track2KBA: An R package for identifying important sites for biodiversity from tracking data

Martin Beal*[1,2], Steffen Oppel[3], Jonathan Handley[2], Elizabeth J. Pearmain[2], Virginia Morera-Pujol[4], Ana P.B. Carneiro[2], Tammy E. Davies[2], Richard A. Phillips[5], Philip R. Taylor[7], Mark G.R. Miller[13], Aldina M. A. Franco[6], Inês Catry[9,10], Ana R. Patrício[1,8], Aissa Regalla[11], Iain Staniland[5], Charlotte Boyd[12], Paulo Catry[1] & Maria P. Dias[1,2]

Running title: Identifying important sites from tracking data

Author details

[1] *MARE - Marine and Environmental Sciences Centre, ISPA - Instituto Universitário, Lisbon, Portugal*
[2] *BirdLife International, The David Attenborough Building, Pembroke Street, Cambridge CB2 3QZ, UK*
[3] *RSPB Centre for Conservation Science, Royal Society for the Protection of Birds, The David Attenborough Building, Pembroke Street, Cambridge CB2 3QZ, UK.*
[4] *Institut de Recerca de la Biodiversitat (IRBio) and Department de Biologia Evolutiva, Ecologia i Ciències Ambientals (BEECA), Universitat de Barcelona, Av Diagonal 643, Barcelona 08028, Spain.*
[5] *British Antarctic Survey, Natural Environment Research Council, High Cross, Madingley Road, Cambridge CB3 0ET, UK.*
[6] *School of Environmental Sciences, University of East Anglia, Norwich Research Park, Norwich NR4 6HN, UK.*
[7] *The Open Seas Trust, 51 Atholl Rd, Pitlochry, Scotland, PH16 5BU*
[8] *Centre for Ecology and Conservation, University of Exeter, TR10 9EZ Penryn, UK*
[9] *CIBIO/InBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, Laboratório Associado, Universidade do Porto, Campus Agrário de Vairão, Rua Padre Armando Quintas, 4485–601 Vairão, Portugal*
[10] *CIBIO/InBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, Laboratório Associado, Instituto Superior de Agronomia, Universidade de Lisboa, 1349-017 Lisbon, Portugal*
[11] *Instituto da Biodiversidade e das Áreas Protegidas, Bissau, Guinea-Bissau*
[12] *International Union for Conservation of Nature (IUCN), North America Office, 1630 Connecticut Ave., NW, Suite 300, Washington, D.C. 20009*
[13] *School of Biology, Faculty of Biological Sciences, University of Leeds, Leeds, LS2 9JT, UK*

** corresponding author: Martin Beal - martinbeal88@gmail.com*

Abstract

1. Identifying important sites for biodiversity is vital for conservation and management. However, there is a lack of accessible, easily-applied tools that enable practitioners to delineate important sites for highly mobile species using established criteria.
2. We introduce the R package ‘track2KBA’, a tool to identify important sites at the population level using tracking data from individual animals based on three key steps: (1) identifying individual core areas, (2) assessing population-level representativeness of the sample, and (3) quantifying spatial overlap among individuals and scaling up to the population.
3. We describe package functionality and exemplify its application using tracking data from three taxa in contrasting environments: a seal, a marine turtle, and a migratory land bird.
4. This tool facilitates the delineation of sites of ecological relevance for diverse taxa and provides output useful for assessing their importance to a population or species, as in the Key Biodiversity Area (KBA) Standard. As such, ‘track2KBA’ can contribute directly to conservation planning at global and regional levels.

Introduction

Site-based conservation is a key strategy for protecting biodiversity worldwide (Watson et al., 2014). However, to be effective, sites designated for protection must represent ecologically meaningful processes. Amid the rush to meet protected area coverage targets, there is a risk that governments protect spaces opportunistically where there are few competing interests, rather than where biodiversity is concentrated and at risk (Venter et al., 2018). Avoiding this scenario requires practical and accessible tools be made available to process ecological data for conservation planning.

