

Congruence in demersal fish, macroinvertebrate, and macroalgal community turnover on shallow temperate reefs

RUSSELL J. THOMSON,^{1,3} NICOLE A. HILL,¹ REBECCA LEAPER,^{1,4} NICK ELLIS,² C. ROLAND PITCHER,²
NEVILLE S. BARRETT,¹ AND GRAHAM J. EDGAR¹

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

²*Commonwealth Scientific and Industrial Research Organization Wealth from Oceans Flagship, P.O. Box 2583, Brisbane, Queensland 4001 Australia*

Abstract. To support coastal planning through improved understanding of patterns of biotic and abiotic surrogacy at broad scales, we used gradient forest modeling (GFM) to analyze and predict spatial patterns of compositional turnover of demersal fishes, macroinvertebrates, and macroalgae on shallow, temperate Australian reefs. Predictive models were first developed using environmental surrogates with estimates of prediction uncertainty, and then the efficacy of the three assemblages as biosurrogates for each other was assessed.

Data from underwater visual surveys of subtidal rocky reefs were collected from the southeastern coastline of continental Australia (including South Australia and Victoria) and the northern coastline of Tasmania. These data were combined with 0.01°-resolution gridded environmental variables to develop statistical models of compositional turnover (beta diversity) using GFM. GFM extends the machine learning, ensemble tree-based method of random forests (RF), to allow the simultaneous modeling of multiple taxa. The models were used to generate predictions of compositional turnover for each of the three assemblages within unsurveyed areas across the 6600 km of coastline in the region of interest.

The most important predictor for all three assemblages was variability in sea surface temperature (measured as standard deviation from measures taken interannually). Spatial predictions of compositional turnover within unsurveyed areas across the region of interest were remarkably congruent across the three taxa. However, the greatest uncertainty in these predictions varied in location among the different assemblages. Pairwise congruency comparisons of observed and predicted turnover among the three assemblages showed that invertebrate and macroalgal biodiversity were most similar, followed by fishes and macroalgae, and lastly fishes and invertebrate biodiversity, suggesting that of the three assemblages, macroalgae would make the best biosurrogate for both invertebrate and fish compositional turnover.

Key words: Australia; biological surrogacy; coastal planning; gradient forest; marine biodiversity; prediction; random forest; species turnover; Tasmania.

INTRODUCTION

Knowledge of the spatial distribution of biodiversity is fundamental to many aspects of theoretical and applied ecology. Yet, such basic knowledge is lacking for many systems, and new tools are needed that make full use of limited biological data. For example, while some continental-scale biogeographical features, such as latitudinal gradients in species richness, are considered well defined, detailed inspection invariably indicates that taxa respond in different and nonlinear ways to such features (Kerswell 2006). The need for broadscale analytical tools that generalize gradients across large

scales is particularly acute for marine studies, given that the marine realm is out of sight and extremely expensive to survey, resulting in ecological data sets that are at best patchy and, in many regions, completely lacking. Additionally, marine conservation and spatial management that attempt to maximize representation of biodiversity targets within regional reserve networks (Ball and Possingham 2000, Margules and Pressey 2000, Ban 2009, Watts et al. 2009) rely on adequate ecological data. Clearly, delays in conservation action while we collect more ecological data are undesirable, as good environmental outcomes may be reduced due to lost management opportunities (Grantham et al. 2009). In this paper, we present a method that uses physical surrogates to predict broadscale spatial patterns of compositional turnover (McArthur et al. 2010). We then use the observed and predicted spatial patterns of compositional turnover to assess the efficacy of biological surrogacy.

Manuscript received 10 September 2012; revised 11 June 2013; accepted 2 July 2013. Corresponding Editor: S. S. Heppell.

³ E-mail: russell.thomson@utas.edu.au

⁴ Deceased.

Both physical (McArthur et al. 2010, Huang et al. 2011, Dunstan et al. 2012, Pitcher et al. 2012) and biological (Ward et al. 1999) surrogates of marine biodiversity have increasingly been applied to fill the ecological data gap in conservation planning. However, marine biodiversity mapping has largely focused on predictions made from the distribution of coarse categorical habitat features (Ward et al. 1999, Jordan et al. 2005, Mumby et al. 2008) such as seagrass beds, rocky reef, coral reef, sand, and mud. A much finer and continuous biodiversity grain is needed for conservation planning because quite different ecological communities can occur within the same mapped habitat categories, depending on variation in such factors as wave exposure, depth, light penetration, or habitat structural complexity (Lindsay et al. 2008). Unless the full range of communities or their surrogates (biological and/or environmental) are considered during planning, complete sets of species will potentially be excluded from protected area networks.

In a recent meta-analysis on the use of biosurrogates for predicting marine biodiversity (Mellin et al. 2011), higher-taxa or subset-taxa surrogates performed better than cross-taxa surrogates. Yet in conservation management, cross-taxa biosurrogacy is frequently applied. In this study, we predict the spatial patterns of compositional turnover among three assemblages (fishes, invertebrates, and macroalgae), based on physical surrogates. By comparing the predictions of the three assemblages, we examine the utility of applying predicted turnover of one of these assemblages for the conservation management of the other assemblages. In other words, we evaluate the effectiveness of biosurrogacy in the context of broadscale prediction of biodiversity. To our knowledge, this is the first study to examine the effectiveness of community-based predictions of marine biodiversity for biosurrogacy at such large scales.

While there have been many predictive biodiversity studies based on physical surrogates, to date the majority have tended to focus on a single species of interest (Guisan and Thuiller 2005) or elements of alpha diversity, such as species richness and abundance (e.g., Leathwick et al. 2006, Mellin et al. 2010). Methods that incorporate multivariate information regarding the composition of communities have received less attention (but see Ferrier et al. [2007]), yet spatial variation in community composition (or beta diversity sensu Whitaker 1972) lies at the heart of many biological and ecological phenomena. Because beta diversity can quantify the turnover of species across space, it has important applications to the scaling of diversity, the delineation of biotic regions, and conservation planning (McKnight et al. 2007). Here we use gradient forest modeling (Ellis et al. 2011a) to map spatial patterns of compositional turnover using environmental covariates across broad scales. Specifically, we map biodiversity in the form of compositional turnover in n -dimensional

space (where n is the number of environmental covariates modeled). Throughout this paper, the term biodiversity refers to compositional turnover.

Gradient forest modeling (GFM) is a tree-based method that is based on random forest (RF) models (Breiman 2001). RF models typically perform equivalent to, or in most cases surpass, other statistical and machine learning methods in comparisons relating ecological with covariate data (e.g., Keller et al. 2011). Key features of RF modeling is that as a classification/regression tree method, nonlinearities and interactions among predictor variables are considered by default and that it copes with highly correlated predictor variables (Strobl et al. 2008). GFM extends RF modeling to incorporate whole assemblages instead of single species. It combines the outputs of single-species models to relate changes in the composition of assemblages to gradients in environmental covariates (Ellis et al. 2011a). The results of GFM are species turnover curves for each environmental covariate. These curves can be used to predict biodiversity in new locations based on the environmental covariates at those locations (Pitcher et al. 2012).

We model spatial patterns of turnover in community composition using GFM for fish, invertebrate, and macroalgal species records using 314 survey sites from an extensive spatiotemporal quantitative survey of a rocky subtidal system on the southern coastline of continental Australia and the northern coastline of Tasmania (Stuart-Smith et al. 2008, Barrett et al. 2009, Leaper et al. 2011). Based on these models, we predict biodiversity at unsurveyed locations across ~6000 km of coastline in southeastern Australia that is of interest biogeographically and hence also to management (Commonwealth of Australia 2006). This area is a hotspot of biodiversity both in terms of the number of benthic species present and the extraordinarily high levels of regional floral endemism (Phillips 2001). Our broadscale study includes three state government jurisdictions: (1) Victoria, where a network of Marine Protected Areas (MPAs) currently exists; (2) South Australia, where an MPA network has been proposed and is presently subject to public consultation; and (3) Tasmania, where a single MPA has been designated within the study region (the Kent Group Marine Park) with the commitment to developing a larger network within the National Representative System of MPAs in the future.

