Vol. 730: 59–78, 2024
https://doi.org/10.3354/meps14538

MARINE ECOLOGY PROGRESS SERIES
Mar Ecol Prog Ser

Published March 7

Modeling the spatial distribution of numbers of coral reef fish species and community types in the Western Indian Ocean faunal province

T. R. McClanahan[1,*], Alan M. Friedlander[2,3], Pascale Chabanet[4,5], J. H. Bruggemann[4,5], J. Wickel[6], M. K. Azali[7]

[1]Wildlife Conservation Society, Global Marine Programs, Bronx, NY 10460, USA
[2]Pristine Seas, National Geographic Society, Washington, DC 20036, USA
[3]Hawai'i Institute of Marine Biology, University of Hawai'i, Kaneohe, HI 96744, USA
[4]UMR 9220 ENTROPIE, Université de La Réunion, Saint Denis, La Réunion 97400, France
[5]Laboratoire d'Excellence CORAIL, Perpignan 66000, France
[6]GIE MAREX, Saint Leu, La Réunion 97400, France
[7]Wildlife Conservation Society, Kenya Marine Program, Mombasa 80107, Kenya

ABSTRACT: Predicting and mapping coral reef diversity at moderate scales can assist spatial planning and prioritizing conservation activities. We made coarse-scale (6.25 km^2) predictive models for numbers of coral reef fish species and community composition starting with a spatially complete database of 70 environmental variables available for 7039 mapped reef cells in the Western Indian Ocean. An ensemble model was created from a process of variable elimination and selectivity to make the best predictions irrespective of human influences. This best model was compared to models using preselected variables commonly used to evaluate climate change and human fishing and water quality influences. Many variables (~27) contributed to the best number of species and community composition models, but local variables of biomass, depth, and retention connectivity were dominant predictors. The key human-influenced variables included fish biomass and distance to human populations, with weaker associations with sediments and nutrients. Climate-influenced variables were generally weaker and included median sea surface temperature (SST) with contributions in declining order from SST kurtosis, bimodality, excess summer heat, SST skewness, SST rate of rise, and coral cover. Community composition variability was best explained by 2 dominant community richness axes of damselfishes—angelfishes and butterflyfishes—parrotfishes. Numbers of damselfish—angelfish species were ecologically separated by depth, and damselfishes declined with increasing depth, median temperature, cumulative excess heat, rate of temperature rise, and chronic temperature stresses. Species of butterflyfish—parrotfish separated by median temperature, and butterflyfish numbers declined with increasing temperature, chronic and acute temperature variability, and the rate of temperature rise. Several fish diversity hotspots were found in the East African Coastal Current Ecoregion centered in Tanzania, followed by Mayotte, southern Kenya, and northern Mozambique. If biomass can be maintained, the broad distributions of species combined with compensatory community responses should maintain high diversity and ecological resilience to climate change and other human stressors.

KEY WORDS: Africa · Biodiversity · Bony fish · Environmental drivers · Species diversity · Spatial modeling

1. INTRODUCTION

Patterns of coral reef fish diversity and their drivers are increasingly being understood at large spatial scales (Parravicini et al. 2013, Pellissier et al. 2014, Williams et al. 2015, Ceccarelli et al. 2023). Many large-scale approaches are based on compilations and mapping of site-based empirical observations.

*Corresponding author: tmcclanahan@wcs.org

Publisher: Inter-Research · www.int-res.com

For example, compiled presence/absence observations of species are used to create polygon distribution maps that estimate total species diversity (Selig et al. 2014, Jenkins & Van Houtan 2016, Bullock et al. 2021). Species diversity patterns arising from this common method produce distributions based on extrapolations across large areas. However, suitable environments and habitats may be lacking in the species' ranges created by this extrapolation method. Therefore, tests of predictive strength of this extrapolative method have found it weak at locating populations of the studied species (Lee-Yaw et al. 2022). Improved predictions could be made if the habitats and environmental conditions that create the species' niches were explicitly modeled across the geographies of interest. The density and distribution of presence/absence field sampling is also expected to influence large-scale diversity maps. Therefore, mapped patterns may be very coarse and potentially misleading where field sampling is sparse, which could adversely influence conservation prioritization decisions. Moreover, these and other conservation prioritization methods often identify large unspecified areas, many too large to be contained in the most common modest-sized protected areas (EAME 2004, Obura et al. 2012, Boonzaier & Pauly 2016). Recently, improvements in satellite coverage, environmental proxies, and spatial modeling methods are increasingly making finer-scale predictions possible (Pilowsky et al. 2022).

Spatially resolved habitat and environmental information has the potential to produce more explicit species and community distribution maps (Stephenson et al. 2020, Mokany et al. 2022). Increasing satellite and shipboard measurement availability can provide environmental data to explore potential environment—taxa associations (Tyberghein et al. 2012, Yeager et al. 2017, Assis et al. 2018). Complex associational and regression algorithms can then map numbers of species and community types at the spatial resolution of the environmental data. Moreover, satellite data availability has encouraged the development of many single and multivariate proxies. For example, there are proxies for thermal stress (Maina et al. 2011), human influences (Maire et al. 2016), water pollution (Andrello et al. 2022), and larval connectedness (Fontoura et al. 2022) that can be used to improve the predictive strength of models.

An advantage of an environmental modeling approach is that many of the associated variables are expected to reflect species' niches, thereby producing more ecologically realistic distribution patterns (Pilowsky et al. 2022). The moderate spatial resolution and wide coverage of these data sources can assist in making finer-scale predictions than overlapping species distribution polygons or cumulative species rarefaction alternatives. Environmental modeling is particularly useful where there is an interest in formally protecting biodiversity but where there is a history of sparse field sampling, high biodiversity, and limited information for prioritization (Parravicini et al. 2013, McClanahan et al. 2015, Kusumoto et al. 2020, UNEP-Nairobi Convention & WIOMSA 2021). Emerging predictive modeling tools can be used broadly, include many potentially predictive variables, and may eventually be able to assist global-scale conservation priorities.

The spatial scale of biodiversity conservation in tropical marine regions is small and limited by trade-offs between conservation and food needs. For example, in Africa and the Western Indian Ocean (WIO), human dependence on natural resources and the subsequent unsustainable use of fish is widespread (Zeller et al. 2021, McClanahan et al. 2023). Therefore, human needs often limit the size and numbers of large-scale fully protective management systems. Conservation increasingly involves more local jurisdictions, implementation at smaller scales, and greater access by resource users (Cinner et al. 2012, Rocliffe et al. 2014, Ban et al. 2023). Knowledge of fish biodiversity at this scale is limited and mostly available for specific sampled sites or management systems, such as marine protected and adjacent fished areas (McClanahan 2019). To assist the development of large-scale species diversity modeling, we used environmental modeling methods to predict the number of fish species and community composition. Here, satellite, shipboard, and multivariate proxies were used to predict numbers of reef fish species and community distributions in the WIO faunal province for 7039 reef cells mapped at a ~6 km^2 scale (Burke et al. 2011).

The research described below was a collaborative effort to explore environmental relationships with a proxy coral reef fish assemblage. Our model made predictions based on an empirical census of 5 commonly sampled species-rich families previously found useful for estimating numbers of species at large scales (Allen & Werner 2002). Two metrics of fish biodiversity were modeled: number of fish species and community composition richness. We focused on the potential associations and responses of these taxonomic metrics to relevant oceanographic data and commonly evaluated climatic change, water quality, and human resource extraction variables. We asked which of the many available environmental variables contribute to making good predictions and how these predictions compare with those commonly preselected for studying human and climate impacts. Moreover, we used spatial clustering methods to identify locations with high densities of biodiverse reefs that may benefit from conservation efforts. The investigation was undertaken to better understand fish–environment associations and the specific strengths of climate change, water quality, and human-use variable relationships.

2. MATERIALS AND METHODS

The study builds on several methodological advances that have made it possible to evaluate biogeographic patterns of reef fishes in the WIO. These include: (1) moderate-resolution mapping of coral reefs, (2) global satellite coverage of environmental variables that are potential proxies for biodiversity, (3) increased scale of field data collection and collaboration, and (4) machine learning algorithms capable of evaluating and calculating partial effects that allow making fair comparisons when evaluating numerous variables and locations. These 4 emergent tools allowed us to evaluate and map our proxy of numbers of fish species in the WIO coral reefs.

2.1. Study region

We used recent iterations of coral reef maps to establish reef distribution patterns (https://data.unep-wcmc.org/datasets/1) (Burke et al. 2011). Specifically, we used the map of the WIO composed of ~7039 cells (each 6.25 km^2) distributed among 9 ecoregions, namely the Northern Monsoon Current Coast, East African Coral Coast, Seychelles, Cargados Carajos/Tromelin Island, Mascarene Islands, Southeast Madagascar, Western and Northern Madagascar, Bight of Sofala/Swamp Coast, and Delagoa. Empirical data were only available for 6 of the more accessible ecoregions, but we used the model from these data to predict values in the 3 unsampled ecoregions. Therefore, predictions for Cargados Carajos/Tromelin Island, Southeast Madagascar, and Bight of Sofala/Swamp Coast should be seen as unconfirmed but testable predictions. The WIO faunal province also includes 9 national governance jurisdictions, with Madagascar, Tanzania, and Mozambique having the most coral reefs cells (43%) (Text S1, Fig. S1 in the Supplement at www.int-res.com/articles/suppl/m730p059_supp.pdf). The scale of this map aligns well with the scale of the environmental data described below.

2.2. Environmental data sources

Environmental data compilations accessed several sources from satellite and shipboard measurements and multivariate compilations (Table 1, Table S1). Environmental oceanographic layers included those expected to influence marine organisms, such as wave energy, photosynthetic active radiation (PAR), light diffusion attenuation, pH, calcite, dissolved oxygen, salinity, net primary productivity, chlorophyll *a*, and phytoplankton carbon (Tyberghein et al. 2012, Yeager et al. 2017). Additionally, several water temperature or thermal stress metrics were used, including sea surface temperature (SST) mean, median, range, standard deviation, skewness, kurtosis, and rate of rise, as well as cumulative degree-heating weeks (DHW) (https://coralreefwatch.noaa.gov/product/5km/index_5km_sst.php). Kurtosis and skewness are metrics that reflect chronic and acute stress on marine organisms and, along with temperature ranges and standard deviations, have been shown to influence coral cover, loss of coral symbionts, and other coral reef community metrics (Ateweberhan & McClanahan 2010, Safaie et al. 2018, McClanahan 2020). Water chemistry metrics of dissolved oxygen, salinity, pH, and calcite concentrations are expected to reflect conditions known to affect fish health and distributions (Gallo et al. 2020, Pinheiro et al. 2021). Two multivariate integrated metrics of thermal stress and water quality were included (Maina et al. 2011, Andrello et al. 2022). Habitats were assigned to grid cells based on the most common habitat observed in each cell from satellite image observations. Finally, for each cell, we used several layers that measure connectivity, including average net flow, indegree, outdegree, and retention metrics that are expected to influence numbers of species (Fontoura et al. 2022). Fish census observers also recorded local site metrics including depth and habitats, recorded as reef edge, reef crest, reef flat, or reef lagoon. Field-based habitat classifications were used to build the model, but predictions were based on the dominant habitat type in each cell.