One method for assessing the importance of a site for biodiversity is to use systematic criteria, such as those of the Key Biodiversity Area (KBA) Standard (IUCN, 2016). Under such criteria, ecological data are used to assess whether a site contributes significantly to the persistence of biodiversity, which is useful for conservation planning and the design and evaluation of protected areas networks (Boucher et al., 2014). However, delineating ecologically relevant boundaries for sites that can be assessed against such criteria remains challenging, particularly for highly mobile species in habitats that are spatially dynamic or otherwise hard to map directly. Although there are many tools available to analyze tracking data (Joo et al., 2019), bespoke tools that enable the use of tracking data to identify important sites for biodiversity are lacking.

We introduce the R package ‘track2KBA’, a tool for analyzing tracking data and deriving ecological information useful for (1) delineating important sites for local animal populations, and (2) assessing the importance of these sites against quantitative criteria of importance for regional or global biodiversity, such as the KBA Standard. The approach underlying ‘track2KBA’ was originally developed to identify marine Important Bird and Biodiversity Areas (mIBAs) for seabirds (Lascelles et al., 2016). Here, we exemplify applications of the method in three different taxa and environments: 1) Antarctic fur seals (*Arctocephalus gazella*) during the breeding season in the South Atlantic; 2) green sea turtles (*Chelonia mydas*) during the post-nesting period off the coast of West Africa, and 3) white storks (*Ciconia ciconia*) during migration between the breeding grounds in southern Europe and the wintering grounds in sub-Saharan Africa.

Methods

Package overview

The R package 'track2KBA', based on the approach of Lascelles et al. 2016, has three key steps: (1) estimating individual core areas, (2) assessing sample representativeness, and (3) quantifying spatial overlap among individuals and scaling up to the population level. Four functions perform these key steps, and nine accessory functions are available for processing tracking data and plotting (Fig. 1, Table S1). See Supplementary Information S1 for a detailed description of the package workflow.

Identify individual core areas

In 'track2KBA', the space used during each independent tracking event is calculated using kernel density estimation (KDE). KDE is a non-parametric technique for deriving a probability surface, known as the utilization distribution (UD), from point data. When the input are animal locations regularly spaced in time, the UD represents the probability of an animal occurring in space (Worton, 1989). KDE was selected as the method to estimate space use given its accessibility and familiarity to a wide range of users, which facilitates the description and communication of the method to non-scientists, e.g., at policy fora (Lascelles et al., 2016).

UDs for several independent tracks can be derived using the function `estSpaceUse`. KDE requires the setting of a smoothing parameter (*h*, or 'scale' in package documentation) that affects kernel width and the resulting spatial estimate (Gitzen et al., 2006; see Supplementary Information S1B for details). When determining important sites for biodiversity, it is important that the results are not under- or over-smoothed, and that the *h* value reflects both the resolution of the available data (i.e., larger *h* for coarser data) and the ecology of the study species (Lascelles et al., 2016). The function `findScale` calculates several candidate *h* values, allowing the user to select the most appropriate for the study organism and objective (Supplementary Information S1B); for central-place foragers, the functions `tripSplit` and `tripSummary` may be used to derive metrics to facilitate comparison between candidate *h* values. Once an *h* value is selected, it is important to consider 1) the resolution of the spatial grid used for KDE and 2) the probability quantile that reflects the core areas of the track (‘UDLev’); 50% is a standard choice, but in some cases other values may be more appropriate (Dias et al., 2018).

Assess sample representativeness

Whether a tracking dataset is representative of the distribution of the source population is fundamental to identifying areas of importance for population persistence (Lascelles et al., 2016). Therefore, a vital step in the ‘track2KBA’ workflow is to assess the degree of representativeness of the tracked sample using the `repAssess` function (Supplementary Information S1C). `repAssess` iteratively selects sub-samples of individual tracks, averages them into a pooled UD and outlines a desired quantile (e.g., 50%), and then calculates the proportion of out-of-sample tracking locations within the resulting area (i.e., ‘inclusion rate’). A non-linear least squares regression is fitted to the relationship between sample size and inclusion rate to project this rate until its asymptote (i.e., the sample size which fully represents the source population distribution) and calculate the degree to which the tracked sample represents the space use of the wider population. The inclusion rate at the maximum sample size should approximate the specified UD quantile when the tracked sample is fully representative (Supplementary Information S1C). `repAssess` returns the percent representativeness, the estimated asymptote, and estimates of the sample sizes needed to achieve 70% and 95% representativeness.

