Methods in Ecology and Evolution

Optimising sample sizes for animal distribution analysis using tracking data

Takahiro Shimada[1,2], Michele Thums[1], Mark Hamann[3], Colin J. Limpus[4], Graeme C. Hays[5], Nancy FitzSimmons[4], Natalie E. Wildermann[6,7], Carlos M. Duarte[2] and Mark G. Meekan[1]

[1]Australian Institute of Marine Science, Crawley, Western Australia 6009, Australia

[2]Red Sea Research Center, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia

[3]College of Science and Engineering, James Cook University, Townsville, Queensland 4811, Australia

[4]Department of Environment and Science, Brisbane, Queensland, 4102, Australia

[5]Deakin University, Geelong, Victoria, Australia

[6]Texas Sea Grant at Texas A&M University, College Station, Texas 77843, USA

[7]Harte Research Institute for Gulf of Mexico Studies, Corpus Christi, Texas 78412, USA

taka.shimada@gmail.com

m.thums@aims.gov.au

mark.hamann@jcu.edu.au

Col.Limpus@des.qld.gov.au

g.hays@deakin.edu.au

nancy.fitzsimmons@des.qld.gov.au

nwildermann@tamu.edu

carlos.duarte@kaust.edu.sa

m.meekan@aims.gov.au

Correspondence: taka.shimada@gmail.com

Running Head

Optimal sample size for animal distributions

Keywords

habitat use, telemetry, overlap probability, kernel density, continuous-time Markov chain, step-selection functions, time spent analysis, sea turtle, power analysis

Abstract

1. Knowledge of the spatial distribution of populations is fundamental to management plans for any species. When tracking data are used to describe distributions, it is sometimes assumed that the reported locations of individuals delineate the spatial extent of areas used by the target population.
2. Here, we examine existing approaches to validate this assumption, highlight caveats, and propose a new method for a more informative assessment of the number of tracked animals (i.e. sample size) necessary to identify distribution patterns. We show how this assessment can be achieved by considering the heterogeneous use of habitats by a target species using the probabilistic property of a utilisation distribution. Our methods are compiled in the R package *SDLfilter*.
3. We illustrate and compare the protocols underlying existing and new methods using conceptual models and demonstrate an application of our approach using a large satellite tracking data-set of flatback turtles, *Natator depressus,* tagged with accurate Fastloc-GPS tags (n = 69).
4. Our approach has applicability for the post-hoc validation of sample sizes required for the robust estimation of distribution patterns across a wide range of taxa, populations and life history stages of animals.

1 | INTRODUCTION

An understanding of the spatial distribution of animals is essential for conservation management of populations (Hays et al., 2019). Data on the location and movements of animals must be generated to achieve this goal, but this task is difficult for many species, especially those that move large distances or are difficult to observe directly (e.g. birds, forest fauna, aquatic animals). Mark-recapture and genetic approaches have been used to define movements and distributions of individuals (Boyle et al., 2009; Strandberg, Klaassen, & Thorup, 2009) but the resolution of these data are usually coarse and thus only useful at large spatial scales (often entire species or population ranges). As new tracking technologies have developed, the analysis of spatial and temporal distributions of wide-ranging, mobile species now typically involves the deployment of tracking devices (Block et al., 2011; Hussey et al., 2015).

Due to the cost and logistics involved in telemetry, only a subsample of relatively few individuals are usually tracked from a target population (Sequeira et al., 2019). In some cases, tracking data may be sufficient to directly assess the space occupied by a population. This is most likely to be an option when tracking certain life history stages of a species (e.g. breeding) that have restricted distributions, as is the case for some birds, marine mammals, fishes and sea turtles (Block et al., 2011; Maxwell et al., 2011; Thiebot, Cherel, Trathan, & Bost, 2012; Lascelles et al., 2016). Alternatively, tracking data can be linked with environmental conditions and used in ecological niche models to predict the overall space use by a population (Guisan & Thuiller, 2005; Kearney & Porter, 2009). This approach reveals the area with suitable conditions for a population but cannot assess if that population occupies all available habitat, which is particularly problematic for species that occur in low numbers such as top-order predators and endangered species (Magg et al., 2016).

Population distributions can be estimated with tracking data using several analytical techniques, with the utilisation distribution (UD) being the most popular of these (Van Winkle, 1975; Worton, 1989; Laver & Kelly, 2008). Central to the estimation of a UD are assumptions about the adequacy of the number of tracked individuals, or sample size hereafter (e.g. Maxwell et al., 2011; Thiebot et al.,

2012). At present there are no 'off the shelf' tools (e.g. R packages) to easily assess adequacy of sample size, which is concerning, particularly if management strategies (e.g. boundaries of marine protected areas) are then based on insufficient sampling (Hebblewhite & Haydon, 2010; Kendall, Klein, & Possingham, 2015). A common approach to address this issue applies a pre-defined volume contour of a UD for each tracked animal (e.g. 95% UD) and calculates collective areas from random permutations of the contours, then repeats this process many hundreds or thousands of times. An estimate of optimal sample size is provided by an asymptote in the relationship between collective areas and sample size (e.g. Hindell, Bradshaw, Sumner, Michael, & Burton, 2003; Soanes, Arnould, Dodd, Sumner, & Green, 2013; Thums, Rossendell, Guinea, & Ferreira, 2018). However, despite the selective use of volume contours to exclude rarely visited areas (e.g. 95% UD), the analysis assumes that animals visit all areas within the individual contour boundaries with equal probability. This assumption is unlikely to hold for most species because individuals are typically distributed in a heterogeneous fashion at nested spatial scales (Kotliar & Wiens, 1990).

Here, we offer a new approach to validate sample size of tracked animals for estimating spatial distribution. The new method uses the probabilistic property of utilisation distributions rather than a user-defined volume contour and allows distribution analyses to more realistically account for variation in space use. We demonstrate the use of this technique using a data set consisting of 69 deployments of accurate Fastloc-GPS tags on a single genetic stock of flatback turtles, *Natator depressus* (FitzSimmons et al., 2020). This species provides an ideal case study since they have been the subject of a number of satellite tracking studies, driven by the need for accurate and robust data on the distribution of these threatened marine reptiles in order to manage anthropogenic impacts, such as climate change, boat strikes, predation, and pollution (Hamann, Fuentes, Ban, & Mocellin, 2013; Lewison, Wallace, Alfaro-Shigueto, Mangel, & Maxwell, 2013; Shimada, Limpus, Jones, & Hamann, 2017; Critchell, Hamann, Wildermann, & Grech, 2019). Our new method is compiled in the package *SDLfilter* for the R environment (R Core Team, 2019). The stable version is available on the Comprehensive R Archive Network (CRAN) at https://cran.r-project.org/package=SDLfilter, and the latest/developmental version can be downloaded from the GitHub repository of the package at https://github.com/TakahiroShimada/SDLfilter (Shimada, 2020). The package can serve as a

collective tool-box for researchers to easily apply these assessments to their own data to objectively assess the adequacy of their sample size when calculating spatial probability distributions, including UDs.

