

Do communities exist? Complex patterns of overlapping marine species distributions

REBECCA LEAPER,¹ PIERS K. DUNSTAN,^{2,3} SCOTT D. FOSTER,² NEVILLE S. BARRETT,¹ AND GRAHAM J. EDGAR¹

¹*Institute for Marine and Antarctic Studies, University of Tasmania, Private Bag 49, Hobart, Tasmania 7001 Australia*

²*CSIRO Wealth from Oceans Flagship, GPO Box 1538, Hobart Tasmania 7001 Australia*

Abstract. Understanding the way in which species are associated in communities is a fundamental question in ecology. Yet there remains a tension between communities as highly structured units or as coincidental collections of individualistic species. We explored these ideas using a new statistical approach that clusters species based on their environmental response: a species archetype, rather than clustering sites based on their species composition. We found groups of species that are consistently highly correlated, but that these groups are not unique to any set of locations and overlap spatially. The species present at a single site are a realization of species from the (multiple) archetype groups that are likely to be present at that location based on their response to the environment.

Key words: Clements; community; composition; Gleason; grouping; prediction; species archetype; structure; subtidal rocky reefs.

INTRODUCTION

Understanding the processes that structure communities has been, and remains, a frustrating challenge for ecologists (e.g., see McIntosh 1995 for review of the topic; Lawton 1999, Ricklefs 2008). That communities have some form of structure is undeniable, but exactly what processes drive that structure and what the consequences are, remains unclear (Vellend 2010, Chase and Myers 2011, Lessard et al. 2012). There is currently a strong school of thought that communities are highly structured and form cohesive units (e.g., Paine 1980, Richardson 1980, Gotelli and McCabe 2002, Chase and Leibold 2003). Contrasting this are the observations that communities are structured by weak interactions (e.g., Fagan and Hurd 1994, Raffaelli and Hall 1995, McCann et al. 1998, Berlow 1999) or are perhaps largely a coincidence (i.e., random aggregation) of individualistic species (e.g., Whittaker 1952, Mills 1969, Ter Braak and Prentice 1988, Bastow Wilson and Gitay 1995).

The ecological community has been here before. Early 20th-century plant ecologists Frederic E. Clements and Henry A. Gleason provide some of the earliest proposals for different approaches to explaining communities, and their ideas have influenced ecological thought since (Eliot 2007, 2011). While Clements developed a theory of the plant community as governed by deterministic laws of succession that led to a highly integrated and self-regulating stable climax community (Clements 1916, 1936), Gleason developed an “individualistic” theory where communities did not have distinct functional

organization, but rather were chance associations of organisms that were particularly adapted to one or another environment (Gleason 1917, 1926).

If the concept of community as a cohesive unit is realistic, it follows that all the species that are members of that community should be associated in a nonrandom way; that is, the species in any location should be highly correlated with other species in the community. The expectation is that species from these types of communities will show the same environmental preferences and would be evidenced by having a very similar species distribution model up to the realized niche. The alternative is a completely random assortment of species that shows no common environmental preferences, and where species are essentially uncorrelated with each other. A recently developed statistical approach to modeling species groups, species archetype models (SAMs), would be informative in understanding the way species are associated in a community (Dunstan et al. 2011).

A SAM is an unsupervised model-based approach that clusters species based on their environmental responses. The approach does *not* cluster sites based on their species composition in contrast to most other analyses. Species that respond similarly to the environment are grouped and termed species archetypes, where one species archetype may represent a single species or many (Dunstan et al. 2011). There are no a priori assumptions about the composition of the archetypes. Our key hypothesis was that if a community is a tightly interacting group of species, then those species should generally be found together and should occupy the same environmental space. However, if communities are comprised of completely disassociated species that

Manuscript received 30 April 2013; revised 15 November 2013; accepted 18 November 2013; final version received 18 December 2013. Corresponding Editor: H. H. Wagner.

³ Corresponding author. E-mail: Piers.Dunstan@csiro.au

appear randomly, then we should see no species groups forming and every species with a different distribution.

Statistically based community models are typically based on distance measures of composition between sites rather than species and their responses to the environment (e.g., Whittaker 1967, Faith et al. 1987, Legendre and Legendre 1998, De'ath 2002, Ferrier et al. 2007). In contrast, species distribution models (SDMs; e.g., Guisan and Thuiller 2005, Elith et al. 2006, Guisan et al. 2006, Elith and Leathwick 2009) estimate the relationship between species records at sites and the environmental and/or spatial characteristics of those sites (Elith et al. 2011), but are problematic to group together prior to undertaking modeling or after multiple fits (Ferrier and Guisan 2006, Dunstan et al. 2011). Multispecies approaches, such as multivariate regression trees (MVPART; De'ath 2002) and multivariate adaptive regression splines (MARS; Friedman 1991) model the distribution of several species simultaneously, but as with distance approaches, they continue to consider the site as the fundamental data unit. These approaches predict the distribution of composition across sites, rather than the distribution groups of species based on the groups' response to the environment into unsurveyed regions, allowing assessment of broad scale multispecies patterns.

We tested our hypothesis of species grouping with the analysis and prediction of species archetypes for demersal fish and macroinvertebrate species records from an extensive spatial and temporal quantitative survey of a well-understood rocky subtidal system on the southern coastline of continental Australia and the northern coastline of Tasmania. Reefs in this region are subject to strong environmental gradients (Butler et al. 2002) and support a rich and biogeographically varied flora and fauna (Edgar 1997, Andrew 1999, Gomon et al. 2008). This work considers species groupings in the context of a sound knowledge of the ecology of many of the subtidal rocky reef species.

MATERIALS AND METHODS

Study reefs

The southern coast of continental Australia, which extends ~5500 km from Cape Naturaliste, Western Australia (33.32° S, 115.01° E), across the South Australian and Victorian coasts to the Victorian/New South Wales border (37.34° S, 149.45° E), is the longest ice-free east–west coastline in the world (Phillips 2001). The island of Tasmania is located to the south of Victoria, between 40° S and 43° S latitude (Fig. 1). Since 1992, rocky subtidal reefs associated with marine protected areas (MPAs) have been monitored regularly (see Plate 1), with major emphasis on reefs in South Australia and Tasmania. Here, we analyzed data for South Australia, Victoria, and northern Tasmania, a region where, additional to the extensive ecological data set, a good (but not complete) geographical coverage of environmental data is available.

