Hidden Markov Models as Tools to Identify Seabird Foraging Areas

Edwin Onuonga

s1628873@ed.ac.uk

THE UNIVERSITY *of* EDINBURGH
School of Mathematics

Overview

Marine protection areas are essential for minimizing human impact on seabird habitats. For governments to establish these, we need to identify foraging areas used by seabirds. The Joint Nature Conservation Committee (JNCC) has provided location and behavioural data for common (*Sterna hirundo*), Arctic (*Sterna paradisaea*), roseate (*Sterna dougallii*) and sandwich (*Sterna sandvicensis*) terns.

Hidden Markov Models (HMMs) are a commonly used method in statistical ecology for modelling the behaviour of animals, which the JNCC wish to apply to terns.

The main goals of the project were to use the data provided by the JNCC to:

- apply HMMs to identify tern foraging/non-foraging areas by predicting labels corresponding to each step of a recorded GPS track (via moveHMM [1]),
- evaluate the effectiveness of the HMMs by comparing predicted labels with known labels derived from recorded continuous and instantaneous tern behaviour,
- conduct and compare year-by-year and species-by-species analyses.

Data

The JNCC data consists of 2010/2011 GPS locations of a boat following terns from each of the four species leaving and returning from foraging trips at Coquet Island near Amble, Northumberland, where there are various seabird colonies.

Additionally, various categories of instantaneous and continuous behaviour were recorded. This data is used to determine whether a tern is foraging or not, and is used for evaluation. Instantaneous behaviour are one-off events such as a plunge dive, and continuous behaviour is recorded at every step of the track.

Our work focuses on the chick-rearing breeding stage, where terns forage for prey to return as food to their chicks which are incapable of flight.

Methodology

Mapping Behaviour to Foraging/Non-Foraging States
We only consider foraging terns to exhibit the active search (AS) continuous behaviour. All continuous behaviour apart from AS is considered to belong to the non-foraging state. For instantaneous behaviour, we expect foraging terns to perform kleptoparasitism (KP), a plunge dive (PD), surface peck (SP) or surface dip (SD) — the latter three being commonly used tactics by many seabird species [2].

Observation Modelling
Rather than using latitude/longitude data from the GPS directly, we represent the path of a tern as a sequence of step lengths and angles as this is simpler to model[a].

Step lengths and angles are modelled with gamma and von Mises [3] distributions.

Hidden Markov Models
For terns, step length and angle observations may reveal information about whether a tern is foraging or not, and a HMM is able to capture this information by modelling transitions between the behavioural states (with an initial state distribution π and transition matrix P), as well as modelling how likely a given observation was generated by each state (with emission distributions E). Each state has one emission distribution for step lengths and one for angles.
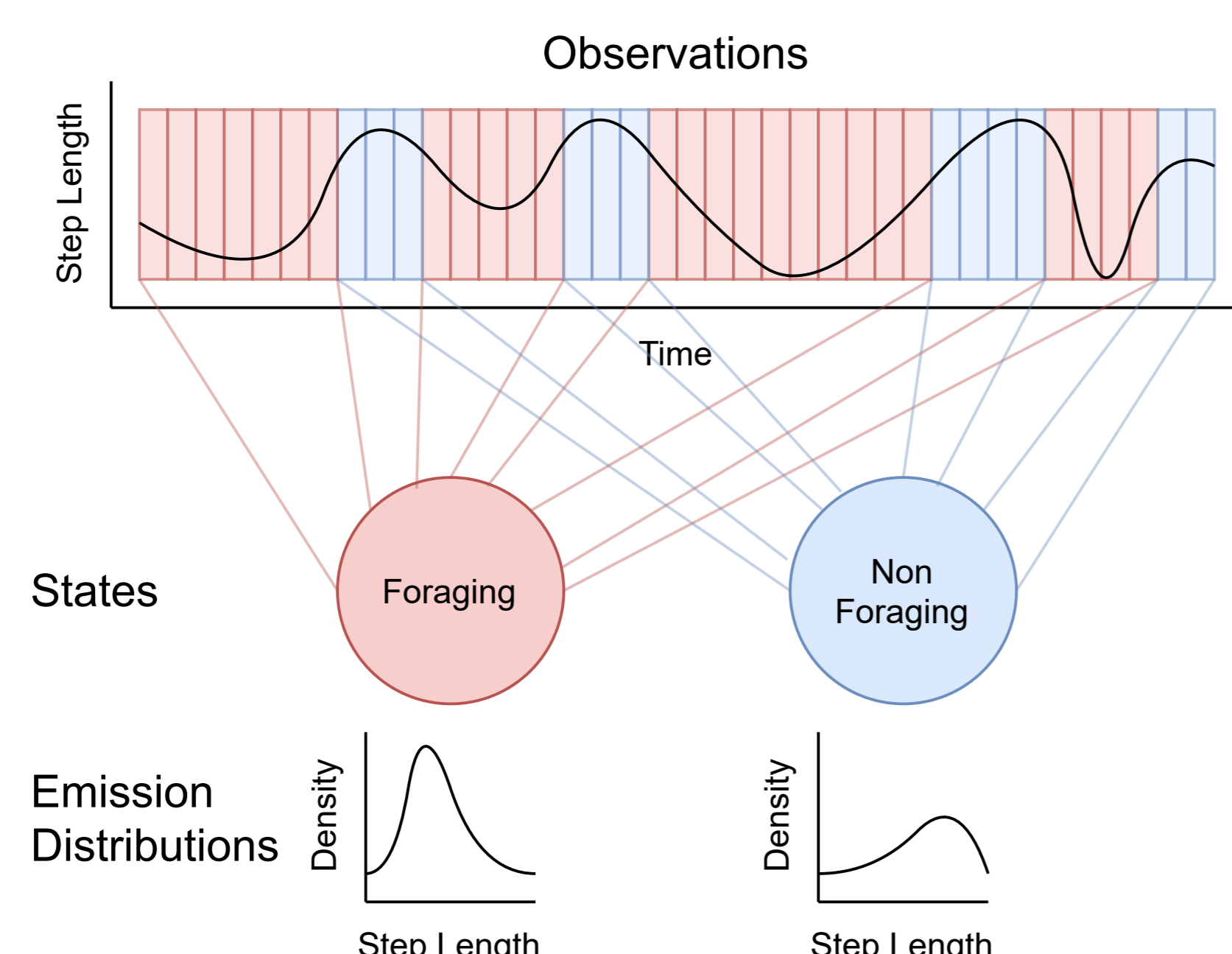

