

Multi-scale marine biodiversity patterns inferred efficiently from habitat image processing

CAMILLE MELLIN,^{1,2,6} LAEL PARROTT,³ SERGE ANDRÉFOUËT,⁴ COREY J. A. BRADSHAW,^{2,5} M. AARON MACNEIL,¹
AND M. JULIAN CALEY¹

¹Australian Institute of Marine Science, PMB No. 3, Townsville MC, Townsville, Queensland 4810 Australia

²The Environment Institute and School of Earth and Environmental Sciences, University of Adelaide, South Australia 5005 Australia

³Complex Systems Laboratory, Department of Geography, University of Montreal, C.P. 6128 Succursale Centre-Ville, Montreal, Quebec H3C 3J7 Canada

⁴Institut de Recherche pour le Développement, UR 227 COREUS 2, BP A5, 98848 Nouméa, New Caledonia

⁵South Australian Research and Development Institute, P.O. Box 120, Henley Beach, South Australia 5022 Australia

Abstract. Cost-effective proxies of biodiversity and species abundance, applicable across a range of spatial scales, are needed for setting conservation priorities and planning action. We outline a rapid, efficient, and low-cost measure of spectral signal from digital habitat images that, being an effective proxy for habitat complexity, correlates with species diversity and requires little image processing or interpretation. We validated this method for coral reefs of the Great Barrier Reef (GBR), Australia, across a range of spatial scales (1 m to 10 km), using digital photographs of benthic communities at the transect scale and high-resolution Landsat satellite images at the reef scale. We calculated an index of image-derived spatial heterogeneity, the mean information gain (MIG), for each scale and related it to univariate (species richness and total abundance summed across species) and multivariate (species abundance matrix) measures of fish community structure, using two techniques that account for the hierarchical structure of the data: hierarchical (mixed-effect) linear models and distance-based partial redundancy analysis. Over the length and breadth of the GBR, MIG alone explained up to 29% of deviance in fish species richness, 33% in total fish abundance, and 25% in fish community structure at multiple scales, thus demonstrating the possibility of easily and rapidly exploiting spatial information contained in digital images to complement existing methods for inferring diversity and abundance patterns among fish communities. Thus, the spectral signal of unprocessed remotely sensed images provides an efficient and low-cost way to optimize the design of surveys used in conservation planning. In data-sparse situations, this simple approach also offers a viable method for rapid assessment of potential local biodiversity, particularly where there is little local capacity in terms of skills or resources for mounting in-depth biodiversity surveys.

Key words: biodiversity; coral reef fish; ecological indicators; Great Barrier Reef, Australia; Landsat; mean information gain; multilevel mixed-effects model; photography; remote sensing; spectral signal.

INTRODUCTION

Rapid global change and the widespread decline of marine resources worldwide argue for developing cost-effective predictors of marine biodiversity as tools for conservation planning and prioritization (Beger and Possingham 2008, Dalleau et al. 2010), particularly in highly threatened environments such as coral reefs (MacNeil et al. 2010, Mellin et al. 2010a, b). Among potential predictors of species diversity, structurally complex environments should provide greater opportunities for variation and adaptation according to the niche theory (Huston 1979, Levin 1999). This link between species diversity and habitat complexity, however, has been difficult to quantify in benthic

environments, mostly due to problems in measuring structural complexity among different habitats. A heterogeneous environment should give rise to a greater variety of niches and thus favor the establishment and maintenance of a greater diversity of species, particularly specialists, and this is the case on coral reefs (Caley and St John 1996, Halford and Caley 2009). Coral carbonate structure provides many crevices and shelter opportunities, as well as large gradients of environmental conditions in light, depth, temperature, and hydrodynamic exposure. This structural complexity of coral habitats is closely linked with fish species diversity at multiple spatial scales and distant locations (e.g., Sale and Douglas 1984, Caley and St John 1996, Chittaro 2004). However, measuring reef complexity using traditional in situ methods is time-consuming and tends to be locally focused (e.g., Frost et al. 2005), limiting the application of such methods over broad biogeographic scales.

Manuscript received 23 November 2011; accepted 7 December 2011. Corresponding Editor: V. C. Radeloff.

⁶ E-mail: c.mellin@aims.gov.au

The habitat complexity–diversity hypothesis has been tested at the site scale (Luckhurst and Luckhurst 1978, Friedlander and Parrish 1998, Attrill et al. 2000) using three-dimensional physical descriptors of reef structure and heterogeneity such as substratum rugosity, topography, and the distribution of hole size and volume, all of which are measured in situ. At broader spatial scales, high-resolution remote-sensing methods are potentially useful, but are limited in their ability to describe habitat complexity as accurately as field-based surveys (e.g., Kuffner et al. 2007). Such limitations, as well as their high costs, have limited their routine use. Therefore, rapidly deployable, cross-scale, and cost-effective proxies for habitat complexity still need to be developed. Here we explore the use of imagery for developing such metrics that effectively characterize the complexity of two-dimensional reef habitats and are applicable across multiple scales. Our approach is distinct from previous methods in using compressed imagery that does not require extensive and complicated processing.

