Marine Mammal Scientific Support Research Programme MMSS/001/11

MR 5.2:

Report

Activity classification using state space modelling

Sea Mammal Research Unit

Report to
Scottish Government

July 2015 [version F1]

Russell, D. J. F

Sea Mammal Research Unit, Scottish Oceans Institute, University of St Andrews, St Andrews, Fife, KY16 8LB.

Editorial Trail			
Main Author	**Comments**	**Version**	**Date**
B. McConnell	collation of individual sub-task reports	V1	25/01/2015
B. McConnell D. Thompson	addition of information to original report	V2	11/03/2015
Marine Scotland	comments	V3	22/04/2015
P. Irving J. Williamson	review	V4	27/04/2015
B. McConnell	review	V5	16/06/2015
B. McConnell	review	VF1	17/07/2015
O. Racu	final copy editing	VF2	12/08/2015

Citation of report

Russell, D. J. F. (2015) Activity classification using state space modelling. Sea Mammal Research Unit, University of St Andrews, Report to Scottish Government, no. MR 5.2, St Andrews, 10pp.

Contents

1 Executive summary

The state-space model developed for defining activity budgets on a coarse resolution (6 hours) was used to define activity budgets at a fine temporal resolution (2 hours) based on both geo-centric and hydro-centric movements. Hydro-centric movements (active movement of seals through the water) were estimated by deleting vectors of current from the geographic movement data. ARGOS data were excluded because the temporal resolution of the location data prohibits the delineation of fine resolution activity budgets.

Activity budgets at a 2 hour resolution were successfully defined using the data from 90% of the 76 GPS/GSM tags considered. Problems apparent when defining activity budgets on the coarse resolution (including the estimation of only one diving state) appeared to be reduced when considering the fine resolution. Although there were some spatial differences in apparent foraging on the two resolutions, the activity budgets defined on the coarse resolution did not appear to be subject to consistent biases.

A significantly higher proportion of time was estimated to be devoted to foraging when hydro-centric rather than geo-centric movements were considered, indicating the importance of incorporating data on water movement when modelling activity budgets in marine animals.

2 Introduction

A framework for defining activity budgets using telemetry data was developed: resting and diving periods were assigned using behavioural summary data and then diving behaviour was assigned to foraging and travelling based on movement characteristics within a state-space model (Russell *et al.*, 2015). Activity budgets were defined for 63 grey (*Halichoerus grypus*) and 126 harbour (*Phoca vitulina*) seals, but for 20% of harbour seals only one diving state could be defined. This was probably because in some areas, harbour seals only engage in short distance trips and in these cases travelling periods may last less than 6 hours. Data from tags deployed in areas of high tidal currents (i.e. Pentland Firth and Strangford Lough) were excluded because they may lead to unreliable movement-based estimates of state (foraging or travelling; Gaspar *et al.*, 2006). For example, if the current is flowing east at a rate of 1 ms^{-1} and a seal is also moving east at a speed of 0.5ms^{-1}, then geographically the seal is moving at 1.5ms^{-1} which is a speed which would often be associated with travelling. However, the seal is only moving through the body of water at 0.5ms^{-1} and thus is actually more likely to be foraging. In contrast, if the seal is moving west at a speed of 1.5ms^{-1} which is a speed associated with travelling, but the eastward movement of the water means geographically the seal is only moving west at 0.5ms^{-1}, the interval would probably be misclassified as foraging. It was assumed that it is more common for activity to be misclassified as travelling rather than foraging, as it is expected that seals rarely move against a current when attempting to travel but it would be expected for them to forage in tidal areas.

Using data from GPS/GSM tags, for which behavioural summary data are available at a 2 hour resolution, attempts were made to define activity budgets at a 2 hour resolution for 76 individuals. Of the 76 individuals, data from 30 were available when the coarse resolution analysis was carried out and thus attempts were also made to define activity budgets at a 6 hour resolution for these individuals; the behavioural summary data constituted aggregated 2 hour summary data. At a 6 hour resolution activity budgets were not successfully defined for five of 30 seals, either due to convergence issues or because only one diving state could be defined.

In this subtask, how estimated activity budgets are affected by temporal scales (two and six hour intervals) and by movement being transformed into hydro-space was examined. Specifically, the following questions were addressed:

- Can activity budgets of individuals, which could not be completely defined at a 6 hour resolution (e.g. convergence issue or only one diving state could be defined), be successfully defined at a 2 hour resolution?
- Is the proportion of time foraging consistently under- or over-estimated at a resolution of 6 hours compared to 2 hours?
- At a 2 hour resolution, is the proportion of time foraging consistently under- or over-estimated using geo-centric compared with hydro-centric movement data?
- Do the intervals estimated to be foraging using geo-centric movement data remain foraging when using hydro-centric movement data?