Survey methods and data collection

Data describing densities of demersal fish and macroinvertebrates were obtained along the southern coastline of continental Australia and northern coastline of Tasmania using underwater visual census protocols. For each of the 298 and 320 site locations surveyed for fish and macroinvertebrates, respectively, censuses of biota were undertaken by divers along a 200-m transect set on reef at 5–10 m water depth. Field survey protocols are described by Edgar et al. (1997). Divers laid four 50-m transect lines along either the 5-m or 10-m isobath at each site. The number and size of all fish species sighted within 5 m of the transect line were recorded for both sides of the transect line, as were rock lobsters (*Jasus edwardsii* Hutton, 1875), sea urchins (class Echinoidea), abalone (*Haliotis* spp.), and other large (>2.5 cm) mobile macroinvertebrate species within 1 m of one side of the line. A total area of 2000 m² was surveyed for fish at each site and 200 m² for macroinvertebrates. The data set analyzed here included surveys conducted in austral summer of 2006–2007.

Environmental data

A set of environmental covariate data was collated across a 0.01° grid at a national scale (Huang et al. 2010) as part of the Commonwealth of Australia's Environment Research Facility (CERF) Marine Biodiversity Hub (information *available online*).⁴ All data sources and references for spatial interpolation are annotated in the Appendix: Table A3. Each site location was assigned the environmental covariate of the closest node on the 0.01° grid. We also included a variable for wave exposure; bathymetric altered fetch (Hill et al. 2010). Wave exposure is a particularly important environmental covariate that structures shallow reef systems, and has recently been made available as different quantitative indices for much of southern Australia (Hill et al. 2010).

Modeling and prediction

We chose to model all species data using the analysis introduced in Dunstan et al. (2011) and extended in Dunstan et al. (2013). This approach is a multivariate statistical model based on finite mixtures of generalized linear models (e.g., McLachlan and Peel 2000). Finite mixture models assume that the complete set of observed data arises from multiple un-observed groups and is a flexible modeling framework that allow for unsupervised clustering of heterogeneous responses. In the SAM framework, the unobserved groups are represented as a generalized linear model (GLM), which is called a "species archetype." It is the archetypal GLMs that relate the species group to the environmental covariates, which are the primary interest. In a naive analysis, the distribution of all the single species may be represented

⁴ <http://www.marinehub.org/>

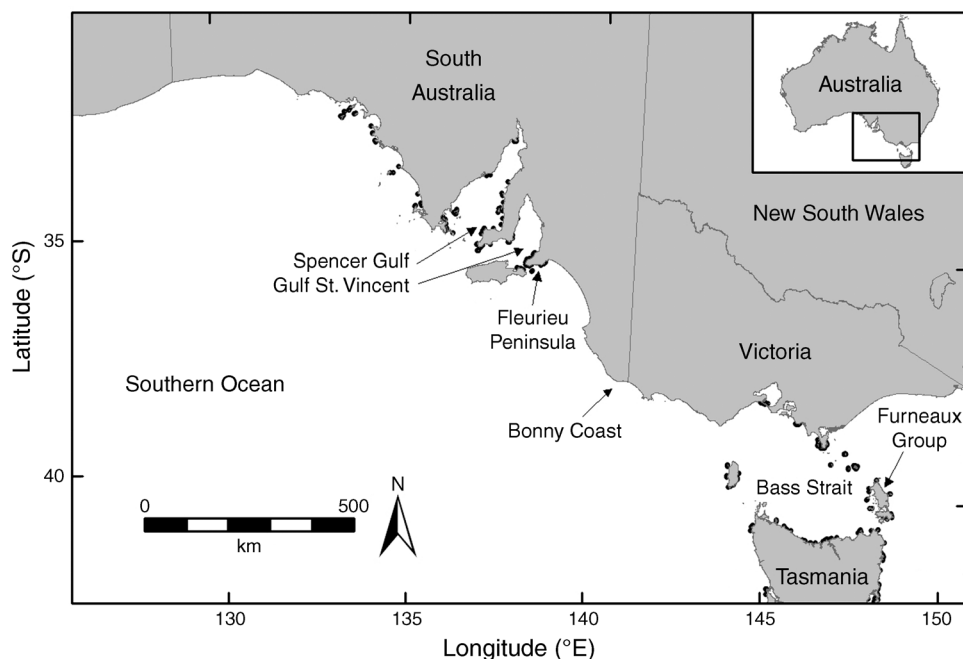


FIG. 1. The southern coastline of continental Australia and northern coastline of Tasmania, with the location of rocky subtidal reefs monitored annually since 1992.

by a single generalized linear model (see McCulloch and Nelder 1989). Each species could be estimated separately in a multispecies situation (similar to redundancy analysis). This would be an inefficient approach as the distributions of many species would be statistically indistinguishable and it would be possible to represent them with a single model. The statistical challenge is in identifying which species could be grouped together into a single model. Our approach relies on the flexibility of finite mixture model to cluster species with similar responses to environmental gradients into a single generalized linear model without any supervision. Each archetype is a group of species that share a statistically indistinguishable environmental response and are described by a single logistic GLM. Multiple archetype GLMs are derived from a single finite mixture model. The models were fitted in R-2.14.2 using the package SpeciesMix_0.2.7 (R Development Core Team 2013).

A key point of the analysis is that the estimation of the group composition occurs simultaneously with the estimation of the shared response, ensuring that uncertainty is propagated from the data. The fitted models describe the probability of a species belonging to a particular species archetype and separate GLMs for each species archetype that describes the response of that entire group to the environment.