2 | MATERIALS AND METHODS

We examined the effects of sample size (i.e. the number of tracked animals) on distribution estimates using two indices derived from utilisation distributions (UD); here referred to as collective areas (i.e. merged $k\%$ volume contours of n individuals) and overlap probability (i.e. probability of an individual within merged habitats of others). Sample size assessment using collective areas generally follows these steps: 1) Delineate the area of interest (e.g. 95% volume contour) from each individual UD; 2) Randomly order individuals; 3) Merge the area of ith individual with all the others that appeared earlier in the sequence (i.e. 1st to i^{th}); 4) Calculate size of the merged area; 5) Repeat steps 2-4 for z times; 6) Plot the collective areas (e.g. means of the repeated values generated in step 5) against sample sizes. The collective areas will initially increase as sample size increases, but because some individuals overlap in their space use, if the sampled individuals visited all the areas occupied by the population, the cumulative plot should eventually reach an asymptote.

Building on this idea, our new method examines the probability of an individual being inside the habitats used by the other individuals (overlap probability). In this way, we can directly quantify the contribution of each new sample to the UDs calculated from the existing samples. For a set of UDs of n individuals, overlap probability relative to sample size can be estimated in the following steps:

1. Randomly take out one UD of an individual j (UD_j) from a set of n UDs.
2. Randomly re-order the remaining UDs.
3. Overlay 1st to i^{th} UDs from the sequence above and merge the $k\%$ volume contours (here we used 100% volume contours). This is a collective area identified by the existing data (A_i).
4. Calculate the probability of individual j to be within the collective area A_i ($P_{i,j}$). Since the UD is a spatial probability distribution function, $P_{i,j}$ is a definite integral of the UD_j overlapping

with A_i. In other words, $P_{i,j}$ is any probability associated with the collective area A_i. Following Fieberg & Kochanny (2005):

$$P_{i,j} = \iint_{A_i} UD_j(x,y)dxdy$$

5. Repeat steps 3 to 4 for $i = [1, n - 1]$.
6. Repeat steps 1 to 5 for z times (with z being the number of random permutations).

The mean overlap probability $(\overline{P_i})$ can be defined as follows: when the number of existing samples is i, on average there is $\overline{P_i}$ probability that an additional individual is located within the collective area used by other individuals. The overlap probability equals one if an individual UD is entirely within the collective area, less than one and greater than zero if it is partially within the collective area, or zero if it is entirely outside the collective area.

Means of collective areas (km^2) and overlap probability associated with each sample size were calculated from 1000 random permutations. To estimate a horizontal asymptote and a curve reaching the asymptote, we fitted a rational function to the means of the indices (i.e. collective areas, overlap probability) against sample sizes (i) with the numerator and denominator containing first-degree polynomials as:

$$f(i) = \frac{ai + c}{bi + d}$$

where i was greater than 0, and the coefficients (a, b) and constants (c, d) were estimated using the R package *pracma* (Borchers, 2019). The horizontal asymptote (h) was then:

$$h = \frac{a}{b}$$

We considered that an asymptote had been attained once the mean index value exceeded 95% of the estimated horizontal asymptote. The sample size linked to this value was deemed to be the minimum number of tracked animals required to represent the general distribution of the group.

To demonstrate an application of this approach using real data, we obtained a data set of satellite telemetry locations from 69 adult female flatback turtles, sampled from a genetically distinct population of the eastern Australian stock (Table S1). The turtles were tracked from four nesting

beaches located in the central Queensland region of Australia including one of the four major rookeries (Avoid Island) and three minor mainland rookeries (Curtis Island, Mackay, Mon Repos). On average, approximately 50-200 and 1-50 turtles nest annually on each of the major and minor rookeries respectively (Limpus C.J., Parmenter C.J., & Chaloupka M., 2013; Limpus et al., 2019). These rookeries encompass effectively the entire geographical spread of nesting by the eastern Australian genetic stock of flatback turtles (Limpus, 2009). During a breeding season that lasts several months, females of this species remain in the vicinity of nesting beaches (inter-nesting habitats) and once breeding is complete, animals re-migrate, returning to their home foraging sites that are widely scattered along the coast of eastern Australia and beyond (Limpus, 2009; Shimada et al., 2020).

Fastloc-GPS Argos tags were deployed on turtles between 2013 and 2018, which allowed researchers to track individuals in their inter-nesting habitats, during migrations, and at their foraging habitats over periods of several months (Table S1). Details of tagging procedures can be found in Sperling & Guinea (2004). Fastloc-GPS data were relayed via the Argos satellite system and decoded into GPS locations using the data processor provided by the tag manufacturer (Wildlife Computers Inc.). Only Fastloc-GPS data were used in the analysis. Although 95% of Fastloc-GPS data are expected to be ≤219 m, the remainder of locations can have an error of >1 km (Shimada, Jones, Limpus, & Hamann, 2012; Dujon, Todd Lindstrom, & Hays, 2014). To improve the overall accuracy of the Fastloc-GPS data-set, we initially eliminated a Fastloc-GPS fix if residual error exceeded 30, or less than four GPS satellites were involved in the process of location estimation (Shimada et al., 2012). We then used the R package *SDLfilter* to remove temporal and spatial duplicates, and erroneous locations that followed biologically unrealistic movement patterns (Shimada et al., 2012; Shimada, Limpus, et al., 2016). The error (mean ± SD) of Fastloc-GPS data treated with these filtering methods has been estimated to be 47 ± 61 m (Shimada et al., 2012).

Locations along each track were segmented into 'inter-nesting' or 'foraging' phases separated by post-breeding migrations, based on the changes in 2-day average of traveling speed and angle by following established procedures (Schofield et al., 2010; Shimada et al., 2020). Some turtles commenced post-breeding migration to their foraging habitats immediately after satellite tag

deployment on their nesting beaches, hence no tracking data were obtained from their inter-nesting habitats. Several tags stopped transmitting while turtles were still migrating back to their foraging habitats and so no data were acquired for these animals during their foraging phases. Together, sixty-four inter-nesting and fifty-seven foraging phases were identified from the tracks of 69 turtles (Table S1). Track durations in the inter-nesting phase ranged from 13 to 86 days (mean = 31 days) and so included at least one cycle of nesting activity. The turtles were tracked during the foraging phases for periods ranging from 6 to 564 days (mean = 115 days). Due to the small number of individuals tagged at Mon Repos (n = 2) we did not attempt to assess the adequacy of sample size during the inter-nesting phase for this location. These tracks were included in the assessment of sample size for the foraging phase together with data from all other tagging locations.

For each turtle we estimated a UD for the inter-nesting phase and/or for the foraging phase using a movement-based kernel density estimator (Benhamou, 2011). A spatial grid of 50 m was used for UD estimation and all other parameters required were adopted from Benhamou (2011) and Shimada et al. (2016). The R package *adehabitatHR* (Calenge, 2006) was used to calculate the UD and delineate the volume contours. To visually inspect a collective UD of multiple individuals in a map, we overlaid each UD, averaged the UD at each 50 m grid cell, and generated a heat map (Shimada et al., 2017).