The development of both in situ digital photography and remote sensing has increased the availability of habitat maps for species distribution modeling. In situ images are typically used to define a typology of habitats relevant for understanding species distributions at scales < 100 m (Dumas et al. 2009), whereas remotely sensed images are used to map these habitats and associated species at scales > 100 m (e.g., Mattio et al. 2008). These habitat maps are increasingly used as spatially explicit layers in habitat suitability models, whereby biodiversity metrics (e.g., species richness, abundance, functional groups) can be mapped indirectly (Garza-Pérez et al. 2004, Mellin et al. 2007). However, these methods are time-consuming, often requiring some form of field validation, are scale-specific, and are rarely transferrable across different study areas. Such nontransferability clearly limits regional and global comparisons (Mellin et al. 2009). In addition, most of the cost of habitat mapping lies in labeling, whereby clusters of same-colored pixels are labeled according to habitat name or some other benthic property such as complexity (Andréfouët 2008). Avoiding the need for qualitative and quantitative labeling would greatly reduce the cost of using such images. Here we present a scale-independent method that does not require labeling, but is instead based directly on image metrics computed automatically by measuring the spatial heterogeneity (e.g., texture) of the signal (radiance or reflectance) in different spectral bands.

Digital images are composed of two interdependent characteristics: tone (i.e., spectral information) and texture (i.e., tonal variability in a given area; Baraldi and Parmiggiani 1995). Therefore, the texture of an image contains valuable information about the spatial and structural arrangements of the objects it represents (St-Louis et al. 2006). Because such images incorporate within-habitat heterogeneity (as opposed to discrete habitat categories) and minimize errors associated with

boundary delineation (Andréfouët et al. 2000, St-Louis et al. 2006), an image's heterogeneity might be a reasonable estimate of habitat heterogeneity at any spatial scale. Therefore, if image-derived spatial heterogeneity is directly linked to the complexity and heterogeneity of habitats across multiple scales, greater species diversity would be expected in areas of greater image heterogeneity.

Few studies to date have demonstrated positive relationships between image-derived spatial heterogeneity and biodiversity metrics, with the exception of, for example, the diversity of vascular plants in temperate forests at scales < 100 m (Proulx and Parrott 2008, 2009), and birds (St-Louis et al. 2006, Bellis et al. 2008), mammals (Estes et al. 2008), and vascular plants (Oldeland et al. 2010) at scales > 100 m. Although these studies' results varied depending on the ecosystem, species, and image-based metrics used, there was a relationship between species diversity and habitat characteristics derived from imagery at relatively small spatial scales (e.g., 57–63% of variation explained in bird species richness in semiarid habitats [St-Louis et al. 2006]; 27% of bongo antelope *Tragelaphus euryceros isaaci* habitat selection [Estes et al. 2008]). To our knowledge, no similar study has been applied to marine ecosystems; as such, it remains unknown whether, how well, and at what scales image-derived spatial heterogeneity can be used as a predictor for marine biodiversity.

Among the many metrics available to describe digital images, mean information gain (MIG) is a well-established measure of the complexity of spatial and temporal patterns (Grassberger 1988, Wackerbauer et al. 1994, Gell-Mann and Lloyd 1996, Lloyd 2001). MIG provides an index for two-dimensional surfaces on a scale from complete order to complete disorder. MIG also has the advantage of being a global indicator of spatial heterogeneity in an image that can be easily applied to both high-resolution digital photographs and lower-resolution satellite images.

Here we test the ability of MIG-based estimates of spatial heterogeneity in coral reef images, taken at different spatial scales, to predict fish biodiversity patterns. We use an extensive, hierarchically stratified coral reef data set, coupling fish census data with in situ benthic and satellite images of reef habitats. Our objectives were to (1) measure variation in heterogeneity of reef habitat images taken at different spatial scales and the covariance among them; (2) evaluate the relative ability of each scale of image-derived spatial heterogeneity to characterize fish biodiversity using hierarchical, multilevel statistical models; and (3) assess the combined performance of MIG across all scales for predicting observed patterns in coral reef fish biodiversity. We explicitly test the hypotheses that increasing habitat complexity, as measured by higher image MIG, mirrors increasing species richness and abundance, and that these relationships are spatially scale-dependent.

