

Supplementary Information

**Differences in foraging range between white-tailed tropicbirds
breeding on inner and outer Seychelles islands**

**Ensanyar-Volle O., Appoo J., Bunbury N., Clucas G., Khan N., Rocamora G.,
Sanchez C., Fayet A.L.**

Table of Contents

Supplementary Information	1
Text S1. Hidden Markov Models	2
Figure S1	2
Figure S2	3
Figure S3.	3
Text S2. Kernel density estimation	4
Text S3. General Additive Models	4
Figure S4	5
Text S4. Trip metrics	6
Table S1	6
Text S5. Trip duration based on nest observations	6
Text S6. Diet	7
Table S2	7
Text S7. Foraging trips	8
Figure S5	8
References	9

Text S1. Hidden Markov Models

We used Hidden Markov models (HMMs) to identify foraging behaviour (Bennison et al., 2018; Dean et al., 2013). HMMs were applied to the interpolated GPS data using the *moveHMM* package (Michelot et al., 2016). HMMs use observable information to deduce unobservable sequences of states. In our case, we wanted to identify behaviours (sequences of states) from the GPS locations (observable information). We calculated the step length and the turning angle between every point, and we assumed the distribution of these two parameters were determined by a mode of movement and therefore a behaviour. We used Gamma and Von Mises distribution to analyse respectively step length and turning angle. In order to avoid abnormal information, every trip was subdivided in different segments of regular points without gaps. To define the optimal number of states, different models were trained with an increasing number of states and the likelihood of each model was calculated. The greatest increase of likelihood was between the 2-state model and the 3-state model (Annexe 1, Figure S1) and revealed the existence of a distinct knee point (Dean et al., 2013). The 3-state model was thus chosen to balance the trade-off between model accuracy and complexity. This model defines three distributions of angles and step lengths (Annexe 1, figures S2 and S3), assigns the probability for every GPS point to be in each state, and keep the most probable. All locations are then split into three states corresponding to three modes of movement/behaviours. The three behaviours were interpreted as resting on the water surface (low turning angle, low step length), foraging (high turning angle, intermediate step length), and relocation/commuting (low turning angle, high step length) as in (Dean et al., 2013; Oppel et al., 2015). Proportion of time spent in every behaviour were then calculated per trip and the mean value was calculated for each colony, breeding stage, and sex. The average was weighted with the number of points per trip. The proportion of activities was also assessed per portion of the day for each colony with steps of 30 minutes.

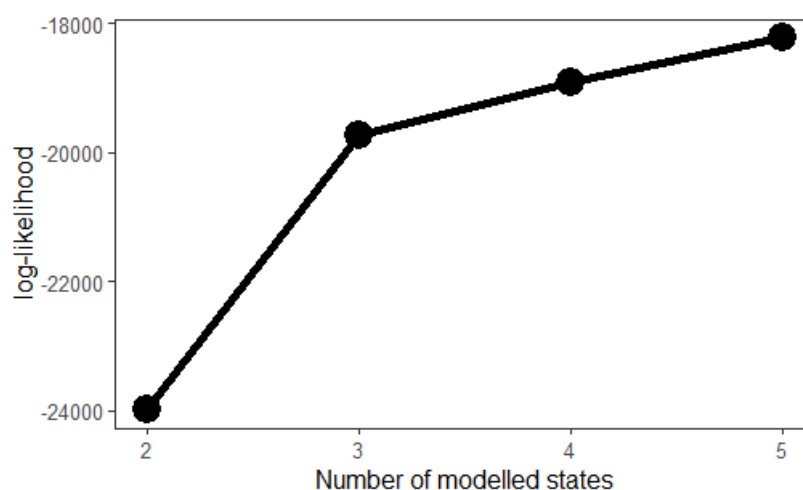


Figure S1. Log likelihood calculated for Hidden Markov Models with between 2 and 5 states. The greatest increasing is between the 2-state model and the 3-state model revealing a distinct knee-point.