3 | RESULTS

3.1 | Conceptual models

We first illustrate how the use of collective UD area versus overlap probability of UDs behaves differently with increasing sample size under two different scenarios and show that assessments of adequate sample size can be achieved more informatively using the latter. In Figure 1a-c we present a situation where both the collective areas and overlap probability approaches lead to the same conclusion. In this scenario, the first three tracks each result in the identification of additional habitat use (white boxes), whereas the 4th and 5th tracks stay within the areas identified by the previous three individual's tracks (grey boxes), indicating an asymptote was attained once the third individual was

tracked (Figure 1b). The overlap probability also asymptotes once three individuals are tracked because there was very low probability (≤0.04) of the 3rd, 4th and 5th tracks occurring outside the areas that were already identified by the 1st and 2nd individuals (Figure 1c).

The second scenario (Figure 1d-f) presents an alternative situation where contrasting conclusions are drawn from the results of the collective areas versus overlap probability approaches. In Figure 1d & e, the collective area increases constantly as each new individual is tracked. However, the overlap probability still asymptotes at more than 95% after three individuals because the probability of the 4th and 5th tracks occurring outside the areas already identified by earlier individuals are very low (only 0.02 and 0.01, respectively). This second scenario highlights how our new method (probabilistic properties of UDs) may indicate adequate sample size whereas existing techniques (collective area) do not, particularly when new tracks identify rarely visited new habitats.

3.2 | Application to flatback sea turtle data

Mean collective areas of inter-nesting and foraging habitats for each rookery increased as more individuals were tracked regardless of the volume contour (Figure 2a-c & Figure 3a). The rate of increase slowed as sample size increased, although from visual inspection it was not clear if an asymptote in the mean area curve occurred for increasing sample size at any site. During the inter-nesting period at Curtis Island, for example, the last addition to the data set (the 29th track) resulted in identification of an additional area of 2.2 km^2 at the 25% and 11.2 km^2 at the 95% volume contour ranges on average (Figure 2a). At the foraging grounds, on average, the last additional track (the 57th individual) identified new areas of 14.2 km^2 at the 25% and 211.7 km^2 at the 95% volume contours (Figure 3a).

Overlap probability increased steeply with increasing sample size for turtles during the inter-nesting phase (Figure 2d-f). Probability curves attained an asymptote at a total of 11 and 6 turtles tracked from Curtis and Avoid islands, respectively, during the inter-nesting phase. However, this was not the case for turtles at the Mackay rookery, where there was a steep and continuous rise in curves for both cumulative area and overlap probability (Figure 2c,f). At this locality the sample size of five tagged

turtles was therefore insufficient to adequately represent the distribution of the inter-nesting habitat. During the foraging phase, both area and overlap probability curves failed to reach an asymptote (Figure 3). This implies that the sample size of 57 individuals was insufficient to characterise the foraging area of the eastern Australian genetic stock, at least at the scale of spatial resolution used in this case study (50 m).

Heat maps of averaged utilisation distributions aligned with the sample size assessments based on overlap probability (Figures 4 & 5). Using the example of turtles tracked from Curtis Island, high density areas or hotspots (darker colours) of habitat occupancy during the inter-nesting phase were more similar between the UD heat maps derived from 11 and 29 samples than the heat map estimated from 4 samples, which is the mid-point in overlap probability between 1 to 11 samples (Figure 4). In these examples, the mean overlap probabilities were 0.8, 0.926 and 0.965 with the respective sample size of 4, 11 and 29 (Figure 2d). For foraging turtles, although there was some consistency in occurrence, the number of hotspots in UDs increased and shifted as sample size grew from 10, 23, and 57 tracks (Figure 5). The overlap probabilities were 0.236, 0.445 and 0.656 at a sample size of 10, 23 and 57 respectively, which is the sample size at approximately 1/3, 2/3 and 3/3 of the maximum mean overlap probability (0.656) estimated from 57 tracks (Figure 3b).

4 | DISCUSSION

Identifying optimum sample size of animals to be tracked is a key consideration in ecological research programs, given the importance of accurate descriptions of the distribution of mobile species for conservation management, and the cost and invasive nature of tracking devices (Jones et al., 2013; Hamelin & James, 2018). We showed that such assessments of sample size can be facilitated by incorporating the probabilistic property of utilisation distributions into analyses, while maintaining the diagnostic cumulative curve plots that are widely accepted by the research community. Our approach allows for objective identification of an asymptote in the curve of probability overlap and thus

estimation of an appropriate sample size for tagging. This also avoids the use of contour designations (e.g. 25%, 95% UD) within which movement heterogeneity is not accounted for.

By applying our new probability-based methods to the case study, we found that only 11 individuals were required to estimate the general distribution of the inter-nesting habitats from one rookery (Curtis Island) as long as the population was randomly sampled. In contrast, existing area-based analytical approaches suggested that a sample size of >29 individuals would be required. The smaller sample size of 11 turtles would represent a considerable cost saving for a study and reduce any adverse impacts of tag deployments on this population. To improve resource management, researchers may assess the adequacy of sampling after each field season. This is a realistic proposition in many tracking studies because tag deployment tends to occur over multiple seasons due to the cost of the devices. For example, in our study the deployment of tags on 29 turtles at Curtis Island took more than 3 years.

Our example analysis also shows that the sample size of tracking data adequate to determine distribution patterns are likely to vary due to the behaviours and life history stages of the study species. Although relatively few tracks were required (≤11) to describe the distribution patterns of turtles during the inter-nesting phase when females tended to use a spatially discrete area, the entire data set of 57 turtles was insufficient to characterise the distribution of foraging grounds, which encompassed much of the coast of Queensland, a distance of >2000 km between Gladstone and Torres Strait. Such temporal variation in space use by different life history stages is typical of many species. For example, many sea-birds rearing chicks tend to forage near the breeding colony. This results in a higher overlap in space use among individuals than during incubation, when the birds have much wider foraging ranges (Dehnhard et al., 2020). Additionally, taxa that forage on spatio-temporally patchy prey (e.g. the whale shark *Rhincodon typus*) are often loosely attached to sites and can travel thousands of km to forage (Sleeman et al., 2010; Reynolds, Norman, Beger, Franklin, & Dwyer, 2017), so that overlap probabilities may only asymptote after some hundreds of individuals are tracked. In this case, ecological niche models may be useful, although they have some limitations (Elith & Leathwick, 2009; Cianfrani, Lay, Hirzel, & Loy, 2010). Any analysis of space use by a

species must therefore consider general movement patterns (residential or nomadic) and temporal changes in distribution on both short (diel, seasonal) and long (ontogenetic, breeding cycles etc.) scales that are relevant to survey design and the questions posed. Furthermore, telemetry transmissions can unexpectedly cease for many reasons (Hays, Bradshaw, James, Lovell, & Sims, 2007) and shorter tracking durations may not provide data representative of the movement of the individual. In such cases, a larger number of tracked animals may be required to compensate for this uncertainty.