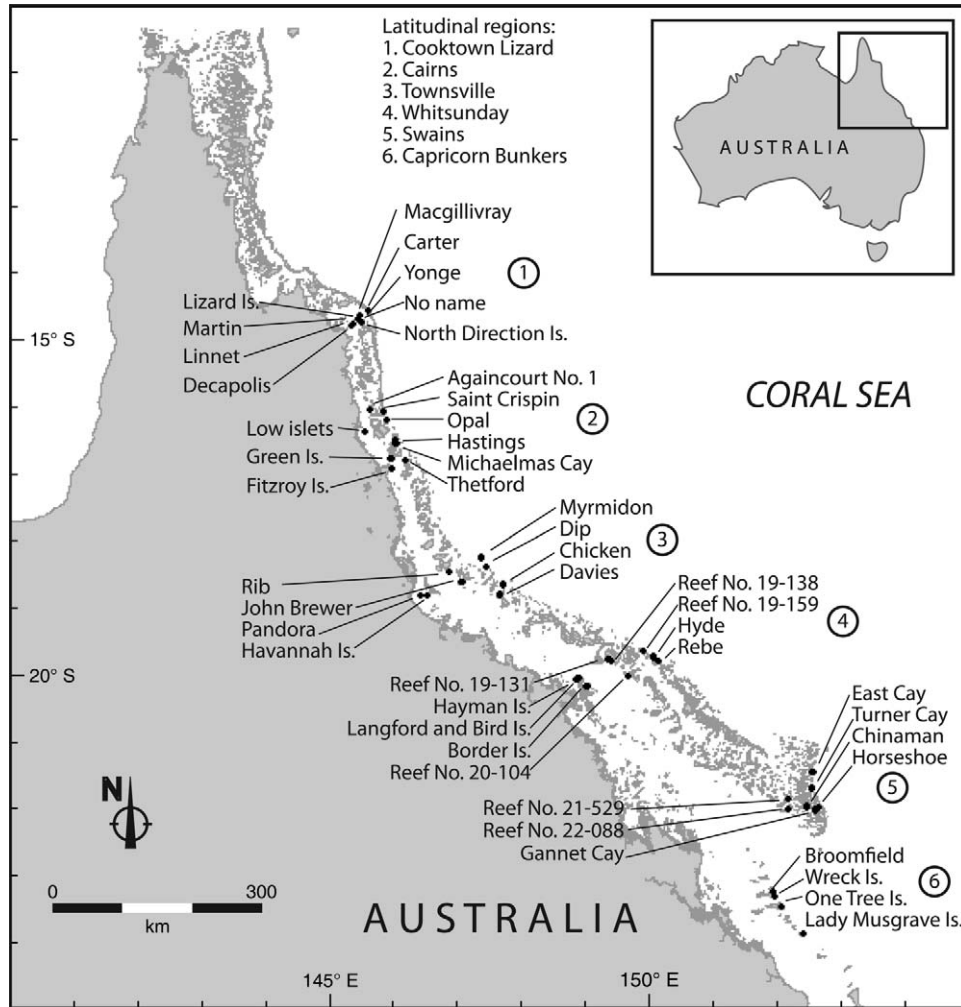


FIG. 1. Map of the Great Barrier Reef, Australia, with the positions of 46 AIMS (Australian Institute of Marine Science) reefs monitored annually since 1993. Circled numbers indicate the location of each latitudinal sector.

METHODS

Study reefs and data collection

The Great Barrier Reef (GBR) consists of more than 2900 reefs extending over 2300 km between 9° and 24° S latitude and covers ~350 000 km². Between 1993 and 2007, fish communities on 46 reefs on the GBR have been monitored annually by the Australian Institute of Marine Science in the Long Term Monitoring Program (Sweetman et al. 2005). These reefs occur in six longitudinal sectors (Cooktown/Lizard Island, Cairns, Townsville, Whitsunday, Swains, Capricorn Bunkers) spanning most of the GBR (Fig. 1). Within each sector, with the exception of the Swains and Capricorn Bunker sectors, at least three reefs were sampled in each of three shelf positions: inner, middle, and outer. In the two southern sectors, reef formation does not occur at all three shelf positions.

On each survey reef, three sites in a single habitat (the first stretch of continuous reef on the northeast flank of

the reef, excluding vertical drop-offs) and separated by >250 m were selected for sampling. Within each site, five randomly selected and permanently marked 50-m transects were deployed roughly parallel to the reef crest, each separated by 10–40 m along the 6–9 m depth contour. Sampling was evenly distributed among years and transects except for 2005. Each year, counts within these transects were recorded for 251 fish species from 10 families. This set of species excludes those that are cryptic or nocturnal and therefore have a low probability of detection. Larger mobile species were counted first along a transect 5 m wide, and smaller, less mobile species such as damselfishes (Pomacentridae) were counted in a 1-m wide strip along the same transect during the return swim (for detailed methods, see Halford and Thompson 1996). Only adult fish (>1 year old) were recorded, these being distinguished from juveniles by their size and coloration. Sites were sampled by different divers within and among years; annual

calibration exercises were done to ensure consistency among divers (Halford and Thompson 1996).