Population-scaling and site delineation

The final step in the ‘track2KBA’ workflow is to delineate areas used by a substantial portion of the population, and produce quantitative information of site importance (Supplementary Information S1D). To delineate a candidate site, the function `findSite` calculates the proportion of individual core areas overlapping each grid cell and multiplies this by the proportional representativeness of the tracked sample. The result is a scaled estimate of the proportion of the source population that predictably uses each grid cell in the study region in the season of interest (Fig. S2). Potential important sites are then delineated by grouping together grid cells used by a threshold percentage of the source population (Fig. S2, Supplementary

Information S1D). If the size of the source population is known or estimated, `findSite` multiplies the estimated proportion of the population using each grid cell by the population size to estimate the number of animals predictably using the candidate site; this is useful for assessing sites against standardized criteria, such as the KBA Standard (IUCN, 2016).

Example applications

We analyze data from three species to illustrate how ‘track2KBA’ can be used to identify important sites for populations of mobile animals. We emphasize that sites identified here are illustrative examples, not proposed sites for conservation. See Supplementary Information S2 for a walk-through with code and Supplementary Information S3 for further details.

Antarctic fur seals

Using tracking data from Antarctic fur seals from Bird Island, South Georgia we illustrate a typical workflow for deriving important sites for a population and their preliminary assessment against global KBA criteria. An estimated 64,545 female seals breed at Bird Island (Boyd, 1993), amounting to 8% of the global adult population. Using data from 117 females tracked during the breeding season, we identify important at-sea sites for this population and assess whether they might meet global KBA criteria (i.e., areas used predictably by ≥ 1% of global population under ‘Demographic aggregations’ Criterion D1a; IUCN, 2016).

During the breeding season, female fur seals forage at sea and regularly return to suckle their pups on land. Using the function `tripSplit`, we split the tracking data into foraging trips, defines as periods of ≥ 12 h away from the colony at a distance of ≥ 5 km (Fig. 2A), and filtered out points falling within this radius (argument ‘rmNonTrip’). We then calculated trip characteristics using the function `tripSummary`, which showed that seals travelled a mean maximum distance of 114 km (max 296 km) and spent a median of 5.6 days (max 19.8 days) at sea per foraging trip. Next, we projected the tracks to a custom-centered equal-area projection with the function `projectTracks`, and estimated core areas (i.e., 50% UDs) for each individual seal using the function `estSpaceUse` with an *h* parameter value of 4.65 km (Fig. 2B). The *h* value was calculated using the function `findScale` and represents the log of the median foraging range (in km); this value was selected as it captured areas representing ecologically realistic estimates of the space in which individuals spent most of their time at sea (Fig. S1B). Using `repAssess`, we estimated the representativeness of this tracked sample for the distribution of the wider Bird Island population, obtaining a value of 96% (Fig. 2C).

Next, we provided the individual core areas, the representativeness estimate, and the population size estimate as input to the function `findSite` and delineated sites used by at least 10% of the population (i.e., the default threshold when representativeness > 90%; Fig. 2D, Fig. S2). We identified an area of 1,576 km^2 to the north-west of South Georgia used by up to 23.6% (16,787 seals) of the female population (Fig. 3D, Fig. 4A). This translates to an estimated 1.7-2% of the 700,000-1,000,000 fur seals globally (Hofmeyr, 2016) that predictably use the site during breeding, potentially meeting the criteria for a global KBA under Criterion D1.