Specifically, we (1) compare the biodiversity predictions for fishes, invertebrates, and macroalgae to assess biological surrogacy across ~6000 km of coastline; (2) assess and quantify the influence of the environmental covariates on biological assemblages and their turnover using the 314 survey sites; and (3) quantify the uncertainty of the predictions of biodiversity by comparing our predictions with the survey data. Quantifying uncertainty is rarely conducted in community modeling.

METHODS

Biological data

Estimates of the abundance of fishes and invertebrates, and the cover of macroalgae were obtained for 314 sites along the South Australian, Victorian, and northern Tasmanian coastline (Fig. 1) using diver-based, underwater visual censuses (Edgar et al. 1997, Barrett et al. 2007, 2009). Surveys were conducted between 2004 and 2009, and each site consisted of a 200-m transect at either 5 or 10 m water depth. Fishes were censused by counting all individuals sighted to a distance of 5 m either side of the transect line (i.e., sampled area $\approx 2000 \text{ m}^2$). Large, mobile invertebrates (echinoderms other than ophiuroids, gastropods, and crustaceans) that were $>2.5 \text{ cm}$ length were counted on a 1-m swath along one side of the transect line (i.e., sampled area $\sim 200 \text{ m}^2$), and macroalgal cover for each taxon was recorded in situ by divers using 20 quadrats of $0.5 \times 0.5 \text{ m}$ placed equidistantly along the transect length. The cover of canopy macroalgal species was recorded initially, and then the canopy pushed aside so that understory species could also be scored. Data were recorded at the species level where possible, with some invertebrate and macroalgal taxa recorded at the genus level and occasionally (for macroalgae) as a broad functional group (e.g., encrusting coralline algae). While these biotic data sets are considered relatively robust to seasonal and interannual variation (Stuart-Smith et al. 2008, Barrett et al. 2009), such variation was minimized as much as possible by using data collected from the most recent survey that was conducted at each survey location and within the period from March to June.

Environmental data

Nineteen environmental covariates describing salinity, temperature, nutrient concentrations, productivity, and exposure were used in the present study. These covariates have been generated across a national scale at a 0.01° resolution as part of the Commonwealth of Australia's Environment Research Facility (CERF) Marine Biodiversity Hub (Huang et al. 2010; data available online).⁵ A detailed description of each covariate, its source, and derivation can be found in Appendix A.

Statistical modeling

Gradient forest modeling (GFM) is based on a "machine learning" approach known as random forest (RF) modeling (Breiman 2001). This ensemble method utilizes regression trees, combining many decision trees to produce a distribution of splits rather than a point estimate. The cumulative distribution of splits, or cumulative importance curves, for a given predictor provides a gradient of biological change, where the

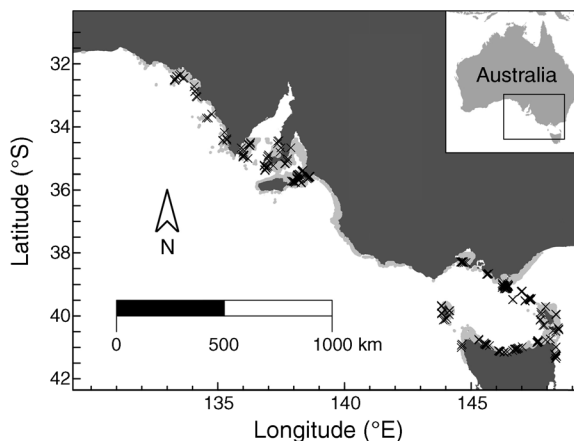


FIG. 1. A map showing $\sim 6600 \text{ km}$ of the southeastern coastline of continental Australia for which compositional turnover of fish, invertebrate, and macroalgal assemblages has been predicted. The 314 survey sites are marked with crosses. The predicted region, made up of 24 817 grid points, is shown using a light gray color.

steeper the gradient, the greater the change in composition (see Fig. 3 and Appendix C for the cumulative importance curves related to this study). These curves provide a measure of compositional turnover in n -dimensional space, where n is the number of predictors used.

While the cumulative importance curves are not the same as traditional measures of turnover (such as Bray-Curtis dissimilarity), in a paper (Leaper et al. 2011) comparing the GFM predictions with predictions based on generalized dissimilarity modeling (Ferrier et al. 2007) there was little difference.

In the present study, a total of 1000 trees were generated for each taxon. The RF technique contains both bootstrapping and cross-validation routines. Each tree is generated using a random partitioning procedure based on a subset of the survey sites (the "in-bag," $\sim 63\%$ by default), and each split is selected from a random subset of six of the 19 predictors. Cross-validation is conducted using the remaining "out-of-bag" data, allowing estimation of prediction performance for each taxon (R^2). Calculation of the gradients is weighted by both taxon importance (R^2) and predictor importance. Predictor importance is estimated by the degradation of performance when randomly permuting each predictor. The standard marginal procedure in the RF methodology can suffer from bias toward correlated predictors (Strobl et al. 2008). Hence, conditional importance of the predictors was used (Ellis et al. 2011a). The gradient forest (GF) methodology adopted for this paper is described in more detail in Ellis et al. (2011a). The analysis was carried out with the "gradientForest" and "extendedForest" packages for R (available online).⁶

⁵ <http://www.nerpmarine.edu.au>

⁶ <https://r-forge.r-project.org/projects/gradientforest>

We used a total of 19 candidate environmental covariates for modeling (see Appendix A). We modeled compositional turnover for fish density, invertebrate density, and macroalgal cover separately. Taxa not recorded at five or more of the 314 survey sites were removed from the analysis. A Box-Cox method was used to select an appropriate transformation for density data for all taxa. The majority of the taxa required an inverse square transformation with a positive offset of one.

For all analyses presented in this paper, the geographic covariates, latitude and longitude, were not included as predictors in the GFM model. When they were included, the predictive ability of the model was improved by a small amount (0.022%, 1.3%, and 1.1% for fishes, invertebrates, and macroalgae, respectively).

Predictions of assemblage diversity

Predictions of the compositional turnover of fishes, invertebrates, and macroalgae were generated using the procedure described in Pitcher et al. (2012). Predictions were made at new unsurveyed locations across 24 817 grid points (0.01° resolution) along 5500 km length of coastline of southern Australia and 1100 km of northern Tasmania (including the islands in the Bass Strait; Fig. 1). The environmental covariates at the grid points were transformed using the cumulative importance curves for each corresponding covariate as empirical functions, resulting in a prediction space consisting of biologically transformed environmental variables (or predicted turnover in assemblage composition associated with each covariate).

Principal Components Analysis (PCA) was used to reduce the dimensions of the 19 biologically transformed environmental variables. The PCA was centered but not scale transformed (to preserve biological importance of the variables). For each of the three assemblages, the difference in assemblage composition between grid points was mapped using the first three principal components (PCs).

Uncertainty in the predictions

To ascertain uncertainty, the procedure described in Ellis and Pitcher (2011) was used. Uncertainty was obtained by calculating the difference between the predicted data and the actual data at the 314 survey sites. Uncertainty at a new unsurveyed location was estimated from a weighted average of these differences.

Specifically, gradient predictions were compared to the observed taxonomic composition at the survey sites. As the predictions were in the form of a transformed 19-dimensional environment space representing compositional variation, the observed taxonomic composition at the survey sites was converted to dissimilarity measures using Bray-Curtis. These Bray-Curtis dissimilarities were then converted into a form that allowed direct comparison at individual sites using ordination via nonmetric multidimensional scaling (MDS), also with 19 dimensions. Uncertainty was measured via the difference between the predicted compositional turnover

and the MDS ordination of the taxonomic composition. To make this comparison, the predicted compositional turnover and the MDS ordinated taxonomic composition were compared using a Procrustes superimposition (where the matrices are rotated to minimize the sum of square of the distances between the points; Peres-Neto and Jackson 2001). The differences between the rotated matrices (Procrustes residuals) were then used as a measure of uncertainty at the survey sites.

