Supplementary Material

Probabilistic bioregion models suggest seabird assemblages are linked to the major boundary current off eastern Australia

2023-10-17

This document is an accompanying file for **XXXXX et al. (2023) Divers. Distr.**. Briefly, here you will find –

- **List of packages** and version used for data wrangling, visualisation and analyses;
- **Table S1**: Summary of sampling effort by voyage;
- **Figure S1**: Number of occurrence, frequency of occurrence (FO), and numeric frequency (NF) by season;
- **Figure S2**: Species richness and total number of seabirds counted by grid/season;
- **Figure S3**: Plots for choosing best RCP group number (`multifit`);
- **Figure S4**: Residual plots from the best fitted models;
- **Figure S5**: Partial plots for covariates;
- **Figure S6**: Probability maps from seasonal predictions;
- **Figure S7**: Species profiles for each season;
- **Figure S8**: Species-richness and sample-coverage curves.

List of packages

We used the following packages 'plyr' 1.8.8 (Wickham, 2011), 'dplyr' 1.1.2 (Wickham, François, et al., 2023), 'tidyr' 1.3.0 (Wickham, Vaughan, et al., 2023), 'readr' 2.1.4 (Wickham, Hester, et al., 2023), 'tibble' 3.2.1 (Müller & Wickham, 2023), 'lubridate' 1.9.2 (Grolemund & Wickham, 2011), 'stringr' 1.5.0 (Wickham, 2022), 'purrr' 1.0.1 (Wickham & Henry, 2023), 'ggplot2' 3.4.2 (Wickham, 2016), 'ggspatial' 1.1.7 (Dunnington, 2022), 'patchwork' 1.1.2 (Pedersen, 2022), 'RColorBrewer' 1.1-3 (Neuwirth, 2022), 'rnaturalearth' 0.3.2 (Massicotte & South, 2023), 'sp' 1.6-0 (Bivand et al., 2013; E. J. Pebesma & Bivand, 2005), 'sf' 1.0-8 (E. Pebesma, 2018), 'mapview' 2.11.0 (Appelhans et al., 2022), 'raster' 3.5-21 (Hijmans, 2022a), 'terra' 1.6-7 (Hijmans, 2022b), 'rerddap' 1.0.2 (Chamberlain, 2023), 'rerddapXtracto' 1.1.4 (Mendelssohn, 2022), 'hadsstR' (Byrnes & Dunic, 2017), 'corrplot' 0.92 (Wei & Simko, 2021), and the ones referenced in the main text.

The code is archived in a repository **(Zenodo repo)**, and you can find a detailed walk-through in it.

TABLE S1. Summary of seabird sampling effort by voyage, off eastern Australia, during Australasian Seabird Group's ship-based surveys between 2016–2021. Start and finish dates and geographic ranges of each voyage, including the number of seabird records and the number of individuals and species recorded

