Diversity and Distributions

WILEY

Diversity and Distributions

RESEARCH ARTICLE OPEN ACCESS

Migratory Connectivity and Non-Breeding Habitat Segregation Across Biogeographical Scales in Closely Related Seabird Taxa

Virginia Morera-Pujol[1,2] | Paulo Catry[3] | Maria Magalhães[4] | Clara Péron[5] | José Manuel Reyes-González[1,2] | José Pedro Granadeiro[6] | Teresa Militão[1,2] | Maria P. Dias[7] | Daniel Oro[8,9] | José Manuel Igual[9] | Giacomo Dell'Omo[10] | Martina Müller[11] | Vitor H. Paiva[12] | Benjamin Metzger[13] | Verónica Neves[14] | Joan Navarro[15] | Georgios Karris[16] | Stavros Xirouchakis[17] | Jacopo G. Cecere[18] | José Manuel Zamora-Marín[19,20,21] | Manuela G. Forero[22] | Isabel Afán[23] | Ridha Ouni[23] | Mohamed Salah Romdhane[24] | Fernanda De Felipe[1,2] | Zuzana Zajková[1,2,8] | Marta Cruz-Flores[1,2] | David Grémillet[25,26] | Jacob González-Solís[1,2] | Raül Ramos[1,2]

[1]Departament de Biologia Evolutiva, Ecologia i Ciències Ambientals, Facultat de Biologia, Universitat de Barcelona (UB), Barcelona, Spain | [2]Institut de Recerca de la Biodiversitat (IRBio), Universitat de Barcelona (UB), Barcelona, Spain | [3]MARE, Marine and Environmental Sciences Centre, ISPA-Instituto Universitário, Lisbon, Portugal | [4]Regional Secretariat for the Sea, Science and Technology, Regional Directorate for Sea Affairs (DRAM), Azores, Portugal | [5]Laboratoire de Biologie Des Organismes et Ecosystèmes Aquatiques (UMR BOREA) MNHN, CNRS, IRD, SU, UCN, UA, Paris, France | [6]CESAM, Departamento de Biologia Animal, Faculdade de Ciências, Universidade de Lisboa, Lisboa, Portugal | [7]Centre for Ecology, Evolution and Environmental Changes (cE3c), Department of Animal Biology, Faculdade de Ciências da Universidade de Lisboa, Lisboa, Portugal | [8]Centre d'Estudis Avançats de Blanes (CSIC), Carrer Accés Cala Sant Francesc, Blanes, Spain | [9]IMEDEA (CSIC-UIB), Esporles, Spain | [10]Ornis Italica, Rome, Italy | [11]Department of Natural Resources Science at the University of Rhode Island, Kingston, Rhode Island, USA | [12]Department of Life Sciences, MARE-Marine and Environmental Sciences Centre/ARNET – Aquatic Research Network, Calçada Martim de Freitas, University of Coimbra, Coimbra, Portugal | [13]BirdLife Malta, Ta'Xbiex, Malta | [14]Department of Oceanography and Fisheries, MARE – Marine and Environmental Sciences Centre, IMAR & Okeanos, University of the Azores, Azores, Portugal | [15]Institut de Ciències del Mar, CSIC, Barcelona, Spain | [16]Department of Environment, Faculty of Environment, Ionian University, Zakynthos, Greece | [17]University of Crete, Natural History Museum of Crete, Crete, Greece | [18]Area per l'Avifauna Migratrice (BIO-AVM) Istituto Superiore per la Protezione e la Ricerca Ambientale (ISPRA), Ozzano Emilia, Italy | [19]Southeast Naturalists Association (ANSE), Murcia, Spain | [20]Department of Zoology and Physical Anthropology, University of Murcia, Espinardo Campus, Murcia, Spain | [21]Department of Applied Biology, Centro de Investigación e Innovación Agroalimentaria (CIAGRO-UMH), Miguel Hernández University, Elche, Spain | [22]Estación Biológica de Doñana CSIC, Sevilla, Spain | [23]Université de Tunis El Manar, Faculté Des Sciences Mathématiques, Physiques et Naturelles de Tunis (FST), Tunis, Tunisia | [24]Université de Carthage Institut National Agronomique de Tunisie, Tunis, Tunisia | [25]Centre d'Ecologie Fonctionnelle et Evolutive, UMR 5175, CNRS – Université de Montpellier – Université Paul-Valéry Montpellier – EPHE, Montpellier, France | [26]Percy Fitzpatrick Institute of African Ornithology, NRF-DST Centre of Excellence at the University of Cape Town, Rondebosch, South Africa

Correspondence: Virginia Morera-Pujol (morera.virginia@gmail.com)

Received: 3 May 2024 | **Revised:** 27 February 2025 | **Accepted:** 1 March 2025

Editor: Christine Howard

Funding: EU H2020 programme through grant 634,495; Seventh Framework Programme (Research Executive Agency) through Marie Curie Career Integration Grant 618,841 (FP7-PEOPLE-2013-CIG); ESFRI LifeWatch Project; LIFE programme of the European Commission through projects LIFE10 NAT/MT090 and LIFE11 NAT/IT/000093; Ministerio de Ciencia e Innovación/Ministerio de Economia y Competitividad (Spain) through projects CGL2009–11278/BOS, CGL2013-42585-P, CGL2013-42203-R, CGL2016-78530-R and CGL2017-85210-P; Organismo Autónomo de Parques Nacionales (Spain) through project 1248/2014; Fundação para a Ciência e a Tecnologia (MCTES, Portugal) through projects MARE-UID/MAR/04292/2019; IF/00502/2013/CP1186/CT0003, PTDC/BIA-ANM/3743/2014, PTDC/MAR-PRO/0929/2014, UID/AMB/50017/2019 and UIDP/50017/2020 + UIDB/50017/2020 (to CESAM); Office Français de la Biodiversité (France), through the Programme PACOMM, Natura2000 en mer; Hellenic Bird Ringing Centre; MSDEC (Malta). VMP was supported by pre-doctoral contract BES-2014-068025 of the Spanish Ministerio de Industria, Economía y Competitividad; MM by grant SFRH/BPD/47047/2008 from the Portuguese Foundation for Science and Technology; JMRG by Ph.D. grant AP2009-2163 from the Spanish Ministerio de Educación; GDO and MMü by Ornis italica and by the Regione Siciliana and Assessorato Risorse Agricole e Alimentari through a grant to the Ringing Unit of Palermo; VHP by grant SFRH/BPD/85024/2012 from the Portuguese Foundation for Science and Technology; VN by grant SFRH/BPD/88914/2012 from the Portuguese Foundation for Science and Technology; and JN by the Spanish National Programme Ramón y Cajal (RYC-2015-17,809); GK and SX were partially funded by the Operational Program

“Environment and Sustainable Development” (EPPERAA) of the National Strategic Reference Framework (NSRF) 2007–2013, co-financed by the ERDF and Greek EDP; FdF by a Ph.D. grant from the Coordination for the Improvement of Higher Education Personnel (CAPES—Brazilian government agency; Bex Process 1307/13–4); ZZ by a PhD grant from the University of Barcelona (APIF/2012); MCF by a PhD grant from the University of Barcelona; and RR by postdoctoral contracts of the PLEAMAR programme from MINECO and Fundación Biodiversidad (2017/2349), and Ministerio de Ciencia, innovación y Universidades (RYC-2017-22,055). This publication is part of the project I + D + i/PID2020-117155GB-I00, funded by MCIN/AEI/10.13039/501100011033.

Keywords: animal movement | metapopulation | migration | migratory connectivity | seabird | species differenciation

ABSTRACT

Aim: In highly mobile species, Migratory Connectivity (MC) has relevant consequences in population dynamics, genetic mixing, conservation and management. Additionally, in colonially breeding species, the maintenance of the breeding geographical structure during the non-breeding period, that is, a strong MC, can promote isolation and population divergence, which ultimately can affect the process of lineage sorting. In geographically structured populations, studying the MC and differences in environmental preferences among colonies, populations, or taxa can improve our understanding of the ecological divergence among them.

Location: Atlantic Ocean and Mediterranean Sea.

Methods: We investigated the MC and non-breeding ecological niche of three seabird taxa from the genus *Calonectris* ($n = 805$ individuals). Using 1346 year-round trips from 34 different breeding colonies, we assess the level (from taxa to colony) at which MC and non-breeding spatial and environmental segregation emerge.

Results: At a taxon level, we found a clear difference in the non-breeding distributions between Cory's (*C. borealis*) and Scopoli's (*C. diomedea*) shearwaters, and a clear ecological divergence between Cory's and Cape Verde (*C. edwardsii*) shearwaters. At an intermediate aggregation level, we found that birds breeding in proximity had similar non-breeding habitat preferences, while birds breeding in very distant colonies (and therefore classified in different populations) had different non-breeding habitat preferences. Furthermore, within each taxon, we found more structure (i.e. stronger MC) and non-breeding divergence at an intermediate aggregation level than at the colony scale, where MC was weak.

Main Conclusions: These results suggest that conspecifics from nearby colonies mix in common non-breeding areas, but not with birds from more distant colonies or different taxa. These results support the need for management and conservation strategies that take into account this structure when dealing with migratory species with high connectivity.

1 | Introduction

Migratory Connectivity (MC), or the link between breeding and non-breeding distributions in migratory species, is defined as the preservation of the spatial structure between the breeding and non-breeding distributions, and is usually classified along a continuum from weak to strong (Webster et al. 2002). Weak MC occurs when the geographic structure of populations in the breeding grounds is not maintained in the non-breeding grounds (i.e. animals from different breeding populations mix in common non-breeding areas), while strong MC occurs when the geographic structure in the non-breeding grounds mirrors that of the breeding grounds (i.e. animals that breed closer to each other also spend the non-breeding period closer to each other, and farther from others, Finch et al. 2017).

In colonially breeding species, such as most seabirds, the breeding populations have a segmented geographic structure, formed by isolated patches of suitable habitat (islands or coastline) in a continuum of habitat unsuitable for breeding the ocean (Hanski and Gilpin 1997). Particularly in some highly mobile species, although the populations are not physically separated by geographical barriers or distance, the structure is still maintained, mostly by behavioural traits, such as natal philopatry and breeding site fidelity (Matthiopoulos et al. 2005; Friesen et al. 2007).

In such geographically structured species, strong MC favours the ecological divergence between populations as, with time, their non-breeding habitat preferences diverge, which eventually increases the genetic structure of the population (Quillfeldt et al. 2017; Hipfner et al. 2020), as the opportunities for encounters between animals of different breeding populations are reduced (Marra et al. 2019). Conversely, in species with weak MC, there are more opportunities for encounters between animals from different breeding populations in the non-breeding grounds, which can facilitate changes in breeding locations (e.g. even though pair bonding happens generally at the colonies, an animal (particularly a young and inexperienced one) sharing non-breeding ground with animals from another breeding population, can follow them back in their prenuptial migration and change breeding location), potentially leading to gene flow among populations, hybridisation and introgression (Tigano et al. 2015; Quillfeldt et al. 2017). Thus, MC has relevant consequences for population dynamics and genetic mixing, but also for conservation and management actions as the geographic structure of a population will dictate where conservation or management actions should be directed (Bauer et al. 2016; Finch et al. 2017).