Green turtles

To demonstrate the utility of the package for a marine species when not foraging from a central place, we analyzed tracking data from green turtles at Poilão Island in the Bijagós archipelago of Guinea-Bissau. Poilão hosts one of the largest rookeries in the Atlantic (Catry et al., 2009), with an estimated laying population of 18,573 females (Supplementary Information 3). After nesting, female green turtles disperse to foraging grounds where they feed and remain resident until the subsequent breeding event (Hamann et al., 2002). We used ‘track2KBA’ to identify core areas for 23 tracked females and assess the degree to which this sample captures the distribution of the adult female population during the post-nesting foraging period. For each individual, we estimated core areas (50% UD) using an *h* parameter of 2.18 km, which was the median of the reference bandwidth across individuals (Fig. S5A) and reflects an ecologically realistic scale for the species when foraging.

Due to the broad area over which turtles disperse in the post-nesting period, and the restricted scale of their movement when foraging, the sample achieved only 32% representativeness and is therefore not considered representative of the population-level distribution. Given the low level of representativeness, no important sites for the source population were delineated. We estimated that 98 turtles would need to be tracked in order to achieve 70% representativeness, a level at which delineating important sites at the population level is more feasible (Lascelles et al., 2016). Nonetheless, we found overlapping core areas among the tracked individuals, indicating that important sites may be identifiable for this population with further data collection (Fig. 3B).

White storks

We used data from the population of white storks in Portugal to exemplify a use-case in a migratory system. Of the 46,027 white storks in Portugal, an estimated 26,196 migrate to sub-Saharan Africa and back each year (see Supplementary Information S3). During migration, white storks often aggregate in large numbers at stopover sites to refuel (Arizaga et al., 2018). We used ‘track2KBA’ to identify important stopover sites for this population on migration to and from sub-Saharan Africa, using the GPS tracks of 76 individuals.

We estimated core areas (50% UD) for each individual using an *h* parameter of 7.5 km, which was determined as the median peak in the variance of the log First-Passage Time across individuals (Fig. S7A). This method identifies the spatial scale at which the birds are spending the most time (Fauchald and Tveraa, 2003), which on migration should represent stopover sites. The estimated representativeness of the sample for migratory Portuguese white storks was 96% (Fig. S7B).

We delineated nine stopover sites, covering areas of between 19 km^2 and 1,150 km^2 in Spain and Morocco, used by at least 10% of the Portuguese population of migrating white storks. Of these sites, four are predictably used by up to 8,600 (2 sites), 9,600, and 11,580 storks, respectively representing 18.7%, 20.9%, and 25.2% of the total Portuguese population (i.e., migratory and resident birds) of white storks.

Conclusion

Sites of importance for avian diversity (i.e., mIBAs) have been identified for seabirds across the world using this framework (Dias et al., 2018; Lascelles et al., 2016). Our example applications illustrate that, given a representative sample of the population-level distribution, this method can be used to identify important sites for species other than seabirds. 'track2KBA' facilitates application of the method across other vertebrate taxonomic groups, which could assist expansion of the taxonomic coverage of important areas and ultimately protected-area networks.

For formal assessment of sites against global criteria, as in the KBA Standard, users must consult with the relevant body and additional steps may be required, such as consulting with relevant stakeholders to ensure that site boundaries reflect the management landscape. Users are encouraged to provide feedback about possible extensions to package functionality that may facilitate yet broader implementation of 'track2KBA'.

Acknowledgments

Thanks to Ben Lascelles and the participants of a workshop in Chizé in 2009 which was a major contribution to developing these methods. This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 766417. This communication reflects only the authors' views, and the Research Executive Agency of the European Union is not responsible for any use that may be made of the information it contains. Further funding included: grants to MARE (UIDB/04292/2020 and UIDP/04292/2020) by FCT, Portugal, funding for stork tracking came from the Natural Environment Research Council (NERC) (NE/K006312/1), the British Trust for Ornithology, (BTO) British Ornithological Union and by the FEDER Funds through the Operational Competitiveness Factors Program — COMPETE and by National Funds through FCT – Foundation for Science and Technology within the scope of the project POCI-01-0145-FEDER-028176. IC was funded by FCT through contract DL57/2016/CP1440/CT0023. The MAVA Foundation funded tracking of sea turtles through various projects granted to IBAP – Guinea-Bissau.

