MS East Australia seabird RCPs

Nicholas W Daudt

2023-03-31

# Intro

# Methods

## Study area

The EAC is the dominant current in East Australia, standing as the major western boundary current in the South Pacific (Ganachaud et al., 2014). It originates from the South Equatorial Current (SEC), a tropical, nutrient-poor water mass. When the SEC’s North Caledonian Jet hits the Queensland Plateau (~18°S), it splits into a north branch running towards the Gulf of Papua, and a south branch where it marks the start of the EAC (Ganachaud et al., 2014; Kessler & Cravatte, 2013). The South Caledonian Jet completes the EAC system at ~22°S after encountering the shelf break in the southern portion of the Great Barrier Reef (**Fig 1 (a)**; Kessler & Cravatte (2013); Ganachaud et al. (2014)). The EAC interacts with topographical features on the seafloor, from which oceanographic features and processes arise, such as smaller currents, upwellings, and persistent dynamic eddies (Oke, Roughan, et al., 2019; Oke & Middleton, 2000; Ridgway & Dunn, 2003). The EAC is mainly attached to the shelf break, but at ~32°S it spawns an area of eddy generation that dissipates throughout the western Tasman Sea (Mata et al., 2006; Oke, Roughan, et al., 2019), the ‘EAC eastern extension’ (Oke, Pilo, et al. (2019); typically mentioned as the ‘Tasman Front’). Although most of the EAC energy is converted into eddies, the EAC still runs southwards following the shelf break towards Tasmania (the ‘EAC southern extension’; **Fig 1 (a)**), possibly carrying eddy-like features as a transient current (Oke, Roughan, et al., 2019).

## Seabird and environmental data sources

Seabird data were gathered while opportunistically aboard RV Investigator and RV Falkor during 15 voyages, between 2016–2021, over ~37° of latitude (**Fig 1 (b)**). One observer recorded seabirds within a 300-m forward quadrant (eye height **18-m?** a.s.l), continuously, throughout daylight hours while ship was under way in constant speed. Species, number of birds and behaviour were logged directly onto a data entry system through a ship-based web server. Individual records received a date/time/geographic coordinate label (multi-species records have the same label, but are logged as individual entries for each species).

To explain and predict species distribution and biogeography, a thoughtful choice of covariates is important (Barbet-Massin & Jetz, 2014). We used seven oceanographic variables, calculated three climatic layers from them, as well as three physiographical variables (**Table 1**). These variables are within the most used in seabird-environment relationship studies (Tremblay et al., 2009), capturing from large-scale water-mass proprieties (e.g. sea surface temperature) to increased predictability of prey availability in mesoscale oceanographic features (e.g. fronts and eddies; Weimerskirch (2007); Scales, Miller, Hawkes, et al. (2014)). Seabird records were matched with best open-access, remote-sensed or interpolated data available, and distance from coast was calculated in R software (**Table 1**; see details below).

## Data treatment

Data wrangling, modelling and visualization were done in R 4.2.0 (**R citation**), using packages **X, Y, …, Z**. Code and (gridded) data are available in **Daudt et al. 2023 (Zenodo repo)**.

Some environmental layers were calculated prior to data extraction. Eddy kinetic energy (EKE) was calculated as , where and are the surface zonal and meridian geostrophic currents, respectively, calculated over a rolling mean of 90 days (**ROB: Ref here?**). Sea surface temperature gradient (SSTgrad) was calculated as the spatial gradient from north-to-south (NS) and west-to-east (WE) neighbourhood average maximum, then the magnitude of such gradient is defined as (e.g. Burrows et al., 2011). Both EKE and SSTgrad calculations were done layer-by-layer, to match seabird data on the finer scale possible. Spatial and temporal persistence of features such as eddies (EKE) and fronts (SSTgrad) may influence seabird movement and distribution by offering higher likelihood to finding food (Scales, Miller, Embling, et al., 2014; Weimerskirch, 2007). Therefore, for EKE and SSTgrad we calculated climatic, seasonal average layers based on 10-year data (2011–2021), as well as EKE standard deviation (EKEsd) over the same period. Slope was calculated using raster::terrain() and distance from coast was calculated through sf::st\_distance() using Australia and Tasmania as mainland (ignoring smaller islands). The most proximate value from each environmental layers was then extracted for each seabird record label using build-in functions from raster, terra, and sf packages.

Seabirds are know to be highly mobile animals and many species migrate during non-breeding periods (Schreiber & Burger, 2002). Therefore, seasonal differences in species composition were expected, as well as their probability of occurrence and total numbers. Given that, data were aggregated data by 1° latitude 1° longitude grid cells (c. 110 110 km) and season (**Fig 1 (b)**). Birds following the boat were previously excluded from analyses. Seabirds from the same species were summed and environmental data were averaged. Considering our aim to test if presence/absence (P/A) and abundance data would result in similar assemblage patterns, we build two species matrices, where the P/A matrix transformed values >1 to 1. In addition, to avoid spurious correlations with assemblage/environment and numerical instabilities in the models prompted from unusual or rare species, we kept only species with at leas six occurrences per season (summer = 12; autumn = 21; winter = 20; spring = 33). Environmental data were scaled and centred prior to modelling.