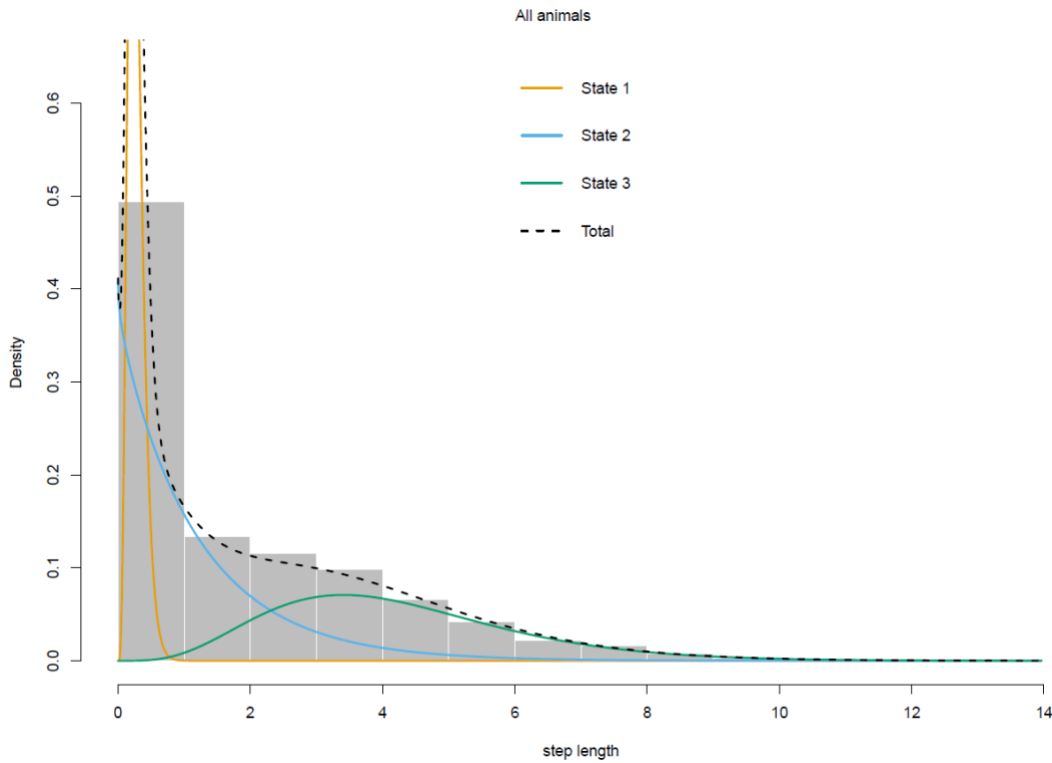


Figure S2. Histogram of step length with fitted distribution of the 3 states of the 3-state HMM State 1 (in orange) represent low step length distribution, state 2 (in blue) intermediate and state 3 (in green) represent high step length distribution. Grey bars are the distribution of all step length.