Uncertainty at the unsurveyed locations was estimated using kernel smoothing of the Procrustes residuals along the 19-dimensional predicted biological space:

$$P_j = \sum_{i=1}^{314} P_i e^{-(1/2)(d_{i,j}/w)^2} / \sum_{i=1}^{314} e^{-(1/2)(d_{i,j}/w)^2}. \quad (1)$$

For Eq. 1, P_i is the Procrustes residual at the 314 survey sites, P_j is the estimated Procrustes residual at $j = 1, \dots, 24817$ grid points, $d_{i,j}$ is the predicted biological distance between the i th survey site and the j th grid point, and w is the weight. Leave-one-out cross-validation at the survey sites was used to obtain an optimal value of $w = 0.005$ (see Appendix B for more detail).

Procrustes correlations, $r_p = \sqrt{1 - SS}$ were calculated for each of the three assemblages, where SS is the symmetric sum of squares of the Procrustes residuals between the MDS ordinated taxonomic composition and the predictions at the survey sites. A permutation test with 1000 permutations was used to test the significance of these correlations (as per Peres-Neto and Jackson [2001]).

Comparing GF-predicted biodiversity among the three biological assemblages

In order to ascertain how well fishes, invertebrates, and macroalgae would act as biological surrogates for each other, a Procrustes rotation was carried out between the predicted turnover of each pair of the three assemblages at the 24817 grid points (Peres-Neto and Jackson 2001). Only one of each pair of assemblage matrices was rotated so that the rotation was not dependent on the order of the matrices (scale = TRUE option for the Procrustes R function). The Procrustes residuals were used to obtain a location-specific measure of relative similarity/dissimilarity or difference between the assemblages. The Procrustes residuals were divided by the maximum Euclidean distance between predicted biodiversity for all pairs of grid points over the three assemblages, so that the measure of difference is in units of maximum possible difference.

Matrix correlations (or Mantel correlations; Legendre and Legendre 1998) were calculated for each pair of assemblages, based on the Euclidean distances between the predicted biodiversity at the 24817 grid points. 95% confidence intervals for the correlations were estimated using bootstrapping without replacement (Politis and

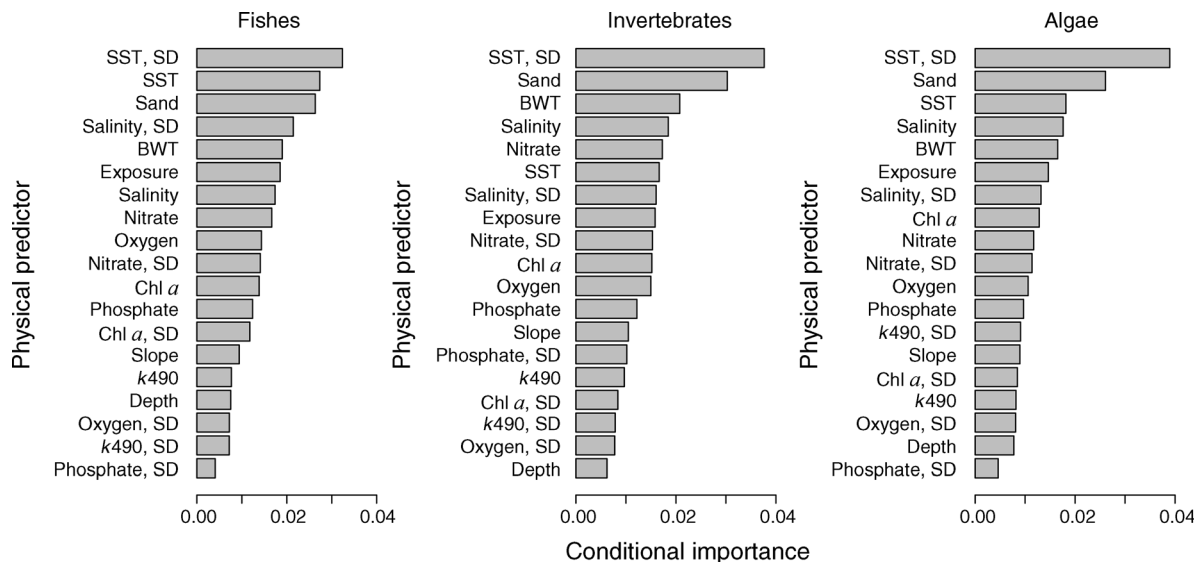


FIG. 2. The importance of the physical predictors for the three different assemblages in the gradient forest model. Conditional importance is estimated by the degradation of performance when randomly permuting each predictor. See Appendix A for a detailed description of each of the physical predictors. Key to abbreviations: SST, sea surface temperature mean; SD, standard deviation; BWT, bottom water temperature; $k490$, mean diffuse attenuation coefficient at wavelength 490 nm; chl a , chlorophyll a .

Romano 1994). Confidence intervals were scaled using the assumption that

$$SE(r_M) \approx SE(r_M^{(sub)}) \sqrt{\frac{b-2}{n-2}}$$

where $r_M^{(sub)}$ is the Mantel correlation of a subset of size b and n is the number of grid points.

Comparing the observed biodiversity among the three biological assemblages

The apparent biosurrogacy measured using the predicted biology has possibly been mediated by the environmental relationships, given that the predicted biology is based on transformed environmental variables. Biosurrogacy was also assessed using the surveyed biology at the 314 survey sites. Biosurrogacy was measured via Mantel correlations (and 95% confidence intervals) between the three pairs of assemblages, based on the Bray-Curtis dissimilarities between the sites.

RESULTS

Totals of 123 fish, 118 invertebrate, and 184 macroalgal taxa were observed at the 314 survey sites. The mean (range) of richness across sites was 15.9 (2–34) for fishes, 12.9 (2–28) for invertebrates, and 19.5 (2–42) for macroalgae. Of all the taxa analyzed, 61 fishes, 59 invertebrates, and 112 macroalgae showed a relationship with the predictors (taxa importance > 0).

Models of biodiversity

The order of importance of the physical predictors in the gradient forest modeling (GFM) was relatively

consistent across fishes, invertebrates, and macroalgae (Fig. 2). The standard deviation of sea surface temperature (SST, SD) was the most important physical predictor of biodiversity for all three assemblages. The conditional importance of SST, SD relative to the other variables was greatest for macroalgae (the percentage of the total conditional importance that was attributed to SST, SD was 11.2%, 12.9%, and 15.2% for fishes, invertebrates, and macroalgae, respectively). Percentage sand and mean sea surface temperature were also consistently important predictors, while nitrate, wave exposure, sea surface and bottom water temperature, salinity, and the standard deviation of salinity were consistently among the top 10 predictors of compositional turnover for all assemblages. Bottom water temperature was more important than sea surface temperature for invertebrates, while for fishes and macroalgae, sea surface temperature was more important.

Cumulative importance curves for each of the physical predictors are shown in Fig. 3 and Appendix C for each assemblage. They represent the rate at which the community composition changes across the range of each environmental covariate. The gradients are remarkably congruent among the three assemblages, with the invertebrate and macroalgal gradients appearing more similar to each other than to fish. From the slope of these cumulative importance curves, regions with high turnover can be identified. For example, fish gradients were noticeably steeper when SST, SD was $\sim 3.0^\circ\text{C}$ and SST $\sim 17^\circ\text{C}$ (indicating greater change in community composition).

