. Princeton University Press: Princeton.
- CAIN, I., J. W., OWEN-SMITH, N., AND MACANDZA, V. A. (2012). The costs of drinking: comparative water dependency of sable antelope and zebra. *Journal of Zoology*, **286**, 58–67.
- CALENGE, C. (2006). The package adehabitat for the R software: tool for the analysis of space and habitat use by animals. *Ecological Modelling*, **197**, 1035.
- CAPPE, O., MOULINES, E., AND RYDEN, T. (2005). *Inference in Hidden Markov Models*. Springer: New York.
- DRAGON, A. C., BAR-HEN, A., MONESTIEZ, P., AND GUINET, C. (2012). Comparative analysis of methods for inferring successful foraging areas from Argos and GPS tracking data. *Marine Ecology Progress Series*, **452**, 253–267.
- EPHRAIM, Y. AND MERHAV, N. (2002). Hidden Markov processes. *IEEE Transactions on Information Theory*, **48** (6), 1518–1569.
- EVERITT, B. S. AND HAND, D. J. (1981). *Finite Mixture Distributions*. Chapman & Hall: London.
- FRANKE, A., CAELLI, T., AND HUDSON, R. J. (2004). Analysis of movements and behaviour of caribou (*Rangifer tarandus*) using hidden Markov models. *Ecological Modelling*, **173**, 259–270.
- FRANKE, A., CAELLI, T., KUZYK, G., AND HUDSON, R. J. (2006). Prediction of wolf (*Canis lupus*) kill-sites using hidden Markov models. *Ecological Modelling*, **197**, 237–246.
- GOODALL, V. L. (2014). *Statistical Approaches Towards Analysing Ungulate Movement Patterns in the Kruger National Park*. Thesis, Ph.D, University of the Witwatersrand, Johannesburg, South Africa.
- GUTENKUNST, R., NEWLANDS, N., LUTCAVAGE, M., AND EDELSTEIN-KESHET, L. (2007). Inferring resource distributions from Atlantic bluefin tuna movements: An analysis based on net displacement and length of track. *Journal of Theoretical Biology*, **245**, 243–257.
- LANGROCK, R., HOPCRAFT, J. G. C., BLACKWELL, P. G., GOODALL, V. L., KING, R., NIU, M., PATTERSON, T. A., PEDERSEN, M. W., SKARIN, A., AND SCHICK, R. S. (2014). Modelling group dynamic animal movement. *Methods in Ecology and Evolution*, **5** (2), 190–199. doi:10.1111/2041-210x.12155.
- LANGROCK, R., KING, R., MATTHIOLPOULOS, J., THOMAS, L., FORTIN, D., AND MORALES, J. M. (2012). Flexible and practical modeling of animal telemetry data: hidden Markov models and extensions. *Ecology*, **93** (11), 2336–2342.
- MCCLINTOCK, B. T., KING, R., THOMAS, L., MATTHIOPoulos, J., MCCONNELL, B. J., AND MORALES, J. M. (2012). A general discrete-time modeling framework for animal movement using multistate random walks. *Ecological Monographs*, **82** (3), 335–349.
- MCLACHLAN, G. AND PEEL, D. (2000). *Finite Mixture Models*. John Wiley & Sons: New York.
- MCLACHLAN, G. J. AND BASFORD, K. E. (1988). *Mixture Models Inference and Applications to Clustering*. Marcel Dekker: New York.

- MORALES, J. M., HAYDON, D. T., FRIAR, J., HOLSINGER, K. E., AND FRYXELL, J. M. (2004). Extracting more out of relocation data: building movement models as mixtures of random walks. *Ecology*, **85**, 2436–2445.
- OWEN-SMITH, N. (2013). Daily movement responses by African savanna ungulates as an indicator of seasonal and annual food stress. *Wildlife Research*, **40** (3), 232–240. doi:<http://dx.doi.org/10.1071/WR13024>.
- OWEN-SMITH, N., FRYXELL, J., AND MERRILL, E. H. (2010). Foraging theory upscaled: the behavioural ecology of herbivore movement. *Philosophical Transactions of the Royal Society B*, **365**, 2267–2278.
- OWEN-SMITH, N. AND GOODALL, V. (2014). Coping with savanna seasonality: comparative daily activity patterns of African ungulates as revealed by GPS telemetry. *Journal of Zoology*, **293**, 181–191. doi:10.1111/jzo.12132.  
URL: <http://dx.doi.org/10.1111/jzo.12132>
- OWEN-SMITH, N., GOODALL, V. L., AND FATTI, L. P. (2012). Applying mixture models to derive activity states of large herbivores from movement rates obtained using GPS telemetry. *Wildlife Research*, **39**, 452–462.
- PATTERSON, T. A., BASSON, M., BRAVINGTON, M. V., AND GUNN, J. S. (2009). Classifying movement behaviour in relation to environmental conditions using hidden Markov models. *Journal of Animal Ecology*, **78**, 1113–1123.
- PATTERSON, T. A., THOMAS, L., WILCOX, C., OVASKAINEN, O., AND MATTHIOPOULOS, J. (2008). State-space models of individual animal movement. *Trends in Ecology and Evolution*, **23** (2), 87–94.
- R CORE TEAM (2016). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.  
URL: <https://www.R-project.org/>
- SCHICK, R. S., LOARIE, S. R., COLCHERO, F., BEST, B. D., BOUSTANY, A., CONDE, D. A., HALPIN, P. N., JOPPA, L. N., MCCLELLLEN, C. M., AND CLARK, J. S. (2008). Understanding movement data and movement processes: current and emerging directions. *Ecology Letters*, **11**, 1338–1350.
- TUCKER, B. C. AND ANAND, M. (2005). On the use of stationary versus hidden Markov models to detect simple versus complex ecological dynamics. *Ecological Modelling*, **185**, 177–193.
- ZUCCHINI, W. AND MACDONALD, I. L. (2009). *Hidden Markov Models for Time Series: An Introduction Using R*. Chapman & Hall/CRC: Boca Raton.