| Voyage	Date start	Date end	Latitudinal range	Longitudinal range	No. of records	No. of birds	No. of species
in2016_t02	2016-08-25	2016-08-28	-43 − -34	147 − 152	344	475	25
in2016_v06	2016-10-29	2016-11-12	-27 − -27	153 − 155	284	2892	14
in2017_v02	2017-03-16	2017-03-27	-47 − -43	142 − 147	911	7122	30
in2017_t01	2017-09-24	2017-10-01	-33 − -9	143 − 154	370	4540	17
in2017_t02	2017-11-24	2017-11-25	-44 − -42	141 − 147	113	11010	17
in2018_t02	2018-05-14	2018-05-20	-43 − -27	148 − 154	214	6168	31
in2018_c01	2018-05-28	2018-06-07	-41 − -39	146 − 149	644	2846	25
in2018_v04	2018-09-11	2018-10-07	-47 − -34	141 − 155	1136	10434	36
in2018_v06	2018-11-22	2018-12-18	-44 − -41	146 − 149	1957	59628	43
in2019_v07	2019-04-10	2019-04-22	-43 − -38	147 − 150	412	1472	26
in2019_t01	2019-04-29	2019-05-02	-44 − -39	141 − 148	140	177	18
in2019_v04	2019-08-07	2019-09-01	-24 − -10	146 − 160	1321	13383	26
in2019_t02	2019-10-03	2019-10-10	-27 − -9	142 − 154	245	1895	28
fk201228	2020-12-27	2021-01-25	-27 − -23	153 − 157	1063	17298	20
fk210206	2021-02-06	2021-03-05	-27 − -19	153 − 157	1107	3306	14
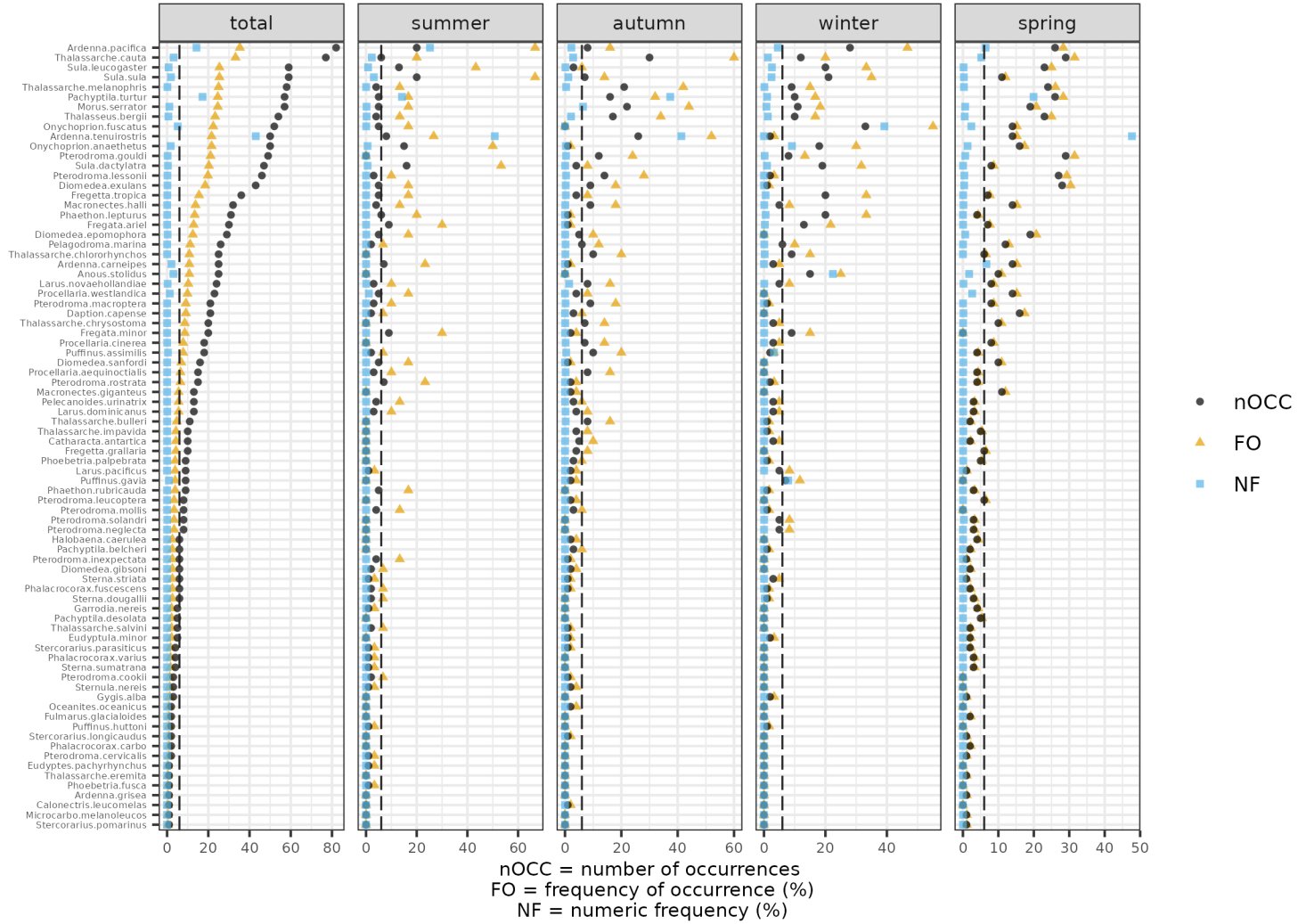
Figure 1: **Figure S1.** Number of occurrences (nOCC), frequency of occurrence (FO) and numeric frequency (NF) of seabirds recorded off eastern Australia during Australasian Seabird Group's ship-based surveys between 2016–2021. The dashed line represents the number of occurrence thresholds (n = 6) each taxon had to match for its inclusion in the seasonal models (see *Methods* in the main text). Species are ordered from the largest to the lowest total number of occurrences.
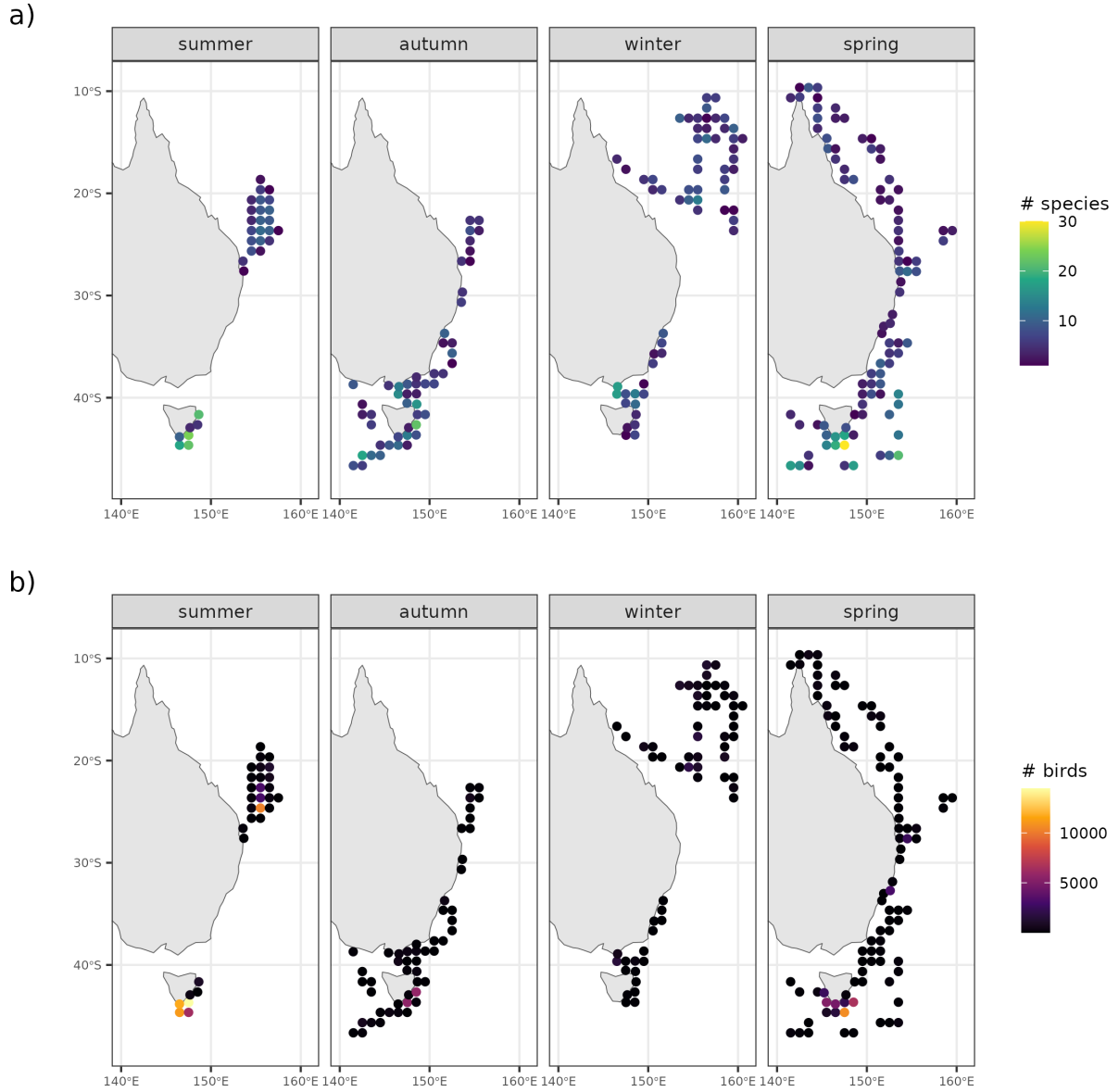
a)

b)