3 Methods

Polpred (http://noc.ac.uk/using-science/products/software/polpred/polpred) provides predicted current direction and speed (water displacement) data. The current movement was predicted at a set of interpolated telemetry locations: (1) between all observed telemetry observations that were less than 15 minutes apart (2) between regularised 15 minute interpolated locations. This ensured that a single current value at the mid-point between two observed locations far apart in time was not assigned to the whole time period during which a seal may have passed through areas of low and high current flow. It was assumed that the current at the midpoint applied to an entire period (a maximum of 15

minutes) and by deleting the tidal vectors from the geographical vectors, the active (hydro-centric) movement for each period was calculated. New locations in X and Y starting at coordinates 0, 0 using these displacements were then generated. If there was no current data available for a period, the non-adjusted geographical displacement data were used. The resulting data set was referred to as the hydro-centric movement dataset. Both the hydro-centric and geo-centric data were interpolated onto a 2 hour resolution. Two-hour intervals were considered non-estimable if no behavioural summary data were available or if there were > 6 hours between observed locations surrounding the interpolated locations. Individuals were only included if ≥ 75% of the intervals were estimable and if ≥ 75% of their time at-sea could be adjusted for the presence of currents. In some areas, e.g. Kyle Rhea, current data are not available so these individuals were excluded.

The state-space models were run twice for each individual, using the geographical (geo-centric) movement data and then the hydro-centric movement data. Two chains starting at different initial values with a burn in of 50,000 iterations were used. Convergence was judged by visual inspection of the chains and using the Gelman-Rubin (gbr) statistic. Usually 50,000 iterations were used for the posterior distributions but 50,000 more iterations were run if the gbr statistic was not 1.0. Please see Russell *et al., (*2015), for details of the state-space model.

4 Results

Appropriate data (as defined above) were available to define both geo- and hydro- centric activity budgets for 76 seals. Eight of these seals were subsequently excluded: there were convergence issues for six of the seals and for two of the seals using hydro-centric movement data there was a very low proportion of diving intervals apportioned to foraging and visually the model appeared to have misclassified intervals as travelling (see Discussion). Thus the final sample size for comparison between geo- and hydro- centric movements was 68.

4.1 Can activity budgets of individuals, which could not be completely defined at a 6 hour resolution (e.g. convergence issue or only one diving state could be defined), be successfully defined at a 2 hour resolution?

Activity budgets were successfully defined for all 30 individuals for which attempts to define activity budgets at a 6 hour resolution had previously been carried out. Given that at a 6 hour resolution, activity budgets could not be completely defined for five of these individuals there is a suggestion that to some degree problems when using a 6 hour resolution were resolved when using the finer 2 hour resolution. The 25 remaining seals for which activity budgets were successfully defined at a 6 and 2 hour resolution comprised six harbour seals and 19 grey seals (17 moulted pups).

4.2 Is the proportion of time foraging consistently under- or over-estimated at a resolution of 6 hours compared to 2 hours?

There were on average 7% more non-estimable intervals at a 2- compared to 6 hour (95% CI: 5.7-9.1%) resolution as a result of the stricter threshold on exclusion of intervals when there was a gap between observed locations (≥ 6 hours for the 2 hour resolution compared to ≥ 12 hours for the 6 hour resolution). Furthermore, at a 6 hour resolution, the 2 hour summary periods are aggregated and often a majority behaviour (i.e. resting or diving) could be defined from only two 2 hour intervals (four hours in total). At a 2 hour resolution, a missing 2 hour summary period would result in a non-estimable interval.

As expected there were areas of foraging predicted using data at a 2 hour resolution that were not apparent at a 6 hour resolution (e.g. Figure 1). However, areas of intense foraging were, for the most part, the same for both resolutions. Indeed, there was not a significant difference in the proportion of foraging on a 6 hour compared to 2 hour resolution (paired t-test: $t_{1,23} = -1.45$, $P = 0.16$). This similarity, despite some spatial differences in foraging areas, is probably because both intervals that

were classified as foraging or travelling at a 6 hour resolution, are now a mixture of foraging and travelling intervals at a 2 hour resolution.