In this context, tools to model the non-breeding distribution and ecological niche of these migratory, geographically structured species, such as Ecological Niche Models (ENMs, Melo-Merino et al. 2020), and to analyse at what scale connectivity patterns emerge (i.e. at colony, population, or taxon scale) can be particularly insightful when dealing with closely related taxa with incomplete lineage separation, for example, when morphological divergence is subtle, reproductive isolation is not absolute or genetic differentiation is still small (Rissler and Apodaca 2007; Fišer et al. 2018; Ferrer Obiol et al. 2023). This is the case for the *Calonectris* complex, a group of four taxa of pelagic seabirds breeding colonially in islands and islets in subtropical waters of the northern hemisphere (Warham 1996). Three of them have a parapatric breeding distribution spanning the North-eastern coast of the Atlantic Ocean and the Mediterranean Sea, and spend the non-breeding season in several discrete areas of the Atlantic and Indian Oceans, with different degrees of inter-specific mixing (González-Solís et al. 2007). The most recent common ancestor of the *Calonectris* species complex is thought to have been 4.3 Mya (Ferrer Obiol et al. 2022), subsequently splitting into the Pacific streaked shearwater (*Calonectris leucomelas*) and the Atlantic clade. From the latter, the Cape Verde shearwater (*C. edwardsii*, CALEDW hereafter) diverged later than 1.0 Mya and is consensually regarded as a full species (Hazevoet 1995), and Cory's (*C. borealis*; CALBOR) and Scopoli's (*C. diomedea*; CALDIO) shearwaters diverged shortly after that, late enough that they have only recently been recognised as separate species (Sangster et al. 2012), and there is still debate as to whether they should be considered as such (Genovart et al. 2013; Ferrer Obiol et al. 2022). Despite their mainly parapatric breeding distribution, there is a small region around the strait of Gibraltar where CALBOR and CALDIO breed in sympatry. Mixed-species pairs and occasional hybridisation have been reported from mixed colonies or nearby colonies of each taxon (Gómez-Díaz et al. 2006; Navarro et al. 2009; Zidat et al. 2017). These characteristics make them a valuable species complex to study ecological divergence through the analyses of non-breeding distributions and MC at different hierarchical levels (i.e. colony, population and taxon).

For this work, we collated a dataset of 1346 year-round tracks of 805 individuals from the Atlantic *Calonectris* species complex, breeding in 34 different colonies distributed along the Mediterranean and Atlantic basins. With this large dataset, we aimed to (1) compare MC values at three different aggregation levels (colony, a distance-based intermediate aggregation level and taxon) and (2) obtain non-breeding distributions and habitat preferences for each taxon and at different aggregation levels, in order to detect differences that can be indicative of ecological and, eventually, evolutionary divergence. First, we expect to find weak MC among breeding colonies of the same taxon, but stronger among different distance-based groups and strongest among different taxa. Second, we expect larger differences in habitat preferences during the non-breeding period among different taxa than among distance-based groups or colonies of a given taxon.

2 | Methods

2.1 | Studied Taxa and Sampling Design

Calonectris shearwaters are medium-sized Procellariforms, with long lifespans, high reproductive investment and low reproductive output (Navarro and González-Solís 2007; Ramos et al. 2012), and they display strong natal and breeding philopatry (Mougin et al. 1999). Females lay a single egg per season, and incubation and chick-rearing duties are shared by both parents (Thibault et al. 1997; Granadeiro et al. 2006). CALBOR breeds mainly on Macaronesia Islands (except Cabo Verde) and off the west coast of the Iberian Peninsula and spends the non-breeding season in different areas of the Atlantic and Indian Oceans, off the coasts of Africa and South America as well as two oceanic areas of the Atlantic, one south of Newfoundland and another in the South Atlantic over the mid-oceanic ridge (González-Solís et al. 2007; Dias et al. 2011). CALDIO breeds in the Mediterranean Sea and spends the non-breeding season in the Atlantic Ocean off the West coast of Africa and East coast of South America (González-Solís et al. 2007; Morera-Pujol et al. 2023), and CALEDW breeds in Cabo Verde Archipelago and spends the non-breeding season off the East coast of South America (Birdlife International 2019). The three taxa have similar phenology, arriving at the colonies for breeding in late February–early March. Laying occurs around the second half of May, and the hatching period is around mid to late July. In late October or early November, chicks fledge and abandon the colony, and they return to the colony after about 6 years of age when they reach breeding age (Granadeiro et al. 1999; Mougin et al. 2000; Paiva et al. 2015).

Our dataset comprises year-round at-sea positions of adult birds breeding in 13 CALBOR colonies (801 year-round tracks), 19 CALDIO colonies (445 year-round tracks) and two CALEDW colonies (100 year-round tracks), representing the entire breeding distribution of the three taxa, and obtained between the summers of 2006 and 2016 (Figure 1).

2.2 | Tracking Data

Positions were obtained with Global Location Sensors (GLS; Wilson et al. 1992), which provide one or two positions per day, using daylight to calculate longitude from the time of twilight and latitude from the length of the day (light period). Detailed deployment information for each colony can be found in Morera-Pujol et al. (2023). Twilight events calculated from the light measurements were visually inspected and corrected when interferences near the twilights were detected. Locations were obtained from the light data using either Intiproc (Migrate Technology Ltd.) or Biotrack (Biotrack Ltd.) software or the *GeoLight* package in R (Lisovski and Hahn 2012). To eliminate biases due to incorrect latitude estimations during or near equinoxes (Ekstrom 2004) we removed position data from 20 days before and after the equinoxes. In addition, we removed unrealistic positions by applying a quadratic speed filter following McConnell et al. (1992).

We visually inspected tracks to assign a phenological state to each position. Based on the distance between consecutive positions and the directionality of movement, we classified periods as 'residency' when for at least 4 days the movement was non-directional (following Dias et al. 2011). The residency periods during the breeding season and around the breeding colonies were classified as 'breeding'. The residency periods outside the breeding season (i.e. between the months of October and the following March) were classified as 'main non-breeding' (the longest) and 'staging' (the rest of them). Movement periods (directional travel) were classified as 'migration'. For our analysis, we used only the residency periods during non-breeding.

Since our dataset contains individuals tracked for up to eight consecutive years and along a 10-year period, we followed Morera-Pujol et al. (2023) to check for the biases caused by individual site fidelity and temporal variability in the spatial distribution of each of the taxa. The colony of Fifla showed a high effect of individual site fidelity, so for it, we randomly selected only one track per individual for the colony of Filfla (Malta, CALDIO), but for all other colonies, no effects of individual site fidelity or

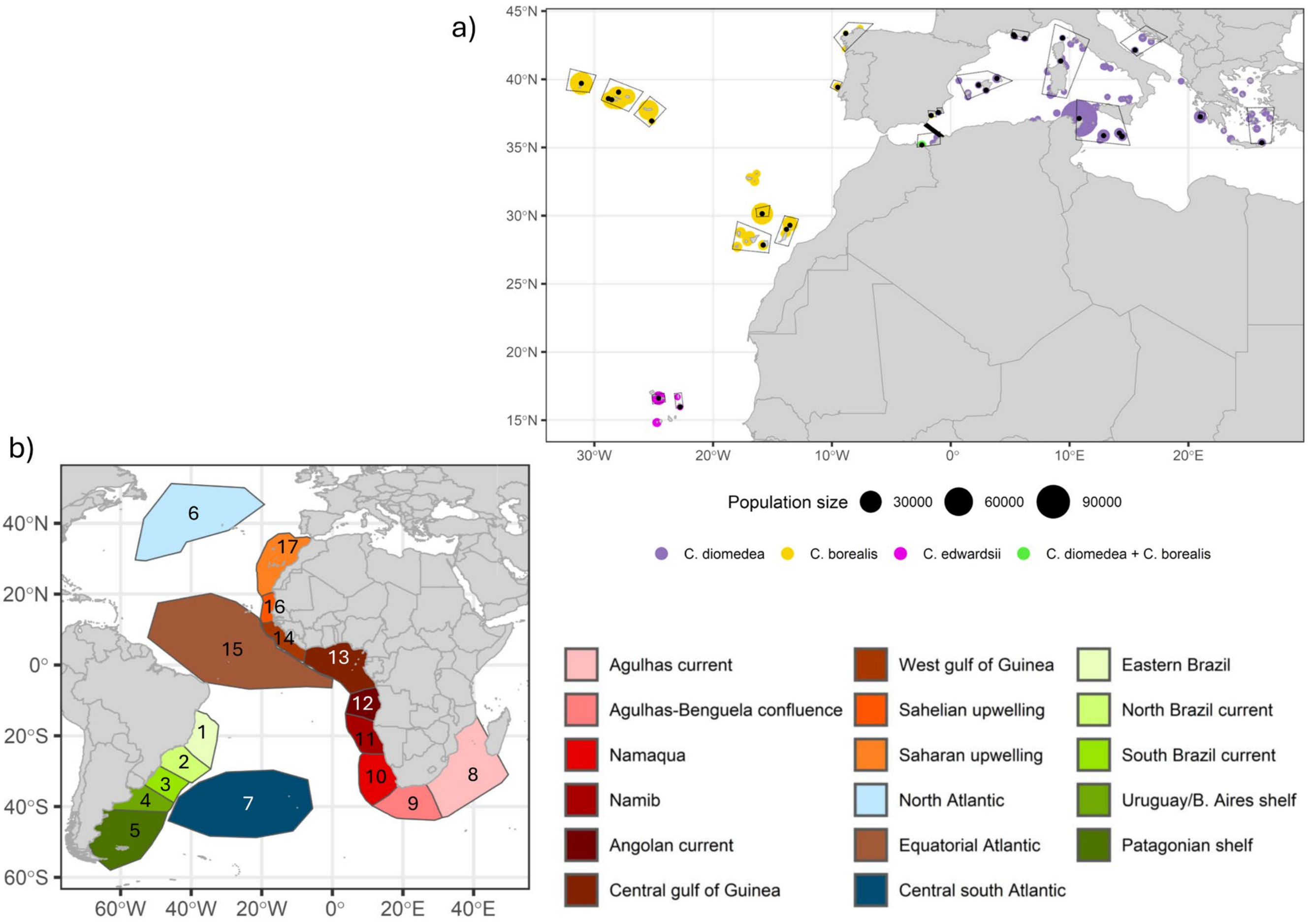

FIGURE 1 | Breeding distribution and potential non-breeding areas of the three *Calonectris* shearwaters studied: (a) known breeding locations of CALBOR (in yellow), CALDIO (in purple) and CALEDW (in pink). The size of the coloured circle is proportional to the number of breeding birds estimated for each colony and the colonies with black full circles inside are the ones where we have tracked birds. Notice that a mixed colony of CALBOR and CALDIO (in green) occurs in Chafarinas Islands, in the area of the Almería—Orán oceanographic front (thick black line). The dashed, thin black lines delimit populations for each taxon. (b) Non-breeding areas of the three Palearctic *Calonectris* species as used for the analyses of this work. The colours and numbers in this figure correspond to those in Figures 2 and 3. Adapted from Spalding et al. 2007; see Section 2).

temporal variability were detected. Additionally, for all colonies, to make sure each of our samples was a good representation of their colonies, we used the function *repAssess* from the package *track2kba* (Beal et al. 2021), which checks the representativeness of each colony sample. This function uses an approach similar to that of species discovery curves to assess whether a sample contains enough variability to be a good representation of the sampled colony. Following Morera-Pujol et al. 2023 we selected a representativeness threshold of 70% and excluded all colonies under that threshold or with a sample size under five tracks from the analyses (CALBOR colonies of Corvo—Central Azores—and Graciosa—West Azores—, and CALDIO colonies of Porquerolles—France—, Gozo—Sicilian Channel—, and Na Pobra and Na Foradada—Balearic Islands—more details in Morera-Pujol et al. 2023).