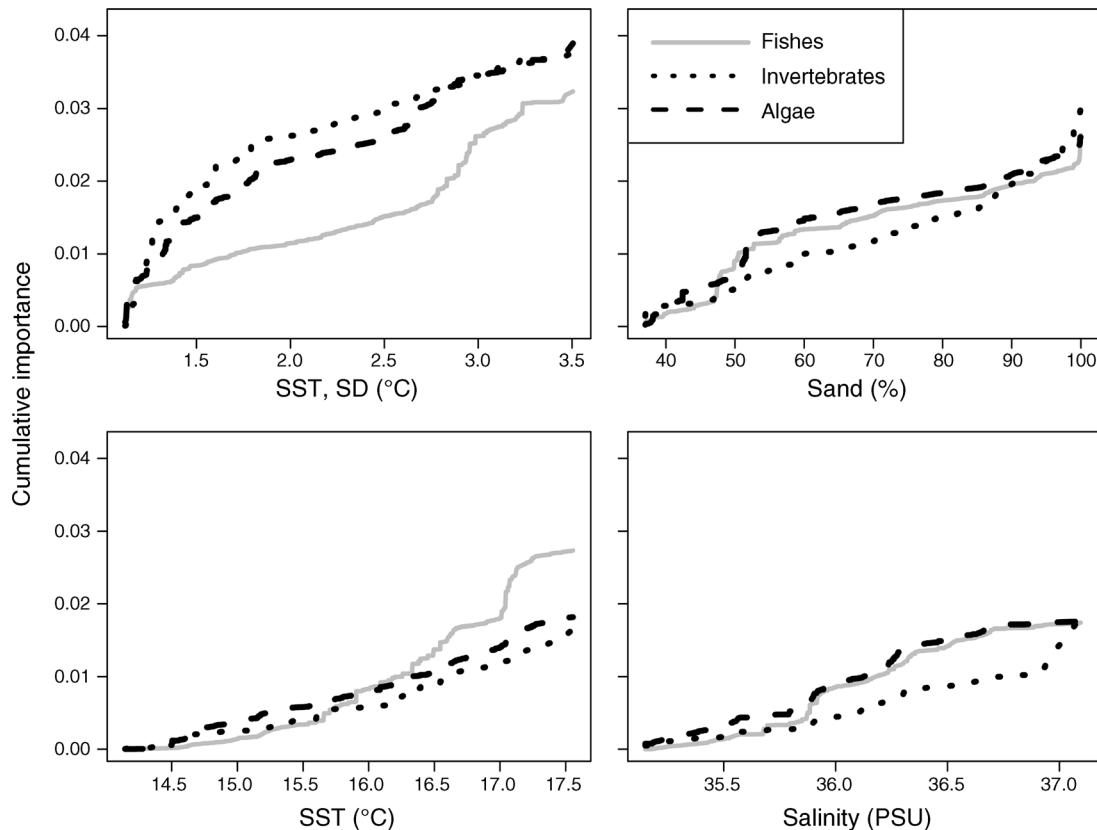


FIG. 3. Cumulative importance curves for four predictor variables (SST, sea surface temperature mean and SD; percentage sand; and salinity). These curves are obtained from the cumulative distribution of tree splits, divided by the density of the predictor data, and scaled so that the maximum cumulative importance is equal to the importance of that predictor (Fig. 1). These curves relate the rate of cumulative compositional change to the environmental gradients (shallow slopes indicate low rates of change in composition, whereas steep slopes indicate high rates) and show contrasting compositional responses along these gradients among the different biological groups.

Predicted patterns of biodiversity

Predictions from the GFM showed a distinctive east–west gradient in compositional turnover of fish, invertebrate, and macroalgal assemblages along the southern Australian coastline (see panels A1, B1, and C1 in Fig. 4). However, changes in composition of all three assemblages were not uniform along the southern Australian coastline. Most noticeably, sheltered areas along the coast stood out as biologically distinct from the rest of the coastline and from each other (i.e., South Australian gulfs [136° E < longitude < 139° E] and Victorian embayments [145° E < longitude < 147° E]).

Panels A3, B3, and C3 in Fig. 4 provide uncertainty associated with predictions of community turnover for fishes, invertebrates, and macroalgae, respectively, as estimated by comparing the predicted data with the actual data. For all assemblages, relatively higher uncertainty in the predictions of compositional turnover was observed in gulfs and bays relative to the open coast. Patterns in uncertainty were not, however, completely consistent across the three assemblages. For example, fish predictions were noticeably more

reliable to the west of 138° longitude (Fig. 4, panel A3), while invertebrates and macroalgae predictions were most reliable on the central Victorian and north Tasmanian coastlines (Fig. 4, panels B3 and C3, respectively). Overall, the predictions were shown to be considerably better than random, as the predictive accuracy (as measured by the correlation between the biological data and the predictions) was highly significant at the original survey sites ($P \leq 0.001$ for all assemblages). The predictive accuracy was lowest for invertebrates, suggesting that of the three assemblages, invertebrate compositional turnover was predicted least successfully ($r_p = 0.54$ for fishes, $r_p = 0.50$ for invertebrates, and $r_p = 0.55$ for macroalgae).

Congruence in GF-predicted biodiversity among the three assemblages

The predicted patterns of compositional turnover are remarkably congruent among fishes, invertebrates, and macroalgae (Fig. 5A–C), a conclusion supported by the high correlation among the predicted compositional turnover of the three assemblages (r_M [95% CI of r_M]: 0.846 [0.844–0.849] for fishes vs. invertebrates, 0.919