Figure 2: **Figure S2.** Species richness and the total number of seabirds counted off eastern Australia during Australasian Seabird Group's ship-based surveys between 2016–2021, by grid/season, after data aggregation (see *Methods* in the main text). Dots represent the grid centroids.
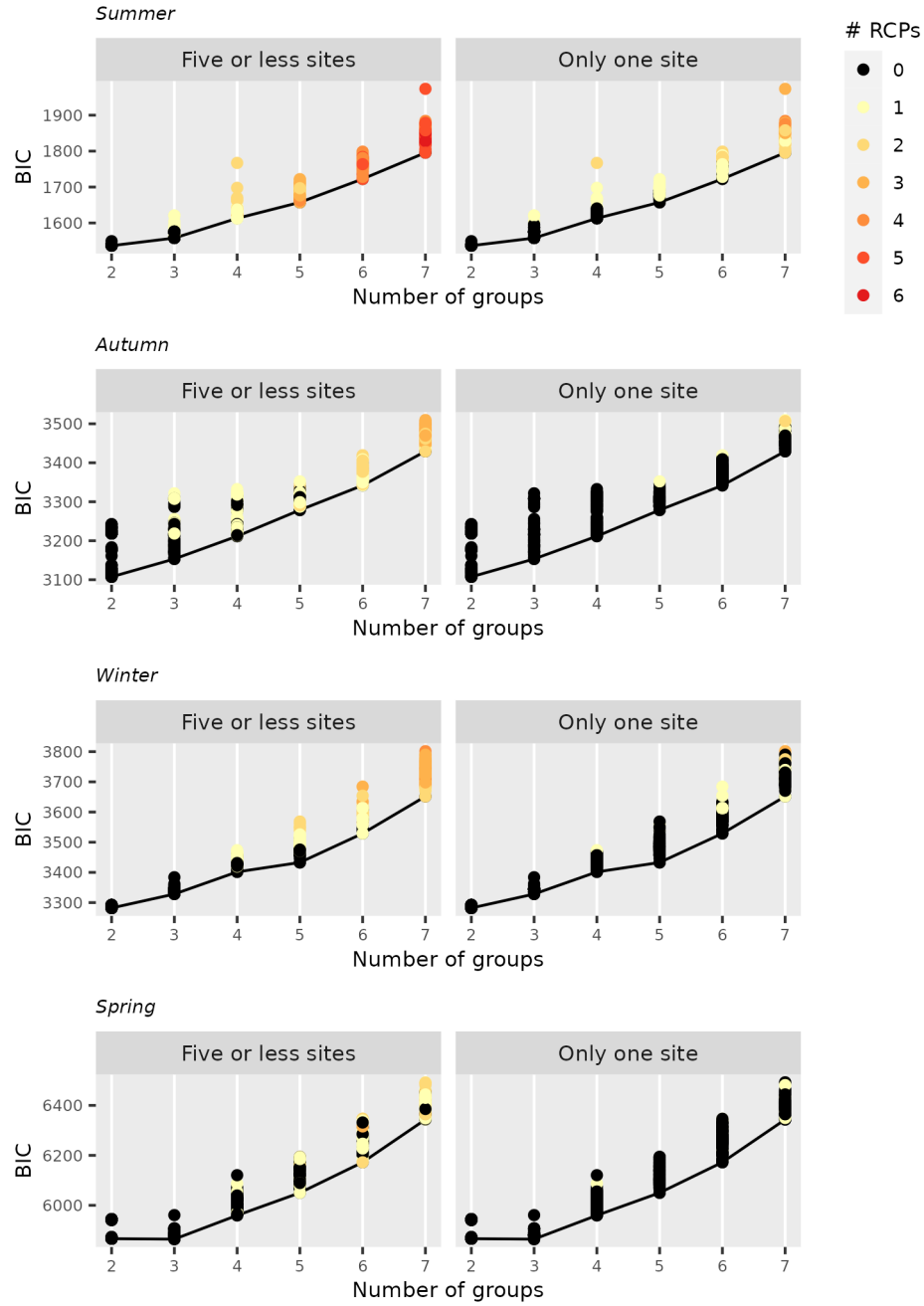
5
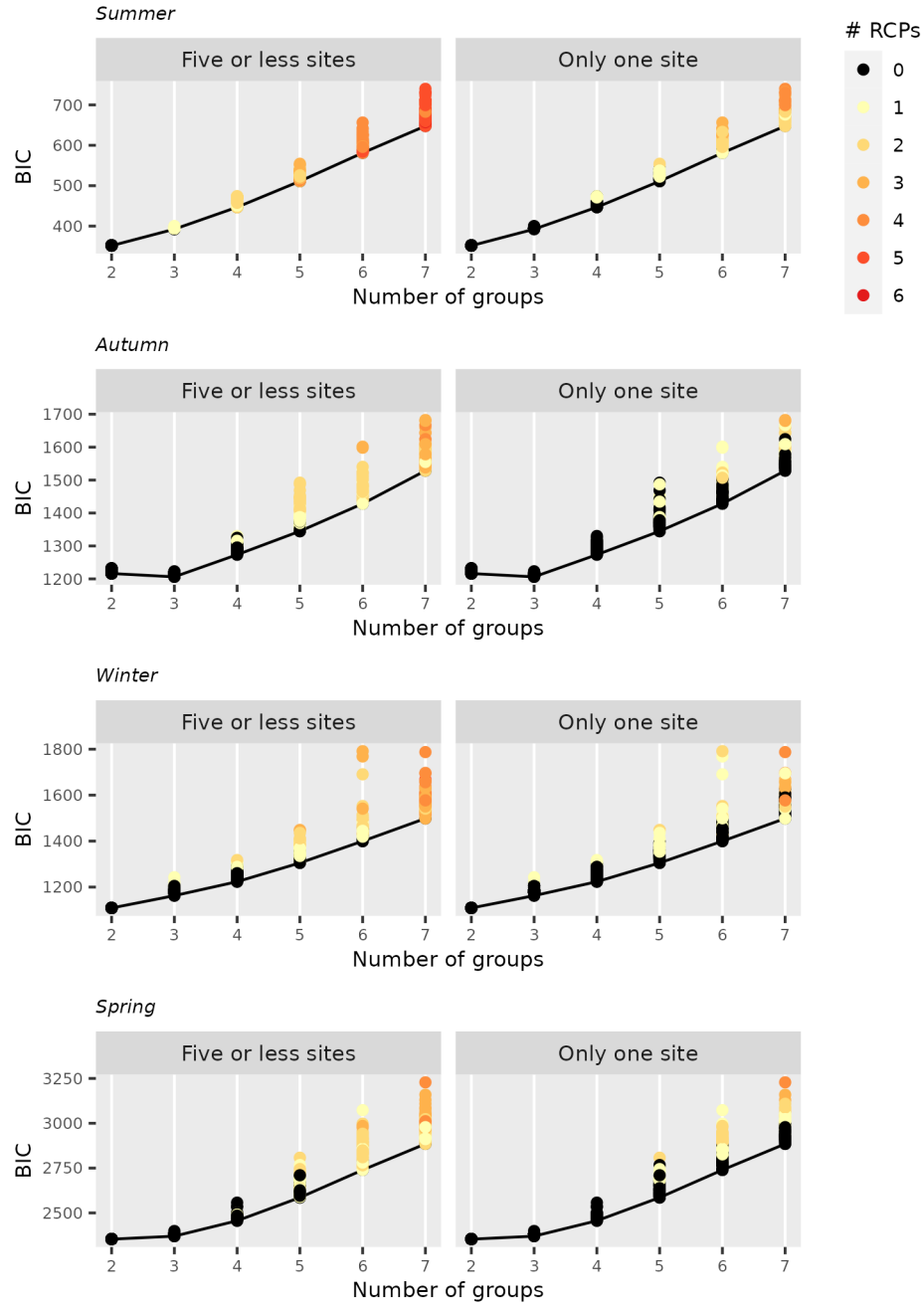
Figure 3: **Figure S3.1.** Multifit plot for Region of Common Profiles (RCP) for each seasonal presence-absence model, applied to seabirds off eastern Australia. The number of groups with the lowest BIC value indicates the best number of groups (assemblages) that describes the data. For each number of groups, we ran 100 models with random starting values to avoid getting stuck in an incorrect 'optima' (see *Methods* in the main text). The resulting plot also shows how many groups were 'empty' (colour scale) with 'five or less' or 'only one' sites assigned to an RCP, i.e. the model was fit with, say, 5 groups, but 3 of them had 'five or less' or 'only one' sites (grids) allocated to an RCP.
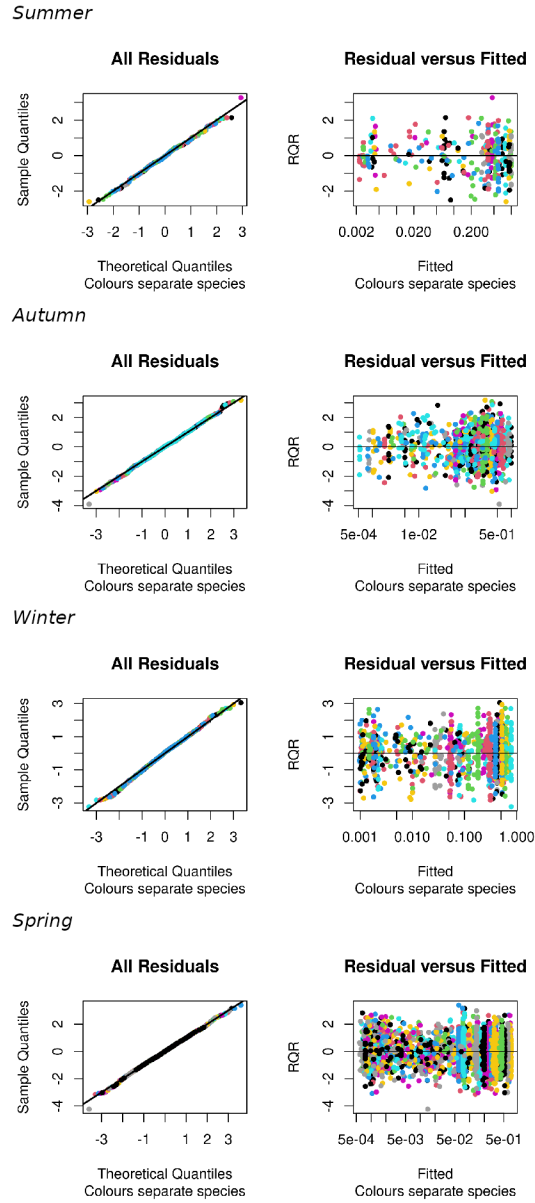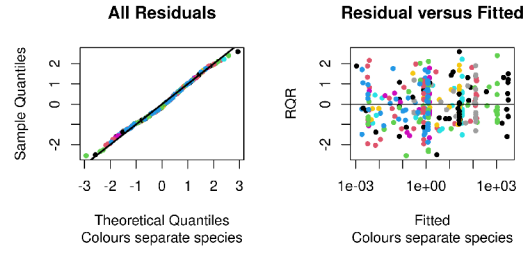
6