Author contributions

MB, SO, RAP, PC, and MD conceived the idea for the package. MB, EJP, JH, VM, and SO wrote the package code. PT and MM developed the original methods and wrote original code. AC, EJP, JH, SO, TD, and MD contributed to method development and earlier code versions. ARP, PC, AR, IS, AF, and IC collected data and contributed to data pre-processing; MB analysed the data and led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

Conflicts of interest

The authors declare no conflicts of interest.

References

Arizaga, J., Resano-Mayor, J., Villanúa, D., Alonso, D., Barbarin, J.M., Herrero, A., Lekuona, J.M., Rodríguez, R., 2018. Importance of artificial stopover sites through avian migration flyways: a landfill-based assessment with the White Stork *Ciconia ciconia*. Ibis 160, 542–553. https://doi.org/10.1111/ibi.12566

Boucher, J., Brooks, T. (Thomas M.), Dudley, N., Cuttelod, A., Langhammer, P.F., 2014. Applications of key biodiversity areas: End-user consultations. IUCN.

Boyd, I., 1993. Pup production and distribution of breeding Antarctic fur seals (*Arctocephalus gazella*) at South Georgia. Antarctic Science 5, 17–24.

Catry, P., Barbosa, C., Paris, B., Indjai, B., Almeida, A., Limoges, B., Silva, C., Pereira, H., 2009. Status, Ecology, and Conservation of Sea Turtles in Guinea-Bissau. Chelonian Conservation and Biology 8, 150–160. https://doi.org/10.2744/CCB-0772.1

Dias, M.P., Carneiro, A.P.B., Warwick-Evans, V., Harris, C., Lorenz, K., Lascelles, B., Clewlow, H.L., Dunn, M.J., Hinke, J.T., Kim, J.-H., Kokubun, N., Manco, F., Ratcliffe, N., Santos, M., Takahashi, A., Trivelpiece, W., Trathan, P.N., 2018. Identification of marine Important Bird and Biodiversity Areas for penguins around the South Shetland Islands and South Orkney Islands. Ecology and Evolution 0. https://doi.org/10.1002/ece3.4519

Fauchald, P., Tveraa, T., 2003. Using First-Passage Time in the Analysis of Area-Restricted Search and Habitat Selection. Ecology 84, 282–288. https://doi.org/10.1890/0012-9658(2003)084[0282:UFPTIT]2.0.CO;2

Gitzen, R.A., Millspaugh, J.J., Kernohan, B.J., 2006. Bandwidth Selection for Fixed-Kernel Analysis of Animal Utilization Distributions. Journal of Wildlife Management 70, 1334–1344. https://doi.org/10.2193/0022-541X(2006)70[1334:BSFFAO]2.0.CO;2

Hamann, M., Limpus, C.J., Owens, D.W., 2002. Reproductive Cycle of Males and Females, in: Lutz, P.L., Musick, J.A., Wyneken, J. (Eds.), The Biology of Sea Turtles, Volume II. CRC Press.

Hofmeyr, G., 2016. Arctocephalus gazella. The IUCN red list of threatened species 2016: e. T2058A66993062. https://dx.doi.org/10.2305/IUCN.UK.2016-1.RLTS.T2058A66993062.en

IUCN, A., 2016. A global standard for the identification of key biodiversity areas. Version 1, 2016–048.

Joo, R., Boone, M.E., Clay, T.A., Patrick, S.C., Clusella-Trullas, S., Basille, M., 2019. Navigating through the R packages for movement. arXiv:1901.05935 [q-bio, stat].