Model selection involves both choosing the number of archetypes and choosing the covariates that describe the environment for the GLM component of the model. Model selection remains a challenge for finite mixture models (McLachlan and Peel 2000) and we opted for a stepwise approach. The number of groups is determined

first by using a full model with all covariate and quadratic terms. We chose the number of groups first, since adding a new group adds a new GLM to the mixture model, with a potentially very large change in likelihood and number of model parameters. The most parsimonious number of archetypes was determined by comparing models with different numbers of archetypes, using Bayesian information criteria (BIC) to determine the most appropriate number (see Dunstan et al. [2011] for the definition of this statistic for SAMs). We plotted $\Delta\text{BIC}(G)$, which is defined as $\text{BIC}(G+1) - \text{BIC}(G)$, with G denoting the number of archetypes. The first increase in $\Delta\text{BIC}(G)$ indicates the most likely number of archetypes is G . Additionally, if the model suggested that there was any group with <1 species, then this was taken to imply too many archetype groups (see Dunstan et al. 2011). This suggests the minimum a priori archetype membership probability, denoted by $\min(\pi)$, should be set to $1/S$ (with S as number of species) and serves as a further model summary.

The second step in model selection included selecting covariates and their functional forms. We considered linear, and quadratic terms for the 16 environmental covariates investigated (see Appendix: Table A3). Models with different combinations of covariates were compared using BIC to select the best model. We used BIC values and parameter standard errors to sequentially remove covariates while holding the number of groups constant.

We used the models with the lowest BIC to predict the probability of presence for each of the species archetypes for both demersal fish and macroinvertebrates. Predic-

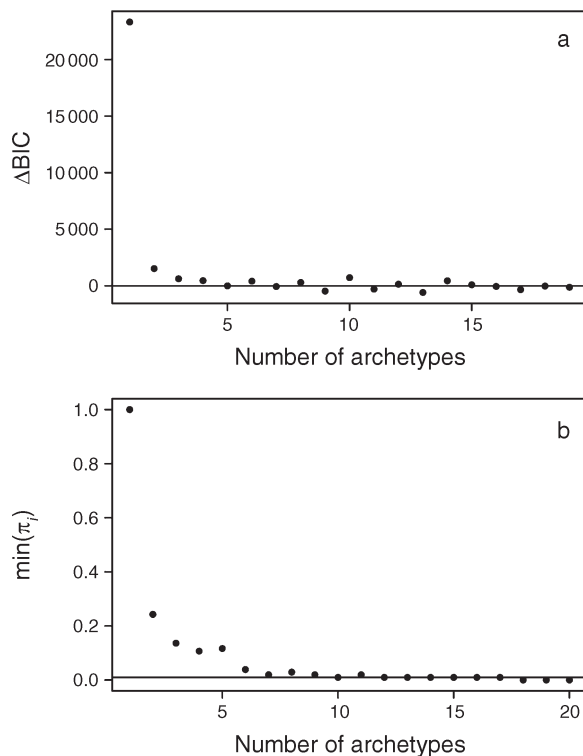


FIG. 2. Plots to aid the choice of the number of species archetypes for demersal fish. Plots are (a) $\Delta\text{BIC}(G)$ vs. G and (b) $\min(\pi_i)$ vs. G . The solid horizontal line in panel (a) is the $\Delta\text{BIC}(G) = 0$ line, and in panel (b), it is the $\min(\pi_i) = 1/S = 1/103$. $\Delta\text{BIC}(G)$ is defined as $\text{BIC}(G+1) - \text{BIC}(G)$. BIC is Bayes Information Criteria, G is the number of species archetypes, S is the number of species, and $\min(\pi_i)$ is the minimum a priori archetype membership probability.

tions were made on the dense grid along the coast of southeastern Australia (bounded by the study region). Standard errors of predictions were also calculated to aid statistical interpretation of the point predictions (see Dunstan et al. 2011).

RESULTS

Demersal fish archetypes

Initial models were fitted across a range of numbers of species archetypes from 1 to 20. The values of BIC as a function of the number of archetypes (G) initially decreased rapidly as the number of archetypes increased from one. We examined the change in BIC as archetype number was increased. Values of BIC started to increase for the first time when the number of archetypes was increased from six to seven (Fig. 2a). Values for $\min(\pi)$ from the model indicated that there were at least four species in every species archetype (Fig. 2b). From these plots (Fig. 2a, b), we concluded that $G = 6$ was the most likely number of species archetypes.

When performing variable selection with $G = 6$ groups, the model with the lowest BIC included both linear and quadratic terms for slope, the average and

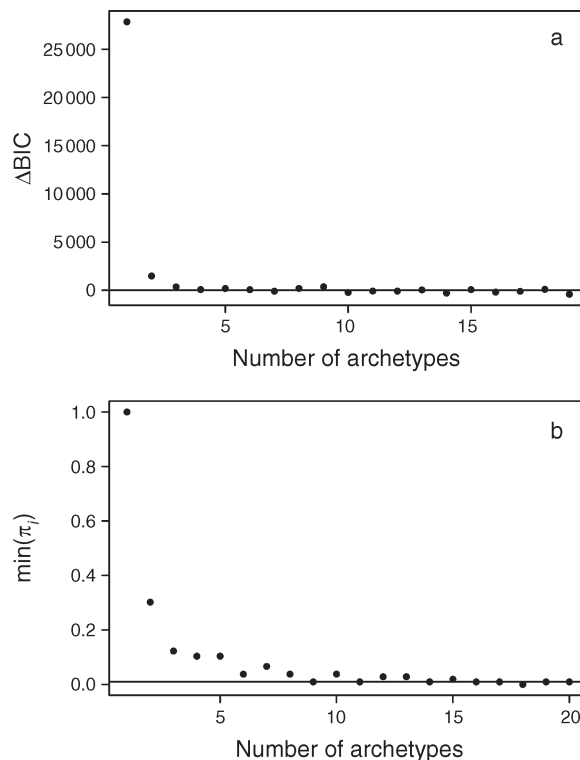


FIG. 3. Plots to aid the choice of the number of species archetypes for macroinvertebrates. Plots are (a) $\Delta\text{BIC}(G)$ vs. G and (b) $\min(\pi_i)$ vs. G . The solid horizontal line in panel (a) is the $\Delta\text{BIC}(G) = 0$ line, and in panel (b), it is the $\min(\pi_i) = 1/S = 1/106$. $\Delta\text{BIC}(G)$ is defined as $\text{BIC}(G+1) - \text{BIC}(G)$.

standard deviation of productivity (i.e., chlorophyll a), and the average, standard deviation, and minimum of sea surface temperature and average exposure (Appendix: Table A1). Examination of the estimated values and standard errors (Appendix: Table A4) of the component GLMs for the model indicated the strength of the relationships between each archetype and covariate. An important covariate will typically have a small SE relative to the size of the estimated value. A detailed interpretation of the relationship between each archetype and the physical covariates is presented in the Appendix. Membership of each species of demersal fish to each archetype is shown in the Appendix: Table A5.