2.3 | Migratory Connectivity

We estimated the MC following Cohen et al. (2017), which improves on the classical MC method (Ambrosini et al. 2009). The method uses transition probabilities between specific breeding and non-breeding sites and can consider uneven sampling and differences in abundance among sites. The resulting MC measure can range between −1 and 1. MC close to 1 indicates that populations that breed close by also spend the non-breeding period close by (strong MC) and MC close to 0 indicates that the relationship between breeding and non-breeding distances is no different from what is expected at random (weak MC). Negative values of MC could theoretically be as low as −1, indicating that populations that breed close together spend the non-breeding period farther apart, but even with simulated datasets designed to obtain negative MC values, the lowest value obtained was −0.06, so it is possible that large negative values are rare or impossible to obtain with natural populations data (Cohen et al. 2017).

The method used for calculating MC requires the a priori delimitation not only of breeding sites but also of non-breeding areas. Thus, we divided the non-breeding distribution of the three taxa using the Marine Ecoregions Of the World (MEOW; Spalding et al. 2007). These regions extend only to 200 nautical miles offshore or to the 200 m isobath (whichever occurs closest to land). Since *Calonectris* shearwaters can spend the non-breeding season further offshore, we extended these limits to include all of the non-breeding positions in each area, but maintaining the limits between regions (Figure 1). In addition, since MEOW does not include offshore areas, but some shearwaters do spend their non-breeding periods in the middle of the Atlantic Ocean, we generated three offshore areas based on prior knowledge of the wintering distribution of these species: North Atlantic (55° W—19° W, 30° N—21° N), Equatorial Atlantic (52° W—0° W, 6° S—20° N), and central South Atlantic (46° W—6° W, 48° S—30° S; Figure 1).

We calculated MC at three different levels: colony, 'population', and taxon. At a colony level, we calculated three MC values, one for each taxon, assigning a small buffer of 1 km around each colony to define breeding sites. To try and detect structure within each taxon, we grouped colonies into what hereafter we will call populations, which are aggregations of colonies based on distance. We delineated these populations by taking all known colonies based on Morera-Pujol et al. 2023 (sampled and unsampled) and grouping them if their 100 km buffers overlapped (Figure 1). The 100 km buffer was chosen via trial and error as a compromise between smaller values that would group colonies in many small populations and larger values that would group almost all colonies of a species in one single population. This provides a good intermediate level of aggregation that has geographical sense as it is based on proximity and will allow us to test whether populations that are close by are more similar to each other. We calculated an MC value at a population level for CALBOR and one for CALDIO (there were not enough sampled colonies of CALEDW to perform this) at a population level. Each population's size was calculated as the sum of the colony size of all colonies (sampled and unsampled) of that population (colony and sample sizes are specified in Morera-Pujol et al. 2023). Finally, we calculated one single overall MC value at a taxon level; that is, grouping all colonies of each taxon in a single origin and using thus the breeding sites of the entire taxon as the breeding area.

MC estimates are reported as median values (95% Confidence Interval, CI) along this study. The functions developed by Cohen et al. (2017) are implemented in the package *MigConnectivity* (Hostetler and Hallworth 2018). We used the function *estMC*, designed to estimate the value of MC taking into account differences in the number of breeding birds among origins (i.e. breeding sites), different sampling efforts, and the position error associated with the GLS data (with estimates of bias and variance obtained from modelling ground-truthing GLS data). For each aggregation level (colonies, populations, and taxa) we ran the *estMC* function with 1000 simulations (value recommended for GLS data) and 1000 samples, with a maximum number of tries of 300 (more details can be obtained in Cohen et al. 2017).

To visually represent the MC, we constructed circular plots that connected breeding sites and non-breeding areas with the package *circlize* (Gu et al. 2014). The first step to construct these plots was to define a single non-breeding area for each individual, but each individual's non-breeding locations were usually distributed among more than one of the pre-defined non-breeding areas. To identify the main non-breeding area, we overlapped all non-breeding locations of each individual with the pre-defined non-breeding areas and assigned as the main non-breeding area that which contained the largest proportion of non-breeding locations. Second, a circumference was drawn, composed of sectors representing the breeding sites (top half) and non-breeding areas (bottom half). The arc lengths of the sectors were proportional to the number of breeding birds of each breeding site (top half) or the number of animals using each non-breeding area (bottom half). The thickness of the links (chords) joining each breeding site to a non-breeding area was designed to be proportional to the number of individuals going from each origin to each destination, corrected by the number of breeding birds at the origin and sampling effort (i.e. dividing by the sample size and multiplying by the population size). Detailed population and sample size numbers for all colonies and species can be found in Appendix S1 of Morera-Pujol et al. 2023.

2.4 | Environmental Data and Ecological Niche Modelling

To model the habitat preferences of shearwaters, we considered environmental variables known to affect seabird at-sea distributions (Wakefield et al. 2009; Louzao et al. 2011). Sources of the environmental variables, units and processing details can be found in Table S1.

We used Bayesian Additive Regression Trees (BARTs, Chipman et al. 2010) through the functions provided in the package *embarcadero* (Carlson 2020). Regression-tree modelling approaches like this one obtain the probability of the desired output variable (likelihood of an individual of a species being present in our case) based on decision trees that use binary-rule sets to split the predictor variables (Carlson 2020). We modelled the non-breeding distributions for each colony, population, for each taxon as a whole, and for the three taxa together (supra-taxon level hereafter). For all multi-colony models (i.e. population, taxon and suprataxon), our non-breeding locations did not accurately represent the real contribution of each colony to the higher-level population size, due to different sampling efforts on each colony and different colony sizes. To correct that, we simulated, from the non-breeding location data, and with the same relative spatial distribution, a number of locations proportional to colony size using the function *simulateDistribution*() from Morera-Pujol et al. 2023. The function creates a kernel density surface from the real at-sea locations and then randomly generates locations following that density to reproduce the real distribution with a known number of locations. From this process, we obtained a set of simulated locations for each population and taxa, and for the supratTaxon level, that were representative of the proportional colony sizes and corrected for sample size. Then, following Carlson (2020) we subsampled our simulated distributions to obtain 2000 locations for each model using the function *rthin* from the package *spatstat* (Baddeley et al. 2015). This function performs independent random thinning, producing a point pattern with the same relative density as the original one but a smaller number of points. For the single-colony models, we did not have to worry about the contribution of different colonies to the model. However, the difference in model accuracy depending on the size of the dataset was still a concern, since our sample sizes ranged between 536 and 27,081 locations per colony. Therefore, to make the process more comparable with the one described above for the multi-colony models, we also simulated non-breeding distributions for each colony and then subset 2000 locations to be used as presences in the model.

After this process, we had a dataset of 2000 simulated non-breeding locations to be used as presences for each colony, population, taxon, and supratanon model. Still following Carlson (2020), 2000 pseudo-absences were generated for each model using the function *randomPoints* from the package *dismo* (Hijmans et al. 2022). These were generated within the spatial range of the covariates (i.e. Atlantic Ocean and Mediterranean Sea) and only at sea (not on land), but otherwise randomly. From the models we ran, we obtained modelled non-breeding distributions that we used in three different ways: (1) we used the three taxon-level models to study spatial segregation among the three taxa and to compare model fit with that of the supratanon model; (2) we used the population-level models of CALBOR and CALDIO to study spatial structure in the non-breeding distribution within each taxon; and (3) we used the colony-level models to study how their pairwise correlation related to the pairwise distances between colonies.

One of the characteristics of using BARTS instead of other non-Bayesian tree-based methods is that it provides a measure of model uncertainty in the predictions; therefore, the modelled non-breeding distributions produced at every iteration are means of the posterior distribution. Additionally, the functions from the *embarcadero* package perform stepwise variable selection by running first a model with a full set of predictors and then iteratively removing the less important one and running the model again with the remaining predictors. It calculates the Root-Mean-Square Error (RMSE) of each of the resulting models and recommends the set of predictors of the model with lower RMSE. For each model, we used the predictors recommended by the stepwise selection.

Model fit was evaluated through the Area Under the Curve (AUC), and predictor importance was calculated as the proportion of total branches used for a given predictor. We use the AUC not only as a measure of model fit but also as a proxy for population structure. The reasoning behind it is that a more structured population will have disparate non-breeding areas that might be different in terms of environmental variables, which will cause the model to struggle in classifying the presences based on the value of the environmental variables, and therefore produce a lower AUC, whereas in an unstructured population all individuals mix in the non-breeding distribution which will have a more uniform environment, so we expect the AUC in these models to be higher. If the three taxa do indeed have different environmental preferences, we expect the AUC of the supra-taxon model to be the lowest. Lastly, the effect of the environmental variables on the distributions was evaluated through partial dependency plots, which calculate the mean effect of the variable of interest (partial effect) measured at all other values of the other predictors for all the values of the variable of interest.

The procedure for all models, from the generation of the simulated dataset to the calculation of variable importance and model projections, was repeated 10 times for each model in order to obtain uncertainty estimates around the variable importance and AUC values. For the modelled non-breeding distributions, we present the mean and SD of all mean distributions in the main text, and the 10 mean distributions for each species can be found in Figure S1.

For the colony levels, and to measure similarity between non-breeding habitats of colonies within each taxon, we calculated pairwise Pearson correlations between all pairs of modelled non-breeding distributions using the function *layerStats* from the package *raster*, and pairwise distances between colonies. We then used linear models—*lm* function from the *stats* package—to calculate the relationship between pairwise colony distances and pairwise Pearson correlations between corresponding model projections. For computational efficiency, and because the uncertainty obtained would not have been propagated to the linear model, we did not run the colony models for 10 iterations. All analyses were performed in R (version 4.2.3, R Core Team). The code is available at https://github.com/VirginiaMorera/Migratory-connectivity, and the data at Dryad at https://doi.org/10.5061/dryad.g79cnp5z7.

3 | Results

3.1 | Migratory Connectivity

MC (which ranges between 0 and 1) was estimated at the colony, population, and taxon levels. Within CALBOR, MC was 0.04 (−0.01–0.11) at the colony level and 0.20 (0.10–0.40) at the population level (Table 1 and Figure S2). Birds from the colony of Veneguera (West Canary Islands), Montaña Clara (East Canary Islands) and Selvagem Grande (Savage Islands) used most of the non-breeding regions (13, 12 and 12, respectively, out of 17; Figure 2). The Namaqua coast and the Agulhas—Benguela confluence non-breeding regions were used by all colonies from all populations, while the central coast of the gulf of Guinea was only used by birds from Chafarinas (Figure 2).