Lascelles, B.G., Taylor, P.R., Miller, M.G.R., Dias, M.P., Oppel, S., Torres, L., Hedd, A., Corre, M.L., Phillips, R.A., Shaffer, S.A., Weimerskirch, H., Small, C., 2016. Applying global criteria to tracking data to define important areas for marine conservation. Diversity and Distributions 22, 422–431. https://doi.org/10.1111/ddi.12411

Venter, O., Magrach, A., Outram, N., Klein, C.J., Possingham, H.P., Marco, M.D., Watson, J.E.M., 2018. Bias in protected-area location and its effects on long-term aspirations of biodiversity conventions. Conservation Biology 32, 127–134. https://doi.org/10.1111/cobi.12970

Watson, J.E.M., Dudley, N., Segan, D.B., Hockings, M., 2014. The performance and potential of protected areas. Nature 515, 67–73. https://doi.org/10.1038/nature13947

Worton, B.J., 1989. Kernel Methods for Estimating the Utilization Distribution in Home-Range Studies. Ecology 70, 164–168. https://doi.org/10.2307/1938423

Figures

Fig. 1. Overview of ‘track2KBA’ R package workflow for identifying important areas from tracking data. ‘Key steps’ are the essential functions for identifying and delineating areas of importance for biodiversity while ‘Optional steps’ are case-dependent.

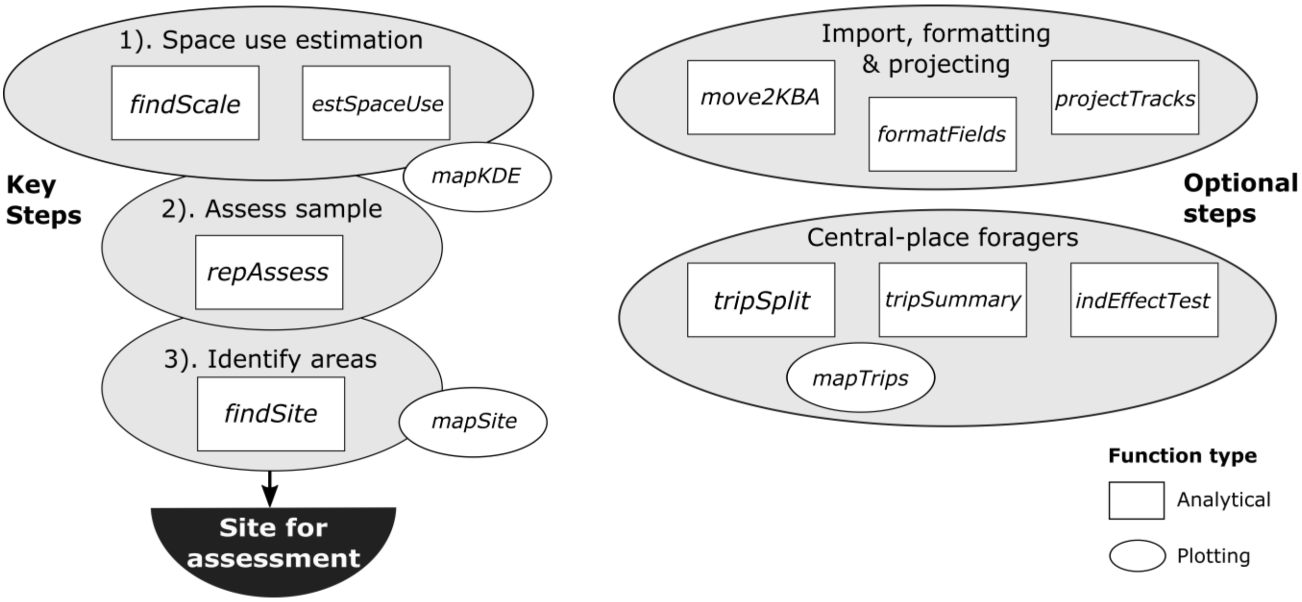

Fig. 2. Example 'track2kba' output from a tracking data set of Antarctic fur seals from South Georgia. (A) Using function `tripSplit`, data from each individual is split into foraging trips. (B) With the `estSpaceUse` function, the core areas are estimated for each individual, with colors representing individuals. (C) The degree to which the tracked sample (n=117) represents the distribution of the source population is estimated using the `repAssess` function. The sample was estimated to achieve 96% representativeness. (D) Using the `findSite` function, a conservative scaling up to the population is made based on the representativeness and the sample-derived pattern of overlap. Areas used by a threshold proportion of the population are delineated; here areas within the red border are used by at least 10% of the local population of fur seals.