Macroinvertebrate archetypes

We assessed the model for invertebrates in the same way as for demersal fish. The values of BIC initially decreased rapidly as the number of archetypes increased from one (Fig. 3a). Values of BIC increased when the number of archetypes was eight, and values for $\min(\pi)$ indicate that there are at least three species in every species archetype (Fig. 3b). From these plots (Fig. 3a, b), we concluded that $G = 8$ was the most likely number of species archetypes.

When performing variable selection with $G = 8$ groups, the model with the lowest BIC included both

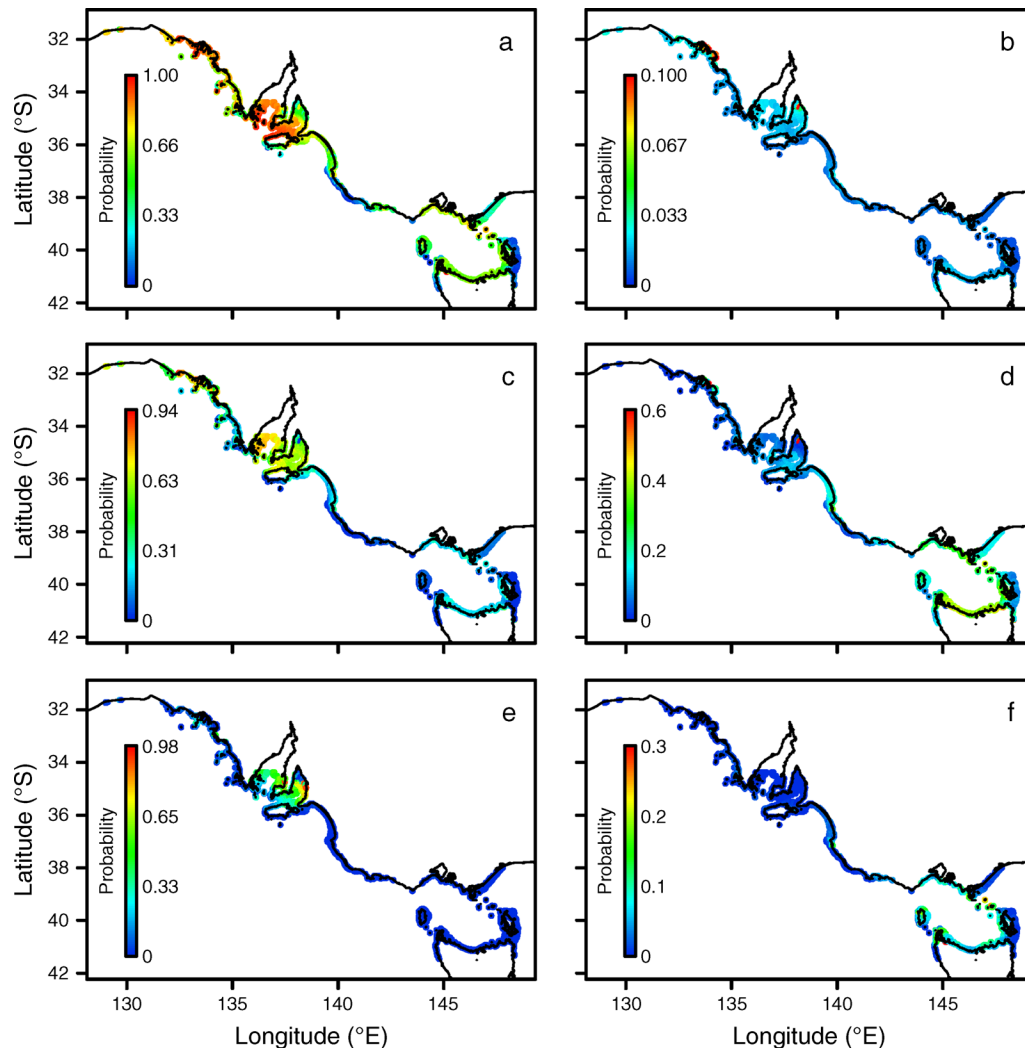


FIG. 4. The probability of presence across southeastern Australia of species archetypes (a) 1, (b) 2, (c) 3, (d) 4, (e) 5, and (f) 6 for demersal fish. Archetype membership is given in Appendix: Table A5.

linear and quadratic terms for longitude, latitude, depth, and the average, standard deviation, and minimum of sea surface temperature (Appendix: Table A2). The estimated values and standard errors are shown in the Appendix: Table A6. A detailed interpretation of the relationship between each archetype and the physical covariates is presented in the Appendix (see Table A7, Archetype membership).

Spatial prediction and patterns of response for demersal fish and macroinvertebrate archetypes

For both demersal fish and macroinvertebrates, it was clear that in most cases there was considerable spatial overlap of different archetypes rather than separation along the coastline. The means of the predicted values for demersal fish archetypes are shown in Fig. 4 and for macroinvertebrate archetypes in Fig. 5 (standard errors in the Appendix: Figs. A1a–f and A2a–h). We present a

more detailed description of the spatial predictions and patterns of response the environmental covariates for both sets of taxa in Appendix: Figs. A1 and A2.