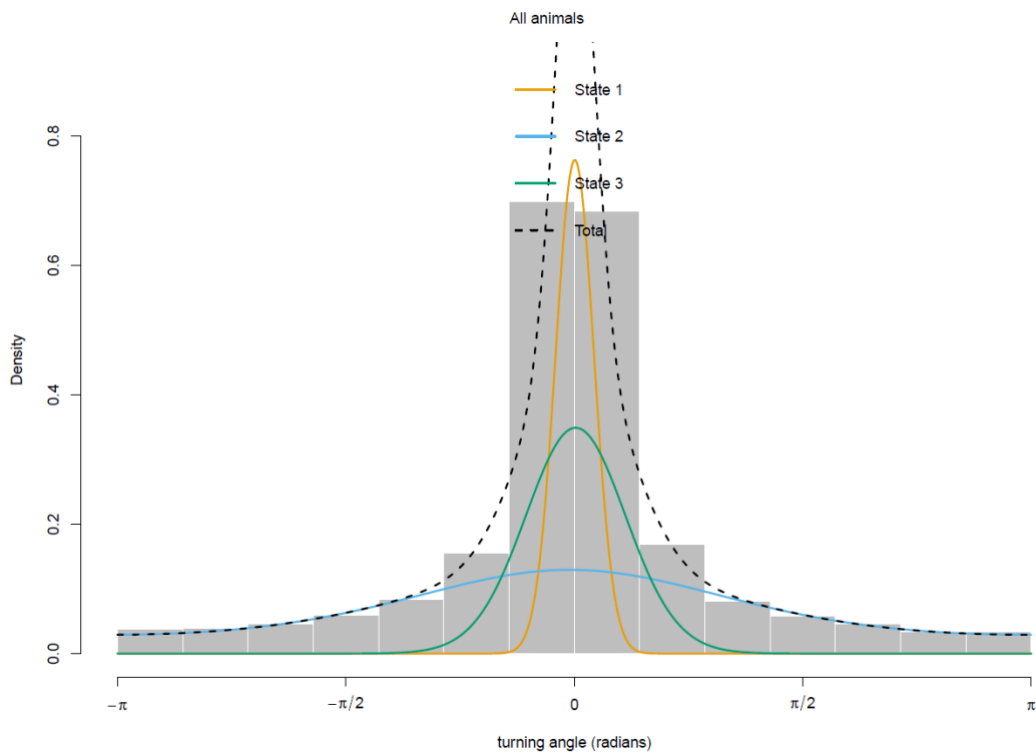


Figure S3. Histogram of turning angle with fitted distribution of the 3 states of the 3-state HMM State 1 (in orange) represent low turning angle distribution, state 2 (in blue) high turning angle distribution and state 3 (in green) represent intermediate distribution. Grey bars are the distribution of all turning angles.

Text S2. Kernel density estimation

Kernel density estimation were used to calculate utilization Distributions (UD) with the package *adehabitatHR* (Calenge, 2019). The function *KernelUD* estimate the UD for the center of each pixel which is define by the grid parameter. The grid controls the resolution of the Kernel density map but barely impact the estimation (Silverman 1986). Different grid values were tested and the value 150 was selected after observing the results graphically. Three methods were tested to estimate the smoothing parameters h (bandwidth) : (1) The bandwidth reference (h_{ref}) is calculated from the distribution of point but tends to overestimate home-range size (Kernohan et al., 2001). (2) The Least Square Cross Validation (h_{lscv}) can help to identify the optimum value of h (Calenge, 2019) but this bandwidth selection failed to converge with our data. (3) An ad hoc (h_{ad_hoc}) smoothing parameters may be usefull (Schuler et al., 2014) to reduce oversmoothing by reducing gradually the value of h and visualising the home range. h is then adjusted manually. In our study, the value of h with the *ad_hoc* method and the *h_ref* method were very similar, we therefore applied the *h_ref* method. The estimation of the bandwidth was $h = 0.367$ for Aldabra and $h = 0.335$ for Aride.

Text S3. General Additive Models

For each trip, 5 pseudo-foraging points for each real location were randomly sampled from a radius around the colony equal to the maximum distance from the nest on that trip, and every point located on land was re-sampled. Environmental data were extracted at each pseudo-foraging location at a date randomly selected between the start and the end of the foraging trip.

Separate models were run for the two populations and the two different breeding stages. For each dataset, we followed a forward stepwise approach (Carneiro et al., 2016; Dehnhard et al., 2020) . All environmental variables were first tested individually, and the best single-variable model was selected (See after for model selection). We then added other environmental covariates to the best model and assessed whether this improved the model. Variables with a Spearman's ranks correlation above 0.5 were not used in the same model to avoid collinearity. When models had correlated variables, the one with the best explanatory power (the highest χ^2) was kept. Furthermore, distance from the nest was included in all models as an additional smoothing parameter to take into consideration the cost of travel. All points with at least one missing value were previously removed. We realised a k-fold cross validation with birds as data-folds to compare GAMs and evaluate the most useful environmental variable. For each of the three datasets, one bird was removed to create the testing set and all the other birds composed the training set.

Mean AUC were calculated on $N = 9$ models for Aldabra incubation (9 birds), $N = 8$ for Aride chick-rearing and $N = 13$ for Aride incubation. We ran GAMs models for each variable on the training set and calculated the area under the receiver curve (AUC, *pROC* package) for each model by testing each model on the testing set. This operation was repeated until each bird had been the testing set once. The mean AUC was next calculated for each model by averaging the AUC values across all train-test sets to assess each model's performance. AUC values <0.7 were considered as poor, $0.7-0.9$ reasonable, and >0.9 as very good (Pearce and Ferrier 2000).