For predicting model results, we built a seasonal 10-year (2011–2021) average layer for each environmental predictor. These variables were averaged within each grid cell, by season. Worth noticing that climatic EKEmean, EKEsd and SSTgrad were processed beforehand as they may provide persistent oceanographic features (as described above). Data for prediction were scaled and centred using the same coefficients from the modelling dataset.

## Statistical modelling

We used the mixture model, Region of Common Profiles (RCP), to identify and describe the seabird assemblages. RCP is a one-stage bioregionalization approach, as it identifies bioregions by specifying a statistical model that simultaneously relate biotic (multi-species) and abiotic (environmental) data (Hill et al., 2020; Woolley et al., 2020). As such, groups derived from RCP models can be directly interpreted as assemblages, i.e. ‘regions’ of similar environmental spaces in which the probability of observing a set of species (‘profiles’) is constant but different from other regions (Foster et al., 2013; Foster et al., 2017). Group memberships are defined as latent (unobserved) and vary as a function of environmental data, assuming sites within each region (group) have constant species profiles. The model also enable predicting assemblage probabilities onto sites where no biotic data exist, given a set of environmental variables. Additionally, as a feature of model-based approaches, RCPs allow the opportunity to compare models through information criteria, present uncertainty around the results, and check for model assumptions through residuals checking (Hill et al., 2020; Warton et al., 2015).

Prior to fitting models, covariates were checked for correlation and collinearity within each season using the modelling dataset. We removed variables (Dormann et al., 2013), favouring *direct measures* (**not sure the best word/term here?**) over the ‘climatic’ layers we have calculated (e.g. if SSTgrad and ‘climatic’ SSTgrad were correlated, we chose SSTgrad). As such, ‘full models’ were different between seasons, although most variables were common.

RCP models were implemented separately for each season and data type (P/A, count), using {ecomix} R package. All models follow the same steps: (i) finding the best number of groups that describes the data, (ii) re-fit the model by fixing the best number of groups and going through covariate selection, (iii) check residuals diagnostics, (iv) bootstrapping results to get uncertainty around parameter estimates, and (v) producing results. For steps (i) and (ii), we used Bayesian Information Criterion (BIC) as the metric for model selection, and diagnostics (iii) were assessed visually.

To identify the best number of groups that describe the data, we ran full-models, with 100 random start numbers (see Foster et al., 2013), from 2–7 groups, and chosen the number of RCPs that minimized BIC. Then, we fixed the number of groups and ran models adding all selected covariates for each season linearly, to assess the best set of environmental predictors (the one with the lowest BIC). We did not consider interactions in our models. (**should we present the ‘full models?’ If yes, do you reckon here is the right place?**). To get parameters uncertainties (confidence intervals), we used default ecomix::regional\_mix.bootstrap() settings, which applies Bayesian bootstrap 1000 times.

# Results

# Discussion

# References

Barbet-Massin, M., & Jetz, W. (2014). A 40-year, continent-wide, multispecies assessment of relevant climate predictors for species distribution modelling. *Diversity and Distributions*, *20*(11), 1285–1295. <https://doi.org/10.1111/ddi.12229>

Burrows, M. T., Schoeman, D. S., Buckley, L. B., Moore, P., Poloczanska, E. S., Brander, K. M., Brown, C., Bruno, J. F., Duarte, C. M., Halpern, B. S., Holding, J., Kappel, C. V., Kiessling, W., O’Connor, M. I., Pandolfi, J. M., Parmesan, C., Schwing, F. B., Sydeman, W. J., & Richardson, A. J. (2011). The Pace of Shifting Climate in Marine and Terrestrial Ecosystems. *Science*, *334*(6056), 652–655. <https://doi.org/10.1126/science.1210288>

Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J. R. G., Gruber, B., Lafourcade, B., Leitão, P. J., Münkemüller, T., McClean, C., Osborne, P. E., Reineking, B., Schröder, B., Skidmore, A. K., Zurell, D., & Lautenbach, S. (2013). Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, *36*(1), 27–46. <https://doi.org/10.1111/j.1600-0587.2012.07348.x>

Foster, S. D., Givens, G. H., Dornan, G. J., Dunstan, P. K., & Darnell, R. (2013). Modelling biological regions from multi-species and environmental data: MODELLING REGIONS OF COMMON PROFILES. *Environmetrics*, *24*(7), 489–499. <https://doi.org/10.1002/env.2245>

Foster, S. D., Hill, N. A., & Lyons, M. (2017). Ecological grouping of survey sites when sampling artefacts are present. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *66*(5), 1031–1047. <https://doi.org/10.1111/rssc.12211>

