Using model-based cluster analysis to describe seabird assemblages off East Australia

Traditional methods of describing species assemblages and bioregions commonly use a two-step approach. Alternatively, probabilistic methods, such as Region of Common Profiles (RCP), enable environmental spaces and species profiles to be modelled simultaneously through specifying a statistical model. This also allows the opportunity to compare models, present uncertainty around the results, and check model assumptions. Australia supports a high diversity of seabirds, yet no attempts to describe seabird assemblages at broad-scales have been made. In particular, the eastern coast of Australia offers wide environmental gradients, spanning from the tropics to the sub-Antarctic, in which a diverse range of seabird assemblages would be expected. Here, we used RCP models to describe seabird assemblages off East Australia. RCP modelling is a type of multivariate model-based cluster analysis, which identifies 'statistical bioregions', 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. Data were collected at-sea, along transects carried out during 15 voyages (2016–2020) over ~37 degrees of latitude between the Coral Sea and Tasmania. Seabirds were recorded continuously and binned into 10-min observations (n = 9990). In total, 142,883 individual seabirds were counted from 73 species. We used oceanographic and physiographic variables to model environmental spaces (sea surface temperature and salinity, eddy kinetic energy, bathymetry, slope, and 8-days averaged chlorophyll-a concentration). Based on the availability of environmental data, 7019 complete observations were used to model RCPs. We filtered the main dataset to include only species exceeding a minimum number of occurrence thresholds (25, 30, 35, 40), then used these datasets to fit the models. For each dataset, we modelled 2 to 12 RCPs and used 63 candidate models derived from all possible linear, additive combinations of the environmental covariates (totalling 2772 models). When using a subset of the main dataset (2018–2019), preliminary analyses suggested two RCPs (coastal and oceanic). We will present the analysis of the whole dataset, aiming to describe the environmental characteristics which define seabird assemblages off eastern Australia.