Activity estimated at a 2 hour resolution frequently shows an increased proportion of apparent foraging near to haul-out sites. This is likely to be because previously some short (< 3 hour) periods of inter haul out diving activity (when tidal haul out sites are unavailable) that would have been classed within a 6 hour haul-out interval are classed as foraging intervals at a 2 hour resolution. If seals are simply waiting for haul out sites to become exposed and not actively foraging, the activity budgets calculated at this finer resolution may result in an over estimate of foraging near haul-out sites.

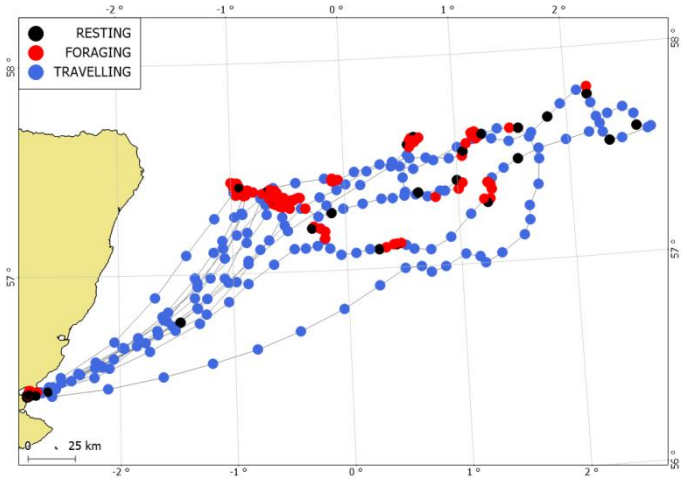
4.3 At a 2 hour resolution, is the proportion of time foraging consistently under- or over-estimated using geo-centric compared with hydro-centric movement data?

The same behavioural summary data are used for the hydro and geo-centric data so the proportion of time spent resting is the same for both models. Between 3.7 and 8.5% (95% CIs) more time was spent foraging when considering hydro-centric rather than geo-centric movements; this difference did not vary with species (LM: $F_{1,66}$ =0.01, P = 0.91).

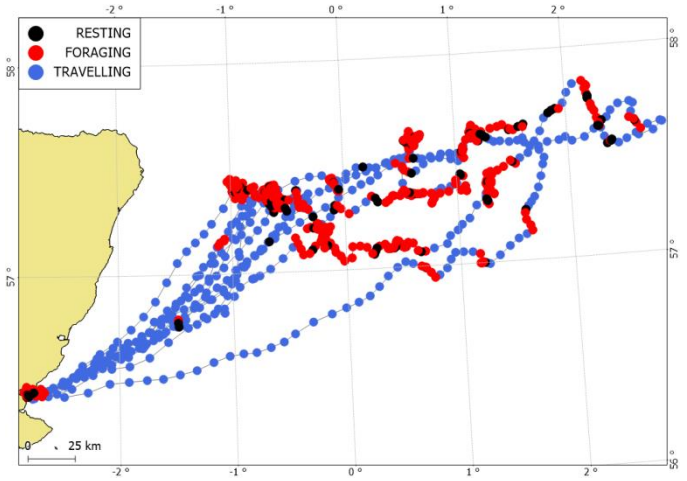
4.4 Do the intervals estimated to be foraging using geo-centric movement data remain foraging when using hydro-centric movement data?

In addition to the differences in the proportion of time spent foraging and travelling using hydro- and geo-centric movement data, it was also interesting to see how these differences were manifested. The majority (median: 96%, 95% CI: 89-99%) of intervals defined as foraging intervals in the geo-centric model were also defined as foraging intervals in the hydro-centric model. Many intervals (median 69%, 95% CI: 52-94%) that were assigned to travelling in the geo-centric model became foraging in the hydro-centric model. Thus important foraging areas defined using geo-centric movement data are likely to be valid, but some important foraging areas may be missed when using geo-centric movement data. These spatial differences in activity were even apparent in areas that do not have particularly fast currents (e.g. an individual seal in south-east Scotland, Figure 1).