Several geographic variables were retained independently of the larger and objective variable selection process to evaluate a common practice of preselecting variables of human and climate change concern. Human-influenced variables included were nation, wilderness (>4 h travel time from a human population), and travel distance to either people living on coastlines or in cities. These are referred to as gravity to a coastal population and city or the number of people living on the nearest shore or cities divided by

Table 1. Statistical results of the boosted regression tree ranking of variables by their relative influence on (% contribution to) the 3 response variables, i.e. number of fish species, canonical correspondence analysis (CCA) axis 1 (fewer Pomacentridae and more Pomacanthidae species) and CCA axis 2 (fewer Chaetodontidae and more Scarinae species), for the 2 models with alternative sets of variables. Blank cells indicate the variable was not used in the specific model. See Table S1 in the Supplement for all variables, further explanations, and their sources. Variable order from most to least important is based on relative importance in predicting the number of species. SST: sea surface temperature; DHW: degree-heating week; PSS: practical salinity scale

Variable	Ecological importance	Number of fish species		Community axis 1 (Pomacentridae–Pomacanthidae)		Community axis 2 (Chaetodontidae–Scarinae)	
		Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Total fish biomass (kg ha^{-1})	Fish community health status	42.6	42.7	17.1	17	17.4	16.4
SST median (°C)	Productivity and metabolism	10.1		2.8		3.3	
Travel time to nearest human population (h)	Coastal fishing pressure	7.5		1.3		3.7	
Retention connectivity	Net fish recruitment	5.7	6.1	1.1	1.1	4.1	4.2
SST kurtosis	Chronic temperature stress	4.2		8.4		2.9	
Net primary productivity (mg C m^{-2} d^{-1})	Food production	3		4.4		2.8	
Depth (m)	Behavioral and escape from fishing effects	2.7	3.1	14.1	14.6	3.8	3.5
Observers	Potential human or methods bias	2.6	4.3	1.5	1.5	6.3	6.1
Travel time to market (h)	Proxy of fishing pressure	2.5		1.4		2.7	
Maximum photosynthetically active radiation (E m^{-2} d^{-1})	Energy availability	2.4	3.2	1.6	1.5	5.2	4.1
Median chlorophyll a (mg m^{-3})	Food availability	2.1		2.8		5.7	
Salinity (PSS)	Environmental chemistry conditions	1.9	1.7	4.3	3.9	6	5.4
Cumulative DHW (°C-weeks)	Excess heat stress	1.6		2.1		3.4	
SST rate of rise (°C yr^{-1})	Speed of climate change	1.5	1.4	19.8	8.8	3.3	3.2
Mean wave energy (kW m^{-1})	Disturbance and production	1.5	1.2	1.9	1.7	2.7	2.8
Reef visitation value (number of tourist visits)	Environmental aesthetic	1.5	1.2	0.7	0.6	3.9	3.4
Hard coral cover (%)	Refuge from predation	1	1.5	3.5	3.1	1.7	1.4
Indegree connectivity	Input of recruits	0.9	1.5	1.5	1.6	1.6	1.6
Nutrients (nitrogen, t km^{-2})	Resource input	0.9		0.9		0.9	
Ecoregion	Biogeographical history	0.7	0.9	0.2	0.2	0.1	0.1
Habitat	Species preferences	0.6	0.8	1.3	1.3	2.7	2.8
Dissolved oxygen (ml l^{-1})	Metabolism influence	0.6	0.7	1	0.9	2.5	2
Net flow connectivity	Recruitment retention metric	0.5	0.8	1.8	1.9	1.3	1.4
Calcite concentration (mol m^{-3})	Environmental chemistry conditions	0.5	0.8	2	2.2	6.9	6.1
Management	Fisheries impact	0.5	0.5	1.5	1.4	1.6	1.6
Outdegree connectivity	Loss of recruits	0.4	0.6	1.1	1	3.5	2.5
pH	Environmental chemistry conditions		6.6		1		1.7
Gravity to nearest population: population/(c + travel time (h)^2)	Coastal fishing population pressure		5.0		1.3		5.4
SST skewness	Acute temperature stress		4.8		2.1		5.6
Current velocity (m s^{-1})	Disturbance and productivity		3.1		1.8		3.4
Diffuse attenuation coefficient (m^{-1})	Available energy		3		2.4		3
SST bimodality	Temperature stress		2.7		10.8		3.9
Climate stress model	Thermal stress		1.1		2.9		3.2
Sediments (t km^{-2})	Visibility and productivity		1		1.6		2.6
Gravity to nearest city: population/(c + travel time (h)^2)	Fisheries market influence		1		1.7		2.6

a constant plus the square of the distance or travel time (Maire et al. 2016). Population and city gravity metrics were log transformed to better visualize patterns. The ecoregion was also evaluated as a potential variable but had low predictive strength. Mapped reef cells were assigned 4 fisheries management categories comprised of unrestricted fishing (42%), restricted fishing (42%), low compliance closures (14%), and high compliance closures (2%). These classifications were based on information in published literature, the experience of the data providers, and discussions with knowledgeable observers (McClanahan et al. 2015).

2.3. Field data collection

Experienced observers counted fishes in designated areas using either circular or belt transect methods of 100, 154, 250, or 500 m^2 described in the authors' foundational methods papers (McClanahan 1994, Friedlander et al. 2014, Chabanet et al. 2016, Graham et al. 2020). Smaller replicates conducted close to each other were pooled such that the final units were number of species per ~ 250, 300, 462, and 500 m^2. Consequently, we evaluated several strategies to account for the differences in sampling units, including the sampled area and using only samples with a similar transect area (Text S2, Figs. S2 & S3). Differences in the 5 tested models were small, but we present results from samples collected at 500 m^2 using the boosted regression tree (BRT) regressions and multivariate analysis of similarity (ANOSIM) of species composition.

We used the number of species from 5 families sampled by all observers (Acanthuridae (surgeonfish), Chaetodontidae (butterflyfish), Labridae (wrasses) — Scarinae (parrotfish), Pomacanthidae (angelfish), and Pomacentridae (damselfish)), which are known to be good proxies for the total numbers of species in regions (Allen & Werner 2002). We separated the herbivorous Scarinae from the other carnivorous Labridae species in our analysis because their feeding and ecological differences affected the community composition analyses. Each observer independently estimated total fish biomass based on the sum of the weights of all individuals. Biomasses were calculated from length estimates and known length—weight relationships compiled at either the species or family level as per the observers' methods. A total of 1201 transects were sampled and 967 500 m^2 used in the final evaluation. Transects were surveyed throughout most of the countries and ecoregions in the WIO province between 1991 and 2022, and the environmental data used for analysis corresponded to the time just prior to field sampling (Fig. S1).

2.4. Data analyses

2.4.1. Species community composition analysis

To distinguish coral reef fish community distributions in the province, standard ecological multivariate ordination was conducted. We used the canonical correspondence analysis (CCA) method to evaluate the distribution of species within the 5 families for their associations with the above environmental variables. We further evaluated the species dissimilarity between ecoregions by a pairwise 1-way ANOSIM using Bray-Curtis distances. Statistical significance was assessed with the Bonferroni correction for multiple testing. Similarity percentages (SIMPER) procedures were performed to evaluate fish families that contributed the most to the average dissimilarity between ecoregions. Similarities were calculated as 1 − overall average dissimilarity between ecoregion pairs (Text S3, Table S2). One-way ANOSIM results indicated an overall significant difference in the per family number of species among ecoregions (ANOSIM $R = 0.11$, $p = 0.0001$), while pooled SIMPER analyses indicated an overall high species similarity of 86% among ecoregions.

2.4.2. Estimated number of species

The number of species was estimated by several methods prior to selecting a final or best model for presentation. Historically, rarefaction methods have been used to account for variable species—area relationships (Gotelli & Colwell 2001). However, BRT machine learning algorithms are an alternative that might uncover associational rather than pre-selected equations for the species—area responses. BRT predictions can control for differences in sampled area and other sampling or environmental factors simultaneously, whereas rarefaction assumes an underlying equation or model structure. Therefore, prior to selecting outputs for presentation, we evaluated 7 potential models to evaluate numbers of species: 2 were based on rarefaction methods and 5 on the machine learning or BRT methods (Text S2). This evaluation was undertaken to account for the effects of observer and sampled area in the BRT model described below. Evaluations with and without country or national jurisdiction were included to evaluate this 'dummy' or non-environmental variable effect in potentially obfuscating underlying patterns of diversity. Comparison of these 7 methods indicate that rarefaction methods produced weaker fits ($r^2 = 0.60$) than the BRT method ($r^2 = 0.79$ to 0.85) (Figs. S2 & S3). Additionally, national jurisdiction, a non-environmental variable, was the top variable in the rarefaction approach. This suggests some weaknesses in rarefaction methods. Rarefaction does not account for the partial effects of important environmental variables and effects that may arise from different observers and methods. The 5 BRT methods produced similar results in their number of species prediction and top variables. The chosen BRT variables were also those expected by ecological theory, such as biomass, depth, connectivity, and temperature. We therefore present the results from the BRT model based on samples collected at 500 m^2.

2.4.3. Environmental association analyses

The BRT analyses evaluated associations between the response variables of numbers of taxa and species community axis 1 and 2 from the multivariate species ordination analyses (Text S4). The number of variables evaluated went through a process of variable elimination to reduce redundancy. Furthermore, a final ensemble model was created to avoid eliminating potentially causative variables (Text S5, Fig. S4). The structure of BRT models ensures they are robust to autocorrelation. Nevertheless, variable selection processes can complicate the interpretation and fail to identify potential causative variables (Pilowsky et al. 2022). For example, it is possible that a potentially causative variable could be eliminated by a small improvement in predictions by a stronger correlational variable. Therefore, we developed ensemble forecast models based on 2 sets of variables as recommended by Araújo & New (2007). Both models included all potentially strong causative variables, but the second model included some correlated variables eliminated in the first model suspected of causative environmental stress relationships (i.e. SST mean, SD, and skewness) (Table 1). By this method, our number of fish species proxy for each cell was the average of the 2 models weighted by their explained deviance. The elimination process reduced the number of variables from 70 to 35 via analysis of redundancy. Of these 35 variables, 26 were used in BRT Model 1 while 27 were used in Model 2. Further, 11 variables were maintained because they were metrics commonly used to study human influences of water quality and fishing impacts. Nine variables were also retained because they were commonly used to evaluate climatic change. This selective variable retention process was undertaken to compare the strength of associations and response relationships of frequently preselected variables relative to a more objective variable selection process.