Habitat image acquisition and treatment

Transect scale.—Still images of benthic sessile communities were extracted from underwater video surveys recorded along each transect subsequent to the 2008 fish surveys (Abdo et al. 2004). A diver swam each transect and recorded benthic communities using a Sony Digital DCR-TRV950E video camera with the zoom set to full wide-angle, the focus to manual, and the focal length to 0.5–1 m. A picture of the data sheet was taken at the beginning of each transect for transect identification and to set white balance. The camera was kept parallel to the substrate at a distance of ~20 cm and moved along the 50-m transect at an approximately constant speed for 4–5 min. One still image (resolution 3072 × 2304 pixels) was extracted from these video records approximately every 1 m, with between 36 and 73 images recorded for each transect, depending on transect topographic complexity. In total, 35 098 images were acquired at the transect scale.

Following Proulx and Parrott (2008), we translated each image from RGB (Red, Green, Blue) color coordinates to HSV (Hue, Saturation, Value) coordinates. The benefit of the HSV system is that it reproduces more effectively how the human brain represents and manipulates color, while preserving within- and among-image variation in the original RGB system. Next, we classified each pixel value (range 0–255) into $M = 10$ evenly distributed classes. We then calculated the mean information gain (MIG), a measure of complexity in temporal and spatial data (Wackerbauer et al. 1994, Andrienko et al. 2000) that previously has been linked to habitat complexity in terrestrial ecosystems (Proulx and Parrott 2008, 2009). MIG provides an index ranging from 0 for uniform patterns across pixels to 1 for completely random patterns. Intermediate values of MIG correspond to irregular patterns composed of objects across a range of sizes, typical of complex natural scenes (Andrienko et al. 2000). MIG is similar to second-order measures of texture (Bellis et al. 2008); it describes the additional information gained by looking at the configuration of values in the neighborhood of each pixel in an image. Thus, we would expect images of uniform, undifferentiated habitats to have low MIG (i.e., low information gain) and images of random, or highly differentiated habitats to have high MIG (i.e., high information gain) (e.g., at the transition between bright sand patches and hard bottoms colonized by benthic organisms, or between deep and shallow areas, the latter being characterized by a brighter signature). Images of irregular patterns typical of complex reef structure (i.e., structural complexity arising from high substratum rugosity and a wide range of hole sizes and distributions at the patch scale, or fractal-like reef forms at the reef scale) should have intermediate MIG.

We computed MIG for each of the three bands in the HSV image as follows:

$$\text{MIG} = \frac{\left[-\sum_{j=1}^{M^k} p(\chi_j) \log p(\chi_j) \right] - \left[-\sum_{i=1}^M p(\gamma_i) \log p(\gamma_i) \right]}{\log(M^k/M)} \quad (1)$$

where $p(\gamma_i)$ is the relative frequency of pixel value γ_i in the image and $p(\chi_j)$ is the relative frequency with which a specific spatial configuration (χ_j) of k values is observed. For M classes of values, the number of possible configurations in a k -pixel neighborhood is M^k . To ensure that each possible configuration has a reasonable probability of occurring in an image, it is generally recommended that the ratio of the total number of pixels in the image to M^k be greater than 100. At all scales, we chose the highest possible value of M to retain as much information as possible from the image. For the transect images, we used $M = 10$ and $k = 4$ (i.e., a pixel neighborhood of 2×2 pixels), giving a ratio of total pixels to M^k of ~700. These parameters maximized the range of scales considered while enabling inter-site comparisons, as well as ensuring the use of a consistent k value for image analyses at all scales.

Calculated in this way, MIG represents the difference between the spatial entropy (calculated using a k -pixel neighborhood) and the aspatial entropy (calculated for individual pixel values irrespective of their location) in the image. Values range from 0 for completely random images to 1 for images of a single, solid color. Intermediate values of MIG are associated with more spatially heterogeneous data and therefore can be correlated with habitat complexity in images taken within particular ecosystems (Parrott 2010).

Site scale.—We obtained a mosaic of 25 Landsat ETM+ images of the GBR and neighboring coastal regions from the GBR Marine Park Authority (GBRMPA). We acquired images for the period between 17 August 1999 and 16 May 2002 (we chose this period because a scan line corrector problem limited the quality of Landsat, creating gaps in the data acquired after 2003). These images were taken mostly during low tide, in clear, offshore shelf waters.

Landsat provides images of the bottom in the blue band down to 30 m. To avoid any geodetic error in fish station locations, images were georectified using a series of ground-truthing points at a precision better than 1 pixel (30 m). To process sites throughout the GBR with the same radiometric quality, GBRMPA radiometrically normalized the 25 images to avoid discontinuities from one image to another. The result is a quasi-cloud-free composite image of the entire GBR without radiometric internal bias. The entire mosaic is distributed by GBRMPA in a low compression rate JPEG format with a pixel size of 30 m. Negligible information loss and changes in MIG values are attributable to JPEG compression (L. Parrott, unpublished data).