There were a number of clear patterns in the spatial prediction of the archetypes for both demersal fish and macroinvertebrates. For example, archetypes 1 and 3 for demersal fish and archetypes 3 and 7 for macroinvertebrates showed a pattern that was typical of species with strong western affinity in the region (i.e., west of 138° E). For demersal fish, they included species such as the zebra fish (*Girella zebra*) and scalyfin (*Parma victoriae*), members of archetype 1, and the Western blue grouper (*Achoerodus gouldii*) and dusky morwong (*Dactylophora nigricans*), members of archetype 3. For macroinvertebrates, archetypes included species such as the pheasant shell (*Phasianella australis*), a member of archetype 3, and the western hollow-spined urchin (*Centrostephanus*

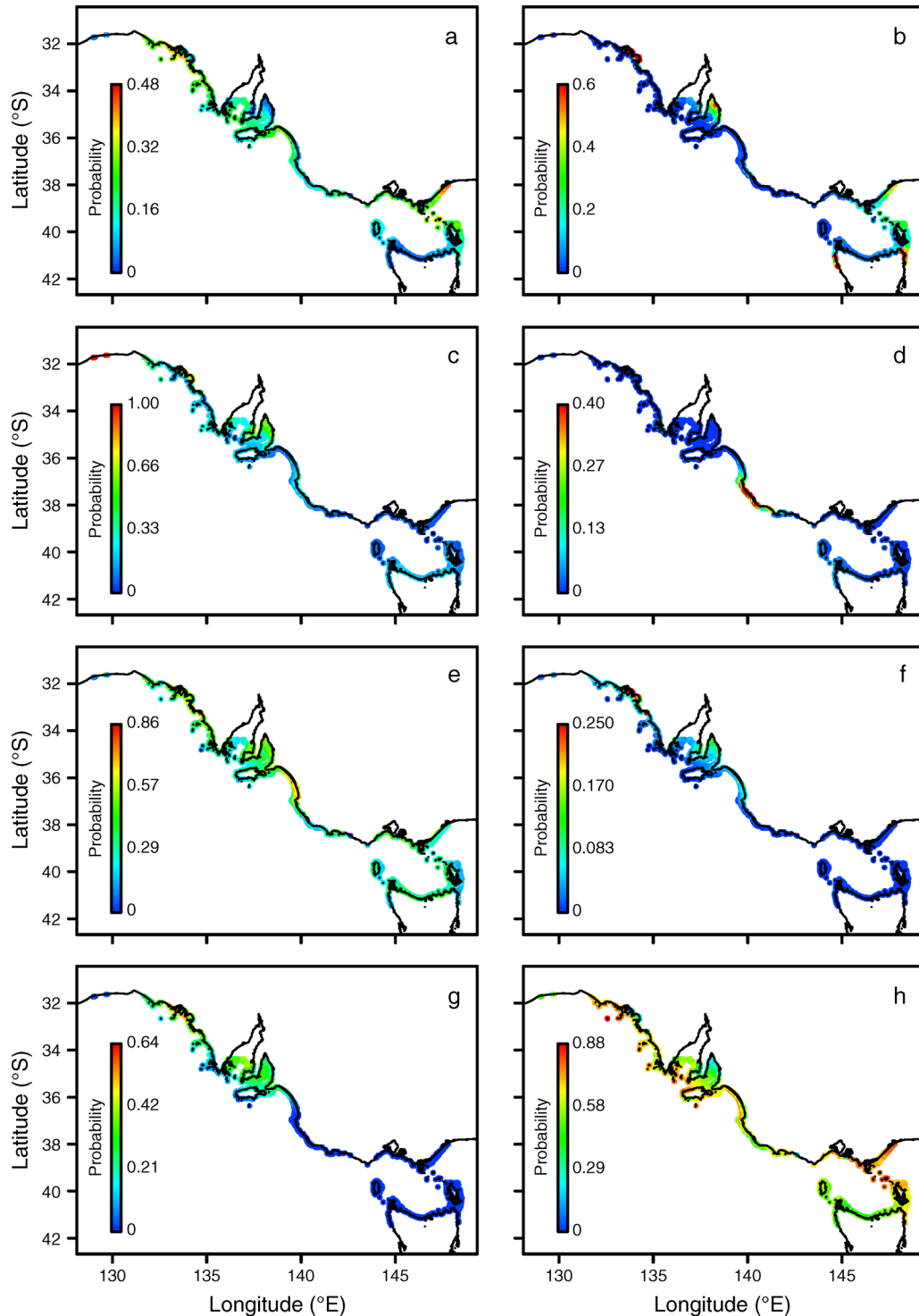


FIG. 5. The probability of presence across southeastern Australia of species archetypes (a) 1, (b) 2, (c) 3, (d) 4, (e) 5, (f) 6, (g) 7, and (h) 8 for macroinvertebrates. Archetype membership is given in Appendix: Table A7.

tenuispinus) and western slate-pencil urchin (*Phyllacanthus irregularis*), members of archetype 7.

Demersal fish archetype 2 and macroinvertebrate archetypes 1 and 6 showed a pattern that was generally

typical of species that are widespread, but also rare in the region. For example, demersal fish archetype 2 included the velvet leatherjacket (*Meuschenia scaber*), while macroinvertebrate archetype 1 included the many-

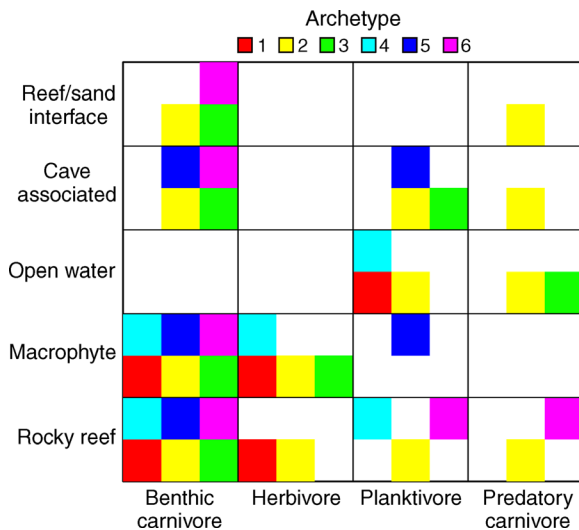


FIG. 6. Different traits (i.e., habitat and trophic level) for each archetype of demersal fish. The colored squares represent the six different archetypes.

pored seastar (*Fromia polypora*), and archetype 6 included the mottled seastar (*Anthaster valvulatus*) and fimbriate helmet (*Cassia fimbriata*) (the latter found only in the South Australian Gulfs).