Model performances were evaluated using the statistical guidelines of randomly splitting the data into a 70% training set and 30% testing set and calculating the Theil's *U*-statistic, percent deviance explained (analogous to R^2), and Pearson's correlation coefficients presented in Table S3 (Kuhn & Johnson 2013). To make fair comparisons of the BRT predictions for each cell, we held fish biomass (600 kg ha^{-1}), coral cover (30%), and sampled area (500 m^2) constant for the final analysis and mapping.

2.4.4. Hotspot predictions

We searched for locations where groups of reefs with high and low diversity exist using an optimized hotspot analysis. We used the Getis-Ord Gi* statistic to identify statistically significant spatial clusters with high (hotspots) or low numbers of species (coldspots) within a defined neighborhood distance (ESRI 2022a). The method used *z*-scores of the predicted number of fish species and p-values corrected for false discovery rates potentially arising from spatial dependence and multiple testing (Ord & Getis 1995). Specifically, a high positive *z*-score with a low p-value ($p < 0.10$) indicated a hotspot, a low negative *z*-score with a low p-value indicated a cold spot, and a score near zero indicated a lack of spatial clustering (Ord & Getis 2001). The optimized hotspot analysis can use several strategies to determine the optimum parameter settings for the analysis, including the neighborhood distance for clustering. The peak incremental spatial autocorrelation, and the average distance to 30 nearest neighbors were among the strategies evaluated (ESRI 2022b). No peaks of spatial autocorrelation were identified in the predicted number of fish species. Therefore, a neighborhood distance of ~25.5 km based on the average distance to the nearest 30 neighbors was used.

3. RESULTS

Below are the results of (1) observed site and ecoregional number of species and composition analyses, (2) local site-, climate-, and human-influenced variable associations with number of species and community composition richness, and (3) predicted number of species and hotspots by ecoregions and nations. Pearson's correlation for the 30% testing of the models was 0.80 for both Model 1 and Model 2 (Text S6). For the full data, the correlation was 0.89.

3.1. Site and ecoregional analysis of empirical data

Species composition of the empirical data was similar (81–89%) among ecoregions when comparing sampled sites (Table S2). The most similar and most central ecoregions in this province were the well sampled East African Coral Coast and the Western and Northern Madagascar ecoregions. The peripheral Mascarene Islands were the least similar to these 2 core ecoregions. The modestly sampled Delagoa was least similar to the Mascarene Islands, but not different from other ecoregions. Individual fish families contributed between 15 and 18% to the overall ecoregional dissimilarities. Pomacanthidae, Pomacentridae, carnivorous Labridae, and herbivorous Scarinae had the highest contribution, while Acanthuridae and Chaetodontidae had the lowest overall contribution to ecoregional dissimilarities.

The unconstrained CCA of the observed number of species per family separated sites by the 6 taxonomic groupings shown in Fig. 1. The East African Coral Coast and Western and Northern Madagascar ecoregions were most similar and therefore the most centrally located ecoregions in the CCA plot. While there were few samples from the Monsoon Coast of northern Kenya, they were associated with higher relative numbers of Scarinae. Madagascar had lower concordance among sites, suggesting high spatial variation. The Mascarene Islands were associated with higher relative numbers of Pomacentridae spe-

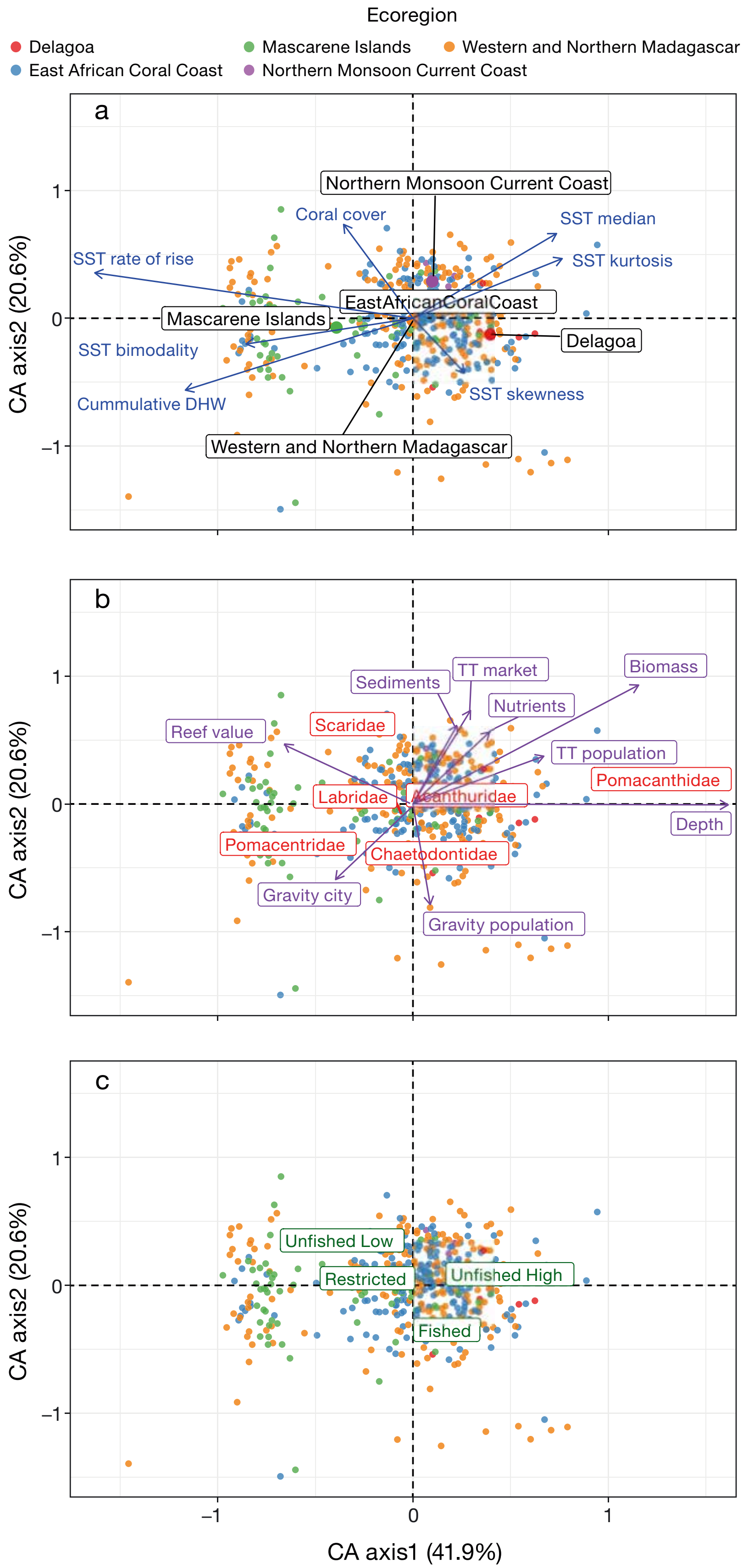


Fig. 1. Canonical correspondence analysis (CCA) of empirical field data for coral reef cells using the square root of the number of species per unit area in the 6 proxy families in the 6 studied ecoregions. (a) Associations with key environmental variables; SST: sea surface temperature; DHW: degree-heating week. (b) Continuous human variables; TT: travel time. Other variables as in Table 1. (c) Human categorical variables

cies and thus distributed furthest left on the first CCA axis. In contrast, the Delagoa ecoregion was distributed to the right side of this axis, having greater numbers of Pomacanthidae species. Therefore, sites adjacent to the Coral Coast were most similar with a strong band of similarity extending seaward from Tanzania towards northwestern Madagascar.

The Acanthuridae and carnivorous Labridae species were the most centrally located families. The first CCA axis separated sites by the numbers of Pomacentridae and Pomacanthidae. The second CCA axis separated sites by the numbers of Chaetodontidae and Scarinae. Therefore, positive values on axis 1 indicated more species of Pomacanthidae and fewer Pomacentridae and for axis 2 more Scarinae and fewer Chaetodontidae.

Numbers of species and the community axes separations were associated with climate- and human-influenced variables (Fig. 2). Biomass was the most influential variable, and total, Pomacanthidae, and Scarinae species numbers increased with biomass (Fig. 3). These patterns were reflected in the fisheries management associations, with numbers of Pomacanthidae and Scarinae species highest in high-compliance fisheries closures; Pomacentridae species numbers were highest in restricted fishing and low-compliance closures, and Chaetodontidae were highest in fished sites.

3.2. Local site- and human-influenced variable associations

Predictive strengths of numbers of species and community composition indicated the importance of the common variables of biomass; depth; SST kurtosis, bimodality, skewness, median, and rate of rise; and coral cover (Table 1). Therefore, these variables were included when evaluating models that retained climate (i.e. DHW) and human impact (i.e. travel time and nutrients) variables (Fig. 2). The human-influenced model had stronger predictive variables than the climate model, especially given that biomass, the strongest variable, was largely associated with fishing impacts. The top non-biomass variables were gravity to people and sediments, with lesser importance of travel time to cities and people, city gravity, nutrients, reef tourism value, and fisheries management (Fig. 2b).

In the climate model, after accounting for the strong partial effect of biomass, the number of species were best predicted by the median SST and kurtosis with lesser importance of bimodality, cumulative DHW, SST skewness, rate of SST rise, and coral cover (Fig. 2a). Therefore, DHW and coral cover were moderate to weak predictors of numbers of species relative to the SST background variables. These variables were also largely selected as important for the 2 community axes, but their relative importance and response relationships varied (Fig. 2c–f).

The Pomacentridae–Pomacanthidae species axis 1 gradient indicated that sediment was the only human-influenced variable with >10% importance. Weaker variables included city gravity, travel time to markets, fisheries management, nutrients, travel time and gravity to coastal population, and tourism value. For the climate model, the SST rate of rise and bimodality had >10% relative importance. Weaker variables included coral cover, SST kurtosis, median, and skewness, and cumulative DHW. The Chaetodontidae–Scarinae species axis 2 gradient indicated human-influenced variables of population and city gravity and sediments had >10% relative importance. Weaker variables included nutrients, tourism value, travel time to markets and coastal population, and fisheries management. For the climate model, SST skewness and rate of rise had >10% relative importance. Weaker variables included cumulative DHW, SST bimodality, median, and kurtosis, and coral cover. Numbers of species of Pomacanthidae, Pomacentridae, Chaetodontidae, and Scarinae were either negatively or not correlated, which explains some of the compensatory patterns (Text S7, Fig. S5).