The AUCs of more complex models were then compared with simpler nested models with paired t-tests to check if the addition of variables significantly increased the quality of the model (as Carneiro et al. 2016 and Dehnhard et al. 2020). For each dataset, the best most parsimonious model was then selected.

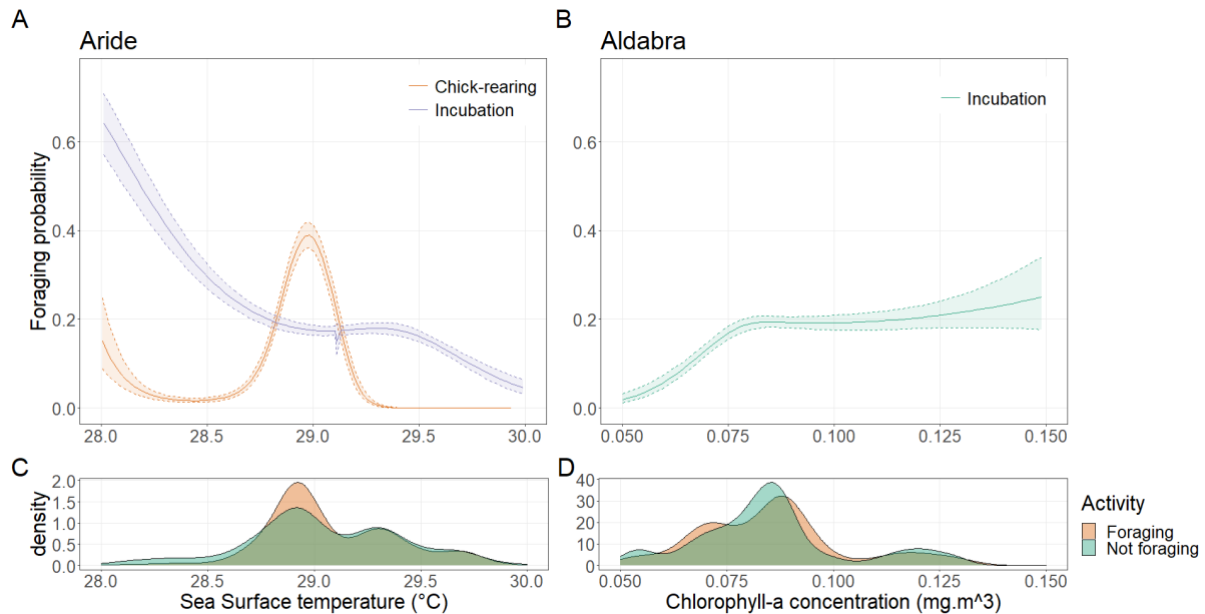


Figure S4. General additive model smoother response curves for the environmental covariate chlorophyll a concentration on Aldabra(A) and Sea Surface Temperature on Aride during chick-rearing and incubation (B). The model chlorophyll-a for Aldabra during incubation was run with chlorophyll-a and distance from nest as predictors. For Aride, GAMs were run with both SST and dist-nest as predictors and separately for each breeding stage. Dashed lines represent the 95% confidence interval. Higher y-axis values mean there is a stronger preference of the bird for the x-axis value. Kernel density estimation of activity point along the x axis for Aride (C) and Aldabra (D).

Text S4. Trip metrics

Table S1. Proportion of time spent in different behaviours on Aride and Aldabra, during chick-rearing and incubation. Each proportion is the proportion per trip, weighted by the number of GPS points per trips. Prop = proportion.

	n	Prop. of time spent foraging (\pm SD)	Prop. of time spent resting (\pm SD)	Prop. of time spent commuting (\pm SD)
Aride	34	0.42 \pm 0.20	0.26 \pm 0.13	0.32 \pm 0.27
Aldabra	26	0.48 \pm 0.31	0.25 \pm 0.14	0.27 \pm 0.29
Chick-rearing	30	0.43 \pm 0.27	0.20 \pm 0.09	0.37 \pm 0.30
Incubation	30	0.44 \pm 0.26	0.28 \pm 0.14	0.28 \pm 0.24
Female	29	0.43 \pm 0.28	0.25 \pm 0.14	0.32 \pm 0.31
Male	31	0.44 \pm 0.27	0.26 \pm 0.13	0.30 \pm 0.28