Figure 4: **Figure S3.2.** Multifit plot for Region of Common Profiles (RCP) for each seasonal abundance (count) model, applied to seabirds off eastern Australia. The number of groups with the lowest BIC value indicates the best number of groups (assemblages) that describes the data. For each number of groups, we ran 100 models with random starting values to avoid getting stuck in an incorrect 'optima' (see *Methods* in the main text). The resulting plot also shows how many groups were 'empty' (colour scale) with 'five or less' or 'only one' sites assigned to an RCP, i.e. the model was fit with, say, 5 groups, but 3 of them had 'five or less' or 'only one' sites (grids) allocated to an RCP.
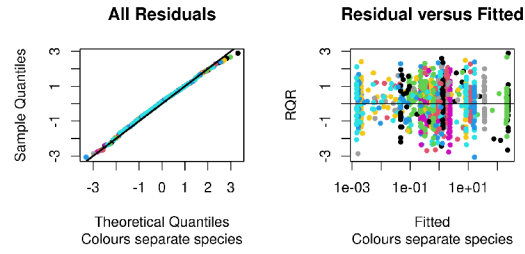
Figure 5: **Figure S4.1.** Best model residuals for each seasonal presence-absence Region of Common Profile model, applied to seabirds off eastern Australia.
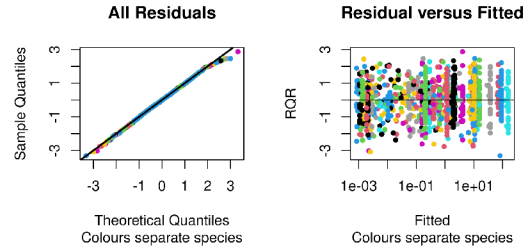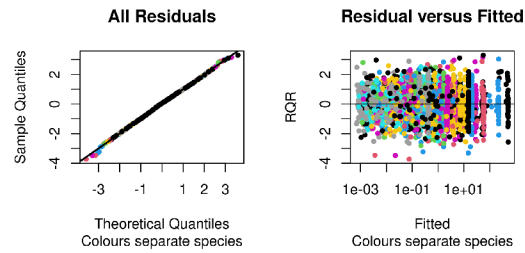
Figure 6: **Figure S4.2.** Best model residuals for each seasonal abundance (count) Region of Common Profile model, applied to seabirds off eastern Australia.
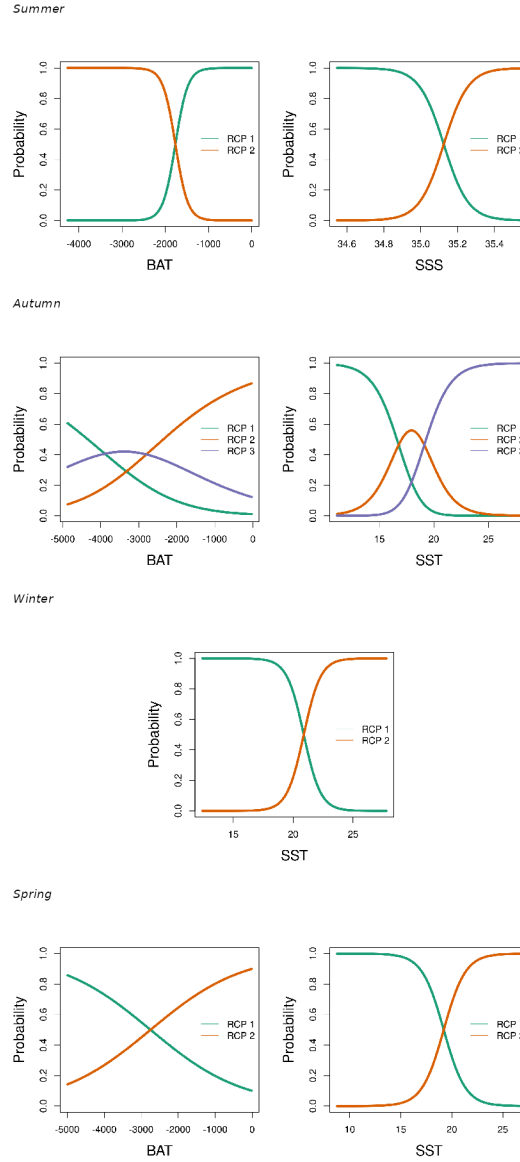
Figure 7: **Figure S5.1.** Partial plots for the retained covariates in the best seasonal models based on presence-absence data. The plot shows the probability of belonging to a Region of Common Profiles (RCP) against the environmental value. Refer to Table 1 in the main text for the environmental data acronyms.
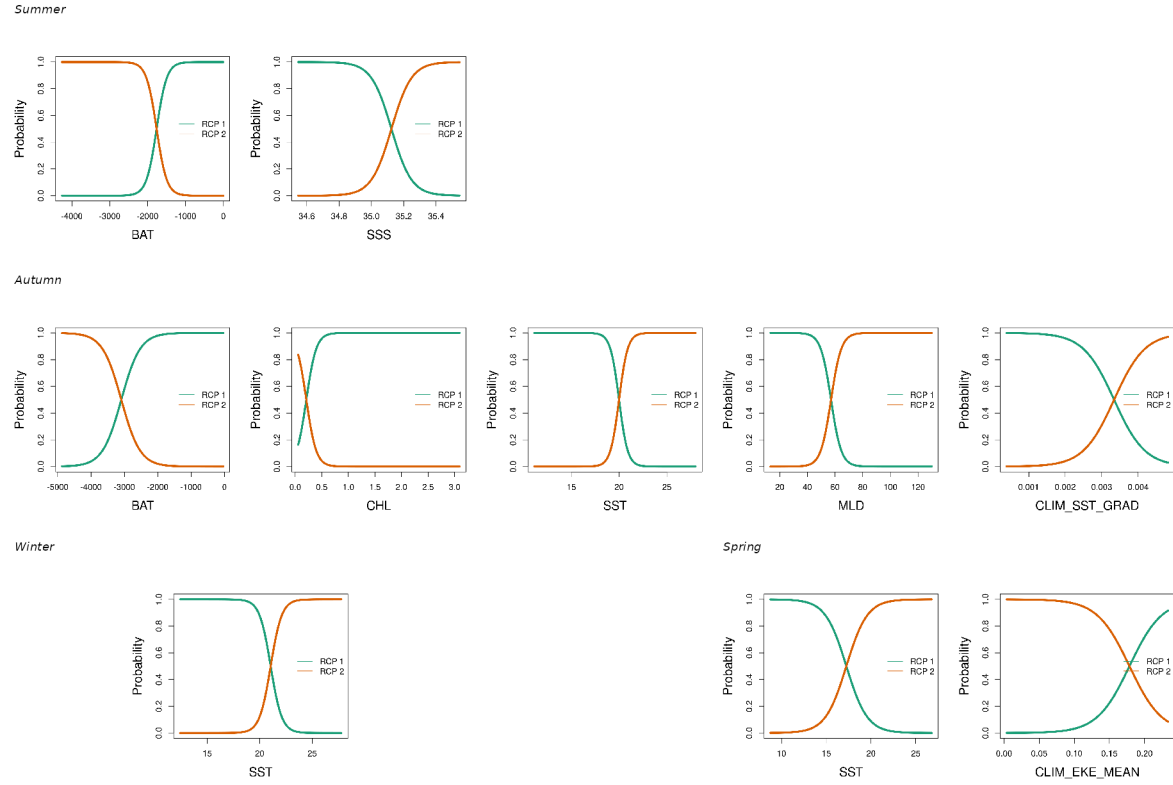
Figure 8: **Figure S5.2.** Partial plots for the retained covariates in the best seasonal models based on abundance data. The plot shows the probability of belonging to a Region of Common Profiles (RCP) against the environmental value. Refer to Table 1 in the main text for the environmental data acronyms.