The accuracy of locations will also influence the assessment of sample size. Locations provided by Argos-only tags are often much less accurate than those by Fastloc-GPS tags, with mean errors ranging from <1 km to >10 km for Argos locations relative to <50 m for Fastloc-GPS locations (Hazel, 2009; Costa et al., 2010; Shimada et al., 2012). Consequently, Argos locations will generate estimates of UDs that cover a far larger area than those provided by Fastloc-GPS technology (Hoenner, Whiting, Hindell, & McMahon, 2012; Thomson et al., 2017). Estimates of UDs based on Argos locations will thus tend to overlap more than those derived from accurate Fastloc-GPS data, with the consequence that probability overlap curves will asymptote at a smaller sample size than those calculated using Fastloc-GPS fixes. Additionally, UDs derived from the same data set can vary depending on the estimator and different parameters (e.g. grid cell size) used in the estimation process (Börger et al., 2006; Walter, Onorato, & Fischer, 2015)(Börger et al., 2006; Walter, Onorato, & Fischer, 2015). Furthermore, non-random sampling likely biases the distribution analysis, especially when the sample size is small (e.g. <5 tracks). For these reasons, in the assessment of sample size it is important that the researcher carefully considers the spatial resolution, UD estimation methods, tracking technology and appropriate random sampling, so that the aims of any tagging program are achieved.

Our method reasonably assumes that important habitats have been identified once a high probability of overlap is achieved (e.g. >0.95 on average). This approach means that some rarely visited habitats may not be recorded. Rare events may be important, such as very distant migrations that may influence population genetic structure across ocean basins (Bonfil et al., 2005; Gore, Rowat, Hall,

Gell, & Ormond, 2008). If relevant, researchers may choose a higher threshold of overlap probability (e.g. 0.99) to ensure that potentially important but rarely visited areas are included in the distribution. For example, this conservative approach may be applicable to a declining or recovering population that needs a maximum protection, or when the sample size is small (e.g. <5 tracks). For our example data-set of flatback turtles (n = 69), 95% of an estimated asymptote appears reasonable to assess the adequacy of sample size given their status as a vulnerable species (Queensland Nature Conservation Act 1992), the broad spatial distribution of the population, and relatively few anthropogenic threats in the areas visited by the tracked individuals.

A key strength of our approach is the ability to probabilistically assess the importance of increases in sample size. We can objectively identify minimum sample size required to estimate realistic distributions based on the likelihood that more tag deployments will identify new areas important to the population. This probabilistic approach is flexible and can be applied regardless of taxa, data accuracy, and methods for distribution analysis. For example, input data could be any UDs that are estimated using either traditional methods such as fixed kernel density estimators (KDE) and local convex hull (Walter et al., 2015), or more recent approaches such as autocorrelated KDE (Fleming et al., 2015), step-selection functions (Signer, Fieberg, & Avgar, 2017), continuous-time Markov chain models (Wilson, Hanks, & Johnson, 2018), and other commonly used estimates such as the proportion of time spent in an area (Soanes et al., 2013). Additionally, alternative overlap indices such as Bhattacharyya's affinity (Fieberg & Kochanny, 2005) are made available as options in the R package *SDLfilter* to enable use of the technique in a variety of situations. Thus, our new approach has wide applicability, and the ability to improve both resource management and the impacts of tracking studies on target species.

Ethics. This research was conducted following all applicable guidelines for the care and use of animals approved by the Ethics Committee of the Queensland Department of Environment and Science (DES) and James Cook University.

Conflict of Interest

The authors declare no conflicts of interest.

Data Accessibility. The data used in this article are available from the Dryad Digital Repository: https://doi.org/10.5061/dryad.x69p8czgh.

References

Benhamou, S. (2011). Dynamic approach to space and habitat use based on biased random bridges. *PLoS ONE*, *6*(1), e14592–e14592. doi:10.1371/journal.pone.0014592

Block, B. A., Jonsen, I. D., Jorgensen, S. J., Winship, A. J., Shaffer, S. A., Bograd, S. J., … Costa, D. P. (2011). Tracking apex marine predator movements in a dynamic ocean. *Nature*, *475*(7354), 86–90. doi:10.1038/nature10082

Bonfil, R., Meÿer, M., Scholl, M. C., Johnson, R., O'Brien, S., Oosthuizen, H., … Paterson, M. (2005). Transoceanic migration, spatial dynamics, and population linkages of white sharks. *Science*, *310*(5745), 100–103. doi:10.1126/science.1114898

Borchers, H. W. (2019). pracma: Practical Numerical Math Functions (Version 2.2.5). Retrieved from https://CRAN.R-project.org/package=pracma

Börger, L., Franconi, N., De Michele, G., Gantz, A., Meschi, F., Manica, A., … Coulson, T. I. M. (2006). Effects of sampling regime on the mean and variance of home range size estimates. *Journal of Animal Ecology*, *75*(6), 1393–1405. doi:10.1111/j.1365-2656.2006.01164.x

Boyle, M. C., FitzSimmons, N. N., Limpus, C. J., Kelez, S., Velez-Zuazo, X., & Waycott, M. (2009). Evidence for transoceanic migrations by loggerhead sea turtles in the southern Pacific Ocean. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1664), 1993–1999. doi:10.1098/rspb.2008.1931

Calenge, C. (2006). The package 'adehabitat' for the R software: A tool for the analysis of space and habitat use by animals. *Ecological Modelling*, *197*(3–4), 516–519. doi:10.1016/j.ecolmodel.2006.03.017

Cianfrani, C., Lay, G. L., Hirzel, A. H., & Loy, A. (2010). Do habitat suitability models reliably predict the recovery areas of threatened species? *Journal of Applied Ecology*, *47*(2), 421–430. doi:10.1111/j.1365-2664.2010.01781.x

Costa, D. P., Robinson, P. W., Arnould, J. P. Y., Harrison, A. L., Simmons, S. E., Hassrick, J. L., … Crocker, D. E. (2010). Accuracy of ARGOS locations of pinnipeds at-sea estimated using Fastloc GPS. *PloS One*, *5*(1), e8677–e8677. doi:10.1371/journal.pone.0008677

Critchell, K., Hamann, M., Wildermann, N., & Grech, A. (2019). Predicting the exposure of coastal species to plastic pollution in a complex island archipelago. *Environmental Pollution*, *252*, 982–991. doi:10.1016/j.envpol.2019.06.031

Dehnhard, N., Achurch, H., Clarke, J., Michel, L. N., Southwell, C., Sumner, M. D., … 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. doi:10.1111/1365-2656.13078

Dujon, A. M., Todd Lindstrom, R., & Hays, G. C. (2014). The accuracy of Fastloc-GPS locations and implications for animal tracking. *Methods in Ecology and Evolution*, *5*(11), 1162–1169. doi:10.1111/2041-210x.12286

Elith, J., & Leathwick, J. R. (2009). Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics*, *40*(1), 677–697. doi:10.1146/annurev.ecolsys.110308.120159

Fieberg, J., & Kochanny, C. O. (2005). Quantifying home-range overlap: The importance of the utilization distribution. *The Journal of Wildlife Management*, *69*(4), 1346–1359. doi:10.2193/0022-541x(2005)69[1346:Qhotio]2.0.Co;2