Text S5. Trip duration based on nest observations

In the discussion, we report some trip durations that we compare with trip durations at the same colonies in the 1960s and 1980s. These trip durations are based on observations at the nest and so are not the same values as given in Table 2, which are based on GPS data. The trip durations from nest-based observations are longer for several reasons. First, they include the duration between the moment the bird leaves the nest and the moment it returns to the nest. In contrast, the numbers in Table 2 come from the GPS data and focus only on the foraging part of the trip, i.e. outside a 2 km radius around the colony. As such, any time spent rafting around the colony before or after the foraging trip is not included. Secondly, trip duration calculated from GPS data does not include the naturally very long trips where the GPS would have died before the bird returned, so, as explained in the discussion, this is also likely an underestimate of true trip duration. We think these nest-based values are best for comparison with older studies as these were also based on nest observations and not on tracking.

Text S6. Diet

Table S2. Prey content of diet samples of WTTBs from Aldabra and Aride. DNA metabarcoding results from eight faecal samples (four from each colony) and visual identification results from seven regurgitate (four from Aldabra, three from Aride). We followed the same method for analysis of faecal samples as in (Fayet et al. (2021)). Note that we only used fish primers and so any non-fish species (e.g. squid) could not have been detected. In the last column, “No” means no other study found the species in WTTBs’ diet, “Yes” means it was found, “genus only” means the genus was found but not the species.

Colony	Sample type	Fish species (Latin name / common name)	Observed in other studies
Aldabra	Faecal	<i>Gymnothorax undulatus</i> / Undulated moray	No
Aldabra	Faecal	<i>Heteropriacanthus cruentatus</i> / Glasseye	No
Aldabra	Faecal	<i>Exocoetus volitans</i> / Tropical two-wing flyingfish	Yes
Aldabra	Faecal	Exocoetidae indet. / Flying fish	Yes
Aldabra	Faecal	<i>Hemiramphus</i> sp. / Halfbeak	Yes
Aldabra	Faecal	<i>Hemiramphus lutkei</i> / Lutke's halfbeak	Genus only
Aldabra	Faecal	<i>Cirripectes</i> sp. / Combtooth blenny	No
Aldabra	Faecal	<i>Gempylus serpens</i> / Snake mackerel	No
Aldabra	Faecal	<i>Psenes cyanophrys</i> / Freckled driftfish	No
Aldabra	Faecal	<i>Oxyporhamphus micropterus</i> Micropterus / Bigwing halfbeak	Yes
Aldabra	Faecal	<i>Mulloidichthys flavolineatus</i> / Yellowstripe goatfish	Yes
Aldabra	Faecal	<i>Parupeneus multifasciatus</i> / Manybar goatfish	Genus only
Aldabra	Faecal	<i>Parupeneus janseni</i> / Jansen's goatfish	Genus only
Aldabra	Regurgigate	<i>Kuhlia</i> sp. / Flagtail sp.	No
Aldabra	Regurgigate	<i>Elagatis bipinnulata</i> / Rainbow runner	Genus only
Aldabra	Regurgigate	Exocoetidae indet. / Flying fish sp.	Yes
Aride	Faecal	<i>Decapterus macarellus</i> / Mackerel scad	Genus only
Aride	Faecal	<i>Decapterus macrosoma</i> / Shortfin scad	Yes
Aride	Faecal	<i>Cantherhines</i> sp. / Filefish sp.	No
Aride	Faecal	<i>Lagocephalus lagocephalus</i> / Oceanic puffer	No
Aride	Faecal	Apogoninae indet. / Cardinalfish	No
Aride	Faecal	<i>Rhabdamia gracilis</i> / Cardinalfish	No
Aride	Faecal	<i>Sargocentron diadema</i> / Crown squirrelfish	Genus only
Aride	Faecal	<i>Cheilopogon</i> sp. / Flyingfish	Yes
Aride	Faecal	<i>Euleptorhamphus viridis</i> / Ribbon halfbeak	Yes
Aride	Faecal	<i>Oxyporhamphus micropterus</i> Micropterus / Bigwing halfbeak	Yes
Aride	Faecal	<i>Euthynnus</i> sp. / Tuna	Yes
Aride	Faecal	<i>Thunnus</i> sp. / Tuna	Yes
Aride	Faecal	<i>Parupeneus janseni</i> / Jansen's goatfish	Genus only
Aride	Regurgigate	Exocoetidae indet. / Flying fish sp.	Yes
Aride	Regurgigate	<i>Cephalopoda</i> / Cephalopod indet.	Yes