3.3. Boosted regression tree associations

Numbers of species and composition responses for human and climate variables often showed similar relationships with key environmental variables, but they were sometimes reversed when evaluated for community composition (Figs. 3 & 4). After accounting for biomass, patterns with depth indicated variable responses, being unimodal for total number of species, positive and saturating for the Pomacentridae–Pomacanthidae axis, and more sinusoidal for the Chaetodontidae–Scarinae axis. Coral cover was a weak variable, but total numbers of species increased with coral cover up to the maximum at 60% cover. However, the 2 community axes were more variable, with the Pomacanthidae species having peaks and troughs but a steep decline at >50% cover. The numbers of species in the Scarinae community displayed a u-shaped distribution, with their lowest numbers at ~40% cover.

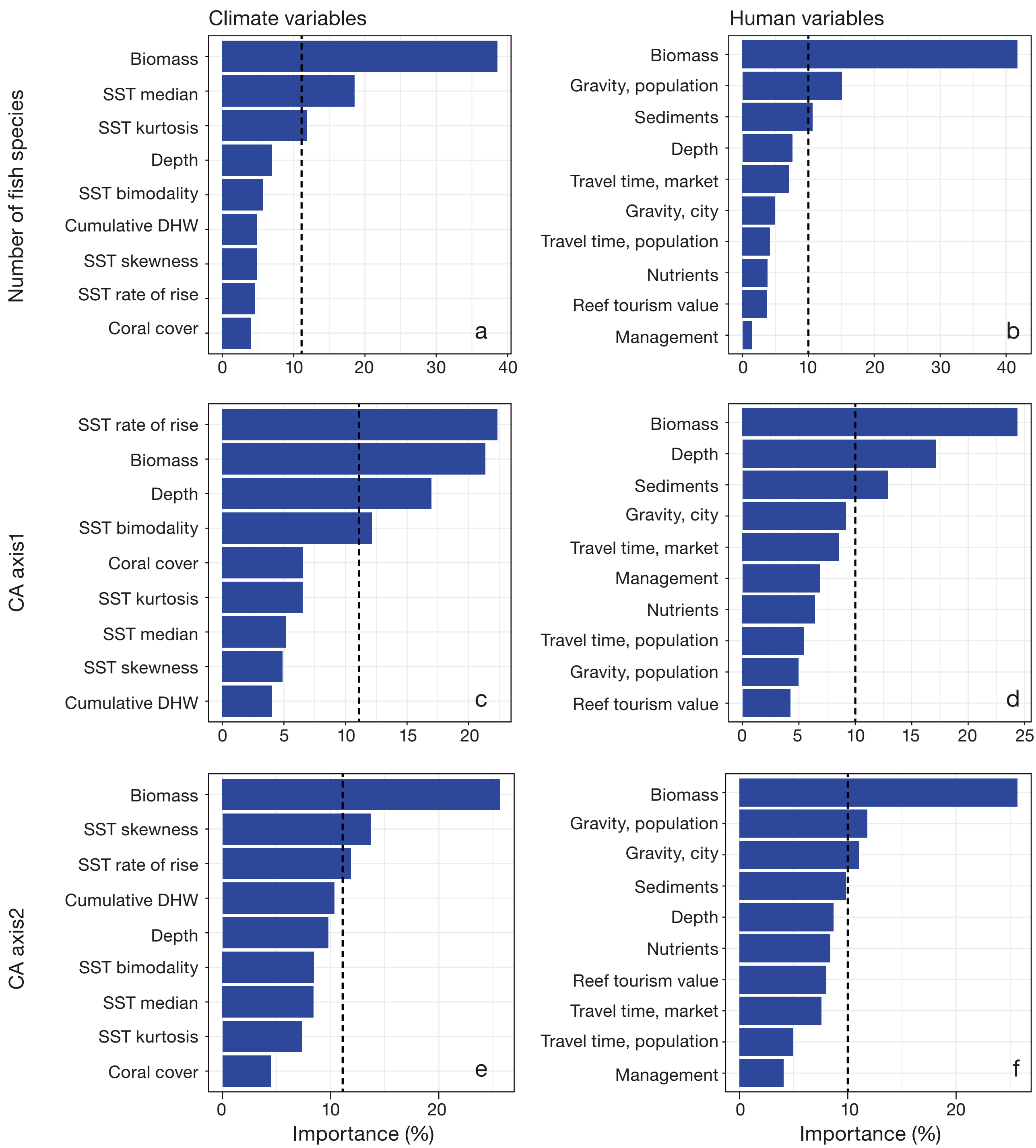
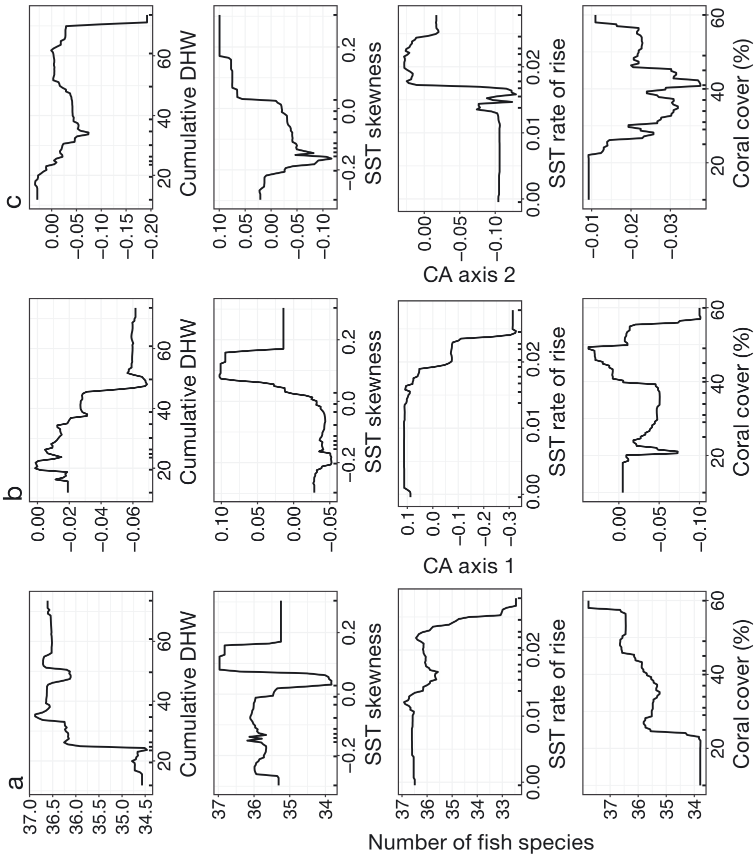


Fig. 2. Relative importance of climate- and human-influenced variables (management, usage, and water quality) for 6 separate boosted regression tree (BRT) models. Model outputs are specific to retained (a,c,e) climate- and (b,d,f) human-influenced variables. Results present the (a,b) BRT variable importance for number of species; (c,d) canonical correspondence analysis (CCA) community axis 1 (positive values are more speciose Pomacanthidae and fewer Pomacentridae-dominated communities); and (e,f) community axis 2 (positive values are more speciose Scarinae and fewer Chaetodontidae). See Table 1 for relative importance for all variables combined in a single model. Sediments and nutrients derived from satellite information is a multivariate value taken from Andrello et al. (2022). SST: sea surface temperature; cumulative DHW: cumulative degree-heating week (excess heat above summer temperature threshold); Gravity, population: gravity to nearest coastal population (nearest coastal population/(c + travel time^2)); Gravity, city: gravity to nearest city (city population/(c + travel time^2)). Vertical dashed line: hypothetical value if all variables made equal contributions to the model



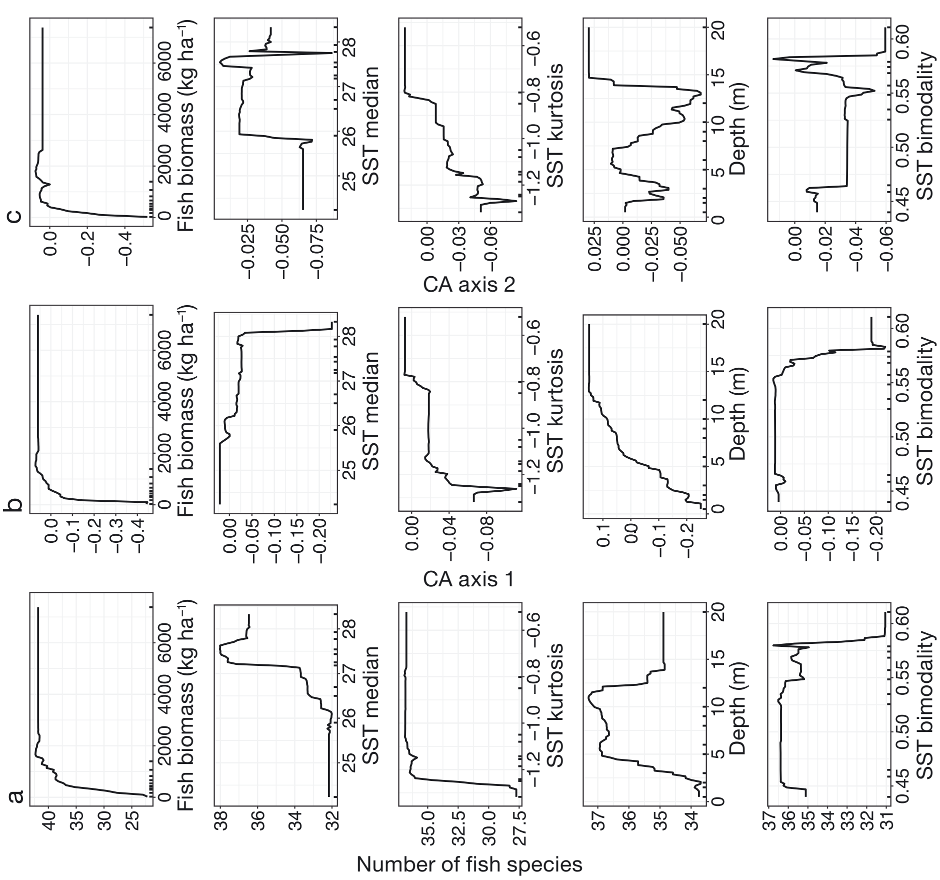


Fig. 3. Boosted regression tree (BRT) response relationships between top climate environmental variables for 3 models: (a) number of fish species, (b) canonical correspondence analysis (CCA) community axis 1 (Pomacentridae (−), Pomacanthidae (+)), and (c) community axis 2 (Chaetodontidae (−), Scarinae (+)). Panels ordered from highest to lowest relative importance based on the number of species. SST: sea surface temperature; DHW: cumulative degree-heating week temperature threshold