FitzSimmons, N. N., Pittard, S. D., McIntyre, N., Jensen, M. P., Guinea, M., Hamann, M., … Whiting, S. (2020). Phylogeography, genetic stocks, and conservation implications for an Australian endemic marine turtle. *Aquatic Conservation: Marine and Freshwater Ecosystems*, *30*(3), 440–460. doi:10.1002/aqc.3270

Fleming, C. H., Fagan, W. F., Mueller, T., Olson, K. A., Leimgruber, P., & Calabrese, J. M. (2015). Rigorous home range estimation with movement data: a new autocorrelated kernel density estimator. *Ecology*, *96*(5), 1182–1188. doi:10.1890/14-2010.1

Gore, M. A., Rowat, D., Hall, J., Gell, F. R., & Ormond, R. F. (2008). Transatlantic migration and deep mid-ocean diving by basking shark. *Biology Letters*, *4*(4), 395–398. doi:10.1098/rsbl.2008.0147

Guisan, A., & Thuiller, W. (2005). Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, *8*(9), 993–1009. doi:10.1111/j.1461-0248.2005.00792.x

Hamann, M., Fuentes, M. M. P. B., Ban, N. C., & Mocellin, V. J. L. (2013). Climate change and marine turtles. In J. Wyneken, K. J. Lohmann, & J. A. Musick (Eds.), *The biology of sea turtles* (Vol. III, pp. 353–378). Florida: CRC Press.

Hamelin, K. M., & James, M. C. (2018). Evaluating outcomes of long-term satellite tag attachment on leatherback sea turtles. *Animal Biotelemetry*, *6*(1), 18. doi:10.1186/s40317-018-0161-3

Hays, G. C., Bailey, H., Bograd, S. J., Bowen, W. D., Campagna, C., Carmichael, R. H., … Sequeira, A. M. M. (2019). Translating marine animal tracking data into conservation policy and management. *Trends in Ecology & Evolution*, *34*(5), 459–473. doi:10.1016/j.tree.2019.01.009

Hays, G. C., Bradshaw, C. J. A., James, M. C., Lovell, P., & Sims, D. W. (2007). Why do Argos satellite tags deployed on marine animals stop transmitting? *Journal of Experimental Marine Biology and Ecology*, *349*(1), 52–60. doi:10.1016/j.jembe.2007.04.016

Hazel, J. (2009). Evaluation of fast-acquisition GPS in stationary tests and fine-scale tracking of green turtles. *Journal of Experimental Marine Biology and Ecology*, *374*(1), 58–68. doi:10.1016/j.jembe.2009.04.009

Hebblewhite, M., & Haydon, D. T. (2010). Distinguishing technology from biology: a critical review of the use of GPS telemetry data in ecology. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1550), 2303–2312. doi:10.1098/rstb.2010.0087

Hindell, M. A., Bradshaw, C. J. A., Sumner, M. D., Michael, K. J., & Burton, H. R. (2003). Dispersal of female southern elephant seals and their prey consumption during the austral summer: relevance to management and oceanographic zones. *Journal of Applied Ecology*, *40*(4), 703–715. doi:10.1046/j.1365-2664.2003.00832.x

Hoenner, X., Whiting, S. D., Hindell, M. A., & McMahon, C. R. (2012). Enhancing the use of Argos satellite data for home range and long distance migration studies of marine animals. *PloS One*, *7*(7), e40713–e40713. doi:10.1371/journal.pone.0040713

Jones, T. T., Van Houtan, K. S., Bostrom, B. L., Ostafichuk, P., Mikkelsen, J., Tezcan, E., … Seminoff, J. A. (2013). Calculating the ecological impacts of animal-borne instruments on aquatic organisms. *Methods in Ecology and Evolution*, *4*(12), 1178–1186. doi:10.1111/2041-210X.12109

Kearney, M., & Porter, W. (2009). Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. *Ecology Letters*, *12*(4), 334–350. doi:10.1111/j.1461-0248.2008.01277.x

Kendall, B. E., Klein, C. J., & Possingham, H. P. (2015). The role of scale in designing protected area systems to conserve poorly known species. *Ecosphere*, *6*(11), art237–art237. doi:10.1890/ES15-00346.1

Kotliar, N. B., & Wiens, J. A. (1990). Multiple scales of patchiness and patch structure: a hierarchical framework for the study of heterogeneity. *Oikos*, *59*(2), 253–260. doi:10.2307/3545542

Lascelles, B. G., Taylor, P. R., Miller, M. G. R., Dias, M. P., Oppel, S., Torres, L., … Small, C. (2016). Applying global criteria to tracking data to define important areas for marine conservation. *Diversity and Distributions*, *22*(4), 422–431. doi:10.1111/ddi.12411

Laver, P. N., & Kelly, M. J. (2008). A critical review of home range studies. *Journal of Wildlife Management*, *72*(1), 290–298. doi:10.2193/2005-589

Lewison, R., Wallace, B., Alfaro-Shigueto, J., Mangel, J. C., & Maxwell, S. M. (2013). Fisheries bycatch of marine turtles: lessons learned from decades of research and conservation. In J.

Wyneken, K. J. Lohmann, & J. A. Musick (Eds.), *The biology of sea turtles* (Vol. III, pp. 329–351). Florida: CRC Press.

Limpus, C. J. (2009). *A biological review of Australian marine turtle species*. (L. Fien, Ed.). Brisbane: Queensland Environmental Protection Agency. Retrieved from https://www.qld.gov.au/environment/library

Limpus, C. J., FitzSimmons, N. N., French, K., Hoffmann, F., Hoffmann, E., Limpus, D. J., … Turner, T. (2019). *Flatback Turtle,* Natator depressus*, 2018-2019 Breeding Season, at Curtis, Peak and Avoid Islands* (p. 73). Brisbane: Department of Environment and Science, Queensland Government. Retrieved from https://www.gpcl.com.au/envirodocs

Limpus C.J., Parmenter C.J., & Chaloupka M. (2013). *Monitoring of Coastal Sea Turtles: Gap Analysis 5. Flatback turtles, Natator depressus, in the Port Curtis and Port Alma Region* (p. 26). Gladstone Ports Corporation. Retrieved from https://www.gpcl.com.au/ermp

Magg, N., Müller, J., Heibl, C., Hackländer, K., Wölfl, S., Wölfl, M., … Heurich, M. (2016). Habitat availability is not limiting the distribution of the Bohemian–Bavarian lynx *Lynx lynx* population. *Oryx*, *50*(4), 742–752. doi:10.1017/S0030605315000411

Maxwell, S. M., Breed, G. A., Nickel, B. A., Makanga-Bahouna, J., Pemo-Makaya, E., Parnell, R. J., … Coyne, M. S. (2011). Using satellite tracking to optimize protection of long-lived marine species: olive ridley sea turtle conservation in central Africa. *PloS One*, *6*(5), e19905–e19905. doi:10.1371/journal.pone.0019905