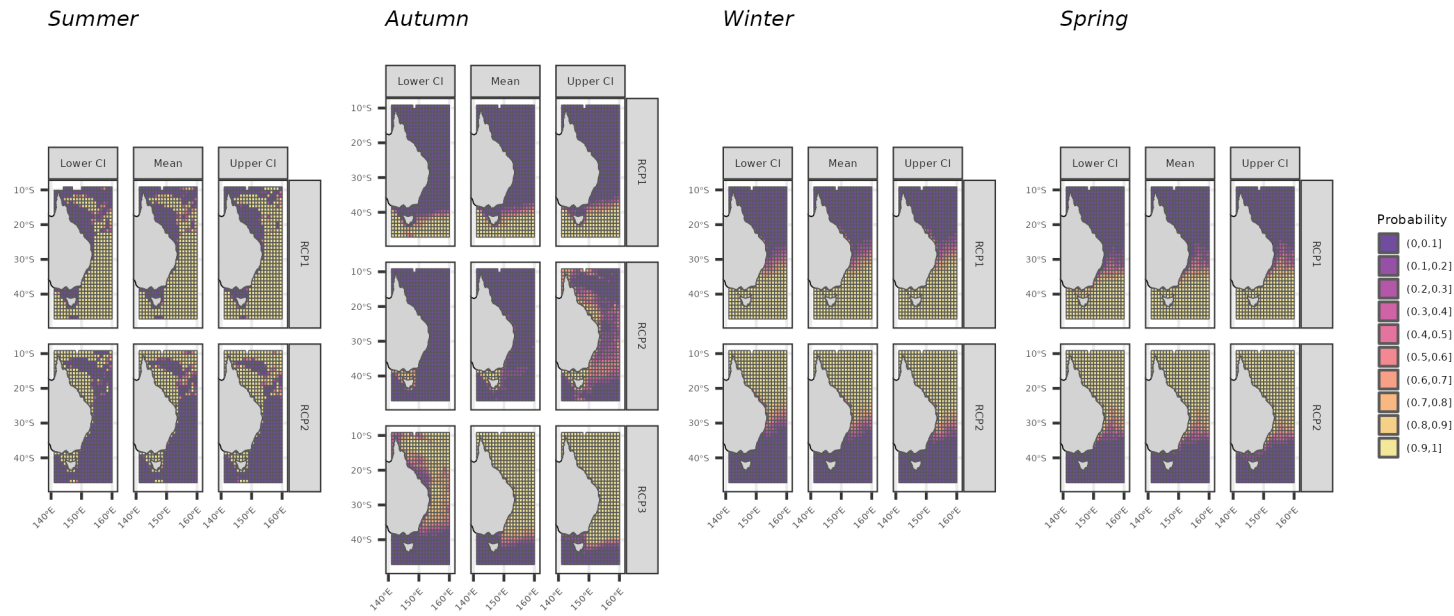
Figure 9: **Figure S6.1.** Predicted probability membership of for each seabird assemblage (Region of Common Profiles; RCP) and grid, off eastern Australia, from presence-absence models. The central column, 'mean', corresponds to the point prediction and Bayesian bootstraped, lower and upper confidence intervals (CI), on its sides.
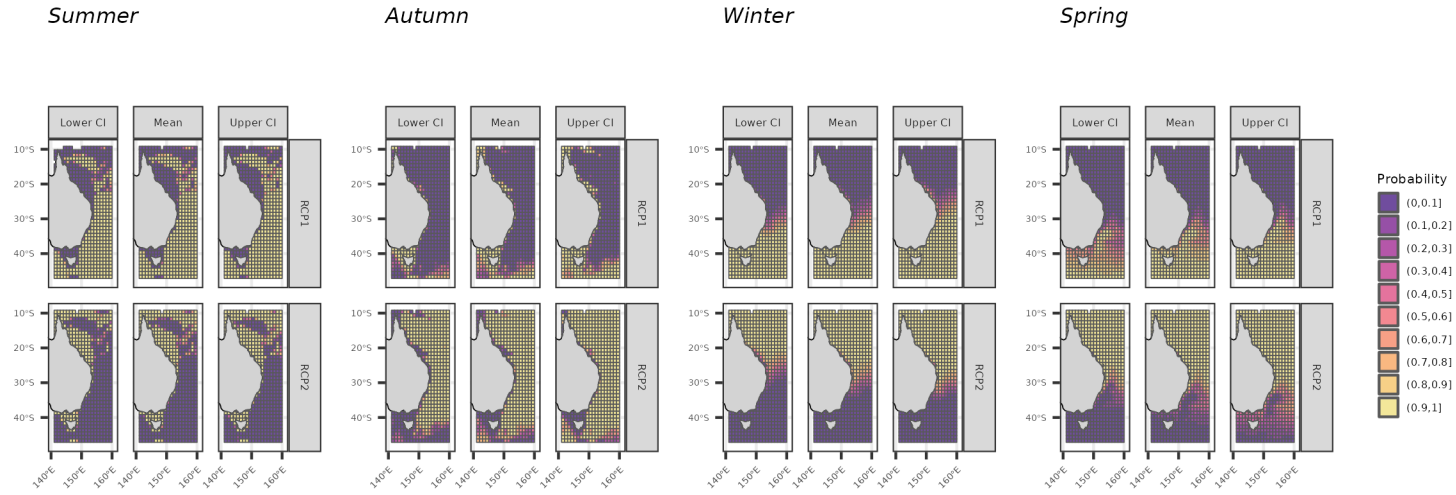
Figure 10: **Figure S6.2.** Predicted probability membership of for each seabird assemblage (Region of Common Profiles; RCP) and grid, off eastern Australia, from abundance (count) models. The central column, 'mean', corresponds to the point prediction and Bayesian bootstraped, lower and upper confidence intervals (CI), on its sides.
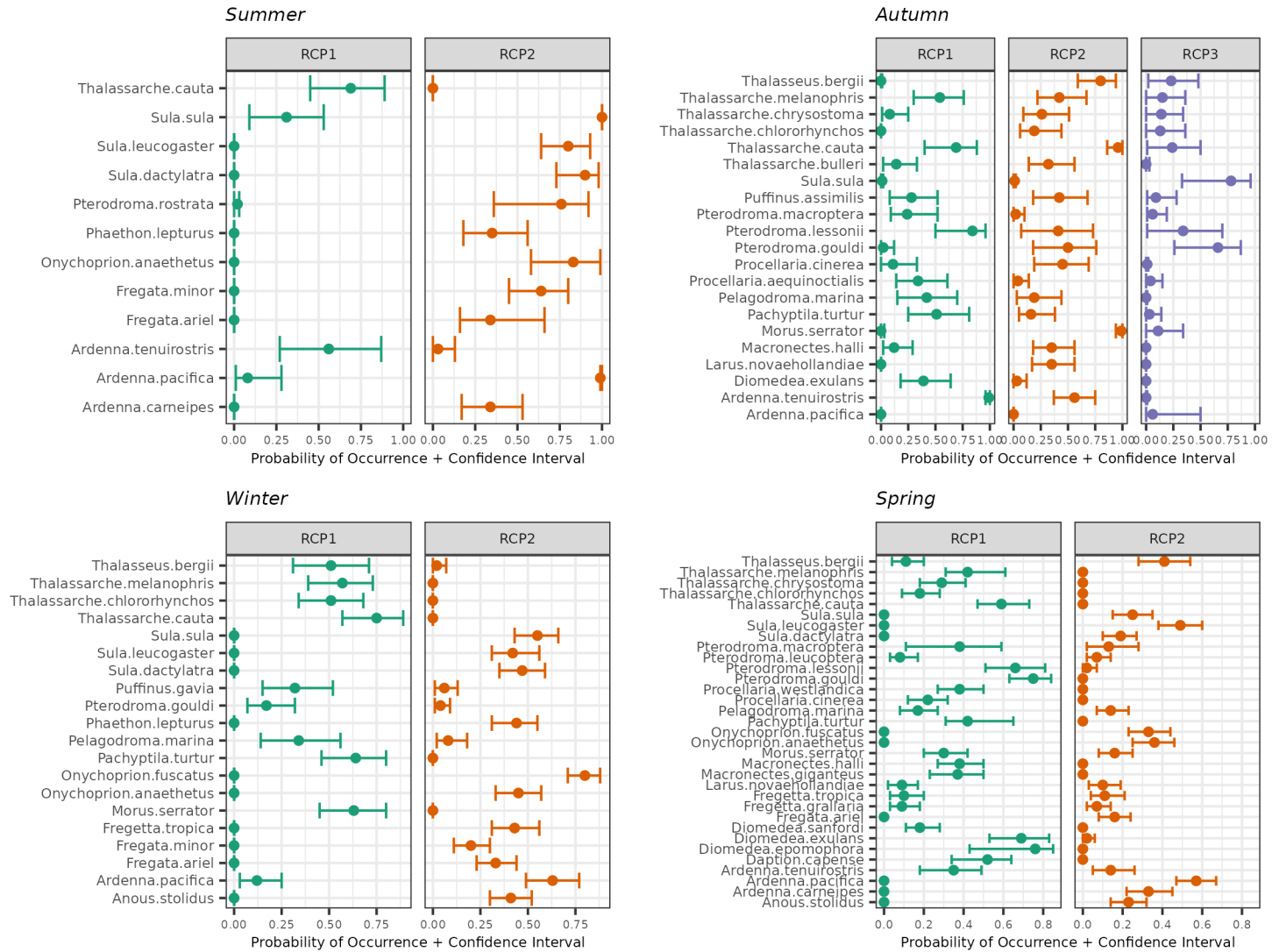