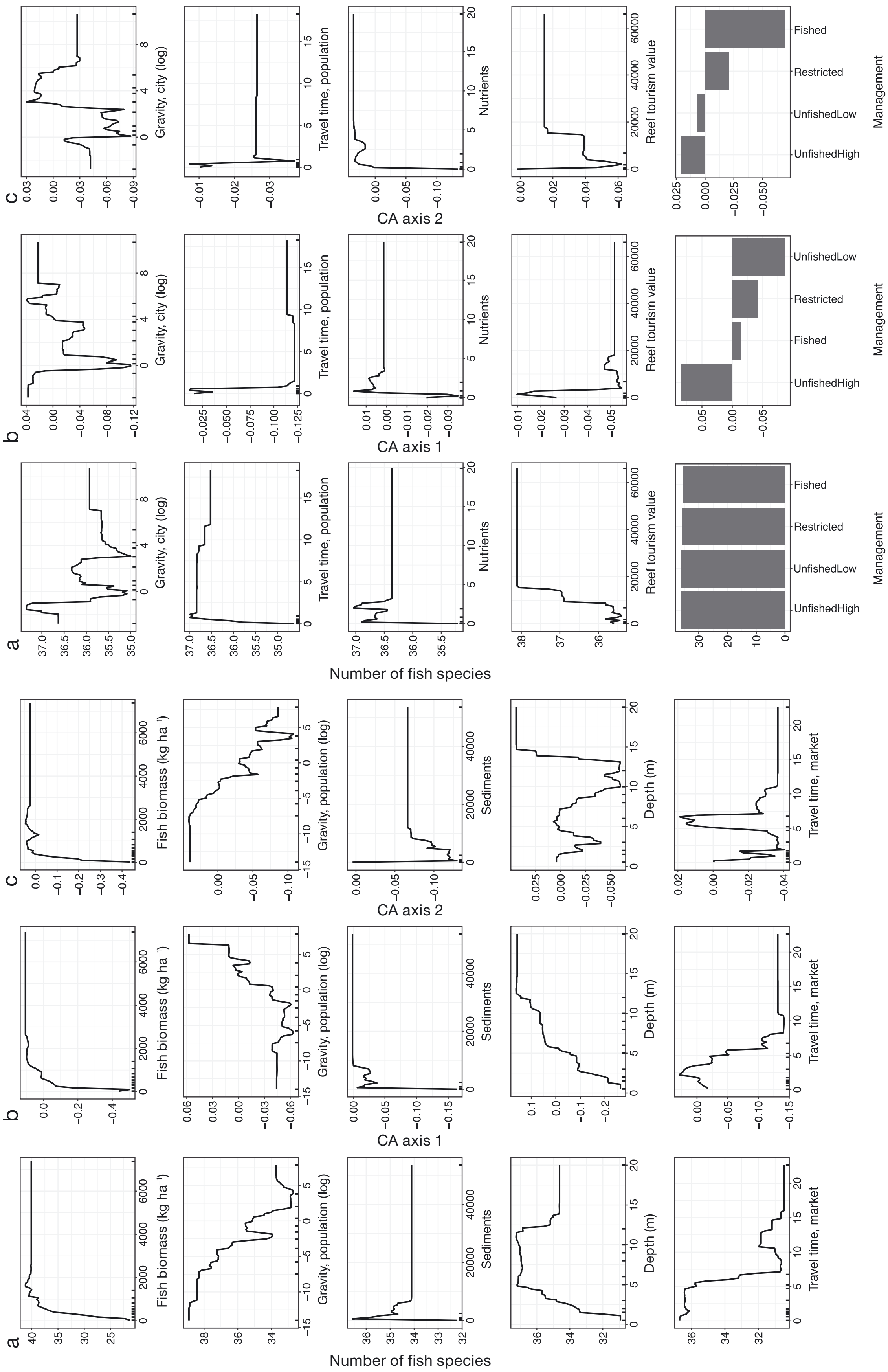


Fig. 4. Boosted regression tree (BRT) response relationships for top human-influenced variables for the 3 models. (a) BRT responses showing relationships between top human-influenced variables and number of species, (b) for community axis 1 (Pomacentridae−Pomacanthidae communities), and (c) community axis 2 (Chaetodontidae−Scarinae). Sediments and nutrients derived from satellite information is a multivariate value taken from Andrello et al. (2022). Panels ordered from highest to lowest relative importance based on the number of species

3.4. Response associations with human-influenced variables

Fisheries management partial effects were weak and displayed minor differences for numbers of species with management (Fig. 4). This suggests that biomass was the stronger driver of number of species rather than the specific management restrictions or their selection of sites. However, the species community axes responses varied with restrictions and suggest more Pomcanthidae and Scarinae species in the stricter fisheries restriction categories. After biomass, the log gravity to coastal people had the strongest effect on numbers of species. Numbers of species of Pomacanthidae increased while Scarinae numbers declined with gravity to people. Gravity to cities was weaker and showed more complex responses but also an overall decline with number of species and more u- and hump-shaped patterns for the Pomcanthidae and Scarinae communities, respectively. Sedimentation showed a peak in total numbers of species at low sediment levels and more tolerance to sediments among the Pomacanthidae community. The Scarinae community had a u-shaped response to high sediments, with a peak at very low levels and modest tolerance among some species. Partial effects of travel time to cities indicated high numbers of all species and of Pomacanthidae species at short travel times. Scarinae numbers had 2 peaks, one at a short time and another at ~5 h. Travel time to humans showed similarly high numbers at short times but the numbers of Pomacanthidae and Scarinae declined with increasing times. The biomass, gravity, and travel time results combined suggest high potential numbers of species close to shore but that human resource extraction and biomass reduction overrides this pattern. All 3 species responses showed increases and tolerance to high nutrient concentrations. Increasing reef tourism value was associated with higher numbers of species but declining Pomacanthidae numbers and a u-shaped relationship with the Scarinae numbers.

3.5. Response associations with climatic variables

Complex species relationships were found among the 8 retained climate variables that suggest multiple climatic influences on species predictions (Fig. 3). Median SST, for example, indicated that the Pomacanthidae-associated numbers of species declined above 28°C whereas Scarinae numbers increased for temperatures between 26 and 27.5°C, with some declines above 27.5°C. The kurtosis metric indicates the benefits of centrally distributed temperatures for all 3 species groups, or losses of Pomacentridae- and Chaetodontidae-associated species with chronic temperature stress (i.e. negative kurtosis). When bimodality was detectable (i.e. >0.55), the number Pomacanthidae species declined, but the Scarinae had a window where their numbers peaked with modest bimodality. Some modest positive skewness (~0.1) was positively associated with numbers of species. Yet, different patterns were found for the Pomacanthidae and Scarinae numbers: more hump-shaped for the Pomacanthidae and u-shaped for Scarinae. Partial effects of increasing cumulative DHW were associated with more species of Pomacanthidae and Scarinae until abrupt declines were observed at 45 and 70 DHW, respectively. Rapid rates of SST rise were associated with declining numbers of Pomacanthidae-associated species, whereas the numbers of Scarinae peaked at ~0.02°C increase per year.

3.6. Regional and hotspot predictions

The predicted number of proxy species for all reef cells per 500 m^2 indicated the highest numbers of species in the East African Coral Coast at 42.6 ± 3.0 (SD) and the lowest in the Mascarene Islands at 28.2 ± 2.2 (Table 2, Fig. 5a). Tanzania was the overall highest ranked nation (43.1 ± 2.4) followed by Mayotte (France) and Mozambique. Madagascar had the most reef cells (n = 2282) but was ranked ninth for numbers of species as a nation. The western and Northern Madagascar ecoregion was ranked fourth, but the total numbers of species was reduced by the seventh ranked Southeast Madagascar ecoregion. Mozambique had 3 ecoregions, and they ranked first, third, and fifth to produce a high overall national ranking. The French Southern Territories and Comoros Islands ranked fourth and fifth in terms of numbers of species but with a low number of reefs cells (n = 376). Kenya also had about the same number of reefs cells, but its 2 ecoregions ranked first and sixth, which results in a seventh overall ranking for all cells combined. Seychelles ranked second at the ecoregion level but fifth as a nation. The more peripheral countries of South Africa, Mauritius, and Reunion were ranked eighth, tenth, and eleventh, respectively. Therefore, the combined countries of Tanzania and Mozambique had the largest reef areas and highest numbers of local species.

Many of the highest-density provincial clusters of species were restricted to the African coastline and

Table 2. Predicted numbers of species per ~500 m^2 for the 6 studied families (mean, SD) by province, ecoregion, nation, and high and low number of species cluster locations. Number of cells are those that fall within the categories of province, ecoregion, nation, and hotspot classifications. See Fig. 5a for distributions at the cell level and Fig. 5b for distribution of hot- and coldspots

	Proxy for predicted number of species	Number of reef cells
Province		
Western Indian Ocean province	38.6 (5.3)	7039
Ecoregion		
East African Coral Coast	42.6 (3.0)	2743
Seychelles	39.2 (4.0)	701
Bight of Sofala/Swamp Coast	38.7 (1.8)	113
Western and Northern Madagascar	36.2 (4.9)	2854
Delagoa	35.5 (1.5)	96
Northern Monsoon Current Coast	34.9 (1.4)	130
Southeast Madagascar	32.6 (3.9)	65
Cargados Carajos/Tromelin Island	31.8 (2.9)	141
Mascarene Islands	28.2 (2.2)	196
Nation		
Tanzania	43.1 (2.4)	1524
Mayotte	41.9 (2.8)	269
Mozambique	41.8 (3.5)	1180
Comoros	40.3 (2.4)	238
French Southern Territories	40.2 (4.2)	138
Seychelles	39.2 (4.0)	701
Kenya	37.1 (3.0)	372
South Africa	36.3 (1.0)	6
Madagascar	34.7 (4.4)	2282
Reunion	27.6 (3.1)	25
Mauritius	29.5 (3.1)	304
Species hot- and coldspots		
Hotspot 99% confidence	43.0 (2.5)	3136
Hotspot 95% confidence	40.9 (2.1)	255
Hotspot 90% confidence	40.2 (1.9)	177
Average hotspots	42.3 (2.5)	3568
Not significant	38.7 (2.1)	965
Coldspot 90% confidence	36.7 (1.7)	71
Coldspot 95% confidence	36.2 (2.3)	189
Coldspot 99% confidence	32.4 (3.3)	2246
Average coldspots	32.8 (3.4)	2506

located in the Coral Coast ecoregion (Fig. 5b). These included a cluster on the Kenya–Tanzania border that extended south to include several high-density reef areas in northern Mozambique. The Western and Northern Madagascar ecoregion and the nation of Madagascar were notable for lacking diversity hotspots despite having high numbers of species per cell. Here, hotspots were limited to a longitudinal band stretching from the Comoros and Mayotte Islands to the Ankivonjy area in NW Madagascar. This resulted from the sparsity of dense neighborhoods of specious cells, which was also observed in other island nations.