(a)



(b)



(c)

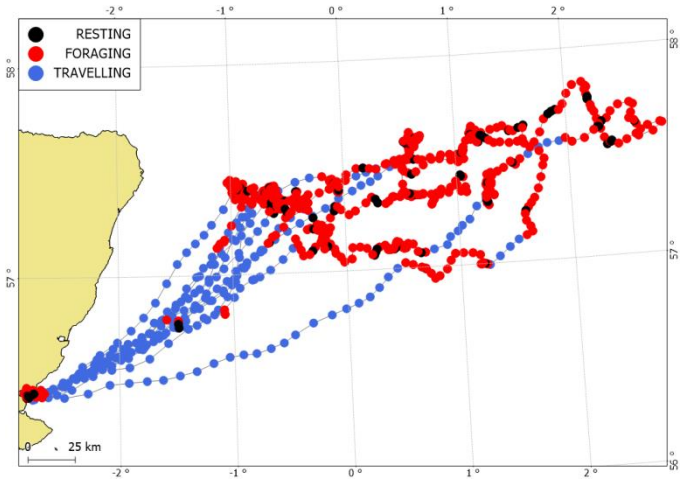


Figure 1. The track of one grey seal with foraging and travelling states defined using geo-centric movement data on a 6 hour (a) and 2 hour (b), and using hydro-centric movement data on a 2 hour resolution (c). The states estimated using the hydro-centric movement data (c) are overlaid onto the geo graphical locations (the hydro-centric movement data are not shown).

5 Discussion

5.1 Caveats

Even at a 2 hour resolution, there were convergence problems for 8% of the individuals (n=6) in the hydro and/or the geo-centric model. However, this included two individuals for which there was less than 10 days data; there may not have been sufficient representation in the data to characterise all three states. Furthermore, for two of the individuals (including one which had less than 10 days of data), convergence problems were only evident in the model using geo-centric movement data. Speed distributions for foraging and travelling may have been difficult to characterise due to the influence of current in some places. For some of the remaining problematic individuals (n=3), convergence issues may be resolved by running more iterations and/or changing the initial values. However, some showed evidence of label switching, an identifiability issue. In the current model, the speed was constrained to be faster when travelling than foraging, but future developments should include constraining the model to have tighter turning angles when foraging compared to travelling. This may solve the identifiability issues.

For two individuals tagged in the Wash, south-east England, although convergence was obtained, they subsequently had to be excluded due to a lack of foraging intervals in hydro-centric space. It is possible that this problem was a result of the creation of three true movement states which were constrained into two available states. One state (assigned to foraging) was associated with very little movement and may have been the result of the individuals actually resting on the bottom (no movement); then true foraging and travelling were assigned into the travelling state. This problem was not apparent when using the geo-centric movement data.

5.2 Further work

It is important to establish whether the short-term inter haul-out behaviour associated with waiting for tidal haul-out sites to become available is actually associated with foraging activity. To do this the diving behaviour should be examined to establish (a) whether the seals are spending prolonged periods on the bottom, often associated with foraging, and (b) whether they are moving in horizontal space or simply resting on the sea bed.

There is a need to establish the traits of the individuals for which there are problems classifying activity budgets. For some individuals, three states within diving may be required if they appear to rest on the sea bed. Furthermore, there is a need to constrain tighter turning angles when foraging compared to travelling and also possibly utilise block updating to assist with identifiability issues.

The hydro-centric movement data are dependent on the accuracy of the current data: current speed is likely to be dependent on the depth of the individual and the spatial resolution of the current data which means that in some places currents may not be accurate or are not available at all (e.g. Kyle Rhea).

6 Conclusion

Activity budgets were successfully defined at a 2 hour resolution for 90% of individuals for which there were appropriate data. This was a higher proportion than for activity budgets defined at a 6 hour resolution, suggesting that some individuals were exhibiting distinct behaviours on a finer time scale than 6 hours. When activity budgets were successfully defined and compared at both temporal resolutions, there did not appear to be any systematic bias in the activity budget estimates at a 6 hour resolution.

The results also demonstrated the importance of accounting for current speed even in areas that are not particularly high energy. Not accounting for such currents is likely to result in an underestimate of the proportion of time foraging.

Using activity data at a 6 hour resolution, predictions of habitats and areas important for foraging are likely to be reliable, but are likely to be an underestimate of the true area important for foraging. Since 2006 almost all tags deployed by SMRU are GPS/GSM rather than ARGOS tags, so finer resolution activity budgets that account for current speed, can be defined for future data.

7 References

Gaspar, P., Georges, J.-Y., Fossette, S., Lenoble, A., Ferraroli, S., & Le Maho, Y. (2006) Marine animal behaviour: neglecting ocean currents can lead us up the wrong track. *Proceedings of the Royal Society B-Biological Sciences*, **273**, 2697–2702.

Russell, D., McClintock, B., Matthiopoulos, J., Thompson, P., Thompson, D., Hammond, P. & McConnell, B. (2015) Intrinsic and extrinsic drivers of activity budgets in sympatric grey and harbour seals. *Oikos*, in early view.