Within CALDIO, MC was 0.05 (−0.02 to 0.17) at a colony level and 0.31 (0.24–0.50) at a population level (Table 1 and Figure S2). The colonies of Pantaleu (Balearic Islands) and Palomas (South-east Spain) used most of the non-breeding areas (seven and eight, respectively), while Paximada (Aegean Sea) and Strofades (eastern Ionian Sea) use only three of them. Of the non-breeding regions, the Agulhas current is exclusive to birds from Palomas and used only by one individual, and the northern Brazil current is used only by birds from two colonies, Frioul (France) and Lavezzi (Corsica—Sardinia; Figure 2). Conversely, the Sahelian coast was used by birds

TABLE 1 | Migratory Connectivity (MC) values (median and 95% CI) estimated at a taxon level, and then at a population and colony level for CALBOR and CALDIO. For CALEDW there is only analysis at a colony level because the two studied colonies belong to the same population.

Aggregation level	Species	Median	95% CI
Taxon		0.45	0.42–0.49
Population	CALBOR	0.20	0.10–to 0.40
Population	CALDIO	0.31	0.24–to 0.50
Colony	CALBOR	0.04	−0.01–0.11
Colony	CALDIO	0.05	−0.02–0.17
Colony	CALEDW	0.01	−0.01–0.04

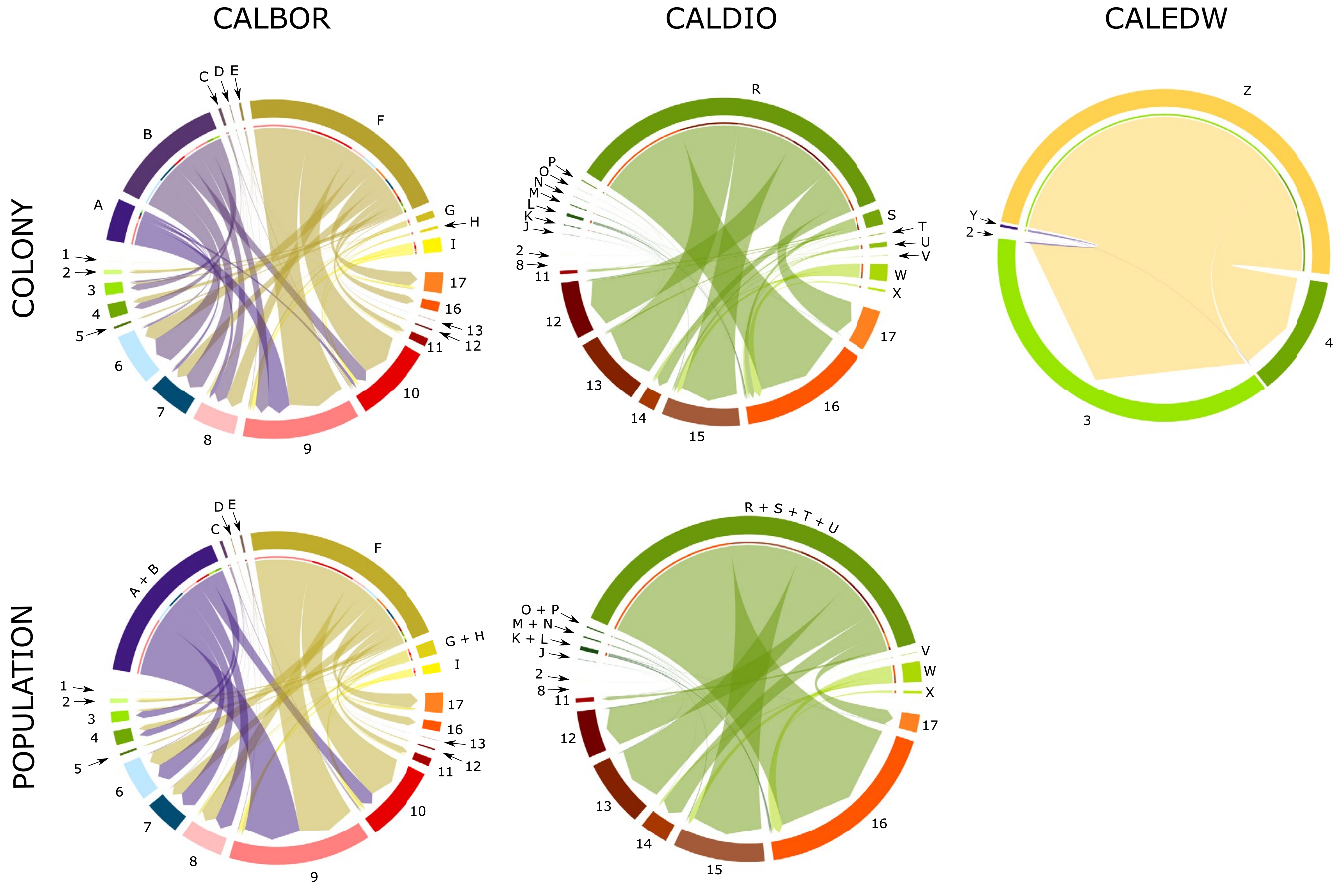

FIGURE 2 | Circular plots showing the links between each breeding (top half) and non-breeding area (bottom half) for colonies (top) and populations (bottom) of CALBOR (left), CALDIO (centre) and CALEDW (right). The arcs of the sectors are proportional to the number of individuals in each area (colony or population, or non-breeding area), and the widths of the links are proportional to the number of individuals travelling to each of the non-breeding areas, after correcting for the different sampling effort in each colony and accounting for different population sizes. Colonies: For CALBOR (A) Faial, (B) Pico, (C) Vila, (D) Chafarinas, (E) Berlenga, (F) Selvagem Grande, (G) Montaña Clara, (H) Timanfaya, and (I) Veneguera; for CALDIO (J) Palomas, (K) Pantaleu, (L) Cala Morell, (M) Riou, (N) Frioul, (O) Giraglia, (P) Lavezzi, (R) Zembra, (S) Linosa, (T) Filfla, (U) Malta, (V) Tremiti, (W) Strofades, and (X) Paximada; and for CALEDW (Y) Curral Velho, (Z) Raso. Populations: For CALBOR (A + B) Central Azores, (C) East Azores, (D) Chafarinas, (E) Berlengas, (F) Savage islands, (G + H) East Canary Islands, (I) West Canary Islands; and for CALDIO (J) Palomas, (K + L) Balearic Islands, (M + N) France, (O + P) Corsica and Sardinia, (R + S + T + U) Sicilian Channel, (V) Middle Adriatic, (W) Strofades, (X) Aegean Sea. Non-breeding areas: (1) Eastern Brazil, (2) Northern Brazil current—South-eastern Brazil, (3) South Brazil current, (4) Uruguay—Buenos Aires shelf, (5) Patagonian shelf, (6) North Atlantic, (7) Central-southern Atlantic, (8) Agulhas current, (9) Agulhas—Benguela confluence, (10) Namaqua, (11) Namib, (12) Angolan current, (13) Central gulf of Guinea, (14) West gulf of Guinea, (15) Equatorial Atlantic, (16) Sahelian upwelling and (17) Saharan upwelling.

breeding in all colonies from all populations, while the Equatorial Atlantic was used by birds of all populations and all colonies but Frioul (Figure 2).

Within CALEDW, as each of the two colonies represented one population, there is only one measure of MC. The median MC value was 0.01 (−0.01 to 0.04; Table 1 and Figure S2). The three non-breeding regions were used in a similar proportion by birds from the two colonies (Figure 2).

At a taxon level (i.e. grouping all individuals by taxon and estimating the MC taking the three taxa as the three only possible origins, and all the non-breeding regions as possible destinations) MC was 0.45 (0.42–0.49; Table 1 and Figure S2). From the 17 non-breeding regions available, CALBOR used a total of 15, of which it shares nine with either of the other two taxa (60.0%, Figure 3). CALDIO used a total of nine, of which seven were shared with at least one of the other two taxa (77.8%) and CALEDW shearwaters used a total of three non-breeding regions, all of them shared (Figure 3).

3.2 | Ecological Niche Modelling

From the Ecological Niche Models (ENMs) we obtained modelled non-breeding distributions at four levels: colony, population, taxon and suprataxon. All models fitted well, with AUC values between 0.886 (±0.001) and 0.994 (±0.000, Table 2). The highest AUC values overall were for CALEDW populations, and the lowest for CALBOR populations, with the lowest value of all being that of the suprataxon model (modelling the three taxa together, AUC = 0.894 (±0.003)). For CALBOR, the AUC for the model at a taxon level was 0.894 (± 0.005), the lowest of the three taxa and equal to that of the suprataxon model (Table 2).

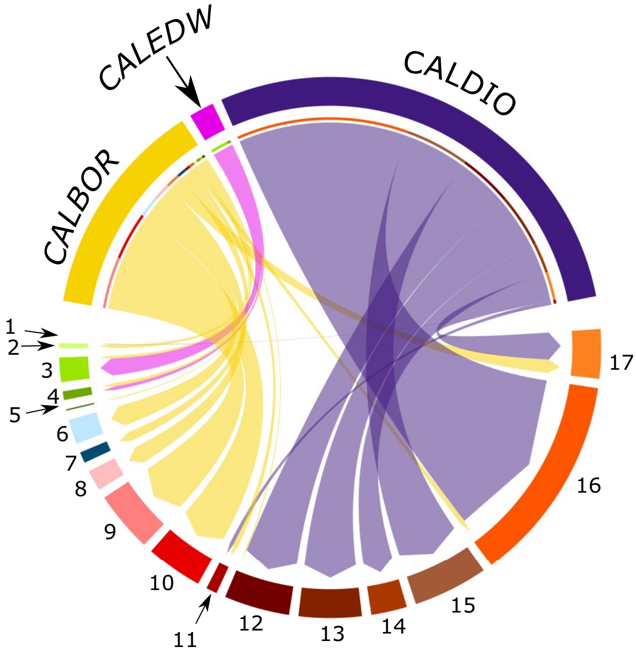

FIGURE 3 | Circular plot showing the links between each taxon (top half) and non-breeding area (bottom half). The arcs of the sectors are proportional to the number of individuals of each taxon (top) and the number of individuals using each non-breeding area (bottom), and the widths of the links are proportional to the number of individuals from each taxon travelling to each of the non-breeding areas, after correcting for the different sampling effort in each colony and accounting for different population sizes. Non-breeding areas: (1) Eastern Brazil, (2) Northern Brazil current—South-eastern Brazil, (3) South Brazil current, (4) Uruguay—Buenos Aires shelf, (5) Patagonian shelf, (6) North Atlantic, (7) Central-southern Atlantic, (8) Agulhas current, (9) Agulhas—Benguela confluence, (10) Namaqua, (11) Namib, (12) Angolan current, (13) Central gulf of Guinea, (14) West gulf of Guinea, (15) Equatorial Atlantic, (16) Sahelian upwelling, and (17) Saharan upwelling.

The stepwise selection process discarded slope as a relevant environmental variable from all population, taxa, and supratexon models in all their iterations, except from the French population model (Table 2). Sea Surface Temperature (SST) gradient was excluded from three of the models as it did not get selected in at least five iterations, and bathymetry and salinity were excluded from one model for the same reason. All the other variables were included in all models, with varying degrees of importance. SST was the most important variable for all population, taxa and supratexon models, with importance values between 0.179 (±0.002) and 0.285 (±0.002). Chlorophyll A (ChlA) and ChlA variability where also highly important for all the models (Table 2).