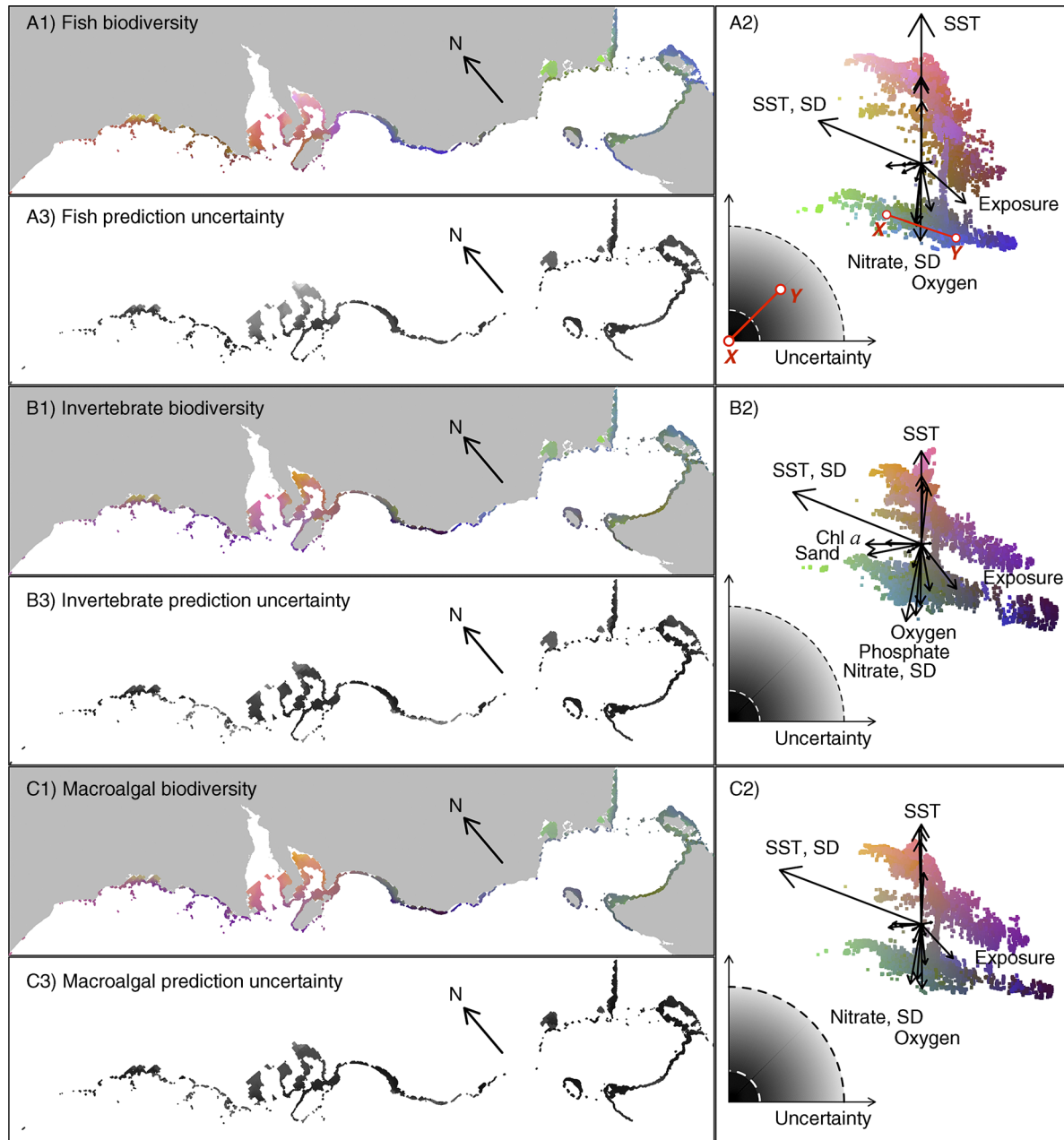


FIG. 4. Biodiversity maps showing patterns of composition for the (A1) fish, (B1) invertebrate, and (C1) macroalgal assemblages. These are developed from a principal components analysis (PCA) of the gradient forest predictions of compositional turnover. Geographical sites with similar colors are predicted to have similar composition. The corresponding biplots from the PCA are given for each of the biodiversity maps for (A2) fish, (B2) invertebrate, and (C2) macroalgal assemblages. These describe how the colors lie within predicted biological space. Colors were assigned by varying the green, red, and blue palettes, corresponding to variation in principal components 1, 2, and 3, respectively. The arrows on the biplots indicate the correlation between the environmental predictors and the principal components. The maps below each biodiversity map show the uncertainty of the predictions (A3) for fishes, (B3) for invertebrates, and (C3) for macroalgae, using a grayscale where the lightest sites are the most uncertain. This uncertainty was obtained from calculating the distance between the predicted biological space and biological data using a Procrustes residual. The key on the biplots shows the relationship between a shade of gray and a distance on the biplot, as illustrated by the x - y segments in panel (A2). The distances are the same in the three panels A2, B2, and C2. The dashed quarter circles on the key indicate the minimum and maximum uncertainty that was observed.

[0.918–0.921] for fishes vs. macroalgae, and 0.9590 [0.9587–0.9593] for invertebrates vs. macroalgae). Invertebrates and macroalgae were more highly correlated with each other than they were with fishes, and fishes were more highly correlated with macroalgae than invertebrates, suggesting that, based on the congruency criterion, macroalgae are the best surrogate for both fishes and invertebrates.

The patterns of correlation between the different assemblages were largely consistent when considered across the full geographical range of the prediction space (Fig. 5; Appendix D). Small exceptions to this are observed in the east (longitude > 146.5° E, where fishes are correlated equally with invertebrates and macroalgae), and along the west coast of Tasmania (~145° E, >40.5° S), where macroalgae are correlated equally with invertebrates and fishes).

*Congruence in observed biodiversity
among the three assemblages*

The Mantel correlations (r_M) for the biological data collected at the 314 survey sites were lower than for the predicted biodiversity: 0.350 (95% CI, 0.293–0.407) for fishes vs. invertebrates, 0.459 (95% CI, 0.410–0.511) for fishes vs. macroalgae, and 0.443 (95% CI, 0.397–0.492) for invertebrates vs. macroalgae. The confidence intervals suggest that the correlation between fishes and invertebrates is significantly smaller than for both fishes and macroalgae and invertebrates and macroalgae, again suggesting that, based on the congruency criterion, macroalgae are the best surrogate for both fishes and invertebrates.

DISCUSSION

While conservation management of communities often assumes strong biological surrogacy (Rodrigues and Brooks 2007), our work represents one of the few marine studies where data are available to test this assumption across broad spatial scales (Beger et al. 2007, O'Hara 2008, Williams et al. 2010, Mellin et al. 2011, Sutcliffe et al. 2012). We have used a novel statistical technique, gradient forest modeling (GFM), to predict patterns of compositional turnover of fishes, invertebrates, and macroalgal assemblages across a substantial section of southern Australian coastline. In addition to showing a moderate degree of congruency in the compositional turnover of demersal fishes, invertebrates, and macroalgae at survey sites, our results show high levels of congruency in predicted patterns, providing empirical support for the efficacy of biological surrogacy in temperate reef systems.

Our work supports the use of biological surrogates for describing and predicting biodiversity patterns at broad scales; however, previous studies have reported mixed success. Consequently, the effectiveness of biological surrogates for a range of purposes has been the subject of some debate (e.g., Sætersdal and Gjerde 2011). Results in the marine realm to date indicate that

biological surrogacy has tended to be more effective when using higher-taxon surrogates at small spatial scales (<10 km), and in relatively homogenous habitats such as soft sediments (Mellin et al. 2011, Sutcliffe et al. 2012). There are, however, two key differences between our study and previous biological surrogacy work in marine ecosystems. First, we examined the efficacy of biological surrogacy for measuring compositional turnover rather than the more traditional measures of species richness. Second, in addition to comparing congruency patterns between assemblages in point-location biodiversity data (as presented in many marine studies to date; Beger et al. 2003, Karakassis et al. 2006, Sutcliffe et al. 2012), we used point-location data combined with environmental data to model and predict spatially explicit turnover patterns. The consistency of the relationships between compositional turnover and environmental covariates for the fish, invertebrate, and macroalgal communities, and the resulting congruency in predicted biodiversity patterns, suggest that where the goal is broadscale prediction of biodiversity, data collection on only one taxonomic group will provide good representation of the biodiversity of other taxa. To confirm this, further studies are needed that compare broadscale prediction of biodiversity for different taxa, also using a complementarity criterion, and with different sites for each taxon to build the prediction model. If broadscale congruency is confirmed, then this has significant implications for the design of future field surveys.

Although overall the congruency of predicted patterns of biodiversity between taxa was high, correlations suggest that some pairs of taxonomic groups performed slightly better as surrogates for each other. For example, macroalgal and invertebrate diversity patterns were the most highly correlated, and fish diversity patterns were better correlated with macroalgae than with invertebrates. Furthermore, these correlations were consistent within smaller geographic units (bioregions; Appendix D), increasing our confidence in the robustness of the surrogate relationships. For a surrogate to be effective, it must also be easy and cost effective to survey and identify (Mellin et al. 2011). Our results indicate that differences in the congruency of predicted biodiversity patterns were small among the three assemblages; consequently, considerations such as the availability of skilled observers for any given taxon group may override other considerations when deciding the most cost-effective form of data collection.

The overall congruency of patterns of biodiversity at the survey sites was not as high as the predicted patterns. This is not surprising since the predictions were based on the same environmental variables, measured at the same survey sites for each of the three assemblages. Conclusions made about the need for collecting biological data for only one assemblage, when the goal is broadscale prediction of biodiversity, are specific to the circumstances where data for building the prediction models