There were also some clear differences in the spatial predictions of the archetypes between the taxa. For example, among demersal fish, archetypes 4 and 6 had a strong eastern affinity and were represented by species either at the edge of their western distribution (including the purple wrasse [*Notolabrus fucicola*] and the southern hulafish [*Trachinops caudimaculatus*], members of archetype 4), or by species that are not found in South Australia (including the crimsonband wrasse [*Notolabrus gymnogenis*], silver sweep [*Scorpius lineolata*], and eastern hulafish [*Trachinops taeniatus*], members of archetype 6). There was also one fish archetype that had a strong affinity to the South Australian Gulfs (archetype 5). The black spotted wrasse (*Austrolabrus maculatus*) largely recruits in the gulfs within the region studied, and the rough headed bullseye (*Pempheris klunzingeri*) and yellow headed hulafish (*Trachinops noarlungae*) are not found east of Kangaroo Island and Gulf St. Vincent (138° E). Macroinvertebrate archetypes 5 and 8 showed a pattern that was typical of species that were ubiquitous and common in the region, a pattern not seen for demersal fish. Macroinvertebrate species included the cartrut shell (*Dicathais orbita*), the green-lip abalone (*Haliotis laevis*), members of archetype 5, and the orange feather star (*Comanthus trichoptera*) and blacklip abalone (*Haliotis rubra*), members of archetype 8.

Three species of macroinvertebrate, the inflated egg urchin (*Holopneustes inflatus*), rasping hermit crab (*Strigopagurus strigimanus*), and long-spined sea urchin (*Centrostephanus rodgersii*), members of macroinvertebrate archetype 2, were predicted to occur outside of

their known range (i.e., east of ~141° E). While these species currently do not occur in South Australia, the high probability of occurrence along the western coast of South Australia (133°–134° E) and St. Vincent Gulf (137.5°–138.5° E) indicates that this may be a niche that these species could potentially occupy.

With respect to the composition of each of the archetypes, for demersal fish, it is clear that each archetype is largely made up of species with different traits; i.e., each archetype contains a combination of different trophic and taxonomic groups (Fig. 6; see Appendix: Table A8). While the same level of information is not available for macroinvertebrates, the different taxonomic groups represented in each archetype (see Appendix: Table A7) also indicate that macroinvertebrate archetypes are likely made up of species with different traits (N. Barrett, *personal communication*).

DISCUSSION

There is an inevitable tension between concepts of communities as cohesive complex webs of interacting species and communities as a random assemblage of species drawn from a pool. Careful examination of the identities of species in most community ecology data sets suggests that communities defined on the basis of location share many species with other communities. The same species can be present in many different communities, suggesting that a community is a fluid concept (for example, Fig. 3b in Dunstan et al. 2012). Our results suggest that there are groups of species, species archetypes, which are consistently highly correlated, but that these groups are unlikely to form complex webs of interacting species between the same species because species identity changes between locations. Within any one site, species will interact, but between sites in the same environmental space, the identities are likely to change. Rather, the species present at a single site are a realization of species from the (multiple) archetype groups that are likely to be present at that location based on their response to the environment. The sites with the same conditions could potentially have a similar assemblage, but equally could have a very different assemblage.

Archetype groups are not members of a single functional group (see Fig. 6 for demersal fish). Rather, each archetype has members of a number of different functional groups (performing different functions in the system) such that the archetypes comprise groups of complementary, rather than similar, species. Each archetype may be a functioning system by itself, but we do not have the evidence to support that idea from these data. Previous work on the structure of the same demersal fish and macroinvertebrate communities by Leaper et al. (2012) clearly showed that structure, like composition, is related to environmental conditions in the region (Butler et al. 2002), and that there are limits to the species number, abundance, and distribution of abundance at each location (Leaper et al. 2012). When



PLATE 1. Diver completing a survey counting fishes and other marine life on a transect in southern Australia. Photo credit: G. J. Edgar.

viewed together, our analyses of community composition and the structure for this region (Leaper et al. 2012) suggest that the species seen at any one location are drawn from the dispersal pool of potential species (modeled as the archetypes that can be present at the location), up to a limit, where there is insufficient resources for more species to exist. Similar patterns of overlapping archetypes have been reported in other marine systems (e.g., inter reefs in the Great Barrier Reef Australia; Dunstan et al. 2011, 2013, Woolley et al. 2013). While we feel confident that these patterns are typical of marine systems (at least in the highly diverse, oligotrophic seas around Australia), we are unsure if they occur in more nutrient-rich seas or terrestrial systems.

Our understanding of the processes that structure communities has been enhanced by the development of new methodological tools to analyze ecological data (e.g., Foster and Dunstan 2010, Dunstan et al. 2011). Similar results could be obtained using R-mode linked to Q-mode (RQL; Dolédec et al. 1996) or fourth-corner analysis (Legendre et al. 1997, Dray and Legendre 2008) if the Q matrices (i.e., the linkages between species) were defined by species environmental correlations. However, RQL and fourth-corner are supervised analyses, meaning that the investigator must define the links between species, whereas our approach is unsupervised and relies on the ecological responses. We feel that ours is a more

robust approach and less likely to fail due to incomplete knowledge.