4. DISCUSSION

Patterns in the numbers of fish species and community composition were predicted for 7039 modest-sized reef cells in the Western Indian Ocean Province. In general, the highest numbers of species were predicted to occur where there was high biomass (>500 kg ha^{-1}), high coral cover (>50%), slow rates of SST rise, low gravity or impact with people, a narrow range of nutrients and sediments, moderate depths (5–12 m), and median SSTs (27–28°C). Human impacts were more influential than climate impacts, particularly considering the central importance of biomass. This diversity exists in the presence of many temperature-modifying factors, including some bimodality, positive skewness, and cumulative excess heat. The community richness of the studied fish families was not positively related. Therefore, variations in biomass and environmental factors cause shifts in species composition to partially compensate and maintain high levels of total diversity. Weak or negative correlations between Pomacentridae and Pomacanthidae and Scarinae and Chaetodontidae species numbers were responsible for this species compensation. Species niches and several environmental changes were often modified by biomass and depth, indicating the important roles of the taxa's life histories for community change and adaptation.

4.1. Provincial patterns

The highest numbers of fish species occurred along the continental band of the East African Coral Coast that included southern Kenya, Tanzania, and northern Mozambique. Both the numbers of species proxy and high-density clusters of reefs were high. Consequently, the top 5–10% of species richness cells largely indicated a single ecoregion with a few additional locations in northwest Madagascar, Mayotte, Comoros, southern Kenya, and Seychelles. The ecoregions to the north on the African continent, or

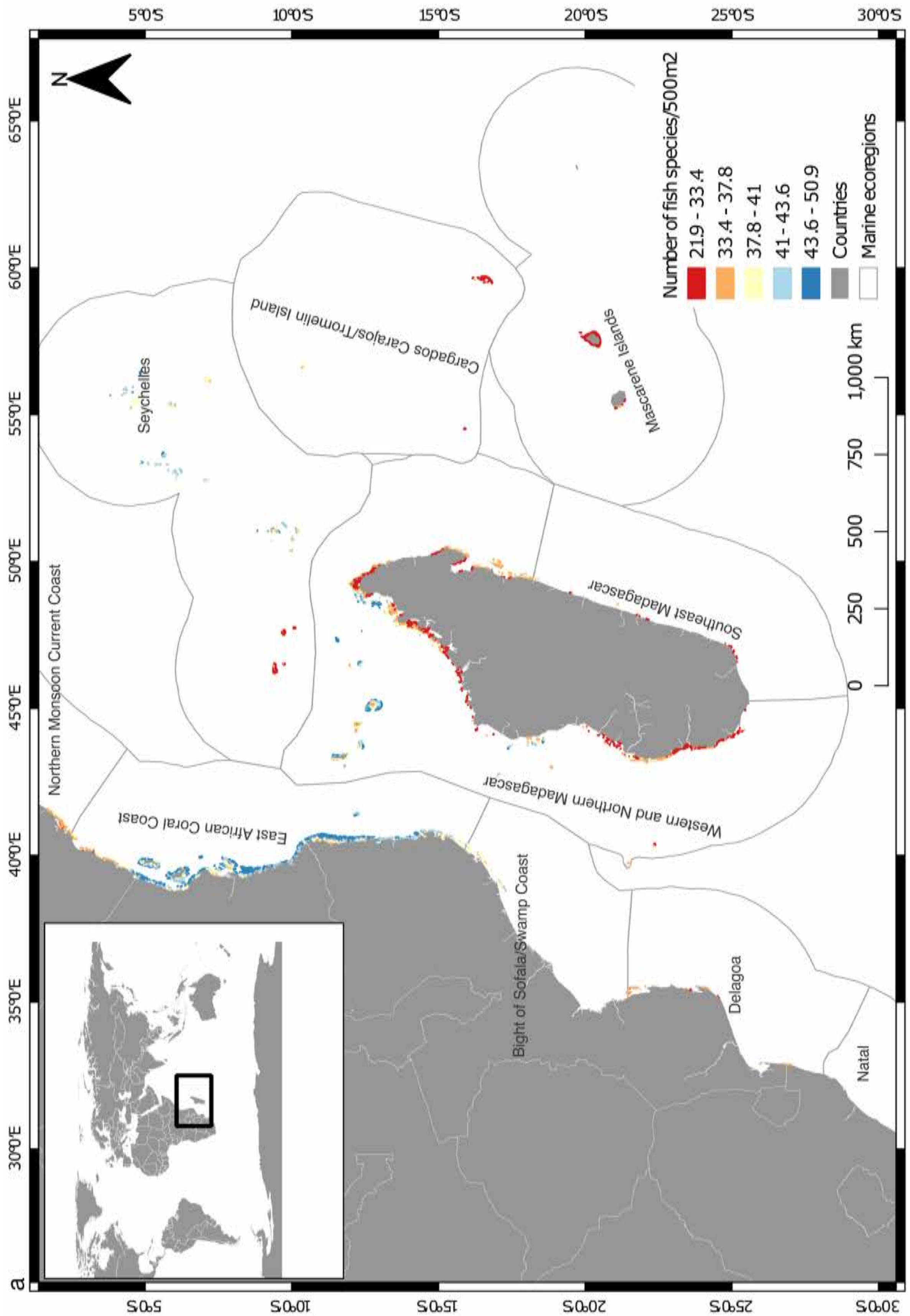


Fig. 5. Regional maps of the predicted numbers of (a) local fish species and (b) location of hot- and coldspots for species numbers based on the 6 proxy families

Figure continues on next page

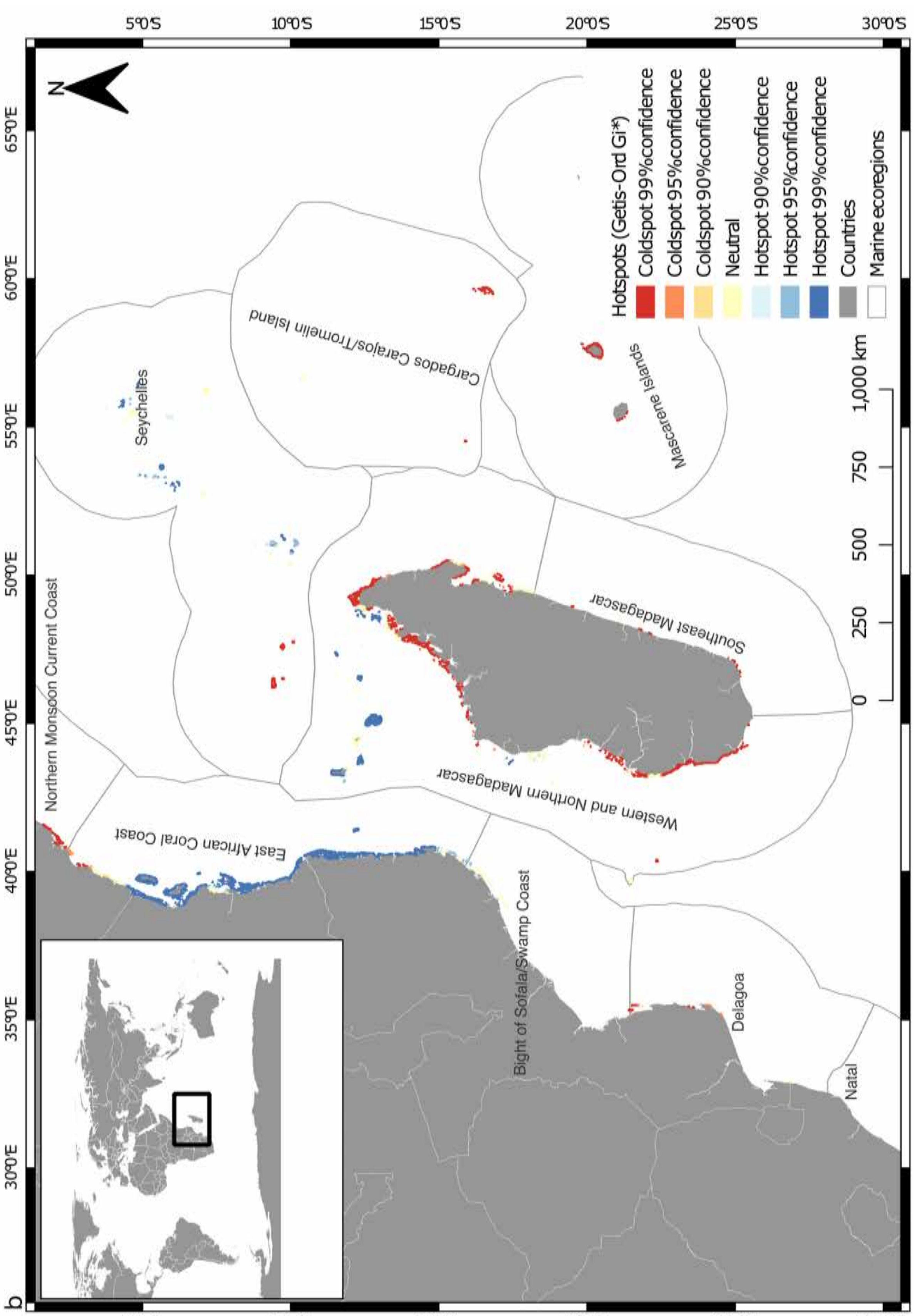


Fig. 5 (continued)

the Monsoon Coast, and south, or Bight of Sofala and Delagoa ecoregions, had high numbers of species and high similarities with the Coral Coast. Previous studies have identified some of these areas, such as the Lamu Archipelago, Pemba Channel, Tanga-Pangani, Mafia-Ruvuma, Mtwara-Quirimbas, and Bazaruto-Tofo, as seascapes of high or unique biodiversity for conservation efforts (van der Elst & Everett 2015). However, the attributes used to select these conservation priorities were charismatic and visible species, including nesting turtles and birds, coelacanths, and marine mammals.

The Western and Northern Madagascar ecoregion and the islands of Comoros, Mayotte, and French Southern Territories represent an off-continent extension of this continental diversity. This appears to create a stepping-stone of islands of diversity and similarity extending to northwest Madagascar (i.e. Mitsio Islands). The ecoregions furthest from the continent or the Mascarene and Cargados Carajos Islands and Southeast Madagascar appeared to be a peripheral fauna, as evidenced by lower numbers of species and lower similarity with the central ecoregions on the African continent. The Mascarene Islands are known for their isolation and high levels of endemism, particularly among the Pomacentridae (Roberts et al. 2003). In the well-sampled Reunion Island, for example, 2.6% of species recorded were endemic, and Mauritius also has several island-specific and ecoregional endemics (Fricke et al. 2009, McClanahan et al. 2022). Predictions for the 3 unsampled ecoregions (Cargados Carajos/Tromelin Island, Southeast Madagascar, and Bight of Sofala/Swamp Coast) should be seen as provisional estimates until field sampling is conducted.