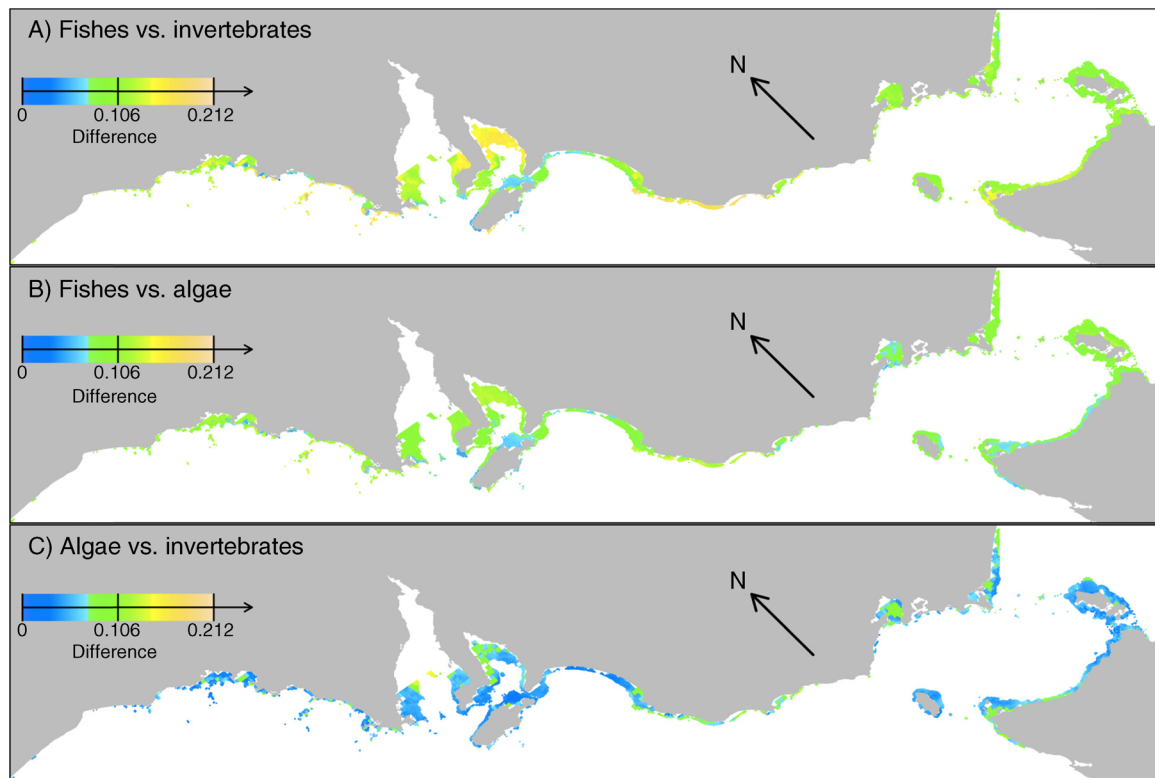


FIG. 5. Maps showing the difference between predicted compositional turnover for the three pairs of assemblages (invertebrates vs. algae, fishes vs. algae, and fishes vs. invertebrates) across the whole prediction space, using the Procrustes residuals as the measure of difference. The difference measure is in units of maximum possible difference (as defined by the greatest distance between two points in the three biplots shown in Fig. 4).

for each of the three assemblages would be collected at the same survey sites. However, the congruency of patterns of biodiversity at the survey sites was still reasonably high, indicating the effectiveness of biosurrogacy in other circumstances. Based on both the predicted and observed patterns of biodiversity, out of the three assemblages examined, macroalgae are the best surrogate for the others.

The ability of macroalgae to act as a suitable surrogate for both fishes and invertebrates probably has a basis in tighter ecological relationships between macroalgae and the other two groups than between fishes and invertebrates. Invertebrates use macroalgae as habitat and commonly also as a food resource, while most reef fishes consistently respond to macroalgae as habitat, although they typically feed opportunistically on a range of invertebrate prey (Davenport and Bax 2002).

High congruency was also observed among the three assemblages in terms of the environmental variables that best predicted compositional turnover. The standard deviation of sea surface temperature was the best predictor of changes in reef communities for all three assemblages. A causal relationship between variability in sea surface temperature and reef communities is ecologically reasonable (e.g., see Perry et al. [2005],

Helmuth et al. [2006]), perhaps through interaction with oxygen stress (Poertner and Knust 2007) or vulnerability to predation through changing patterns of behavior (Biro et al. 2007). Environmental control of marine populations is more likely to depend on the occurrence of extreme temperature events than on mean temperature values (e.g., see Harley [2008]). Such congruency between ecological theory and covariate importance provides some confidence for planning purposes that underlying ecosystem variation is largely captured through the GFM process.

Sites showing the highest variability in sea surface temperature occur in shallow embayments, locations with extreme diel and seasonal fluctuations in temperature, whereas sites with the lowest variability in sea surface temperature are located along the open coast in regions where cold-water upwelling occurs in summer. Assuming that species with low physiological tolerances to temperature fluctuations are most susceptible to impacts of climate change, as is hypothesized for terrestrial organisms (Deutsch et al. 2011), then species that are adapted to open coast are arguably more at risk through climate change, including the anomalously rapid ocean warming currently occurring across the southeastern Australian region (Poloczanska et al. 2007, Wernberg et al. 2011), than species assemblages

restricted to sheltered embayments. Alternatively, the effects of climate change may be more pronounced in shallow embayments as they may not receive the dampening effect of deeper layers. Our results show a strong relationship between community turnover and the standard deviation of sea surface temperature for all three assemblages, particularly macroalgae, indicating distinct differences in community composition between open water and shallow embayment environments. Searching for macroalgal species (and others) that have the strongest (positive or negative) relationship with the standard deviation in sea surface temperature may elucidate specific species that are potentially vulnerable to climate change.

We also found that sediment composition (covariate “sand”) had consistently high relative importance in predicting compositional turnover for all three assemblages. This predictive importance is presumably due to dependencies on unexamined colinear factors, given that ecological surveys were undertaken on reef systems, whereas values for sand were compiled as the mean percentage of sand-size particles in sediments within the nearest $0.01^\circ \times 0.01^\circ$ grid cell. As is evident from the distribution of GFM breakpoints across the range of sand values (Fig. 2), the primary ecological change with sand occurs at very low values, indicative of a basic environmental dichotomy between grid cells with sandy sediments and those dominated by finer silts and clays or coarse gravels. The close coupling between sediment characteristics and the reef community possibly relates to changes in water flow that affect sediments and also biota, or perhaps translocation of materials or organisms between reef and soft-sediment habitats.

The strong relationships evident between sand and reef communities highlight the important point that the nineteen environmental covariates used in the GFM are not necessarily the causal environmental variables. Rather, they are estimates derived from available data that, in an often unknown way, relate to truly causal environmental variables. By including all available variables, we were able to improve the predictions of biodiversity (without identifying the unknown causal variables). The advantage of using a random forest model over traditional regression techniques is that many variables, regardless of colinearity, can be included in the model in order to improve the predictive outputs.

The importance of the physical surrogate “sand” and the biological surrogate “macroalgae” indicate that measures of habitat are good predictors of biodiversity, a result found by others (e.g., Mumby et al. 2008) using a variety of prediction methods.

The environmental variables examined have been applied in a range of research projects aimed at addressing questions of biological surrogacy and prediction of Australia’s marine biota (Pitcher et al. 2011). Predictions of biodiversity described here were recorded at a grid scale of $0.01^\circ \times 0.01^\circ$ so they will not capture

species turnover at scales $<1 \text{ km}^2$. In addition, observational errors were added through biases in field sampling. Some sampling errors may have varied between sites, such as interactions between fish counts, water clarity, and substratum type. However, in a study based on the same survey methodology (Edgar et al. 2004), no interaction between observational bias and different habitat types was observed, and variation in density estimates between divers was extremely low compared to variation between sites and between months, indicating that observational error is dwarfed by the uncertainty in prediction. Fig. 4 displays relationships between the environment and prediction error.

By incorporating multivariate biological data, we were able to model the relationships between environmental variables and community composition, rather than single species or alpha diversity measures. However, the less common species did not show any relationship between the environmental variables and community composition (most likely due to lack of power). For the purposes of management of less common species, an approach that targets rare species is needed.

Maps of biodiversity predictions are often presented without robust estimates of their reliability. While univariate measures of marine diversity have been predicted previously using tree-based classification methods (e.g., Leathwick et al. 2006), the methods used here represent one of the first applications of a tree-based method for predicting the distribution of marine reef communities with estimates of uncertainty (Ellis and Pitcher 2011). Distinct patterns of uncertainty were observed in all three assemblages, although Port Phillip Bay ($\sim 38^\circ \text{ S}$, $\sim 145^\circ \text{ W}$) was found to have consistently high uncertainty. This may be due to human impact in this region (Stuart-Smith et al. 2008) and a lack of availability of environmental variables that relate directly to human impacts. However, relationships between human impacts and species turnover were probably modeled indirectly in part through correlated variables. For example, wave exposure is likely to be correlated with port influences. Inclusion in future analyses of additional data sets that describe variability in pollution (Fabricius et al. 2005, Hewitt et al. 2005), port influences (Hewitt 2002, Stuart-Smith et al. 2008), fisheries management restrictions (including locations of Marine Protected Areas; Barrett et al. 2009), trawl effort (Ellis et al. 2011b), and indicators of introduced species should prove extremely interesting and topical.