Figure 1. An example of how a two-state HMM can be used to decode a sequence of step lengths into a sequence of foraging/non-foraging states.

Results

For each tern track, we compare the predicted state at each 2s interval with the corresponding known foraging/non-foraging label.

Year	Metric	Species			
		Common	Arctic	Sandwich	Roseate
2010		$(n=10)$	$(n=5)$	$(n=3)$	$(n=1)$
	Sensitivity	0.800	0.532	0.337	0.741
	Specificity	0.868	0.885	0.174	0.702
2011		$(n=42)$	$(n=49)$	$(n=70)$	$(n=31)$
	Sensitivity	0.674	0.735	0.564	0.573
	Specificity	0.884	0.844	0.888	0.840

Table 1. Sensitivity and specificity of foraging/non-foraging predictions. n represents the number of terns for a given species and year.

Apart from the 2010 sandwich tern model, all models can accurately predict non-foraging areas (high specificity between 0.7–0.9). The 2010 Sandwich tern model failed to achieve a good fit despite attempting various initial parameters.

We report sensitivity and specificity, but focus on sensitivity as it tells us the proportion of observations that we correctly predicted as foraging (out of the known foraging observations).

As most models have lower sensitivity than specificity, this suggests that the HMMs may be slightly over-predicting non-foraging states.

Nevertheless, the sensitivity is within the 0.5–0.8 range for all but the 2010 sandwich tern model, suggesting that HMMs are effective, but can be improved.