Modelled non-breeding distributions obtained from the Ecological Niche Modelling (ENM) showed optimal habitat for CALBOR mainly in areas of the Atlantic between the 20° and 50° latitude (both North and South), and with high suitability values in coastal as well as in oceanic waters. For CALDIO, high values of habitat suitability were contained mainly in areas between 20° North and South, predominantly in coastal waters but also in some oceanic waters. Lastly, for CALEDW, highly suitable areas are concentrated only in coastal waters, mainly off the coast of southern Brazil and Uruguay, but also in some narrow strips along the coasts of the South and the North-east of the African continent (Figure 4, top). The 95% credible interval (CI) widths show low levels of uncertainty for all taxa but especially for CALEDW. We only observe very large CI widths (i.e. high uncertainty) inside the estuary of the Río de la Plata, where extreme levels of salinity and productivity make model predictions very uncertain (Figure 4, bottom).

Partial effects plots showed the different environmental preferences of each species, with CALDIO and CALEDW showing preference for warmer waters than CALBOR, and CALDIO showing a preference for more dynamic waters (higher suitability at high values of ChlA variability). The three species show similar preferences for waters with high concentrations of ChlA (Figure 5).

Finally, pairwise correlations (representing similarity) among colony-level projections ranged from 0.27 (between Chafarinas and Pico) to 0.91 (between Berlenga and Selvagem Grande) for CALBOR and from 0.29 (between Chafarinas and Strofades) to 0.92 (between Giraglia and Riou) for CALDIO. The linear models showed that the similarity between model projections was negatively related to the distance between colonies for the two taxa, with significant correlations for both, and R-squared values of 0.39 and 0.34, respectively (Figure S3).

4 | Discussion

In this study, we combined Migratory Connectivity (MC) and Ecological Niche Modelling (ENM) to analyse ecological divergence between three closely related taxa of seabirds. A strong MC indicated the presence of structure between the three taxa, and moderate MC values indicated the presence of some structure at a population level within each taxon, which was consistent with the different non-breeding distributions and habitat preferences found through the ENM. These results have important implications for population dynamics and structure, as well as the underlying evolutionary processes of the species.

4.1 | Migratory Connectivity and Structure at Different Aggregation Levels

MC and ENM analyses revealed a clear geographic segregation among the three taxa: CALBOR and CALDIO clearly selected different non-breeding areas, which were reflected in the relatively high MC at a taxon level, and visible in the geographical projections of the ENM. That is, optimal habitat for CALDIO predominantly fell between 30° N and 30° S, exhibiting a clear tropical distribution, whereas optimal habitat for CALBOR extended beyond these limits, showing a clear sub-tropical distribution (Figure 4). Although CALEDW partly shared non-breeding areas with CALBOR, the two species diverged in their habitat use. CALEDW tended to occupy warmer, more productive and dynamic waters than CALBOR. This suggested an ecological, rather than spatial, divergence between the two.

The decreasing values of MC and increasing AUC values from taxon level to populations to colonies evidence the presence of

TABLE 2 | Model evaluation parameters (AUC and variable importance) for each of the BART models at a population, taxon and supratraxon level.

Group	Population	AUC	SST	ChlA	ChlA variability	Salinity	Bathymetry	SST gradient	Slope
CALBOR	Central Azores Is.	0.886 ± 0.001	**0.201 ± 0.007**	*0.178 ± 0.010*	0.176 ± 0.011	0.157 ± 0.009	0.151 ± 0.003*	0.152 ± 0.012	
	East Azores Is.	0.892 + 0.001	**0.217 ± 0.018**	0.177 ± 0.012	*0.183 ± 0.012*	0.166 ± 0.012	0.160 ± 0.014	0.138 ± 0.004[a]	
	Chafarinas Is.	0.917 ± 0.002	**0.208 ± 0.011**	*0.171 ± 0.009*	0.162 ± 0.010	0.166 ± 0.010	0.162 ± 0.008	0.145 ± 0.002[b]	
	Berlengas Is.	0.914 ± 0.001	**0.236 ± 0.003**	0.193 ± 0.002	*0.197 ± 0.002*	0.176 ± 0.002	0.197 ± 0.004		
	Savage Is.	0.911 ± 0.001	**0.208 ± 0.010**	0.166 ± 0.010	*0.171 ± 0.009*	0.160 ± 0.010	0.161 ± 0.011	0.149 ± 0.003[b]	
	East Canary Is.	0.944 ± 0.001	**0.202 ± 0.013**	*0.175 ± 0.007*	0.154 ± 0.009	*0.175 ± 0.009*	0.157 ± 0.012	0.153 ± 0.003[b]	
	West Canary Is.	0.932 ± 0.002	**0.227 ± 0.014**	0.188 ± 0.013	0.170 ± 0.015	*0.189 ± 0.015*	0.182 ± 0.016		
	Taxon level model	0.894 ± 0.005	**0.201 ± 0.003**	0.161 ± 0.005	*0.171 ± 0.006*	0.162 ± 0.005	0.153 ± 0.005	0.152 ± 0.004	
CALDIO	Palomas	0.946 ± 0.001	**0.285 ± 0.003**	*0.246 ± 0.00451*	0.240 ± 0.004	0.230 ± 0.005			
	Balearic Is.	0.948 ± 0.002	**0.217 ± 0.016**	*0.185 ± 0.014*	0.174 ± 0.010	0.169 ± 0.016	0.161 ± 0.011	0.134 ± 0.003[a]	
	France	0.952 ± 0.001	**0.179 ± 0.003**	*0.156 ± 0.003*	0.155 ± 0.003	0.124 ± 0.001	0.130 ± 0.002	0.126 ± 0.003	0.130 ± 0.002
	Corsica—Sardinia	0.958 ± 0.001	**0.203 ± 0.003**	*0.176 ± 0.003*	0.174 ± 0.003	0.146 ± 0.003	0.149 ± 0.002	0.152 ± 0.003	
	Sicilian Channel	0.941 ± 0.001	**0.213 ± 0.005**	*0.168 ± 0.005*	*0.168 ± 0.005*	0.160 ± 0.003	0.154 ± 0.002	0.136 ± 0.002	
	Middle Adriatic	0.936 ± 0.001	**0.197 ± 0.005**	*0.184 ± 0.004*	0.168 ± 0.005	0.163 ± 0.003	0.153 ± 0.003	0.136 ± 0.004	
	Eastern Ionian Sea	0.968 ± 0.001	**0.221 ± 0.012**	0.192 ± 0.013	*0.196 ± 0.013*		0.198 ± 0.012	0.164 ± 0.013	
	Aegean Sea	0.957 ± 0.001	**0.196 ± 0.005**	0.169 ± 0.003	*0.175 ± 0.005*	0.157 ± 0.002	0.155 ± 0.002	0.148 ± 0.003	
	Taxon level model	0.931 ± 0.012	**0.217 ± 0.024**	0.175 ± 0.022	*0.177 ± 0019*	0.169 ± 0.030	0.148 ± 0.003	0.143 ± 0.010	
CALEDW	East Cabo Verde	0.994 ± 0.000	**0.201 ± 0.002**	*0.171 ± 0.003*	0.163 ± 0.002	0.151 ± 0.002	0.149 ± 0.003	0.166 ± 0.007	
	West Cabo Verde	0.991 ± 0.000	**0.200 ± 0.003**	*0.165 ± 0.003*	0.164 ± 0.002	0.153 ± 0.003	0.154 ± 0.003	0.164 ± 0.003	
	Taxon level model	0.993 ± 0.001	**0.200 ± 0.006**	*0.167 ± 0.007*	0.162 ± 0.008	0.149 ± 0.007	0.148 ± 0.009	0.162 ± 0.007	
Suprataxon		0.894 ± 0.003	**0.204 ± 0.004**	0.164 ± 0.003	*0.169 ± 0.005*	0.158 ± 0.007	0.152 ± 0.004	0.152 ± 0.005	

Note: Each model was iterated 10 times and means and standard deviations for each parameter are presented. Only variables selected in at least 5 of the 10 iterations are shown. Every other variable was selected in all 10 iterations. Row wise, values in bold indicate the most important variable for each model, and values in italic the second most important variable.

[a] Variables selected in seven out of 10 iterations.

[b] Values selected in nine out of 10 iterations.

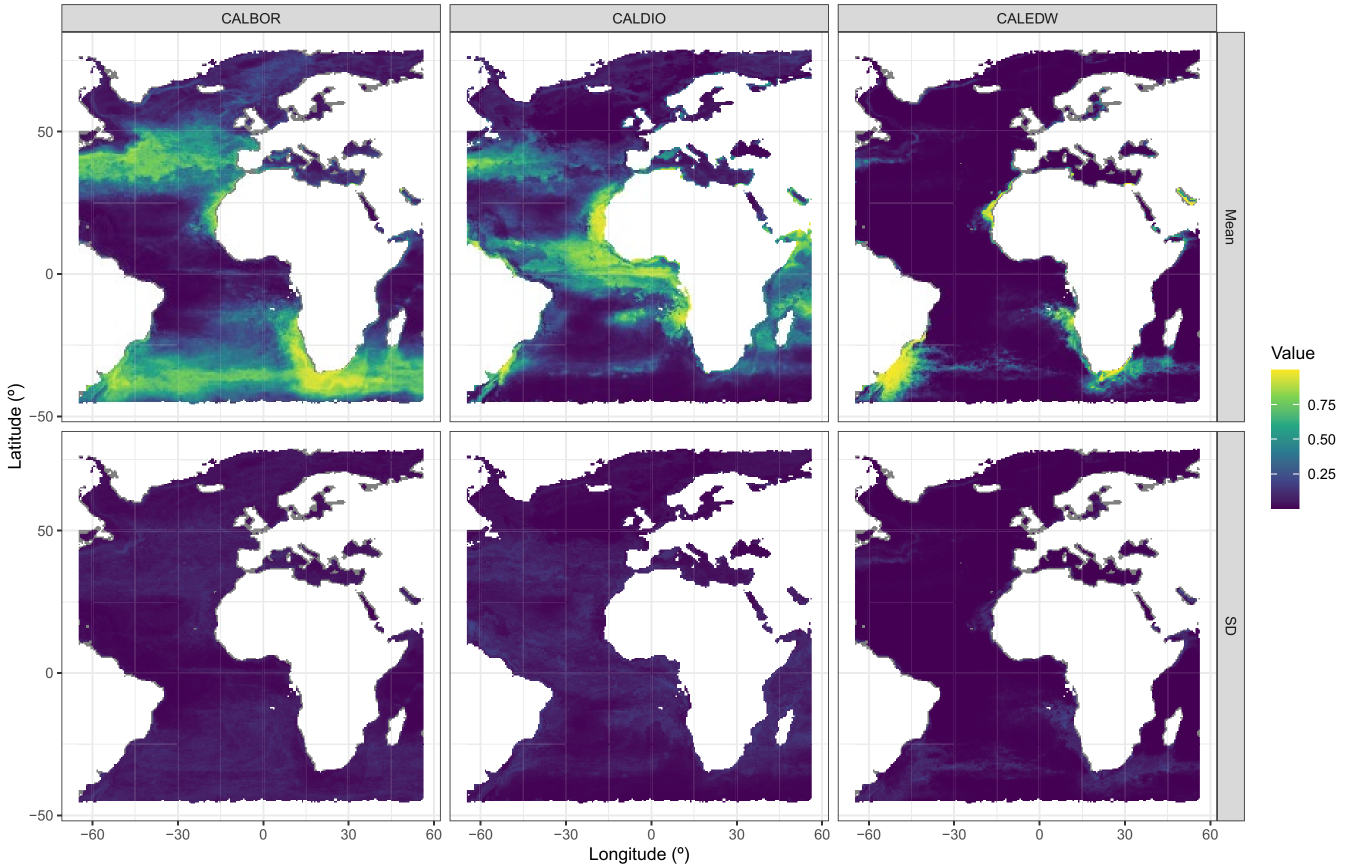

FIGURE 4 | Mean (top row) and standard deviation (bottom row) of the mean modelled distributions obtained from the BARTS model. Results are presented at a taxon level for CALBOR (left), CALDIO (centre) and CALEDW (right). Suitability values range from 0 (not suitable habitat) to 1 (most suitable habitat).

structure between the three taxa, with little mixing among them, but the structure dissipates as we move towards lower levels of aggregation. The presence of MC at a population level, but its absence at a colony level indicated that while there was a high degree of spread and mixing between colonies (Finch et al. 2017), birds from a specific population favoured a subset of non-breeding areas among all those used by the taxon. This pointed out some degree of population structuring within each taxon.