If there was no overlap in the spatial distribution of species archetypes, then we would expect similar results using more traditional statistical methods (i.e., MDS [multidimensional scaling], ANOSIM, PERMANOVA). However, when the archetypes overlap to a significant degree, the concordance between distance-based approaches and species-based approaches will break down. Distance-based approaches show the differences between sites, but cannot reveal how the species at each site are distributed. Approaches such as redundancy analysis (RDA) and canonical-correlation analysis (CCA) are fundamentally similar to our approach, being based on linear regression, but cannot provide the statistical grouping of species present in this analysis. They will provide an estimation of individual species distributions, but not any sort of grouping.

We argue that species-based approaches can provide important additional information about the structure and composition of communities and can be more statistically robust than distance-based approaches (see Warton et al. 2012). We found no evidence that the demersal fish and macroinvertebrate communities we sampled are strongly structured with a single community (i.e., archetype group) at each location that does not overlap with another community type. We also found that the species present at sites are not random (e.g., see Gotelli and McCabe 2002), there are groups of species

that are positively correlated, but these archetype groups can overlap, where the assemblage that is extant is determined by the probabilities of the species archetypes. We suggest that SAMs may offer new insights into the processes that structure communities. We challenge the concept of a community as a discrete entity or a random aggregation of species and suggest that the species found at a site are a realization of the species archetypes that are likely to be present at a site, up to a limit set by the environmental limit to species richness (Leaper et al. 2012). Communities are less discrete entities and more collections of groups of species that exist independently of each other and this challenges the long held intuitive view of an ecological community as single integrated unit at a locality.

ACKNOWLEDGMENTS

We dedicate this manuscript to Rebecca Leaper, who passed away before publication. She will be missed. This work has been funded through the Commonwealth Environment Research Facilities (CERF) program, an Australian Government initiative supporting world class, public-good research. The CERF Marine Biodiversity Hub is a collaborative partnership between the University of Tasmania, CSIRO Wealth from Oceans Flagship, Geoscience Australia, Australian Institute of Marine Science and Museum Victoria. R. Leaper initiated the idea of this study. N. S. Barrett and G. J. Edgar collected the field data. P. K. Dunstan and S. D. Foster developed the modeling approach. R. Leaper formatted the data and performed the modeling work. R. Leaper, N. S. Barrett, and G. J. Edgar analyzed output data. R. Leaper and P. K. Dunstan wrote the first draft of the manuscript. All authors substantially contributed to revisions.

LITERATURE CITED

- Andrew, N. 1999. Under southern seas: the ecology of Australia's rocky reefs. UNSW Press, Sydney, Australia.
- Bastow Wilson, J., and H. Gitay. 1995. Community structure and assembly rules in a dune slack: Variance in richness, guild proportionality, biomass constancy and dominance/diversity relations. *Vegetatio* 16:93–106.
- Berlow, E. 1999. Strong effects of weak interactions in ecological communities. *Nature* 398:330–334.
- Butler, A., F. Althaus, D. Furlani, and K. Ridgway. 2002. Assessment of the conservation values of the Bonney upwelling area. A component of the Commonwealth Marine Conservation Assessment Program 2002–2004. Report to Environment Australia, Canberra, Australia.
- Chase, J. M., and M. A. Leibold. 2003. Ecological niches: linking classical and contemporary approaches. University of Chicago Press, Chicago, Illinois, USA.
- Chase, J. M., and J. A. Myers. 2011. Disentangling the importance of ecological niches from stochastic processes across scales. *Philosophical Transactions of the Royal Society B* 366:2351–2363.
- Clements, F. E. 1916. Plant succession: An analysis of the development of vegetation. Carnegie Institution of Washington, Washington, D.C., USA.
- Clements, F. E. 1936. Nature and structure of the climax. *Journal of Ecology* 24:252–284.
- De'ath, G. 2002. Multivariate regression trees: a new technique for modeling species–environment relationships. *Ecology* 83: 1105–1117.
- Doledec, S., D. Chessel, C. J. F ter Braak, and S. Champely. 1996. Matching species traits to environmental variables: a new three-table ordination method. *Environmental and Ecological Statistics* 3:143–166.
- Dray, S., and P. Legendre. 2008. Testing the species traits–environment relationships: the fourth-corner problem revisited. *Ecology* 89:3400–3412.
- Dunstan, P. K., F. Althaus, A. Williams, and N. J. Bax. 2012. Characterising and predicting benthic biodiversity for conservation planning in deepwater environments. *PLoS ONE* 7(5):e36558.
- Dunstan, P. K., S. D. Foster, and R. Darnell. 2011. Model based grouping of species across environmental gradients. *Ecological Modelling* 222:955–963.
- Dunstan, P. K., S. D. Foster, F. K. C. Hui, and D. I. Warton. 2013. Finite mixture of regression modeling for high-dimensional count and biomass data in ecology. *Journal of Agricultural, Biological, and Environmental Statistics* 18: 1085–1117.
- Edgar, G. J. 1997. Australian marine life: the plants and animals of temperate waters. Reed Books, Melbourne, Australia.
- Edgar, G. J., J. Moverley, N. S. Barrett, D. Peters, and C. Reed. 1997. The conservation-related benefits of a systematic marine biological sampling program: the Tasmanian reef bioregionalisation as a case study. *Biological Conservation* 79:227–240.
- Eliot, C. 2007. Method and metaphysics in Clements's and Gleason's ecological explanations. *Studies in History and Philosophy of Biology and Biomedical Sciences* 38:85–109.
- Eliot, C. 2011. The legend of order and chaos: communities and early community ecology. Pages 49–107 in D. M. Gabbay, B. Brown, P. Thagard, K. De Laplante, K. Peacock, and J. Woods, editors. *Handbook of the philosophy of science. Volume 11. Philosophy of Ecology*. Elsevier BV, Kidlington Oxford, UK.
- Elith, J., et al. 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 29: 129–151.
- Elith, J., and J. Leathwick. 2009. Species distribution models: Ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics* 40: 677–697.
- Elith, J., S. Phillips, T. Hastie, M. Dudik, Y. Chee, and C. Yates. 2011. A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions* 17:43–57.
- Fagan, W. F., and L. E. Hurd. 1994. Hatch density variation of a generalist arthropod predator: population consequences and community impact. *Ecology* 75:2022–2032.
- Faith, D. P., P. R. Minchin, and L. Belbin. 1987. Compositional dissimilarity as a robust measure of ecological distance. *Vegetatio* 69:57–68.
- Ferrier, S., and A. Guisan. 2006. Spatial modelling of biodiversity at the community level. *Journal of Applied Ecology* 43:393–404.
- Ferrier, S., G. Manion, J. Elith, and K. Richardson. 2007. Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. *Diversity and Distributions* 13:252–264.
- Foster, S. D., and P. K. Dunstan. 2010. The analysis of biodiversity using rank abundance distributions. *Biometrics* 66:186–195.
- Friedman, J. 1991. Multivariate adaptive regression splines. *Annals of Statistics* 19:1–67.
- Gleason, H. A. 1917. The structure and development of the plant association. *Bulletin of the Torrey Botanical Club* 43: 463–481.
- Gleason, H. A. 1926. The individualistic concept of the plant association. *Bulletin of the Torrey Botanical Club* 53:7–26.
- Gomon, M. F., D. J. Bray, and R. H. Kuiter, editors. 2008. The fishes of Australia's South Coast. State Print, Adelaide, Australia.
- Gotelli, N. J., and D. J. McCabe. 2002. Species co-occurrence: a meta-analysis of J. M. Diamond's assembly rules model. *Ecology* 83:2091–2096.