Based on 2011 data which is more reliable, sandwich and roseate terns are more difficult to correctly predict foraging areas for (low sensitivity, 0.5–0.6).
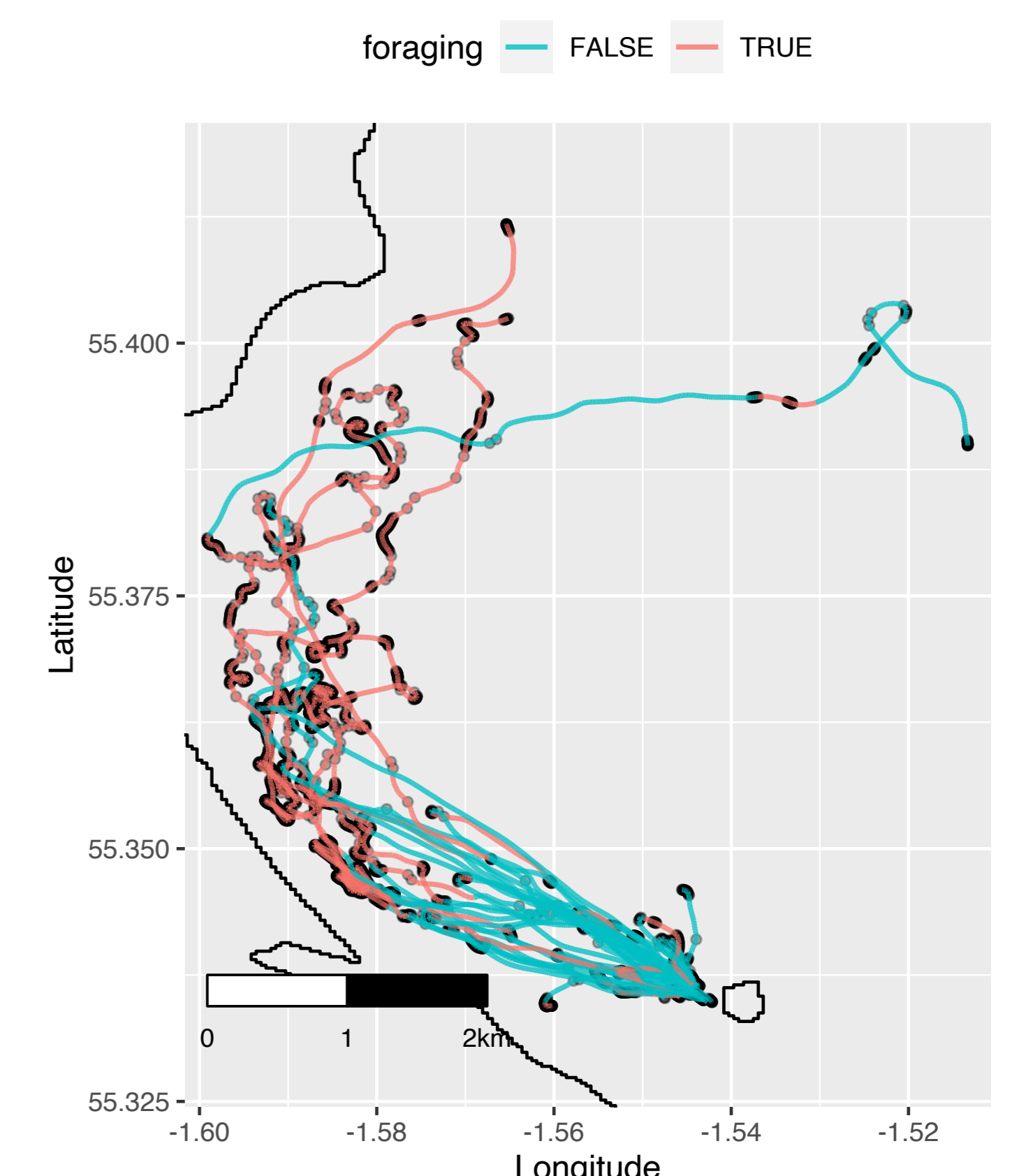

Figure 2. Roseate tern tracks in 2011. Black circles are foraging predictions, red/blue indicate foraging/non-foraging states based on observed behaviour.

By inspection of plots such as Figure 2 for each species and year, there were some differences between species, such as Arctic, sandwich and roseate terns tending to travel further distances to forage in 2010, but these differences were small. In particular for 2010, it is difficult to draw conclusions due to small sample sizes.

Roseate terns in 2011 tended to spend more time foraging than the other species (as shown by the relatively high proportion of red in Figure 2), and also almost always followed the same route to forage along the coast of Amble.

Conclusion

HMMs are an effective approach for identifying foraging areas for the four tern species, but appear to be less effective for sandwich and roseate terns in 2011.

While there are differences between 2010 and 2011, the comparison is inconclusive due to limited data in 2010 as well as a poorly fitted 2010 sandwich tern model. There also appear to be small behavioural differences by species, but this may also be due to prey availability or even weather conditions, as suspected in Figure 2.

Improvements: Experiment with covariates, better data collection (fairer sample sizes across species/years, GPS location of birds rather than boat e.g. by tagging), compare with incubation to see effects of breeding stage.

References

[1] T. Michelot, R. Langrock, and T. Patterson. moveHMM: an R package for the statistical modelling of animal movement data using hidden Markov models. *Methods in Ecology and Evolution*, 7:1308–1315, 2016.

[2] Atlantic Fleet Training and Testing. Environmental impact statement/overseas environmental impact statement. pages 3.9–51, August 2018. https://media.defense.gov/2018/Aug/17/2001955656/-1/-1/1/3.09_AFTT_DRAFT%20EIS_OEIS_BIRDS%20AND%20BATS.PDF.

[3] von Mises distribution. *Wikipedia* (Article). https://en.wikipedia.org/wiki/Von_Mises_distribution. Accessed: 2021-08-14.