Previous studies using rarefaction methods for coral taxa identified the northern Mozambique Channel for high within- and between-site diversity (Obura et al. 2012, Ateweberhan & McClanahan 2016). This result is confirmed here for fish species, but more species were predicted in the northern than the southern Mozambique Channel (Chabanet et al. 2016), notably in Mayotte, where 759 marine fish species have been recorded for a small volcanic island (Wickel et al. 2014). Regional currents that transport larvae have been shown to facilitate inter-reef connectivity in the north (Crochelet et al. 2016, Maina et al. 2020). Previous predictions for fish have been at a coarse scale and not clearly associated with environmental conditions (Jenkins & Van Houtan 2016, Bullock et al. 2021). Some patterns here suggest that environmental factors are more important than ecoregion, biogeographic distance, size, and connectivity metrics. Therefore, failure to use finer-scale environmental information is expected to produce coarse and imprecise biodiversity and regionalization mapping. For example, there was a high overall ANOSIM similarity between the Mascarene Islands and Delagoa, despite their large distance apart and locations on the opposite sides of Madagascar. Therefore, environmental conditions may be more important for faunal similarity than distance and connectivity metrics.

4.2. Human-influenced variables

Biomass, which is largely expected to represent fishing pressure, was the single overriding predictor of numbers of species and taxonomic composition. The other evaluated variables, such as distance to people, travel times, water quality, tourism economic value, and management, had lesser predictive influence. Studies of the distribution of biomass and environmental factors in this province have found that depth, distance to deep water, travel time, management, nation, SST metrics, reef area, and net primary productivity are the major predictors of biomass (McClanahan et al. 2016, 2023, McClanahan 2019). However, biomass is accounted for or normalized here by holding it constant and using partial effect methods when making the broader distributional predictions and comparisons of numbers of species and community structure. Overall, managing biomass is likely to be one of the most effective approaches to protect fish biodiversity regardless of the specific means to accomplish it. Methods to manage biomass are likely to vary with the social and governance aspects of the jurisdictions and various aspects of culture, the fisheries management authority, and human dependence on fish resources (McClanahan & Abunge 2016). The travel time partial effect patterns suggest that in the absence of human pressures, the highest diversity would be located close to shore.

4.3. Climate-influenced variables

After accounting for biomass and depth, the strongest climate predictors of our species proxy were median SST and metrics of SST variability (kurtosis, bimodality, cumulative excess heat, skewness, and rate of rise). These variables are often associated with climate-change studies of coral bleaching and changes in coral cover (McClanahan 2022a). Here, coral cover is subordinate to the temperature variability metrics, which suggests a direct rather than an indirect effect of temperature on fish via coral. Moreover, variability metrics of kurtosis, bimodality, and skewness are influenced by locations and modify climate change impacts on corals in this province (McClanahan & Azali 2021). Studies of coral—environment interactions suggest complex taxa—environment interaction responses expected to create heterogeneous responses among reef species along the African coastline (McClanahan et al. 2020). A fisheries management implication of these rankings and future climate models is that protection of biomass >500 kg ha^{-1} could partially compensate for the near-term climate-change impacts on fishes. In general, fishes appeared to respond to environmental stress and coral losses with shifts in species composition that can compensate to stabilize overall diversity and ecological processes (Chabanet 2002, McClanahan et al. 2014).

4.4. Community composition

Patterns reflected in species composition sometimes differed from our total numbers of species proxy. Fish communities differed based on their exposures to the centralization and bimodal aspects of the SSTs, rare warm and cold water, and rates of SST rise. Specifically, the numbers of species in the Pomacanthidae community were associated with more centralized SST distributions, slow rate of rise, and occasional weak spikes of warm water, while the Scarinae species numbers showed less centralized distributions, associated with higher rates of SST rise, and occasional cold and hot water. Relationships among these communities and coral cover were complex, with peak numbers of Pomacanthidae species at ~50% cover and a low trough for the Scarinae at ~40% cover. Therefore, there are likely to be changes in species composition within these taxonomic groups as benthic cover and reef calcification rates change. Species composition also changed with depth and temperature variability metrics. Specifically, species of Pomacanthidae increased while the Pomacentridae decreased with depth and Chaetodontidae declined and Scarinae increased with median temperature. For the Chaetodontidae—Scarinae community axis, SST rate of rise was strong and negative for the Pomacanthidae, while the Scarinae community showed more tolerance (Table 1). Similar patterns were seen for cumulative excess heat (DHW), which is the dominant variable used to predict climate change. It was shown here to be a weak predictor for fish, and species appear to tolerate moderate to high levels of excess heat. Nevertheless, reefs undergoing rapid temperature rises and experiencing excess heat are predicted to change their species composition towards fewer Pomacanthidae and Chaetodontidae and more Pomacentridae and Scarinae species. An increase in herbivorous fish catch has been documented from studies of climate-degraded Seychellois reefs (Robinson et al. 2019). Complex interactions among species were evident between depth, temperature variability metrics, human resource extraction, and coral, thus making predictions based on a single climate change variable problematic.

4.5. Conclusions

Our study provided a more spatially refined view of provincial fish biodiversity than previous global, regional, and conservation prioritization efforts. Past efforts have used presence/absence data and overlapping polygons or rarefaction methods (Selig et al. 2014, Jenkins & Van Houtan 2016). These methods are sensitive to the density of data, and this affects subsequent spatial patterns and overlap in species distributions, particularly when identifying centers of biodiversity (Kusumoto et al. 2020). The environmental modeling methods used here provided an alternative approach by accounting for habitats and environmental processes and their variability at smaller scales. Environmental modeling methods should therefore provide information to locate small to modest-sized areas for conservation, often preferred or politically feasible for protected area designations (Boonzaier & Pauly 2016, McClanahan 2023). Spatial environmental predictions are useful for planning and should eventually augment the current practices of mapping diversity from sparse presence/absence and extrapolated species range information. The cellular map predictions combined with the hotspot algorithm provided further information to identify reef neighborhoods of high numbers of species and not just local diversity. High fish diversity, reef clusters, and connectivity were positively associated and likely to interact to maintain high numbers of species. Clustering of reefs and common biogeographic factors, such as continents, islands, reef area, and retention connectivity explained some of the variability (Crochelet et al. 2016, Maina et al. 2020). However, the local-scale environmental and human resource use factors were among the strongest predictors.

Species changes, compensation, and high similarity among ecoregions suggests a high resilience in numbers of species. The remote Mascarene and

Carajos/Tromelin ecoregions would be exceptions to species resilience in this faunal province. Some past studies suggest that climate change impacts may be equal or less influential on fish species than ecological changes attributed to reductions in biomass by fishing (McClanahan & Muthiga 2016, Fredston et al. 2023). Fortunately, long-term fisheries yields are maximized above or close to these biomass–ecological change points, which, if achieved, can reduce stark yield gains–biodiversity loss tradeoffs (McClanahan 2018, 2022b). To date, prioritizations in this region for protected area investment has focused on the distribution of charismatic species or connectivity (van der Elst & Everett 2015, Crochelet et al. 2016, Maina et al. 2020). Yet, the relationship between connectivity or the presence of charismatic species with subtidal taxa of economic importance has not been evaluated. Connectivity appeared to contribute to numbers of species but was subordinate to several other environmental factors. The fish assemblage and environmental association undertaken here provides additional information for conservation decision making. The results indicate the importance of management of fisheries stocks near or above maximum sustained yield thresholds to maintain biodiversity and community resilience at large scales.

Data availability. Data are available at https://knb.ecoinformatics.org/view/doi:10.5063/F19885HT.

Acknowledgements: Research undertaken to support this science was made possible by the generous support of the American people through the United States Agency for International Development (USAID) and the US Department of the Interior's International Technical Assistance Program (DOI-ITAP). The findings do not necessarily reflect the views of the USAID or the US Government. The original fieldwork was supported by the Wildlife Conservation Society, USA, through grants from John D. and Catherine T. MacArthur, The Tiffany and Co. Foundations, and the Bloomberg Vibrant Ocean's Initiative. Kenya's Commission for Science, Technology, and Innovation and Kenya Wildlife Services approved the research. The Reunion Marine Reserve and TAAF supported the field monitoring (via France and Europe) and the scientific experts who participated in the fish sampling (C. Cadet, P. Durville, T. Mulochau, E. Tessier, T. Rungassamy). We appreciate the logistical support of J. Maina and R. Oddenyo, and data contributions from B. Wendling, Stephanie D'agata, N.A.J. Graham, and Arielle Hoamby.

LITERATURE CITED

Allen GR, Werner TB (2002) Coral reef fish assessment in the 'coral triangle' of southeastern Asia. Environ Biol Fishes 65:209–214

Andrello M, Darling ES, Wenger A, Suárez-Castro AF, Gelfand S, Ahmadia GN (2022) A global map of human pressures on tropical coral reefs. Conserv Lett 15:e12858

Araújo MB, New M (2007) Ensemble forecasting of species distributions. Trends Ecol Evol 22:42–47

Assis J, Tyberghein L, Bosch S, Verbruggen H, Serrão EA, De Clerck O (2018) Bio-ORACLE v2.0: extending marine data layers for bioclimatic modelling. Glob Ecol Biogeogr 27:277–284

Ateweberhan M, McClanahan TR (2010) Relationship between historical sea-surface temperature variability and climate change-induced coral mortality in the western Indian Ocean. Mar Poll Bull 60:964–970

Ateweberhan M, McClanahan TR (2016) Partitioning scleractinian coral diversity across reef sites and regions in the Western Indian Ocean. Ecosphere 7:e01243

Ban NC, Darling ES, Gurney GG, Friedman W and others (2023) Effects of management objectives and rules on marine conservation outcomes. Conserv Biol 37:e14156

Boonzaier L, Pauly D (2016) Marine protection targets: an updated assessment of global progress. Oryx 50:27–35

Bullock R, Ralph GM, Stump E, Al Abdali F and others (2021) Conservation status of marine biodiversity of the Western Indian Ocean. IUCN, Gland. http://eprints.cmfri.org.in/id/eprint/15157

Burke L, Reytar K, Spalding M, Perry A (2011) Reefs at risk revisited. World Resources Institute, Washington, DC

Ceccarelli DM, Evans RD, Logan M, Jones GP and others (2023) Physical, biological and anthropogenic drivers of spatial patterns of coral reef fish assemblages at regional and local scales. Sci Total Environ 904:166695