For each reef, we extracted a 51×51 pixel JPEG from the Landsat image, centered on the centroid of the three Long Term Monitoring Program sampling sites per reef to capture a section of the reef edge and water around the sites. We calculated the spatial heterogeneity (MIG; Eq. 1) of each of the site scale images in the same manner as described previously for the transect scale using $M = 2$ and $k = 4$.

Reef scale.—We used a fixed-size rectangular window of 251×351 pixels to extract reef-scale images from the overall Landsat image. We selected the size of the observation window to have a width as long as the widest (east to west) reef in the data set and a height as long as the longest (north to south) reef in the data set. We extracted five images for each reef: one with the window centered on the geographical coordinates of the reef centroid and four other images shifted in each of the four cardinal directions by $\sim 20\%$ from the center coordinates. Five images were necessary to test whether the exact position of the window used for extraction had an effect on the MIG obtained. We tested the null hypothesis (H_0) that image differences among reefs were not greater than within reefs using a multivariate analysis of variance (MANOVA; Anderson 2001) based on distance matrices and 1000 permutations. The fixed-size window was necessary to avoid biasing the results due to the effect of image size on MIG; however, images for small reefs therefore contained a larger proportion of water. We analyzed the five images for each reef for spatial heterogeneity as before, using Eq. 1 with $M = 4$ and $k = 4$.

Data organization

We were primarily interested in spatial patterns in fish diversity and how they related to the heterogeneity of reef images, in particular, the still images of benthic communities acquired in 2008. Accordingly, we only used fish data recorded between 2003 and 2007 (discarding fish data collected during 2005 because many reefs were not sampled that year due to inclement weather). Restricting the fish data in this way also minimized any potential effects of past disturbance affecting benthic communities; in particular, storms in 1988 removed a large amount of live coral that did not fully recover until 2003 (Sweatman et al. 2008). We defined fish species richness (R) as the total number of fish species sampled on each transect obtained by pooling species across the four years. For each transect, we defined total fish abundance (N) as the sum of individual species abundances across the four years to avoid potential autocorrelation due to counting the same individuals in multiple years.

For each transect and in each of the HSV bands, we averaged MIG across all images taken along it. At the reef scale, there was no evidence of an effect of the exact position of the observation window on MIG values (MANOVA with 1000 permutations; probability of concluding an effect $P = 0.26$). Therefore, for each reef

and in each of the HSV bands, we averaged MIG across the five observation windows. This procedure ensured that MIG values obtained for each reef did not depend on the exact position of the observation window, but rather accounted for the complexity of the reef of interest, as well as (to a lesser extent) that of its immediate neighborhood. The final data set used in our analyses therefore consisted of a matrix in which fish data at the transect level (i.e., individual species abundances, total species richness, and total abundance) were associated with MIG values at the transect, site, and reef scales (Fig. 2).

ANALYSIS

We first partitioned the variation in MIG at each spatial scale and estimated the relative proportion of each source of variation. We used a MANOVA based on distance matrices and 1000 permutations to investigate patterns of variation in MIG at each scale as a function of the latitudinal and cross-shelf ordinal factors; distances to the coast and seaward extent of the Great Barrier Reef at that latitude (Mellin et al. 2010a), and reef size and isolation (Mellin et al. 2010b).

We then used hierarchical linear models (Gelman and Hill 2007, MacNeil et al. 2009) to quantify the relationship between fish species richness or abundance (i.e., R or N) and MIG at each spatial scale, while accounting for the hierarchical structure of the data on this relationship (i.e., reef and site-within-reef correlation). The hierarchical structure resulted from ecological processes occurring at larger scales and influencing fish species richness and abundance at smaller scales. This hierarchical structure was quantified by including a reef- and/or site-level intercept (which we specify) in our linear models, thereby allowing an estimate of the importance of reef- and site-scale effects in structuring the biological data observed while simultaneously improving the accuracy of parameter estimates relative to non-hierarchical models (MacNeil et al. 2009). We evaluated relative, bias-corrected model support using Akaike's information criterion corrected for small sample sizes (AIC_c).

We constructed two sets of random-intercept, fixed-slope, Poisson hierarchical, linear models to predict R and N as a function of MIG at the reef and transect scales, and at the site and transect scales, respectively. Within each set, we estimated effects on R and N resulting from (1) the hierarchical structure of the data set only, (2) the hierarchical structure of the data set and MIG (each scale separately, then combined), (3) the hierarchical structure of the data set, MIG and their interaction, and (4) MIG only. Scenario (1) was achieved by fitting a model with no fixed effects but random intercepts among reefs (or sites), and comparing it to a single-intercept (NULL) model to determine if any hierarchical structure was present in the data. Scenario (2) was achieved by fitting three models with both random reef/site level intercepts and fixed effects