Ganachaud, A., Cravatte, S., Melet, A., Schiller, A., Holbrook, N. J., Sloyan, B. M., Widlansky, M. J., Bowen, M., Verron, J., Wiles, P., Ridgway, K., Sutton, P., Sprintall, J., Steinberg, C., Brassington, G., Cai, W., Davis, R., Gasparin, F., Gourdeau, L., … Send, U. (2014). The Southwest Pacific Ocean circulation and climate experiment (SPICE). *Journal of Geophysical Research: Oceans*, *119*(11), 7660–7686. <https://doi.org/10.1002/2013JC009678>

Hill, N., Woolley, S. N. C., Foster, S., Dunstan, P. K., McKinlay, J., Ovaskainen, O., & Johnson, C. (2020). Determining marine bioregions: A comparison of quantitative approaches. *Methods in Ecology and Evolution*, *11*(10), 1258–1272. <https://doi.org/10.1111/2041-210X.13447>

Kessler, W. S., & Cravatte, S. (2013). Mean circulation of the Coral Sea. *Journal of Geophysical Research: Oceans*, *118*(12), 6385–6410. <https://doi.org/10.1002/2013JC009117>

Mata, M. M., Wijffels, S. E., Church, J. A., & Tomczak, M. (2006). Eddy shedding and energy conversions in the East Australian Current. *Journal of Geophysical Research: Oceans*, *111*(C9). <https://doi.org/10.1029/2006JC003592>

Oke, P. R., & Middleton, J. H. (2000). Topographically Induced Upwelling off Eastern Australia. *Journal of Physical Oceanography*, *30*(3), 512–531. <https://doi.org/10.1175/1520-0485(2000)030<0512:TIUOEA>2.0.CO;2>

Oke, P. R., Pilo, G. S., Ridgway, K., Kiss, A., & Rykova, T. (2019). A search for the Tasman Front. *Journal of Marine Systems*, *199*, 103217. <https://doi.org/10.1016/j.jmarsys.2019.103217>

Oke, P. R., Roughan, M., Cetina-Heredia, P., Pilo, G. S., Ridgway, K. R., Rykova, T., Archer, M. R., Coleman, R. C., Kerry, C. G., Rocha, C., Schaeffer, A., & Vitarelli, E. (2019). Revisiting the circulation of the East Australian Current: Its path, separation, and eddy field. *Progress in Oceanography*, *176*, 102139. <https://doi.org/10.1016/j.pocean.2019.102139>

Ridgway, K. R., & Dunn, J. R. (2003). Mesoscale structure of the mean East Australian Current System and its relationship with topography. *Progress in Oceanography*, *56*(2), 189–222. <https://doi.org/10.1016/S0079-6611(03)00004-1>

Scales, K. L., Miller, P. I., Embling, C. B., Ingram, S. N., Pirotta, E., & Votier, S. C. (2014). Mesoscale fronts as foraging habitats: composite front mapping reveals oceanographic drivers of habitat use for a pelagic seabird. *Journal of The Royal Society Interface*, *11*(100), 20140679. <https://doi.org/10.1098/rsif.2014.0679>

Scales, K. L., Miller, P. I., Hawkes, L. A., Ingram, S. N., Sims, D. W., & Votier, S. C. (2014). On the Front Line: frontal zones as priority at-sea conservation areas for mobile marine vertebrates. *Journal of Applied Ecology*, *51*(6), 1575–1583. <https://doi.org/10.1111/1365-2664.12330>

Schreiber, E. A., & Burger, J. (Eds.). (2002). *Biology of marine birds*. CRC Press.

Tremblay, Y., Bertrand, S., Henry, R., Kappes, M., Costa, D., & Shaffer, S. (2009). Analytical approaches to investigating seabirdenvironment interactions: a review. *Marine Ecology Progress Series*, *391*, 153–163. <https://doi.org/10.3354/meps08146>

Warton, D. I., Foster, S. D., De’ath, G., Stoklosa, J., & Dunstan, P. K. (2015). Model-based thinking for community ecology. *Plant Ecology*, *216*(5), 669–682. <https://doi.org/10.1007/s11258-014-0366-3>

Weimerskirch, H. (2007). Are seabirds foraging for unpredictable resources? *Deep Sea Research Part II: Topical Studies in Oceanography*, *54*(3-4), 211–223. <https://doi.org/10.1016/j.dsr2.2006.11.013>

Woolley, S. N. C., Foster, S. D., Bax, N. J., Currie, J. C., Dunn, D. C., Hansen, C., Hill, N., O’Hara, T. D., Ovaskainen, O., Sayre, R., Vanhatalo, J. P., & Dunstan, P. K. (2020). Bioregions in marine environments: Combining biological and environmental data for management and scientific understanding. *BioScience*, *70*(1), 48–59. <https://doi.org/10.1093/biosci/biz133>