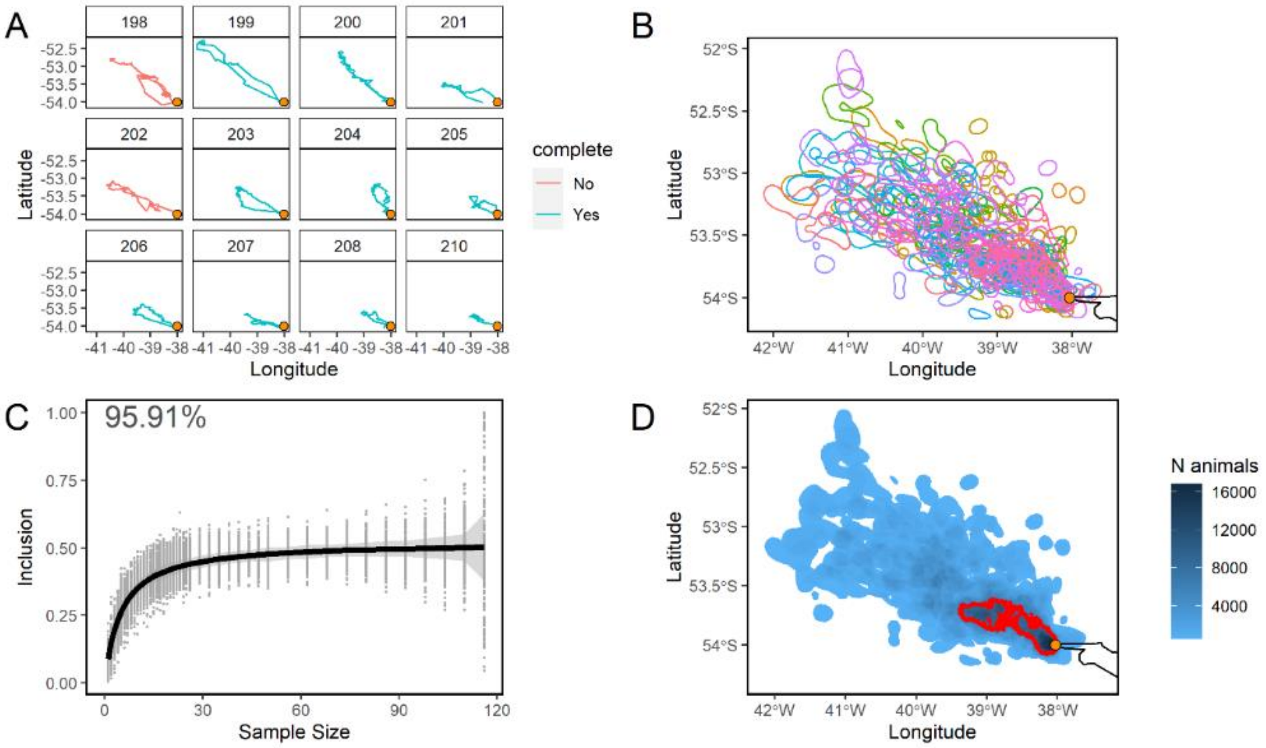

Fig. 3. Mapped results of ‘track2KBA’ tracking data analysis. (A) Areas used by at least 10% of the population of female Antarctic fur seals on Bird Island, South Georgia. (B) Areas used by female green turtles during post-nesting foraging in West Africa; areas shown reflect only overlap of the tracked individuals. (C) Stopover-sites used by 10% of white storks which migrate between Portugal and sub-Saharan Africa. Red diamonds signify the breeding colony or nesting beach.