Chabanet P (2002) Coral reef fish communities of Mayotte (western Indian Ocean) two years after the impact of the 1998 bleaching event. Mar Freshw Res 53:107–114

Chabanet P, Bigot L, Nicet JB, Durville P and others (2016) Coral reef monitoring in the Iles Eparses, Mozambique Channel (2011–2013). Acta Oecol 72:62–71

Cinner JE, Daw TM, McClanahan TR, Muthiga N and others (2012) Transitions toward co-management: the process of marine resource management devolution in three east African countries. Glob Environ Change 22:651–658

Crochelet E, Roberts J, Lagabrielle E, Obura D, Petit M, Chabanet P (2016) A model-based assessment of reef larvae dispersal in the Western Indian Ocean reveals regional connectivity patterns — potential implications for conservation policies. Reg Stud Mar Sci 7:159–167

EAME (WWF Eastern African Marine Ecoregion) (2004) Towards a Western Indian Ocean dugong conservation strategy: the status of dugongs in the Western Indian Ocean region and priority conservation actions. WWF, Dar Es Salaam

ESRI (2022a) Optimized hot spot analysis (spatial statistics). ArcGIS Pro 3.0. ESRI, Redlands, CA

ESRI (2022b) How optimized hot spot analysis works. ArcGIS Pro 3.0. ESRI, Redlands, CA

Fontoura L, D'agata S, Gamoyo M, Barneche DR and others (2022) Protecting connectivity promotes successful biodiversity and fisheries conservation. Science 375:336–340

Fredston AL, Cheung WWL, Frölicher TL, Kitchel ZJ and others (2023) Marine heatwaves are not a dominant driver of change in demersal fishes. Nature 621:324–329

Fricke R, Mulochau T, Durville P, Chabanet P, Tessier E, Letourneur Y (2009) Annotated checklist of the fish species (Pisces) of La Réunion, including a Red List of threatened and declining species. Stuttg Beitr Natkd A Neue Ser 2:1–168

Friedlander AM, Obura D, Aumeeruddy R, Ballesteros E, Church J, Cebrian E, Sala E (2014) Coexistence of low coral cover and high fish biomass at Farquhar Atoll, Seychelles. PLOS ONE 9:e87359
Gallo ND, Beckwith M, Wei CL, Levin LA, Kuhnz L, Barry JP (2020) Dissolved oxygen and temperature best predict deep-sea fish community structure in the Gulf of California with climate change implications. Mar Ecol Prog Ser 637:159−180
Gotelli NJ, Colwell RK (2001) Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. Ecol Lett 4:379−391
Graham NAJ, Robinson JPW, Smith SE, Govinden R, Gendron G, Wilson SK (2020) Changing role of coral reef marine reserves in a warming climate. Nat Commun 11: 2000
Jenkins CN, Van Houtan KS (2016) Global and regional priorities for marine biodiversity protection. Biol Conserv 204:333−339
Kuhn M, Johnson K (2013) Applied predictive modeling. Springer, New York, NY
Kusumoto B, Costello MJ, Kubota Y, Shiono T, Wei CL, Yasuhara M, Chao A (2020) Global distribution of coral diversity: biodiversity knowledge gradients related to spatial resolution. Ecol Res 35:315−326
Lee-Yaw JA, McCune JL, Pironon S, Sheth SN (2022) Species distribution models rarely predict the biology of real populations. Ecography 2022:e05877
Maina J, McClanahan TR, Venus V, Ateweberhan M, Madin J (2011) Global gradients of coral exposure to environmental stresses and implications for local management. PLOS ONE 6:e23064
Maina JM, Gamoyo M, Adams VM, D'agata S, Bosire J, Francis J, Waruinge D (2020) Aligning marine spatial conservation priorities with functional connectivity across maritime jurisdictions. Conserv Sci Pract 2:e156
Maire E, Cinner J, Velez L, Huchery C and others (2016) How accessible are coral reefs to people? A global assessment based on travel time. Ecol Lett 19:351−360
McClanahan TR (1994) Kenyan coral reef lagoon fish: effects of fishing, substrate complexity, and sea urchins. Coral Reefs 13:231−241
McClanahan TR (2018) Multicriteria estimate of coral reef fishery sustainability. Fish Fish 19:807−820
McClanahan TR (2019) Coral reef fish communities, diversity, and their fisheries and biodiversity status in East Africa. Mar Ecol Prog Ser 632:175−191
McClanahan TR (2020) Coral community life histories and population dynamics driven by seascape bathymetry and temperature variability. Adv Mar Biol 87:291−330
McClanahan TR (2022a) Coral responses to climate change exposure. Environ Res Lett 17:073001
McClanahan TR (2022b) Fisheries yields and species declines in coral reefs. Environ Res Lett 17:044023
McClanahan TR (2023) Local heterogeneity of coral reef diversity and environmental stress provides opportunities for small-scale conservation. Divers Distrib 29: 1324−1340
McClanahan TR, Abunge CA (2016) Perceptions of fishing access restrictions and the disparity of benefits among stakeholder communities and nations of south-eastern Africa. Fish Fish 17:417−437
McClanahan TR, Azali MK (2021) Environmental variability and threshold model's predictions for coral reefs. Front Mar Sci 8:778121
McClanahan TR, Muthiga NA (2016) Similar impacts of fishing and environmental stress on calcifying organisms in Indian Ocean coral reefs. Mar Ecol Prog Ser 560: 87−103
McClanahan TR, Graham NAJ, Darling ES (2014) Coral reefs in a crystal ball: predicting the future from the vulnerability of corals and reef fishes to multiple stressors. Curr Opin Environ Sustain 7:59−64
McClanahan TR, Graham NAJ, MacNeil MA, Cinner JE (2015) Biomass-based targets and the management of multispecies coral reef fisheries. Conserv Biol 29: 409−417
McClanahan TR, Darling ES, Maina JM, Muthiga NA and others (2020) Highly variable taxa-specific coral bleaching responses to thermal stress. Mar Ecol Prog Ser 648: 135−151
McClanahan TR, Friedlander AM, Wantiez L, Graham NAJ, Bruggemann JH, Chabanet P, Oddenyo RM (2022) Best-practice fisheries management associated with reduced stocks and changes in life histories. Fish Fish 23:422−444
McClanahan TR, D'Agata S, Graham NAJ, Kodia MA, Maina JM (2023) Multivariate environment-fish biomass model informs sustainability and lost income in Indian Ocean coral reefs. Mar Policy 152:105590
Mokany K, Ware C, Woolley SN, Ferrier S, Fitzpatrick MC (2022) A working guide to harnessing generalized dissimilarity modelling for biodiversity analysis and conservation assessment. Glob Ecol Biogeogr 31:802−821
Obura DO, Church JE, Gabrié C (2012) Assessing Marine World Heritage from an ecosystem perspective: the Western Indian Ocean. United Nations Education SaCOU, Paris
Ord JK, Getis A (1995) Local spatial autocorrelation statistics: distributional issues and an application. Geogr Anal 27:286−306
Ord JK, Getis A (2001) Testing for local spatial autocorrelation in the presence of global autocorrelation. J Reg Sci 41:411−432
Parravicini V, Kulbicki M, Bellwood DR, Friedlander AM and others (2013) Global patterns and predictors of tropical reef fish species richness. Ecography 36:1254−1262
Pellissier L, Leprieur F, Parravicini V, Cowman PF and others (2014) Quaternary coral reef refugia preserved fish diversity. Science 344:1016−1019
Pilowsky JA, Colwell RK, Rahbek C, Fordham DA (2022) Process-explicit models reveal the structure and dynamics of biodiversity patterns. Sci Adv 8:eabj2271
Pinheiro JPS, Windsor FM, Wilson RW, Tyler CR (2021) Global variation in freshwater physico-chemistry and its influence on chemical toxicity in aquatic wildlife. Biol Rev Camb Philos Soc 96:1528−1546
Roberts CM, Andelman S, Branch G, Bustamante RH and others (2003) Ecological criteria for evaluating candidate sites for marine reserves. Ecol Appl 13(Suppl): S199−S214
Robinson JPW, Wilson SK, Robinson J, Gerry C and others (2019) Productive instability of coral reef fisheries after climate-driven regime shifts. Nat Ecol Evol 3:183−190
Rocliffe S, Peabody S, Samoilys M, Hawkins JP (2014) Towards a network of locally managed marine areas (LMMAs) in the Western Indian Ocean. PLOS ONE 9: e103000
Safaie A, Silbiger NJ, McClanahan TR, Pawlak G and others (2018) High frequency temperature variability reduces the risk of coral bleaching. Nat Commun 9:1671

Selig ER, Turner WR, Troëng S, Wallace BP and others (2014) Global priorities for marine biodiversity conservation. PLOS ONE 9:e82898

Stephenson F, Goetz K, Sharp BR, Mouton TL and others (2020) Modelling the spatial distribution of cetaceans in New Zealand waters. Divers Distrib 26:495–516

Tyberghein L, Verbruggen H, Pauly K, Troupin C, Mineur F, De Clerck O (2012) Bio-ORACLE: a global environmental dataset for marine species distribution modelling. Glob Ecol Biogeogr 21:272–281

UNEP-Nairobi Convention, WIOMSA (2021) Western Indian Ocean Marine Protected Areas Outlook: towards achievement of the Global Biodiversity Framework Targets. United Nations Environment Programme/Nairobi Convention Secretariat, Nairobi

van der Elst R, Everett B (eds) (2015) Offshore fisheries of the Southwest Indian Ocean: their status and the impact on vulnerable species. Oceanographic Research Institute (ORI) and the Western Indian Ocean Marine Sciences Association (WIOMSA), Durban

Wickel J, Jamon A, Pinault M, Durville P, Chabanet P (2014) Composition et structure des peuplements ichtyologiques marins de l'île de Mayotte (sud-ouest de l'océan Indien). Cybium Rev Int Ichtyol 38:179–203

Williams GJ, Gove JM, Eynaud Y, Zgliczynski BJ, Sandin SA (2015) Local human impacts decouple natural biophysical relationships on Pacific coral reefs. Ecography 38: 751–761

Yeager LA, Deith MC, McPherson JM, Williams ID, Baum JK (2017) Scale dependence of environmental controls on the functional diversity of coral reef fish communities. Glob Ecol Biogeogr 26:1177–1189

Zeller D, Vianna GMS, Ansell M, Coulter A and others (2021) Fishing effort and associated catch per unit effort for small-scale fisheries in the Mozambique Channel region: 1950–2016. Front Mar Sci 8:707999

Editorial responsibility: Myron Peck, Den Burg, The Netherlands
Reviewed by: 3 anonymous referees

Submitted: June 14, 2023
Accepted: January 22, 2024
Proofs received from author(s): March 4, 2024