R Core Team. (2019). R: a language and environment for statistical computing (Version 3.6.2). Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.r-project.org

Reynolds, S. D., Norman, B. M., Beger, M., Franklin, C. E., & Dwyer, R. G. (2017). Movement, distribution and marine reserve use by an endangered migratory giant. *Diversity and Distributions*, *23*(11), 1268–1279. doi:10.1111/ddi.12618

Schofield, G., Hobson, V. J., Fossette, S., Lilley, M. K. S., Katselidis, K. A., & Hays, G. C. (2010). Fidelity to foraging sites, consistency of migration routes and habitat modulation of home range by sea turtles. *Diversity and Distributions*, *16*(5), 840–853. doi:10.1111/j.1472-4642.2010.00694.x

Sequeira, A. M. M., Heupel, M. R., Lea, M.-A., Eguíluz, V. M., Duarte, C. M., Meekan, M. G., … Hays, G. C. (2019). The importance of sample size in marine megafauna tagging studies. *Ecological Applications*, *29*(6), e01947. doi:10.1002/eap.1947

Shimada, T. (2020). SDLfilter: R package for filtering and assessing the sample size of tracking data. doi:10.5281/zenodo.3631115

Shimada, T., Jones, R., Limpus, C., Groom, R., & Hamann, M. (2016). Long-term and seasonal patterns of sea turtle home ranges in warm coastal foraging habitats: Implications for conservation. *Marine Ecology Progress Series*, *562*, 163–179. doi:10.3354/meps11972

Shimada, T., Jones, R., Limpus, C., & Hamann, M. (2012). Improving data retention and home range estimates by data-driven screening. *Marine Ecology Progress Series*, *457*, 171–180. doi:10.3354/meps09747

Shimada, T., Limpus, C. J., Hamann, M., Bell, I., Esteban, N., Groom, R., & Hays, G. C. (2020). Fidelity to foraging sites after long migrations. *Journal of Animal Ecology*, *89*(4), 1008–1016. doi:10.1111/1365-2656.13157

Shimada, T., Limpus, C., Jones, R., & Hamann, M. (2017). Aligning habitat use with management zoning to reduce vessel strike of sea turtles. *Ocean & Coastal Management*, *142*, 163–172. doi:10.1016/j.ocecoaman.2017.03.028

Shimada, T., Limpus, C., Jones, R., Hazel, J., Groom, R., & Hamann, M. (2016). Sea turtles return home after intentional displacement from coastal foraging areas. *Marine Biology*, *163*(1), 8. doi:10.1007/s00227-015-2771-0

Signer, J., Fieberg, J., & Avgar, T. (2017). Estimating utilization distributions from fitted step-selection functions. *Ecosphere*, *8*(4), e01771. doi:10.1002/ecs2.1771

Sleeman, J. C., Meekan, M. G., Wilson, S. G., Polovina, J. J., Stevens, J. D., Boggs, G. S., & Bradshaw, C. J. A. (2010). To go or not to go with the flow: Environmental influences on whale shark movement patterns. *Journal of Experimental Marine Biology and Ecology*, *390*(2), 84–98. doi:10.1016/j.jembe.2010.05.009

Soanes, L. M., Arnould, J. P. Y., Dodd, S. G., Sumner, M. D., & Green, J. A. (2013). How many seabirds do we need to track to define home-range area? *Journal of Applied Ecology*, *50*(3), 671–679. doi:10.1111/1365-2664.12069

Sperling, J. B., & Guinea, M. L. (2004). A harness for attachment of satellite transmitters on flatback turtles. *Marine Turtle Newsletter*, (103), 11–13.

Strandberg, R., Klaassen, R. H. G., & Thorup, K. (2009). Spatio-temporal distribution of migrating raptors: a comparison of ringing and satellite tracking. *Journal of Avian Biology*, *40*(5), 500–510. doi:10.1111/j.1600-048X.2008.04571.x

Thiebot, J.-B., Cherel, Y., Trathan, P. N., & Bost, C.-A. (2012). Coexistence of oceanic predators on wintering areas explained by population-scale foraging segregation in space or time. *Ecology*, *93*(1), 122–130. doi:10.1890/11-0385.1

Thomson, J. A., Börger, L., Christianen, M. J. A., Esteban, N., Laloë, J.-O., & Hays, G. C. (2017). Implications of location accuracy and data volume for home range estimation and fine-scale movement analysis: comparing Argos and Fastloc-GPS tracking data. *Marine Biology*, *164*(10), 204. doi:10.1007/s00227-017-3225-7

Thums, M., Rossendell, J., Guinea, M., & Ferreira, L. C. (2018). Horizontal and vertical movement behaviour of flatback turtles and spatial overlap with industrial development. *Marine Ecology Progress Series*, *602*, 237–253. doi:10.3354/meps12650

Van Winkle, W. (1975). Comparison of several probabilistic home-range models. *The Journal of Wildlife Management*, *39*(1), 118–123. doi:10.2307/3800474

Walter, W. D., Onorato, D. P., & Fischer, J. W. (2015). Is there a single best estimator? Selection of home range estimators using area-under-the-curve. *Movement Ecology*, *3*(1), 10. doi:10.1186/s40462-015-0039-4

Wilson, K., Hanks, E., & Johnson, D. (2018). Estimating animal utilization densities using continuous-time Markov chain models. *Methods in Ecology and Evolution*, *9*(5), 1232–1240. doi:10.1111/2041-210X.12967

Worton, B. J. (1989). Kernel methods for estimating the utilization distribution in home-range studies. *Ecology*, *70*(1), 164–168. doi:10.2307/1938423