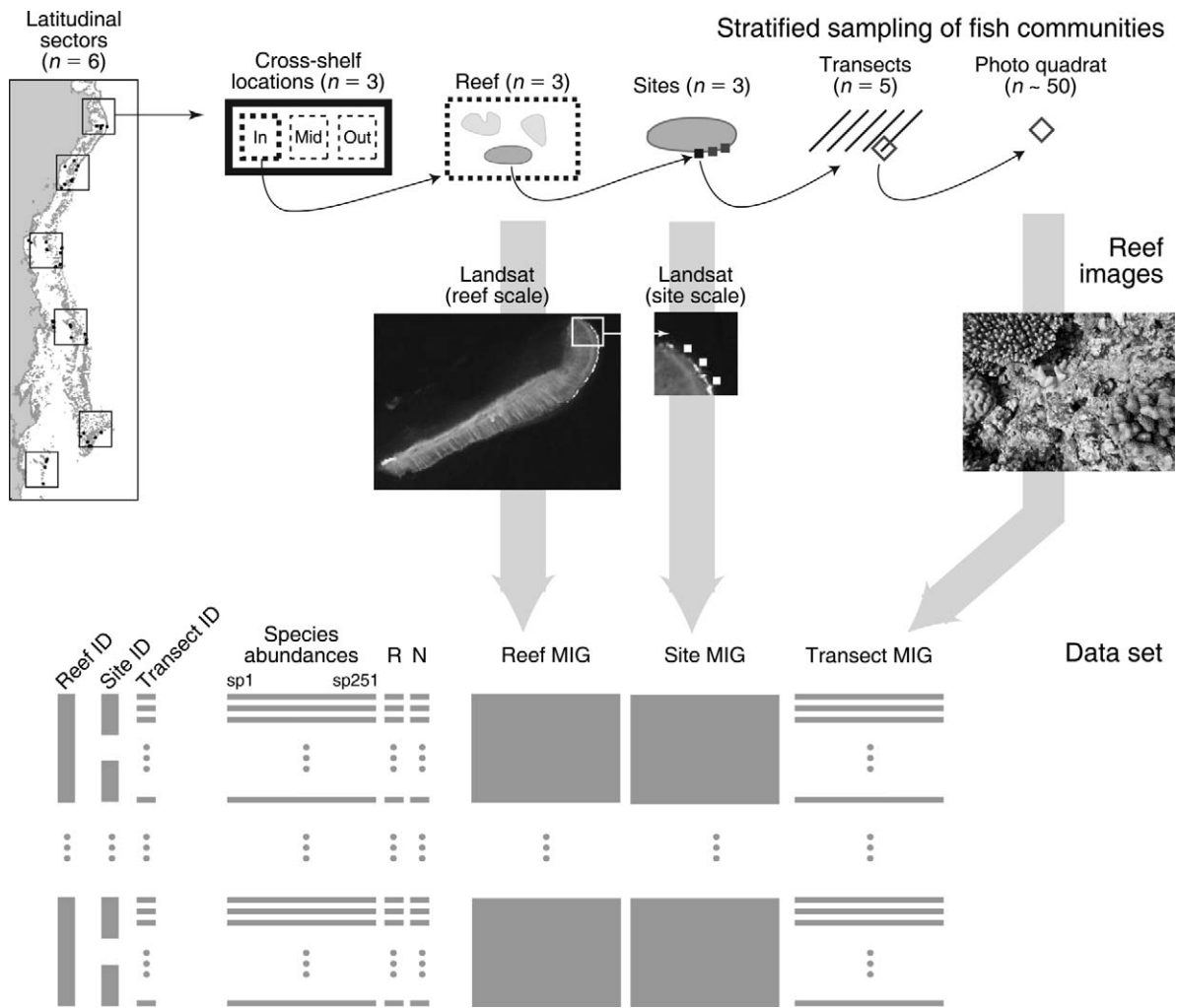


FIG. 2. Schematic representation of the stratified sampling design, reef images available at each spatial scale, and structure of the resulting data set. MIG is the mean information gain, ID is identity, R is fish species richness, and N is total fish abundance. The diagrams are indicative only and the spatial arrangement of transects is not as depicted.

including either the reef- (or site-) scale MIG, or transect-scale MIG, or both. Scenario (3) was achieved by adding to the analysis described for scenario (2), an interaction between fixed effects, allowing the effects of MIG at the transect scale to vary depending on MIG at the reef (or site) scale. Scenario (4) was achieved by randomizing the hierarchical structure, which was done by refitting each model set after randomization transects-within-reefs, simulating a total of 1000 different hierarchical structures (Cornell et al. 2007, MacNeil et al. 2009).