Figure 11: **Figure S7.1.** Species profiles for each assemblage (Region of Common Profiles; RCP) for each seasonal presence-absence model. Values are the average and confidence intervals of probability of occurrence for each species, based on 1000 Bayesian bootstraps.
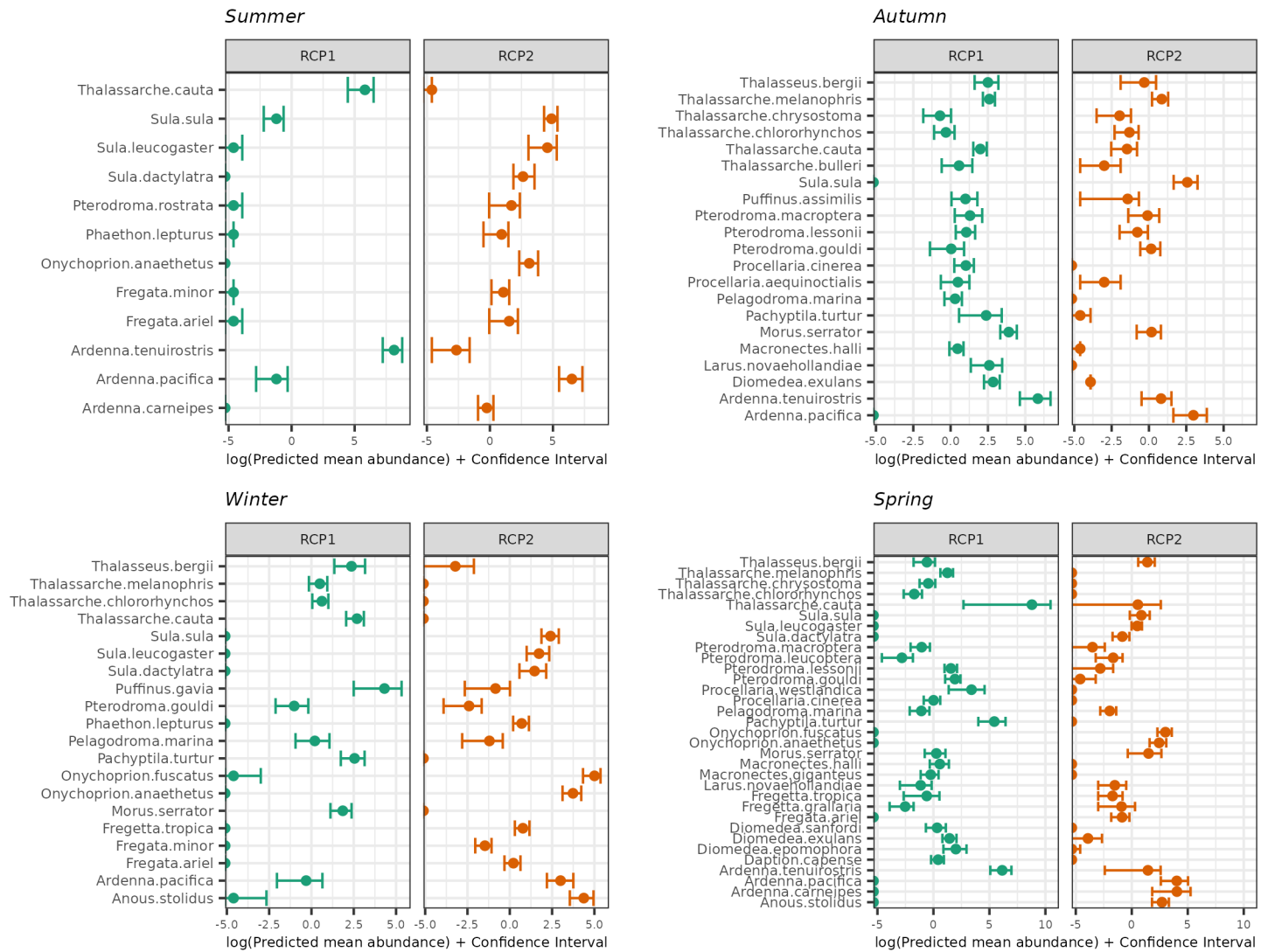
Figure 12: **Figure S7.2.** Species profiles for each assemblage (Region of Common Profiles; RCP) for each seasonal abundance (count) model. Values are the average and confidence intervals of predicted mean abundance for each species, based on 1000 Bayesian bootstraps. Values were log10-transformed to accommodate the high variation between species.
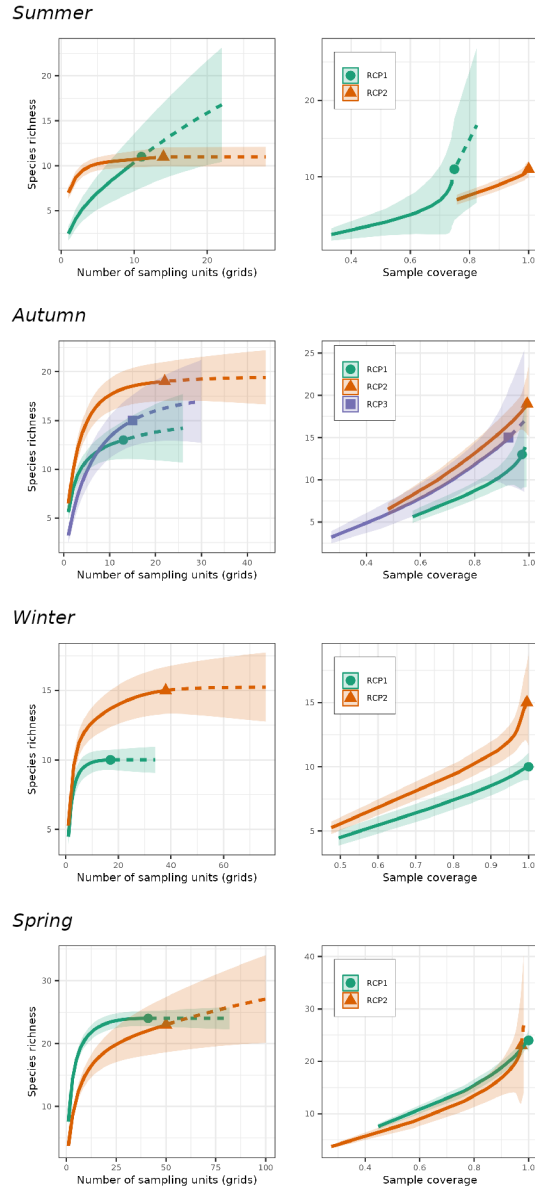
Figure 13: **Figure S8.** Diversity curve (alpha diversity) and sample coverage for each assemblage (Region of Common Profile; RCP) from each presence-absence seasonal model.