In addition, the modelled non-breeding distributions of the population-level models have shown there are some populations of CALBOR that use the non-breeding areas in the northern African coast without sharing them with individuals from other colonies, namely birds from Savage, eastern and western Canary Islands (Figure 3 and Figure S4), and this structure is reflected in the fact that CALBOR has the lowest AUC of the three taxa. The body size of birds from Savage is larger than those breeding in Berlengas and the Azores (Granadeiro 1993), the few studies of the genetic structure within CALBOR populations showed conflicting results (Genovart et al. 2013; Munilla et al. 2016), making further research necessary to elucidate the patterns of population structure within the taxon.

For CALDIO, our results showed that non-breeding habitat from the eastern populations differed from that of central and western populations as the former used almost exclusively the northernmost non-breeding areas (Figure 4 and Figure S5). Some studies showed a body size gradient within CALDIO populations, with a substantial decrease in size with longitude (Massa and Valvo 1986; Gómez-Díaz et al. 2006). In addition, evidence of genetic structure among colonies indicated divergence between eastern and western Mediterranean populations (Gómez-Díaz et al. 2009). Both morphological and genetic results could be a consequence of population structuring that can be linked to the lack of mixing we found in the non-breeding areas between individuals of the eastern and western populations.

For CALEDW, although we only had data from one colony from each population, the use of non-breeding areas of the two sampled colonies was almost identical (Figure 3), which was reflected by a very high AUC in the model at a taxon level. Although no genetic evidence is available for CALEDW, since there are no phenotypic differences between colonies, neither in morphology, foraging areas during breeding (González-Solís pers. comm.), choice of non-breeding area, nor preferred environmental variables, it is reasonable to assume there is no population structuring in this locally restricted taxon endemic to Cabo Verde.

4.2 | When the Drift Is Subtle: The Case of Neighbouring and Mixed Colonies of Different Taxa

Despite the clear non-breeding allopatry between CALBOR and CALDIO, there were a few 'contact zones' between them:

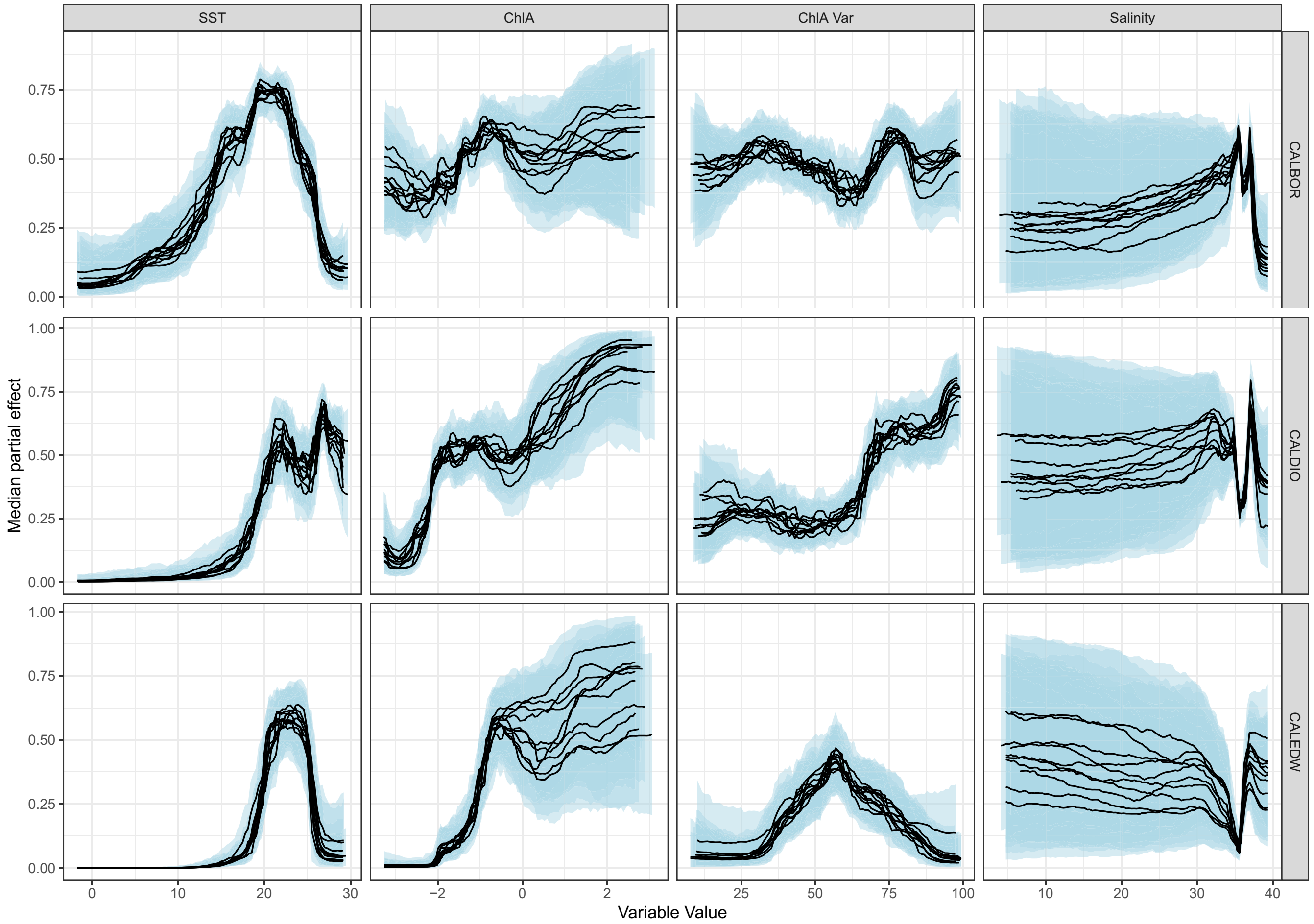

FIGURE 5 | Partial dependency plots displaying environmental preferences. These plots show the effect of the environmental covariates so higher values of the partial effect (in the y axis) indicate covariate values (in the *x* axis) that are more likely to be selected by the non-breeding birds according to the models ran. This is displayed at a taxon level for CALBOR (left), CALDIO (centre) and CALEDW (right). Only the partial effect of the three more relevant variables are shown (i.e. those with a higher importance value for all species). Each of the black lines represents the median of one iteration of the same model, and the shaded blue ribbon represents the 95% credible interval. Ribbons are half transparent so more shaded areas represent more agreement among iterations.

in two of the main non-breeding grounds, both taxa co-occurred in two transition areas, one between the Sahelian and Saharan upwellings and the other along the Namib coast. In addition, CALDIO also frequented subtropical regions of the Brazil current and the North Atlantic region during their migratory movements, thus sharing the areas with CALBOR. The existence of these 'contact zones' favours potential events of dispersion. For instance, a bird of one population wintering in a shared non-breeding ground might follow in the return migration a bird from another population, or even a bird from a different taxon (Friesen et al. 2007). Indeed, there are records of CALBOR individuals ringed as chicks in Atlantic colonies and found years later breeding in Mediterranean colonies (Lo Valvo and Massa 1988; Thibault and Bretagnolle 1998; Gómez-Díaz et al. 2009; Gabirot et al. 2016) including some cases of successful inter-taxa pairing (Martínez-Abraín et al. 2002). In contrast to such occasional dispersals of CALBOR individuals to CALDIO breeding sites, there are no reported dispersal events in the opposite direction, that is., CALDIO birds breeding in CALBOR sites. In addition, although CALEDW and CALBOR completely overlapped in their non-breeding areas, there are only a handful of documented cases of CALEDW individuals seen in CALBOR colonies, all of them in the Canary and Savage Islands (Fagundes et al. 2012; Copete et al. 2015; Gil-Velasco et al. 2019), but none of them with confirmed breeding. Finally, when we modelled the non-breeding distribution of the three taxa together, the model showed a worse fit than when modelled separately. This suggested the three taxa occupy different ecological niches with very little overlap among them, which might evidence lineage divergence in the three *Calonectris* taxa (Raxworthy et al. 2007; Ferrer Obiol et al. 2022).

The Almería—Orán Oceanographic Front serves as an ecological divide between Atlantic and Mediterranean distributions of many marine taxa (Pascual et al. 2017) and also appears to be the divide between CALBOR (breeding in the Atlantic) and CALDIO (breeding in the Mediterranean, Gómez-Díaz et al. 2009). However, there are nearby colonies of both taxa on both sides of the front, with at least one mixed colony (Gómez-Díaz et al. 2006; Zidat et al. 2017). This partial sympatry and breeding synchrony open the possibility of hybridization and introgression, which certainly would produce gene flow between the species (Brown et al. 2015). Indeed, reproductive isolation may not yet be complete: genetic studies in the mixed colony of Chafarinas Islands, based on both mitochondrial DNA (mtDNA, Gómez-Díaz et al. 2009) and microsatellites (Zidat et al. 2017), reported the existence of introgressed individuals which could indicate relatively recent hybridisation events.

Our study included tracks from three colonies in the Front area: Terreros, Palomas, and Chafarinas. Modelled non-breeding distributions for CALDIO from Palomas and CALBOR from Chafarinas showed common areas for both taxa (Figures S4 and S5). Unfortunately, the non-breeding distribution for CALBOR from Terreros and CALDIO from Chafarinas could not be modelled due to their small sample sizes (see Morera-Pujol et al. 2023). However, most of the non-breeding locations of CALBOR from Terreros fell within the Sahelian and Saharan upwellings, which are predominantly used by CALDIO populations (Figure 2). The non-breeding locations of CALDIO from Chafarinas occupied mainly tropical non-breeding areas of the eastern Atlantic coast, typical of their taxon, but also the Brazil coast and North Atlantic pelagic area, which are predominantly used by CALBOR populations. Despite the small sample size constraints of these two colonies, the results seemed to suggest that birds breeding in sympatry do experience mixing in the non-breeding areas (even though the two taxa have clearly segregated non-breeding ecological niches), which could allow locally restricted genetic mixture (as had already been reported in Gómez-Díaz et al. 2006 and Flood and Gutiérrez 2019), producing an intermediate CALBOR–CALDIO phenotype for individuals breeding in this area. Although our data do not allow conclusive remarks on this, it does suggest a way to test, with larger sampling sizes, whether the non-breeding distributions of birds in the intermediate breeding zone between CALBOR and CALDIO also have intermediate non-breeding preferences.