CONCLUSION

Gradient forest modeling is a useful tool for predicting patterns of turnover in assemblage composition across broad geographical scales and in providing uncertainty for these predictions. The congruency of predictions between fishes, invertebrates, and macroalgae (in the importance of environmental covariates

and in terms of the predictions themselves) suggests that outcomes do not differ greatly when based on any of these three assemblages. Of the three assemblages, macroalgae would make the best surrogates for both invertebrate and fish compositional turnover. This taxonomic group is more similar to the other assemblage types than the other assemblage types are with each other. Outcomes described here have a broadscale ecological basis and thus should be relevant to other temperate reef systems.



PLATE 1. While this paper was in submission our coauthor, Rebecca Leaper, passed away. She was an intelligent and enthusiastic scientist and will be greatly missed. Photo credit: Kate Blokker.

ACKNOWLEDGMENTS

This work has been funded through the National Environmental Research Program (NERP) and the Australian Research Council. The NERP Marine Biodiversity Hub is a collaborative partnership among the University of Tasmania, CSIRO Wealth from Oceans Flagship, Geoscience Australia, the Australian Institute of Marine Science, Museum Victoria, the University of Western Australia, and Charles Darwin University. We thank Piers Dunstan and Scott Foster for statistical advice; Bryan McDonald (South Australia), Stefan Howel (Victoria), and Peter Mooney and Colin Buxton (Tasmania) for facilitating field surveys; and Peter Walsh and Carolina Zagal for database development.

LITERATURE CITED

- Ball, I. R., and H. P. Possingham. 2000. MARXAN V1.2: marine reserve design using spatially explicit annealing. <http://www.ecology.uq.edu.au/marxan.htm>
- Ban, N. C. 2009. Minimum data requirements for designing a set of marine protected areas, using commonly available abiotic and biotic datasets. *Biodiversity and Conservation* 18:1829–1845.
- Barrett, N. S., C. D. Buxton, and G. J. Edgar. 2009. Changes in invertebrate and macroalgal populations in Tasmanian marine reserves in the decade following protection. *Journal of Experimental Marine Biology and Ecology* 370:104–119.
- Barrett, N. S., G. J. Edgar, C. D. Buxton, and M. Haddon. 2007. Changes in fish assemblages following 10 years of protection in Tasmanian marine protected areas. *Journal of Experimental Marine Biology and Ecology* 345:141–157.
- Beger, M., G. P. Jones, and P. L. Munday. 2003. Conservation of coral reef biodiversity: a comparison of reserve selection procedures for corals and fishes. *Biological Conservation* 111:53–62.
- Beger, M., S. A. McKenna, and H. P. Possingham. 2007. Effectiveness of surrogate taxa in the design of coral reef reserve systems in the Indo-Pacific. *Conservation Biology* 21:1584–1593.
- Biro, P. A., J. R. Post, and D. J. Booth. 2007. Mechanisms for climate-induced mortality of fish populations in whole-lake experiments. *Proceedings of the National Academy of Sciences USA* 104:9715–9719.
- Breiman, L. 2001. Random forests. *Machine Learning* 45:5–32.
- Commonwealth of Australia. 2006. A guide to the integrated marine and coastal regionalisation of Australia. Version 4.0. Department of the Environment and Heritage, Canberra, Australia.
- Davenport, S. R., and N. J. Bax. 2002. A trophic study of a marine ecosystem off southeastern Australia using stable isotopes of carbon and nitrogen. *Canadian Journal of Fisheries and Aquatic Sciences* 59:514–530.
- Deutsch, C. A., J. J. Tewksbury, R. B. Huey, K. S. Sheldon, C. K. Ghalambor, D. C. Haak, and P. R. Martin. 2011. Impacts of climate warming on terrestrial ectotherms across latitude. *Proceedings of the National Academy of Sciences USA* 105:6668–6672.
- Dunstan, P. K., N. J. Bax, S. D. Foster, A. Williams, and F. Althaus. 2012. Identifying hotspots for biodiversity management using rank abundance distributions. *Diversity and Distributions* 18:22–32.
- Edgar, G. J., N. S. Barrett, and A. J. Morton. 2004. Biases associated with the use of underwater visual census techniques to quantify the density and size-structure of fish populations. *Journal of Experimental Marine Biology and Ecology* 308:269–290.
- 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.
- Ellis, N., and C. R. Pitcher. 2011. Uncertainty for biodiversity prediction. Pages 57–60 *in* N. J. Bax, editor. CERF report on Transition and Extension Program. Marine Biodiversity Hub, Hobart, Tasmania. <http://www.nerpmarine.edu.au/document/cerf-report-transition-and-extension-program>
- Ellis, N., S. J. Smith, and C. R. Pitcher. 2011a. Gradient forests: calculating importance gradients on physical predictors. *Ecology* 93:156–168.
- Ellis, S. L., et al. 2011b. Four regional marine biodiversity studies: approaches and contributions to ecosystem-based management. *PLoS One* 6:e18997.
- Fabricius, K., G. De'ath, L. McCook, E. Turak, and D. M. Williams. 2005. Changes in algal, coral and fish assemblages