16

References

Appelhans, T., Detsch, F., Reudenbach, C., & Woellauer, S. (2022). *Mapview: Interactive viewing of spatial data in r.* https://CRAN.R-project.org/package=mapview

Bivand, R. S., Pebesma, E., & Gomez-Rubio, V. (2013). *Applied spatial data analysis with R, second edition.* Springer, NY. https://asdar-book.org/

Byrnes, J., & Dunic, J. (2017). *hadsstR: R library for working with HadSST data.* https://github.com/jebyrnes/hadsstR

Chamberlain, S. (2023). *Rerddap: General purpose client for 'ERDDAP' servers.* https://CRAN.R-project.org/package=rerddap

Dunnington, D. (2022). *Ggspatial: Spatial data framework for ggplot2.* https://CRAN.R-project.org/package=ggspatial

Grolemund, G., & Wickham, H. (2011). Dates and times made easy with lubridate. *Journal of Statistical Software, 40*(3), 1–25. https://www.jstatsoft.org/v40/i03/

Hijmans, R. J. (2022a). *Raster: Geographic data analysis and modeling.* https://CRAN.R-project.org/package=raster

Hijmans, R. J. (2022b). *Terra: Spatial data analysis.* https://CRAN.R-project.org/package=terra

Massicotte, P., & South, A. (2023). *Rnaturalearth: World map data from natural earth.* https://CRAN.R-project.org/package=rnaturalearth

Mendelssohn, R. (2022). *rerddapXtracto: Extracts environmental data from 'ERDDAP' web services.* https://CRAN.R-project.org/package=rerddapXtracto

Müller, K., & Wickham, H. (2023). *Tibble: Simple data frames.* https://CRAN.R-project.org/package=tibble

Neuwirth, E. (2022). *RColorBrewer: ColorBrewer palettes.* https://CRAN.R-project.org/package=RColorBrewer

Pebesma, E. (2018). Simple features for R: standardized support for spatial vector data. *The R Journal, 10*(1), 439–446. https://doi.org/10.32614/RJ-2018-009

Pebesma, E. J., & Bivand, R. S. (2005). Classes and methods for spatial data in R. *R News, 5*(2), 9–13. https://CRAN.R-project.org/doc/Rnews/

Pedersen, T. L. (2022). *Patchwork: The composer of plots.* https://CRAN.R-project.org/package=patchwork

Wei, T., & Simko, V. (2021). *R package 'corrplot': Visualization of a correlation matrix.* https://github.com/taiyun/corrplot

Wickham, H. (2011). The split-apply-combine strategy for data analysis. *Journal of Statistical Software, 40*(1), 1–29. https://www.jstatsoft.org/v40/i01/

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis.* Springer-Verlag New York. https://ggplot2.tidyverse.org

Wickham, H. (2022). *Stringr: Simple, consistent wrappers for common string operations.* https://CRAN.R-project.org/package=stringr

Wickham, H., François, R., Henry, L., Müller, K., & Vaughan, D. (2023). *Dplyr: A grammar of data manipulation.* https://CRAN.R-project.org/package=dplyr

Wickham, H., & Henry, L. (2023). *Purrr: Functional programming tools.* https://CRAN.R-project.org/package=purrr

Wickham, H., Hester, J., & Bryan, J. (2023). *Readr: Read rectangular text data.* https://CRAN.R-project.org/package=readr

Wickham, H., Vaughan, D., & Girlich, M. (2023). *Tidyr: Tidy messy data.* https://CRAN.R-project.org/package=tidyr