For both response variables, we validated the assumed Poisson error distribution based on the Gaussian distribution of model residuals, using the normalized scores of standardized residual deviance (Q-Q plots). We assessed the predictive ability of the top-ranked model according to AIC_c using a 10-fold cross-validation (Davison and Hinkley 1997). This bootstrap resampling procedure estimates the mean prediction

error, and the goodness of fit (i.e., coefficient of determination, R^2) of the relationship between predicted and observed values for 10% of observations randomly omitted from the calibration data set. This procedure was iterated 1000 times. We generated spatial correlograms assessing autocorrelation in R and N (raw data and model residuals) as a function of the distance between transects using Moran's I (Diggle and Ribeiro 2007). We chose distance classes for the autocorrelation analysis to reflect the nested design of the data set and included first-order neighboring transects (class 1), or all transects within a same site (class 2), reef (class 3), cross-shelf location (class 4), and latitudinal sector (class 5). We assessed evidence for spatial autocorrelation after a Bonferroni correction (Legendre and Legendre 1998).

We did a third set of analyses to partition variation in the multivariate fish assemblage structure (i.e., species abundance matrix) explained by the hierarchical structure of the data set and by MIG at each scale. We used a

TABLE 1. Summary of the hierarchical linear models of mean information gain variables (HSV coordinates) at the transect (tr) scale and as a function of the reef and site-within-reef hierarchical structures (i.e., random intercepts).

Model	<i>k</i>	LL	wAIC _c	<i>D</i> (%)
Hue (H)				
tr-H ~ (1 reef/site)	4	1490.78	1	46.38
tr-H ~ (1 reef)	3	1336.01	0	31.19
NULL	3	1018.42	0	0
Saturation (S)				
tr-S ~ (1 reef/site)	4	1606.78	1	86.47
tr-S ~ (1 reef)	3	1468.77	0	70.46
NULL	3	861.67	0	0
Value (V)				
tr-V ~ (1 reef/site)	4	1886.09	1	21.01
tr-V ~ (1 reef)	3	1747.61	0	12.13
NULL	3	1558.57	0	0

Notes: Abbreviations are: *k*, number of parameters; LL, log likelihood; wAIC_c, weight given by the Akaike's information criterion corrected for small sample size; *D*, percentage deviance explained. Model notation follows that of the lme4 package in R: (1 | reef) denotes a "reef" random effect and (1 | reef/site) denotes a "site-within-reef" random effect.

constrained distance-based redundancy analysis with 1000 permutations (Legendre and Anderson 1999). We computed a Bray-Curtis distance matrix based on the fish assemblage matrix and a Euclidian distance matrix based on MIG (all scales). Transects were then clustered based on each distance matrix successively and using the *k*-means algorithm (Legendre and Legendre 1998). We calculated the similarity between the two classifications, defined as the proportion of transects falling in the same cluster using either distance matrix, and compared this proportion to the distribution of proportions expected under a null model given by 1000 randomizations of transects-within-clusters based on the fish distance matrix.

RESULTS

Hierarchical linear models of MIG at the transect scale revealed substantial hierarchical structure in the data, with clear AIC_c-based evidence for high site-within-reef dependence among all three response variables (i.e., MIG of benthic images calculated for each of the three [HSV] color bands) accounting for up to 86% of deviance in the saturation of benthic images (Table 1). We also found evidence for latitudinal sector effects on MIG at the transect scale, distance to the coast on MIG at the site scale, and reef size on MIG at the reef scale (Appendix A). Because the effect of reef size on MIG at the reef scale is likely to be an artifact of the fixed-size window used to calculate MIG, we included reef area as an offset in subsequent hierarchical linear models.

Reef hierarchical structure alone accounted for 71.2% and 75.1% of null deviance in *R* (Table 2) and *N* (Table 3), respectively. Even though models including image indices were top-ranked according to AIC_c, they reduced model deviance by <1% relative to the random intercept-only models. However, when transects were permuted within reefs to randomize the hierarchical structure, image indices reduced model deviance by 27% (*R*) and 31% (*N*), respectively, relative to intercept-only models (the latter only reduced null deviance by 2% and 7%, respectively; Tables 2 and 3), thereby demonstrating the importance of habitat structure in explaining a considerable percentage of the variation in biodiversity patterns beyond the influence of hierarchical reef structure alone. A combination of MIG at the transect and reef scales was present in the AIC_c top-ranked models for both *R* and *N* (Appendices B and C).

The 10-fold cross-validation showed that the top-ranked fitted models according to AIC_c resulted in a mean prediction error of 7.6% for *R* and 10.6% for *N*,

TABLE 2. Summary of hierarchical linear models of fish species richness (*R*) as a function of mean information gain variables (HSV, coordinates for hue; saturation, and value) at the reef (rf) and transect (tr) scales, on original and randomized data sets.