Altogether, our results suggested that the three *Calonectris* taxa adapted to the environmental conditions present in their non-breeding distributions. This adaptation might be to a particular range of SST, but most likely the preference for a certain range of SST and ChlA values reflects a preference for prey that favours those particular conditions. In any case, this local adaptation, coupled with the strong philopatry and breeding site fidelity, possibly resulted in a process of ecological divergence, despite the three taxa breeding sympatrically. Together with the divergence on other biological traits such as morphology, vocalisations, and odour (Zidat et al. 2017), this divergence adds evidence to the species differentiation of the three taxa.

4.3 | Implications for Population Dynamics, Conservation and Management

This structure at a population and taxon but not colony level has obvious implications in population dynamics: our results indicate the presence of mixing in the non-breeding areas within populations, and as populations were defined based on distance among colonies, this hinted to the possibility of a distance-based structure (i.e. neighbouring colonies had more similar migratory behaviour than distant colonies because there was more gene flow among them). The negative correlation between pairwise similarity of non-breeding distributions and pairwise colony distance provided further support for this hypothesis. It demonstrated that the closer two colonies are, the more similar the non-breeding habitats of their individuals are. This "isolation by distance" model had already been proposed as a driver of population structure within CALDIO (Ramos 2019), and also, as a driver of the differentiation between CALBOR and CALDIO (Gómez-Díaz et al. 2009).

In terms of management and conservation, understanding what constitutes a given population and how different populations relate to each other demographically and genetically is key to quantifying the effect that local (nest availability, predation at colony) or global (climate change, food availability) perturbations might have on the survival of their populations (Esler 2000; Matthiopoulos et al. 2005). In addition, understanding the degree of MC among populations can also help us to understand the effects of perturbations in the non-breeding distribution. In situations of high spread and mixing among populations (i.e. weak MC), any perturbation in a given non-breeding region is likely to cause a global effect on the entire species, while the effect would be more localised in a species with strong MC (Webster et al. 2002; Buckingham et al. 2022). Finally, the degree of intra-population spread in the non-breeding distribution can affect the ability of the population to respond to non-breeding range shifts and reductions in the context of climate change (Finch et al. 2017).

4.4 | Concluding Remarks

To our knowledge, this is the first time that non-breeding distributions have been used to identify the degree of ecological divergence in a species complex of migratory birds of this magnitude (but see Quillfeldt et al. 2017), extending previous attempts of applying ENM to support lineage divergence in resident species (Raxworthy et al. 2007; Rissler and Apodaca 2007). Although the genetic differences between the three taxa seem to be clear now (Ferrer Obiol et al. 2022), phenotype does not always immediately reflect the differences in genotype. Our results on major habitat segregation among the three taxa during the non-breeding period increase the list of evidences supporting the phenotypic divergence between CALBOR and CALDIO that have been documented so far: morphological (Gómez-Díaz et al. 2009), acoustic (Bretagnolle and Lequette 1990), chemical (i.e. odour, Zidat et al. 2017), and now, ecological (this study). In addition, our results preliminarily showed that although CALBOR and CALEDW partly overlap in their non-breeding distribution, their environmental preferences seem to differ, which confirms the ecological divergence between both taxa. Our data also seem to suggest the existence of intermediate habitat preferences in the non-breeding distributions of birds breeding in the contact zone between taxa, which is consistent with previous evidences of hybridization and introgression (Gómez-Díaz et al. 2009). This requires a future effort in data collection and further analyses to ascertain more details in the ecological traits of such populations. Finally, our analyses of MC and modelling of non-breeding habitats detected the presence of within-taxon structure at a population level, rather than at a colony level, thus suggesting that separate managements and conservation plans might be necessary for species with this type of internal structure.

Author Contributions

V.M.-P., J.G.-S. and R.R. conceived the ideas; V.M.-P., J.G.-S. and R.R. designed the methodology; V.M.-P., C.P., M.M., P.C., J.M.R.-G., J.P.G., T.M., M.P.D., D.O., J.M.I., G.D.O., M.Mü., V.H.P., B.M., V.N., J.N., G.K., S.X., J.G.C., J.M.Z.-M., M.G.F., I.A., R.O., M.S.R., F.d.F., Z.Z., M.C.-F., D.G., J.G.-S. and R.R. collected and/or provided data; V.M.-P., J.M.R.-G., T.M. and F.d.F. analysed data; V.M.-P. led the writing of the manuscript; J.G.-S. and R.R. contributed to the writing process; C.P., M.M., P.C., J.M.R.-G., J.P.G., T.M., M.P.D., D.O., J.M.I., G.D.O., M.Mü., V.H.P., B.M., V.N., J.N., G.K., S.X., J.G.C., J.M.Z.-M., M.G.F., I.A., R.O., M.S.R., F.d.F., Z.Z., M.C.-F. and D.G. reviewed and contributed to the final draft.

Acknowledgements

We would like to thank all students, volunteers and technicians who helped us at several stages of fieldwork. We are grateful to the following institutions for permission and logistical support in conducting fieldwork in the respective countries: Direção Nacional do Ambiente (Cabo Verde); Cabildo de Lanzarote, Cabildo de Gran Canaria, Timanfaya National Park, Gobierno de Canarias, Xunta de Galicia, Cabrera National Park, Govern Balear, Consejería de Agua, Agricultura y Medio Ambiente de la Región de Murcia, Consejería de Medio Ambiente y Ordenación Territorial de la Junta de Andalucía and Organismo Nacional de Parques Nacionales (Spain); and Direção Regional do Ambiente dos Açores, Instituto das Florestas e da Conservação da Natureza (IFCN—Madeira), Instituto da Conservação da Natureza e Florestas (ICNF), Reserva Natural das Berlengas and Parque Natural da Graciosa (Portugal), Parc National de Port-Cros, Conservatoire d'Espaces Naturels de Provence-Alpes-Côte d'Azur, Réserve Naturelle Nationale de l'archipel de Riou, Parc Maritime des îles du Frioul, Réserve Naturelle des Bouches de Bonifacio and Direction Départementale de la Protection des Populations (France); Regione Siciliana, Assessorato Agricoltura e Foreste, the Wild Bird Regulation Unit and the Environment and Resources Authority (Malta), Management Agency of the National Marine Park of Zakynthos, Ministry of Environment (Greece).

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The novel code developed for this manuscript and the tracking data necessary to perform the analyses can be downloaded via Dryad at https://datadryad.org/stash/share/4hv2ejVbZcgyRvgBUj9Ah_34pIg2m_D0lxjRXxLZnao. The data on colony locations and breeding populations is available in Appendix S1 of Morera-Pujol et al. (2023). Methods to detect spatial biases in tracking studies caused by differential representativeness of individuals, populations and time. Diversity and Distributions, 29, 19–38. https://doi.org/10.1111/ddi.13642.

Peer Review

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/ddi.70013.

References

Ambrosini, R., A. P. Møller, and N. Saino. 2009. "A Quantitative Measure of Migratory Connectivity." *Journal of Theoretical Biology* 257: 203–211.

Baddeley, A., E. Rubak, and R. Turner. 2015. *Spatial Point Patterns: Methodology and Applications With R*. Chapman and Hall/CRC.

Bauer, S., S. Lisovski, and S. Hahn. 2016. "Timing Is Crucial for Consequences of Migratory Connectivity." *Oikos* 125: 605–612.

Beal, M., S. Oppel, J. Handley, et al. 2021. "track2KBA: An R Package for Identifying Important Sites for Biodiversity From Tracking Data." *Methods in Ecology and Evolution* 12: 2372–2378.

Birdlife International. 2019. "Birdlife International Seabird Tracking Database | Tracking Ocean Wanderers."

Bretagnolle, V., and B. Lequette. 1990. "Structural Variation in the Call of the Cory's Shearwater (*Calonectris diomedea*, Aves, *Procellariidae*)." *Ethology* 85: 313–323.

Brown, R. M., N. M. S. M. Techow, A. G. Wood, and R. A. Phillips. 2015. "Hybridization and Back-Crossing in Giant Petrels (*Macronectes giganteus* and *M. halli*) at Bird Island, South Georgia, and a Summary of Hybridization in Seabirds." *PLoS One* 10: e0121688.

Buckingham, L., M. I. Bogdanova, J. A. Green, et al. 2022. "Interspecific Variation in Non-breeding Aggregation: A Multi-Colony Tracking Study of Two Sympatric Seabirds." *Marine Ecology Progress Series* 684: 181–197.

Carlson, C. J. 2020. "Embarcadero: Species Distribution Modelling With Bayesian Additive Regression Trees in R." *Methods in Ecology and Evolution* 11: 850–858.

Chipman, H. A., E. I. George, and R. E. McCulloch. 2010. "BART: Bayesian Additive Regression Trees." *Annals of Applied Statistics* 4: 266–298.

Cohen, E. B., J. A. Hostetler, M. T. Hallworth, C. S. Rushing, T. S. Sillett, and P. P. Marra. 2017. "Quantifying the Strength of Migratory Connectivity." *Methods in Ecology and Evolution* 9, no. 3: 513–524.

Copete, J. L., J. A. Lorenzo, E. Amengual, et al. 2015. "Observaciones de Aves Raras en España, 2012 y 2013." *Ardeola* 62: 453–508.

Dias, M. P., J. P. Granadeiro, R. A. Phillips, H. Alonso, and P. Catry. 2011. "Breaking the Routine: Individual Cory's Shearwaters Shift Winter Destinations Between Hemispheres and Across Ocean Basins." *Proceedings of the Royal Society B* 278: 1786–1793.

Ekstrom, P. A. 2004. "An Advance in Geolocation by Light." *Memoirs of National Institute of Polar Research* 58: 210–226.

Esler, D. 2000. "Applying Metapopulation Theory to Conservation of Migratory Birds." *Conservation Biology* 14: 366–372.

Fagundes, I., C. Pérez, and F. Moniz. 2012. "Cape Verde Shearwater in Selvagens in April 2012." *Dutch Birding* 34: 310–311.

Ferrer Obiol, J., J. M. Herranz, J. R. Paris, et al. 2023. "Species Delimitation Using Genomic Data to Resolve Taxonomic Uncertainties in a Speciation Continuum of Pelagic Seabirds." *Molecular Phylogenetics and Evolution* 179: 107671.

Ferrer Obiol, J., H. F. James, R. T. Chesser, et al. 2022. "Palaeoceanographic Changes in the Late Pliocene Promoted Rapid Diversification in Pelagic Seabirds." *Journal of Biogeography* 49: 171–188.

Finch, T., S. J. Butler, A. M. A. Franco, and W. Cresswell. 2017. "Low Migratory Connectivity Is Common in Long-Distance Migrant Birds." *Journal of Animal Ecology* 86: 662–673.

Fišer, C., C. T. Robinson, and F. Malard. 2018. "Cryptic Species as a Window Into the Paradigm Shift of the Species Concept." *Molecular Ecology* 27: 613–635.

Flood, R. L., and R. Gutiérrez. 2019. "Status of Cory's Shearwater in the Western Mediterranean." *Dutch Birding* 41: 159–165.

Friesen, V. L., T. M. Burg, and K. D. McCoy. 2007. "Mechanisms of Population Differentiation in Seabirds." *Molecular Ecology* 16: 1765–1785.