Text S7. Foraging trips

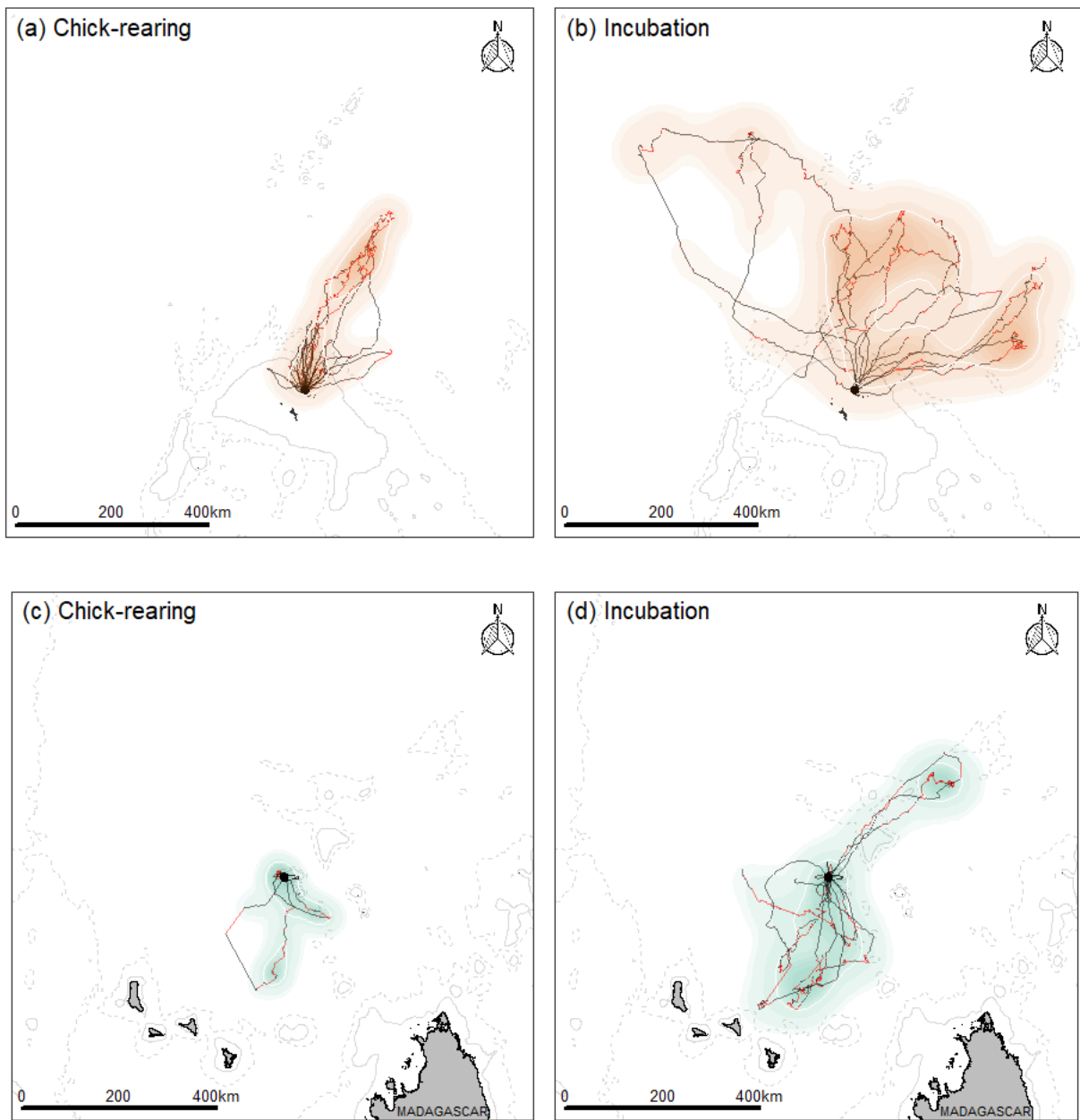


Figure S5. Foraging distribution and trips of white-tailed tropicbirds on Aride during chick-rearing (a) and incubation (b) and on Aldabra during chick-rearing (c) and incubation (d). Densities are shaded from the lightest to the darkest occupancy (from 95% to 10% occupancy) with core foraging area (50% occupancy) marked by a white line. Black lines show foraging trips with foraging position in red. Solid grey and dashed grey lines represent the -1000 m and -3000 m water depth contours, respectively. Black circles represent colonies.

References

- Bennison, A., Bearhop, S., Bodey, T. W., Votier, S. C., Grecian, W. J., Wakefield, E. D., Hamer, K. C., & Jessopp, M. (2018). Search and foraging behaviors from movement data: A comparison of methods. *Ecology and Evolution*, 8(1), 13–24. <https://doi.org/10.1002/ece3.3593>
- Calenge, C. (2019). Home Range Estimation in R: the adehabitatHR Package.
- Carneiro, A., Manica, A., Clay, T., Silk, J., King, M., & Phillips, R. (2016). Consistency in migration strategies and habitat preferences of brown skuas over two winters, a decade apart. *Marine Ecology Progress Series*, 553, 267–281. <https://doi.org/10.3354/meps11781>
- Dean, B., Freeman, R., Kirk, H., Leonard, K., Phillips, R. A., Perrins, C. M., & Guilford, T. (2013). Behavioural mapping of a pelagic seabird: Combining multiple sensors and a hidden Markov model reveals the distribution of at-sea behaviour. *Journal of The Royal Society Interface*, 10(78), 20120570. <https://doi.org/10.1098/rsif.2012.0570>