- along water quality gradients on the inshore Great Barrier Reef. *Marine Pollution Bulletin* 51:384–398.
- 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.
- Grantham, H. S., K. A. Wilson, A. Moilanen, T. Rebelo, and H. P. Possingham. 2009. Delaying conservation actions for improved knowledge: How long should we wait? *Ecology Letters* 12:293–301.
- Guisan, A., and W. Thuiller. 2005. Predicting species distribution: offering more than simple habitat models. *Ecology Letters* 8:993–1009.
- Harley, C. D. G. 2008. Tidal dynamics, topographic orientation, and temperature-mediated mass mortalities on rocky shores. *Marine Ecology Progress Series* 371:37–46.
- Helmuth, B., B. R. Broitman, C. A. Blanchette, S. Gilman, P. Halpin, C. D. G. Harley, M. J. O'Donnell, G. E. Hofmann, B. Menge, and D. Strickland. 2006. Mosaic patterns of thermal stress in the rocky intertidal zone: implications for climate change. *Ecological Monographs* 76:461–479.
- Hewitt, C. L. 2002. Distribution and biodiversity of Australian tropical marine bioinvasions. *Pacific Science* 56:213–222.
- Hewitt, J. E., M. J. Anderson, and S. F. Thrush. 2005. Assessing and monitoring ecological community health in marine systems. *Ecological Applications* 15:942–953.
- Huang, Z., B. P. Brooke, and P. T. Harris. 2011. A new approach to mapping marine benthic habitats using physical environmental data. *Continental Shelf Research* 31:S4–S16.
- Huang, Z., B. Brooke, N. Whitta, A. Potter, M. Fuller, J. Dunn, and R. Pitcher. 2010. Australian marine physical environmental data—descriptions and metadata. *Geoscience Australia Record* 2010/32, Geoscience Australia, Canberra, Australia.
- Jordan, A., M. Lawler, V. Halley, and N. Barrett. 2005. Seabed habitat mapping in the Kent Group of islands and its role in marine protected area planning. *Aquatic Conservation: Marine and Freshwater Ecosystems* 15:51–70.
- Karakassis, I., A. Machais, P. Pitta, K. N. Papadopoulou, C. J. Smith, E. T. Apostolaki, M. Giannoulaki, D. Koutsoubas, and S. Somarakis. 2006. Cross-community congruence of patterns in a marine ecosystem: Do the parts reflect the whole? *Marine Ecology Progress Series* 310:47–54.
- Keller, R. P., D. Kocov, and S. Dzeroski. 2011. Trait-based risk assessment for invasive species: high performance across diverse taxonomic groups, geographic ranges and machine learning/statistical tools. *Diversity and Distributions* 17:451–461.
- Kerswell, A. P. 2006. Global biodiversity patterns of benthic marine algae. *Ecology* 87:2479–2488.
- Leaper, R., N. A. Hill, G. J. Edgar, N. Ellis, E. Lawrence, C. R. Pitcher, N. S. Barrett, and R. Thomson. 2011. Predictions of beta diversity for reef macroalgae across southeastern Australia. *Ecosphere* 2:art73.
- Leathwick, J. R., J. Elith, M. P. Francis, T. Hastie, and P. Taylor. 2006. Variation in demersal fish species richness in the oceans surrounding New Zealand: an analysis using boosted regression trees. *Marine Ecology Progress Series* 321:267–281.
- Legendre, P., and L. Legendre. 1998. *Numerical ecology*. Elsevier, Amsterdam, The Netherlands.
- Lindsay, M. J., H. M. Patterson, and S. E. Swearer. 2008. Habitat as a surrogate measure of reef fish diversity in the zoning of the Lord Howe Island Marine Park, Australia. *Marine Ecology Progress Series* 353:265–273.
- Margules, C. R., and R. L. Pressey. 2000. Systematic conservation planning. *Nature* 405:243–253.
- McArthur, M. A., B. P. Brooke, R. Przeslawski, D. A. Ryan, V. L. Lucieer, S. Nichol, A. W. McCallum, C. Mellin, I. D. Cresswell, and L. C. Radke. 2010. On the use of abiotic surrogates to describe marine benthic biodiversity. *Estuarine, Coastal and Shelf Science* 88:21–32.
- McKnight, M. W., P. S. White, R. I. McDonald, J. F. Lamoreux, W. Sechrest, R. S. Ridgely, and S. N. Stuart. 2007. Putting beta-diversity on the map: broad-scale congruence and coincidence in the extremes. *PLoS Biology* 5:2424–2432.
- Mellin, C., C. J. A. Bradshaw, M. G. Meekan, and M. J. Caley. 2010. Environmental and spatial predictors of species richness and abundance in coral reef fishes. *Global Ecology and Biogeography* 19:212–222.
- Mellin, C., S. Deelan, J. Caley, G. Edgar, M. G. Meekan, R. Pitcher, R. Przeslawski, A. Williams, and C. J. A. Bradshaw. 2011. Effectiveness of biological surrogates for predicting patterns of marine biodiversity: a global meta-analysis. *PLoS One* 6:e20141.
- Mumby, P. J., K. Broad, D. R. Brumbaugh, C. P. Dahlgren, A. R. Harborne, A. Hastings, K. E. Holmes, C. V. Kappel, F. Micheli, and J. N. Sanchirico. 2008. Coral reef habitats as surrogates of species, ecological functions, and ecosystem services. *Conservation Biology* 22:941–951.
- O'Hara, T. 2008. Bioregionalisation of Australian waters using brittle stars (Echinodermata: Ophiuroidea), a major group of marine benthic invertebrates. Department of the Environment, Water, Heritage and Arts, Canberra, Australia.
- Peres-Neto, P. R., and D. A. Jackson. 2001. How well do multivariate data sets match? The advantages of a Procrustean superimposition approach over the Mantel test. *Oecologia* 129:169–178.
- Perry, A. L., P. J. Low, J. R. Ellis, and J. D. Reynolds. 2005. Climate change and distribution shifts in marine fishes. *Science* 308:1912–1915.
- 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.
- Pitcher, C. R., et al. 2011. Prediction program. Pages 38–50 in N. E. A. Bax, editor. Report to Department of Sustainability, Environment, Water, Population and Communities. Final report 2007–2010. Marine Biodiversity Hub, Commonwealth Environment Research Facilities, Canberra, Australia. <http://www.nerpmarine.edu.au/document/cerf-marine-biodiversity-hub-final-report-2007-2010>
- Pitcher, C. R., N. Ellis, S. J. Smith, L. S. Incze, C.-L. Wei, M. E. Greenlaw, N. H. Wolff, J. A. Sameoto, and P. V. R. Snelgrove. 2012. Exploring the role of environmental variables in shaping patterns of seabed biodiversity composition in regional-scale ecosystems. *Journal of Applied Ecology* 49:670–679.
- Poertner, H. O., and R. Knust. 2007. Climate change affects marine fishes through the oxygen limitation of thermal tolerance. *Science* 315:95–97.
- Politis, D. N., and J. P. Romano. 1994. Large-sample confidence-regions based on subsamples under minimal assumptions. *Annals of Statistics* 22:2031–2050.
- Poloczanska, E. S., R. C. Babcock, A. Butler, A. J. Hobday, O. Hoegh-Guldberg, T. J. Kunz, R. Matear, D. A. Milton, T. A. Okey, and A. J. Richardson. 2007. Climate change and Australian marine life. *Oceanography and Marine Biology: An Annual Review* 45:407–478.
- Rodrigues, A. S. L., and T. M. Brooks. 2007. Shortcuts for biodiversity conservation planning: the effectiveness of surrogates. *Annual Review of Ecology, Evolution, and Systematics* 38:713–737.
- Saetersdal, M., and I. Gjerde. 2011. Prioritising conservation areas using species surrogate measures: consistent with ecological theory? *Journal of Applied Ecology* 48:1236–1240.
- Strobl, C., A. L. Boulesteix, T. Kneib, T. Augustin, and A. Zeileis. 2008. Conditional variable importance for random forests. *BMC Bioinformatics* 9.

- Stuart-Smith, R. D., N. S. Barrett, C. M. Crawford, S. D. Frusher, D. G. Stevenson, and G. J. Edgar. 2008. Spatial patterns in impacts of fishing on temperate rocky reefs: Are fish abundance and mean size related to proximity to fisher access points? *Journal of Experimental Marine Biology and Ecology* 365:116–125.
- Sutcliffe, P. R., C. R. Pitcher, M. J. Caley, and H. P. Possingham. 2012. Biological surrogacy in tropical seabed assemblages fails. *Ecological Applications* 22:1762–1771.
- Ward, T. J., M. A. Vanderklift, A. O. Nicholls, and R. A. Kenchington. 1999. Selecting marine reserves using habitats and species assemblages as surrogates for biological diversity. *Ecological Applications* 9:691–698.
- Watts, M. E., I. R. Ball, R. S. Stewart, C. J. Klein, K. Wilson, C. Steinback, R. Lourival, L. Kircher, and H. P. Possingham. 2009. Marxan with zones: software for optimal conservation based land- and sea-use zoning. *Environmental Modelling and Software* 24:1513–1521.
- Wernberg, T., B. D. Russell, M. S. Thomsen, F. D. Gurgel, C. J. A. Bradshaw, E. S. Poloczanska, and S. D. Connell. 2011. Seaweed communities in retreat from ocean warming. *Current Biology* 21:1828–1832.
- Whittaker, R. H. 1972. Evolution and measurement of species diversity. *Taxon* 21:213–251.
- Williams, A., F. Althaus, P. K. Dunstan, G. C. B. Poore, N. J. Bax, R. J. Kloser, and F. R. McEnnulty. 2010. Scales of habitat heterogeneity and megabenthos biodiversity on an extensive Australian continental margin (100–1100 m depths). *Marine Ecology—an Evolutionary Perspective* 31:222–236.

SUPPLEMENTAL MATERIAL

Appendix A

A table describing the 19 physical covariates ([Ecological Archives A024-017-A1](#)).

Appendix B

Description of the method of obtaining an optimal value of w for use in Eq. 1 ([Ecological Archives A024-017-A2](#)).

Appendix C

Plots of the cumulative importance for the environmental covariates that were not included in Fig. 3 ([Ecological Archives A024-017-A3](#)).

Appendix D

A map containing comparisons of the predicted biodiversity among the three assemblages, using 13 boxplots for each of the bioregions within the study region ([Ecological Archives A024-017-A4](#)).