Figures

FIGURE 1 Conceptual diagrams of collective UD areas versus overlap probabilities of UDs in relation to determining adequate sample size (n) under two scenarios. The boxes (a, d) denote the areas identified by the tracks of n individuals. The new calculated area (white) identified by the addition of the last individual's tracking data (i.e. n^{th} track) is distinguished from the area used by the previously tracked individuals (i.e. $n - 1$ tracks; grey). The value in each box denotes the probability of the new individual (i.e. n^{th} track) being located inside (grey boxes) or outside (white boxes) the previously identified collective area. Superficial probability values were assigned in each box for illustration purposes. The bootstrapping process is not included in these conceptual scenarios for simplicity. Horizontal grey lines in (c, f) indicate overlap probability of 0.95. In the first scenario (a-c), the first three tracks identify new areas used by individuals, but the additional two tracks stay within the areas used by earlier individuals. The probability of a new track overlapping the collective areas asymptotes when new tracks stay within that area. In the second scenario (d-f), the collective area increases after the addition of each new track, whereas the overlap probability flattens after the 3rd individual because the probability of the new 4th and 5th individuals to be outside the known areas increases only by 0.02 and 0.01 respectively. In this case, the contribution of the new tracks may be considered insignificant so that the first 3 tracks are sufficient to describe the general space use of the group.
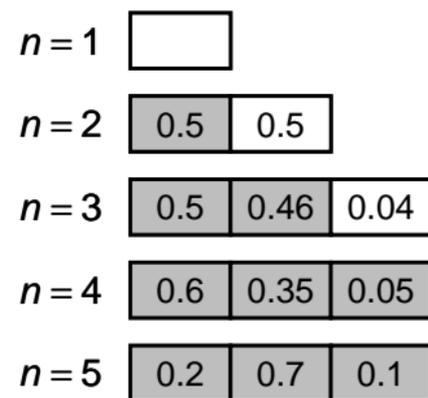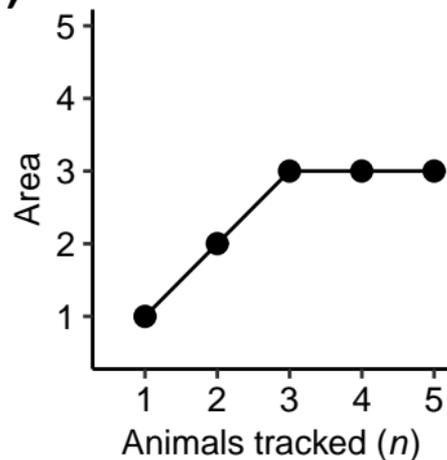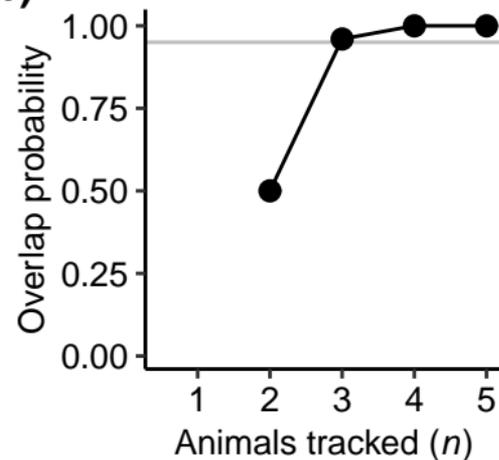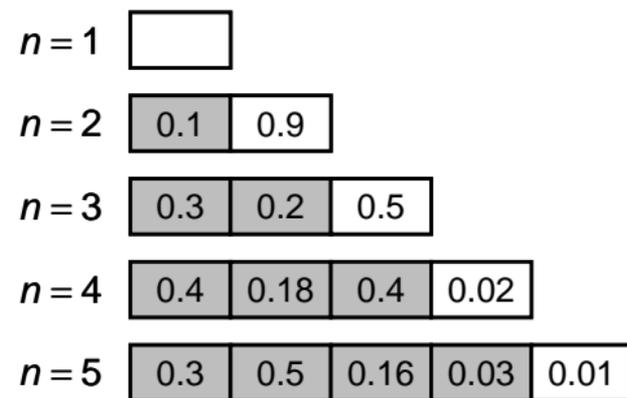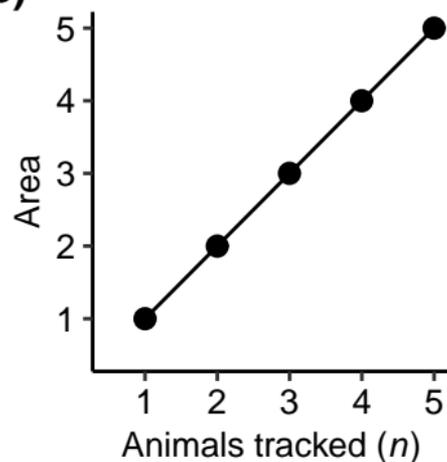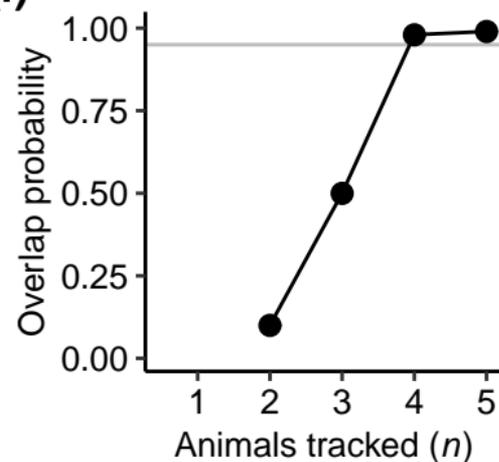