- Dehnhard, N., Achurch, H., Clarke, J., Michel, L. N., Southwell, C., Sumner, M. D., Eens, M., & Emmerson, L. (2020). High inter- and intraspecific niche overlap among three sympatrically breeding, closely related seabird species: Generalist foraging as an adaptation to a highly variable environment? *Journal of Animal Ecology*, 89(1), 104–119. <https://doi.org/10.1111/1365-2656.13078>
- Fayet, A. L., Clucas, G. V., Anker-Nilssen, T., Syposz, M., & Hansen, E. S. (2021). Local prey shortages drive foraging costs and breeding success in a declining seabird, the Atlantic puffin. *Journal of Animal Ecology*, 90(5), 1152–1164. <https://doi.org/10.1111/1365-2656.13442>
- Kernohan, B. J., Gitzen, R. A., & Millspaugh, J. J. (2001). Analysis of Animal Space Use and Movements. In *Radio Tracking and Animal Populations* (pp. 125–166). Elsevier. <https://doi.org/10.1016/B978-012497781-5/50006-2>
- Michelot, T., Langrock, R., & Patterson, T. A. (2016). moveHMM: An R package for the statistical modelling of animal movement data using hidden Markov models. *Methods in Ecology and Evolution*, 7(11), 1308–1315. <https://doi.org/10.1111/2041-210X.12578>
- Oppel, S., Beard, A., Fox, D., Mackley, E., Leat, E., Henry, L., Clingham, E., Fowler, N., Sim, J., Sommerfeld, J., Weber, N., Weber, S., & Bolton, M. (2015). Foraging distribution of a tropical seabird supports Ashmole's hypothesis of population regulation. 13. <https://doi.org/10.1007/s00265-015-1903-3>
- Schuler, K. L., Schroeder, G. M., Jenks, J. A., & Kie, J. G. (2014). Ad hoc smoothing parameter performance in kernel estimates of GPS-derived home ranges. *Wildlife Biology*, 20(5), 259–266. <https://doi.org/10.2981/wlb.12117>
- Silverman, B. (1986). Density estimation for statistics and data analysis. Chapman and Hall, London.