Gabirot, M., L. Raux, G. Dell'Ariccia, et al. 2016. "Chemical Labels Differ Between Two Closely Related Shearwater Taxa." *Journal of Avian Biology* 47: 540–551.

Genovart, M., J.-C. Thibault, J. M. Igual, M. D. M. Bauzà-Ribot, C. Rabouam, and V. Bretagnolle. 2013. "Population Structure and Dispersal Patterns Within and Between Atlantic and Mediterranean Populations of a Large-Range Pelagic Seabird." *PLoS One* 8: e70711.

Gil-Velasco, M., M. Rouco, M. García-Tarrasón, et al. 2019. "Observaciones de aves raras en España, 2017." *Ardeola* 66, no. 1: 69–204.

Gómez-Díaz, E., J. González-Solís, and M. A. Peinado. 2009. "Population Structure in a Highly Pelagic Seabird, the Cory's Shearwater *Calonectris diomedea*: An Examination of Genetics, Morphology and Ecology." *Marine Ecology Progress Series* 382: 197–209.

Gómez-Díaz, E., J. González-Solís, M. A. Peinado, and R. D. M. Page. 2006. "Phylogeography of *Calonectris* Shearwaters Using Molecular and Morphometric Data." *Molecular Phylogenetics and Evolution* 41: 322–332.

González-Solís, J., J. P. Croxall, D. Oro, and X. Ruiz. 2007. "Trans-Equatorial Migration and Mixing in the Wintering Areas of a Pelagic Seabird." *Frontiers in Ecology and the Environment* 5: 297–301.

Granadeiro, J. P. 1993. "Variation in Measurements of Cory's Shearwater Between Populations and Sexing by Discriminant Analysis." *Ringing & Migration* 14, no. 2: 103–112.

Granadeiro, J. P., M. D. Burns, and R. W. Furness. 1999. "Food Provisioning to Nestling Shearwaters: Why Parental Behaviour Should Be Monitored?" *Animal Behaviour* 57: 662–671.

Granadeiro, J. P., M. P. Dias, R. Rebelo, C. D. Santos, and P. Catry. 2006. "Numbers and Population Trends of Cory's Shearwater *Calonectris diomedea* at Selvagem Grande, Northeast Atlantic." *Waterbirds* 29: 56–60.

Gu, Z., L. Gu, R. Eils, M. Schlesner, and B. Brors. 2014. "Circlize Implements and Enhances Circular Visualization in R." *Bioinformatics* 30: 2811–2812.

Hanski, I., and M. E. Gilpin. 1997. *Metapopulation Biology: Ecology, Genetics, and Evolution*. Academic Press.

Hazevoet, C. J. 1995. *The Birds of the Cape Verde Islands. An Annotated Check List*. British-Ornitologist Union.

Hijmans, R. J., S. Phillips, J. Leathwick, and J. Elith. 2022. dismo: Species Distribution Modeling. R package version 1.3–9.

Hipfner, J. M., M. M. Prill, K. R. Studholme, et al. 2020. "Geolocator Tagging Links Distributions in the Non-breeding Season to Population Genetic Structure in a Sentinel North Pacific Seabird." *PLoS One* 15, no. 11: e0240056.

Hostetler, J. A., and M. T. Hallworth. 2018. "MigConnectivity: Estimate Strength of Migratory Connectivity for Migratory Animals. R package."

Lisovski, S., and S. Hahn. 2012. "GeoLight—Processing and Analysing Light-Based Geolocator Data in R." *Methods in Ecology and Evolution* 3: 1055–1059.

Lo Valvo, M., and B. Massa. 1988. "Considerations on a Specimen of Cory's Shearwater Ringed at Selvagem Grande and Recovered in the Central Mediterranean." *Bocagiana* 25, no. 124: 1–5.

Louzao, M., D. Pinaud, C. Péron, K. Delord, T. Wiegand, and H. Weimerskirch. 2011. "Conserving Pelagic Habitats: Seascape Modelling of an Oceanic Top Predator." *Journal of Applied Ecology* 48: 121–132.

Marra, P. P., E. B. Cohen, A.-L. Harrison, C. E. Studds, and M. S. Webster. 2019. *Migratory Connectivity*, edited by K. R. Crooks and M. A. Sanjayan, 643–654. Elsevier.

Martínez-Abraín, A., A. Sánchez, and D. Oro. 2002. "Atlantic Cory's Shearwaters Breeding in a Colony of Mediterranean Cory's Shearwaters." *Waterbirds* 25: 221–224.

Massa, B., and M. L. Valvo. 1986. "Biometrical and Biological Considerations on the Cory's Shearwater *Calonectris diomedea*." In *Mediterranean Marine Avifauna: Population Studies and Conservation*, 293–313. Springer Berlin Heidelberg.

Matthiopoulos, J., J. L. Harwood, and L. Thomas. 2005. "Metapopulation Consequences of Site Fidelity for Colonially Breeding Mammals and Birds." *Journal of Animal Ecology* 74: 716–727.

McConnell, B. J., C. Chambers, and M. A. Fedak. 1992. "Foraging Ecology of Southern Elephant Seals in Relation to the Bathymetry and Productivity of the Southern Ocean." *Antarctic Science* 4: 393–398.

Melo-Merino, S. M., H. Reyes-Bonilla, and A. Lira-Noriega. 2020. "Ecological Niche Models and Species Distribution Models in Marine Environments: A Literature Review and Spatial Analysis of Evidence." *Ecological Modelling* 415: 108837.

Morera-Pujol, V., P. Catry, M. Magalhães, et al. 2023. "Methods to Detect Spatial Biases in Tracking Studies Caused by Differential Representativeness of Individuals, Populations and Time." *Diversity and Distributions* 29: 19–38.

Mougin, J. L., J. P. Granadeiro, C. Jouanin, and F. Roux. 1999. "Philopatry and Faithfulness to Nest Site in Cory's Shearwaters *Calonectris diomedea* at Selvagem Grande." *Ostrich* 70: 229–232.

Mougin, J.-L., C. Jouanin, F. Roux, and F. Zino. 2000. "Fledging Weight and Juvenile Survival of Cory's Shearwaters *Calonectris diomedea* on Selvagem Grande." *Ringing & Migration* 20: 107–110.

Munilla, I., M. Genovart, V. H. Paiva, and A. Velando. 2016. "Colony Foundation in an Oceanic Seabird." *PLoS One* 11, no. 2: e0147222.

Navarro, J., M. G. Forero, J. González-Solís, J. M. Igual, J. Bécares, and K. A. Hobson. 2009. "Foraging Segregation Between Two Closely Related Shearwaters Breeding in Sympatry." *Biology Letters* 5, no. 4: 545–548.

Navarro, J., and J. González-Solís. 2007. "Experimental Increase of Flying Costs in a Pelagic Seabird: Effects on Foraging Strategies, Nutritional State and Chick Condition." *Oecologia* 151: 150–160.

Paiva, V. H., P. Geraldes, I. Rodrigues, T. Melo, J. Melo, and J. A. Ramos. 2015. "The Foraging Ecology of the Endangered Cape Verde Shearwater, a Sentinel Species for Marine Conservation Off West Africa." *PLoS One* 10: e0139390.

Pascual, M., B. Rives, C. Schunter, and E. MacPherson. 2017. "Impact of Life History Traits on Gene Flow: A Multispecies Systematic Review Across Oceanographic Barriers in the Mediterranean Sea." *PLoS One* 12: 1–20.

Quillfeldt, P., Y. Moodley, H. Weimerskirch, et al. 2017. "Does Genetic Structure Reflect Differences in Non-breeding Movements? A Case Study in Small, Highly Mobile Seabirds." *BMC Evolutionary Biology* 17: 160.

Ramos, R. 2019. "Crossing the Pillars of Hercules: Understanding Transoceanic Migrations of Seabirds Throughout Their Breeding Range." *Ecology and Evolution* 9, no. 8: 4760–4771.

Ramos, R., J. P. Granadeiro, M. Nevoux, J.-L. Mougin, M. P. Dias, and P. Catry. 2012. "Combined Spatio-Temporal Impacts of Climate and Longline Fisheries on the Survival of a Trans-Equatorial Marine Migrant." *PLoS One* 7: 1–12.

Raxworthy, C. J., C. M. Ingram, N. Rabibisoa, and R. G. Pearson. 2007. "Applications of Ecological Niche Modeling for Species Delimitation: A Review and Empirical Evaluation Using Day Geckos (*Phelsuma*) From Madagascar." *Systematic Biology* 56: 907–923.

Rissler, L. J., and J. J. Apodaca. 2007. "Adding More Ecology Into Species Delimitation: Ecological Niche Models and Phylogeography Help Define Cryptic Species in the Black Salamander (*Aneides flavipunctatus*)." *Systematic Biology* 56: 924–942.

Sangster, G., J. M. Collinson, P.-A. Crochet, A. G. Knox, D. T. Parkin, and S. C. Votier. 2012. "Taxonomic Recommendations for British Birds: Eighth Report." *Ibis* 154: 874–883.

Spalding, M. D., H. E. Fox, G. R. Allen, et al. 2007. "Marine Ecoregions of the World: A Bioregionalization of Coastal and Shelf Areas." *Bioscience* 57: 573–583.

Thibault, J.-C., and V. Bretagnolle. 1998. “A Mediterranean Breeding Colony of Cory's Shearwater *Calonectris diomedea* in Which Individuals Show Behavioural and Biometric Characters of the Atlantic Subspecies.” *Ibis* 140: 523–528.

Thibault, J.-C., V. Bretagnolle, and C. Rabouam. 1997. *Cory's Shearwater*, 75–98. Oxford University Press.

Tigano, A., M. Damus, T. P. Birt, J. A. Morris-Pocock, Y. B. Artukhin, and V. L. Friesen. 2015. “The Arctic: Glacial Refugium or Area of Secondary Contact? Inference From the Population Genetic Structure of the Thick-Billed Murre (*Uria lomvia*), with Implications for Management.” *Journal of Heredity* 106: 238–246.

Wakefield, E. D., R. A. Phillips, and J. Matthiopoulos. 2009. “Quantifying Habitat Use and Preferences of Pelagic Seabirds Using Individual Movement Data: A Review.” *Marine Ecology Progress Series* 391: 165–182.

Warham, J. 1996. *The Behaviour, Population Biology and Physiology of the Petrels*. Academic Press Ltd.

Webster, M. S., P. P. Marra, S. M. Haig, S. Bensch, and R. T. Holmes. 2002. “Links Between Worlds: Unravelling Migratory Connectivity.” *Trends in Ecology & Evolution* 17: 76–83.

Wilson, R. P., J. J. Ducamp, W. G. Rees, B. M. Culik, and K. Niekamp. 1992. “Estimation of Location: Global Coverage Using Light Intensity.” In *Wildlife Telemetry: Remote Monitoring and Tracking of Animals*, 131–134. Ellis Horwood.

Zidat, T., G. Dell'Ariccia, M. Gabirot, et al. 2017. “Reproductive Isolation Maintains Distinct Genotypes, Phenotypes and Chemical Signatures in Mixed Colonies of the Two European *Calonectris* Shearwaters (*Procellariiformes*: *Procellariidae*).” *Zoological Journal of the Linnean Society* 181: 711–726.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.