FIGURE 2 Effects of sample size (n = number of tracks) on estimation of spatial distributions of flatback turtles during the inter-nesting phase tracked from three nesting sites in eastern Australia showing: (a-c) curves of mean collective areas at increasing sample size calculated from volume contours of utilisation distributions (UD) ranging from 25-95%; (d-f) curves of mean overlap probability at increasing sample size calculated from the 100% UD. Each point on the curve is a mean value calculated from 1000 random permutations and curves are rational functions fitted to these mean values. Once the fitted curve reached 95% of the estimated horizontal asymptotes (horizontal grey line), the sample size (as indicated by a dashed line) was deemed to be sufficient to characterise spatial distributions during the inter-nesting phase.
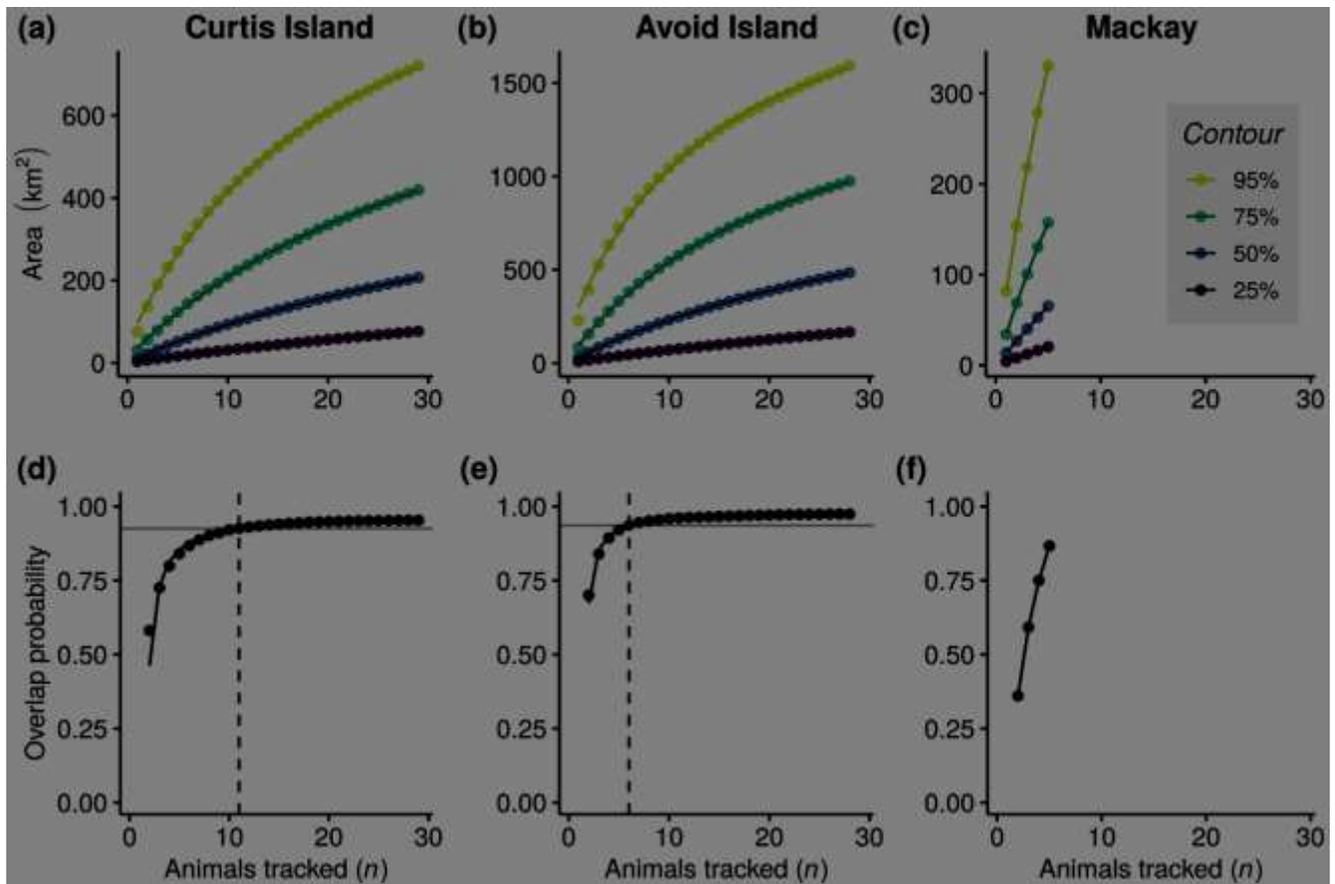
FIGURE 3 Effects of sample size (n = number of tracks) on estimation of spatial distribution of flatback turtles during the foraging phase in eastern Australian waters showing: (a) curves of mean collective areas at increasing sample size calculated from volume contours of utilization distributions (UD) ranging from 25-95%; (b) curve of mean overlap probability at increasing sample size calculated from the 100% UD distribution. Points on curves are mean values calculated from 1000 random permutations and curves are rational functions fitted to the mean values.
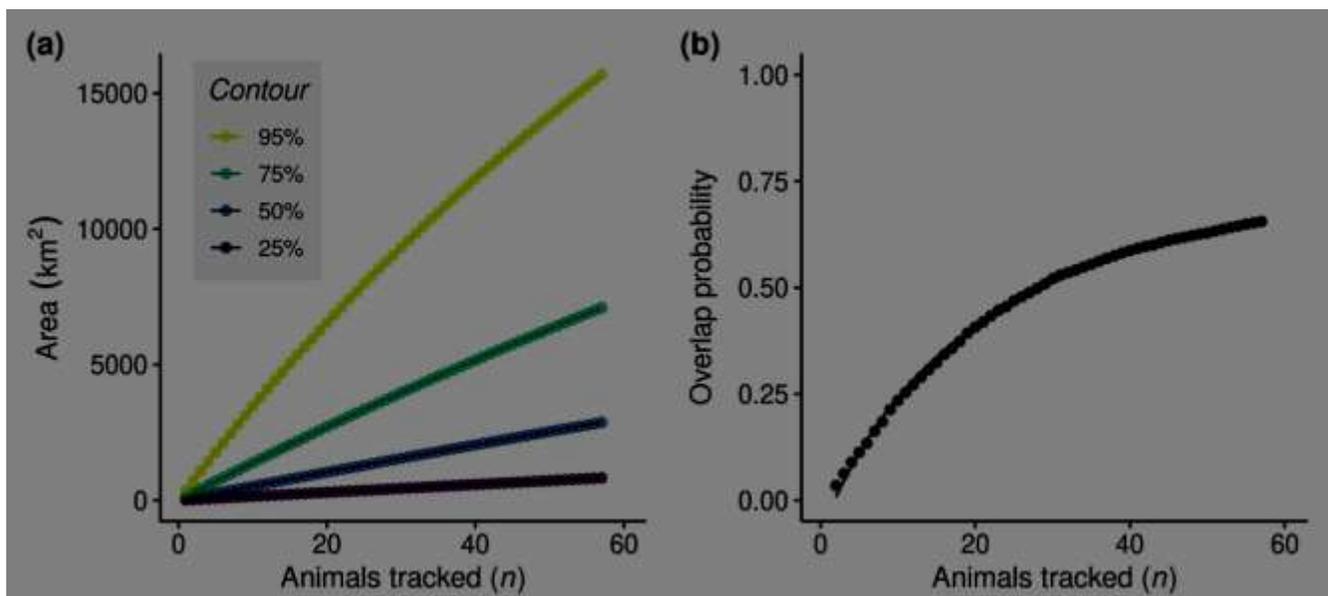
FIGURE 4 Utilisation distributions during inter-nesting period estimated from a sample size of 4, 11, and 29 flatback turtles tracked from Curtis Island, Australia. Analysis of overlap probability indicated 11 samples or greater would be sufficient to describe the general inter-nesting distribution of this population.
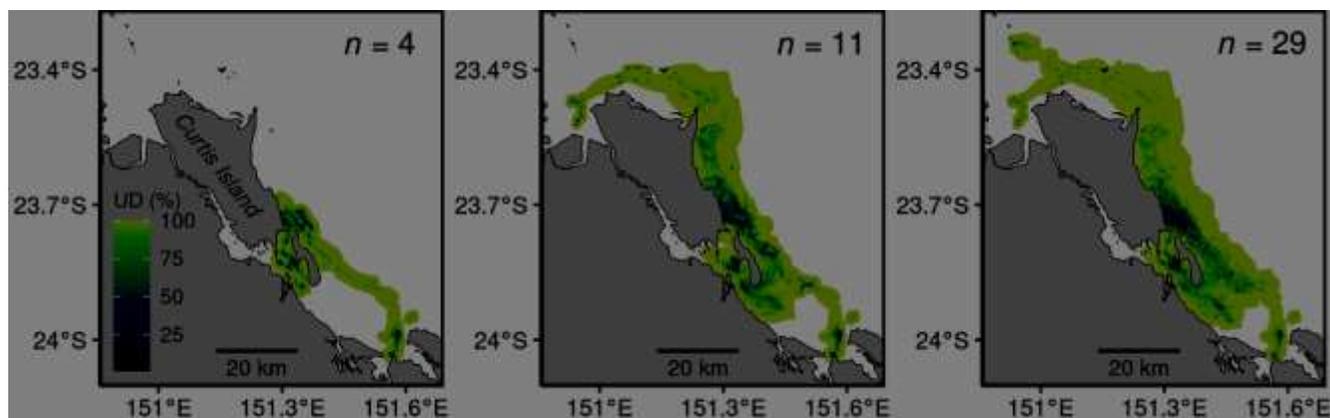
FIGURE 5 Combined (from all rookeries) utilisation distributions of flatback turtles at foraging grounds on the Queensland coast estimated from a sample size of 10, 23, and 57 animals tracked with Fastloc-GPS Argos tags. Figures on left show the estimates of the entire distribution and those on the right show details of the boxed area.