- Guisan, A., A. Lehmann, S. Ferrier, M. Austin, J. Overton, R. Aspinall, and T. Hastie. 2006. Making better biogeographical predictions of species' distributions. *Journal of Applied Ecology* 43:386–392.
- Guisan, A., and W. Thuiller. 2005. Predicting species distribution: offering more than simple habitat models. *Ecology Letters* 8:993–1009.
- Hill, N. A., A. R. Pepper, M. L. Puotinen, M. G. Hughes, G. J. Edgar, N. S. Barrett, R. D. Stuart-Smith, and R. Leaper. 2010. Quantifying wave exposure in shallow temperate reef systems: applicability of fetch models for predicting algal biodiversity. *Marine Ecology Progress Series* 417:83–95.
- Huang, Z., B. Brooke, N. Whitta, A. Potter, M. Fuller, J. Dunn, and R. C. Pitcher. 2010. Australian marine physical environmental data: descriptions and metadata. *Geoscience Australia Record* 2010/32. Geoscience Australia, Canberra, Australia.
- Lawton, J. H. 1999. Are there general laws in ecology? *Oikos* 84:177–192.
- Leaper, R., P. K. Dunstan, S. D. Foster, G. J. Edgar, and N. S. Barrett. 2012. Comparing large scale bioregions and fine scale community level biodiversity predictions from subtidal rocky reefs across southeastern Australia. *Journal of Applied Ecology* 49:851–860.
- Legendre, P., and L. Legendre. 1998. *Numerical ecology. Developments in environmental modelling*. Second edition. Elsevier Science, Amsterdam, The Netherlands.
- Lessard, J. P., J. Belmaker, J. A. Myers, J. M. Chase, and C. Rahbek. 2012. Inferring local ecological processes amid species pool influences. *Trends. Ecology and Evolution* 27: 600–607.
- McCann, K., A. Hastings, and G. Huxel. 1998. Weak trophic interactions and the balance of nature. *Nature* 395:794–798.
- McCulloch, C. E., and J. A. Nelder. 1989. *Generalised linear models*. Chapman and Hall, New York, New York, USA.
- McIntosh, R. P. 1995. H. A. Gleason's individualistic concept and theory of animal communities: A continuing controversy. *Biological Reviews* 70:317–357.
- McLachlan, G., and D. Peel. 2000. *Finite mixture models*. Wiley, New York, New York, USA.
- Mills, E. L. 1969. The community concept in marine zoology, with comments on continua and instability in some marine communities: a review. *Journal of the Fisheries Research Board of Canada* 26:1415–1428.
- Paine, R. T. 1980. Food webs: linkage, interaction strength and community infrastructure. *Journal of Animal Ecology* 49: 667–685.
- Phillips, J. A. 2001. Marine macroalgal biodiversity hotspots: why is there high species richness and endemism in southern Australian marine benthic flora? *Biodiversity and Conservation* 10:1555–1577.
- R Development Core Team. 2013. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>
- Raffaelli, D. G., and S. J. Hall. 1995. Assessing the relative importance of trophic links in food webs. Pages 185–191 *in* G. Polis and K. Winemiller, editors. *Food webs: integration of patterns and dynamics*. Chapman and Hall, New York, New York, USA.
- Richardson, L. 1980. The organismic community: resilience of an embattled ecological concept. *BioScience* 30:465–471.
- Ricklefs, R. E. 2008. Disintegration of the ecological community. *American Naturalist* 172:741–750.
- Ter Braak, C. J. F., and I. C. Prentice. 1988. A theory of gradient analysis. *Advances in Ecological Research* 18:271–317.
- Vellend, M. 2010. Conceptual synthesis in community ecology. *Quarterly Review of Biology* 85:183–206.
- Warton, D. I., S. T. Wright, and Y. Wang. 2012. Distance-based multivariate analyses confound location and dispersion effects. *Methods in Ecology and Evolution* 3:89–101.
- Whittaker, R. H. 1952. A study of summer foliage insect communities in the Great Smoky Mountains. *Ecological Monographs* 22:1–44.
- Whittaker, R. H. 1967. Gradient analysis of vegetation. *Biological Reviews* 49:207–264.
- Woolley, S. N. C., A. W. McCallum, R. Wilson, T. D. O'Hara, and P. K. Dunstan. 2013. Fathom out: biogeographical subdivision across the Western Australian continental margin: a multispecies modelling approach. *Diversity and Distributions* 19:1506–1517.

SUPPLEMENTAL MATERIAL

Appendix

Tables showing fitted values and species affinity to each archetype group for fish and invertebrates. Figures show the predicted probabilities and standard errors for each group, along with a brief description of the species attributes of that group ([Ecological Archives E095-176-A1](#)).