Model	<i>k</i>	LL	wAIC _c	<i>D</i>
Original				
$R \sim A + \text{rf-HSV} + (1 \text{reef})$	7	-266.87	0.58	71.75
$R \sim A + \text{rf-HSV} + \text{tr-HSV} + (1 \text{reef})$	10	-264.79	0.21	71.97
$R \sim (1 \text{reef})$	3	-272.38	0.14	71.17
$R \sim \text{tr-HSV} + (1 \text{reef})$	6	-270.31	0.05	71.39
$R \sim A + \text{tr-HSV} + \text{tr-HSV} \times \text{rf-HSV} + (1 \text{reef})$	10	-267.18	0.02	71.72
NULL	3	-944.66	0	0
Randomized				
$R \sim A + \text{rf-HSV} + \text{tr-HSV} + (1 \text{reef})$	10	-670.27	1	29.05
$R \sim A + \text{tr-HSV} + \text{tr-HSV} \times \text{rf-HSV} + (1 \text{reef})$	10	-690.6	0	26.89
$R \sim A + \text{rf-HSV} + (1 \text{reef})$	7	-786.75	0	16.72
$R \sim \text{tr-HSV} + (1 \text{reef})$	6	-798.53	0	15.47
$R \sim (1 \text{reef})$	3	-924.87	0	2.09
NULL	3	-944.66	0	0

Notes: Reef area (*A*) is used as a controlling factor to account for the fixed size window of Landsat images. Other abbreviations are *k*, number of parameters; LL, log likelihood; wAIC_c, weight given by the Akaike's information criterion corrected for small sample size; *D*, percentage deviance explained. Model notation follows that of the lme4 package in R.

TABLE 3. Summary of hierarchical linear models of fish abundance (N) as a function of mean information gain variables (HSV, coordinates for hue; saturation, and value) at the reef (rf) and transect (tr) scales, on original and randomized data sets.

Model	k	LL	$wAIC_c$	D
Original				
$N \sim A + \text{tr-HSV} + \text{tr-HSV} \times \text{rf-HSV} + (1 \text{reef})$	10	-33 810	1	75.43
$N \sim A + \text{rf-HSV} + \text{tr-HSV} + (1 \text{reef})$	10	-34 078	0	75.24
$N \sim \text{tr-HSV} + (1 \text{reef})$	6	-34 088	0	75.23
$N \sim A + \text{rf-HSV} + (1 \text{reef})$	7	-34 250	0	75.11
$N \sim (1 \text{reef})$	3	-34 259	0	75.1
NULL	3	-137 608	0	0
Randomized				
$N \sim A + \text{rf-HSV} + \text{tr-HSV} + (1 \text{reef})$	10	-92 441	1	32.82
$N \sim A + \text{rf-HSV} + (1 \text{reef})$	7	-95 604	0	30.52
$N \sim A + \text{tr-HSV} + \text{tr-HSV} \times \text{rf-HSV} + (1 \text{reef})$	10	-95 844	0	30.35
$N \sim \text{tr-HSV} + (1 \text{reef})$	6	-124 988	0	9.17
$N \sim (1 \text{reef})$	3	-128 259	0	6.79
NULL	3	-137 608	0	0

Notes: Reef area (A) is used as a controlling factor. Other abbreviations are: k , number of parameters; LL , log likelihood; $wAIC_c$, weight given by the Akaike's information criterion corrected for small sample size; D , percentage deviance explained. Model notation follows that of the lme4 package in R.

with 74.4% and 69.2% of the variance in the predictions explained by the observations (R^2), respectively (Appendix D). We found evidence for autocorrelation at all distance classes in observations of both response variables, but only at the first distance class for model residuals (Appendix E), indicating that our hierarchical models accounted for most spatial dependence via the nested spatial design.

Similar to the univariate results for R and N , reef hierarchical structure accounted for 60% of the variation in the fish community matrix, including 25% from MIG at the reef (7%), site (6%), and transect (12%) scales, with <1% overlap among these variance components (Fig. 3). The hierarchical clustering of transects based on fish data vs. MIG from all scales resulted in 53% of transects being classified in the same cluster for a three-cluster classification, 49% for four clusters (Fig. 4), 54% for five clusters, and 41% for six clusters. There was strong evidence that these percentages were all higher than those expected from a null model ($P < 0.001$).

DISCUSSION

Our results provide the first multi-scale evidence that patterns of coral reef fish biodiversity can be reasonably predicted from a combination of reef habitat complexity indices derived from the spectral signal of digital (camera and satellite) images, together with information on the hierarchical structure of the data set. Andréfouët et al. (2010) previously established partial links between unprocessed red-green-blue bands and fish communities in New Caledonia. We have developed this method further using a hierarchical framework to account for the complex spatial structure of coral reef ecosystems (MacNeil et al. 2009, Mellin et al. 2009).

Having incorporated hierarchical spatial effects into our models, habitat complexity as estimated by MIG accounted for an additional 29% of the null deviance in

fish species richness, 33% in fish abundance, and 25% in fish assemblage structure. At least for community structure, this accords well with recent work reporting that ~22% of variance is explained for these same fish communities by reefs nested within habitat (i.e., cross-shelf position) and region (i.e., latitudinal sector; Burgess et al. 2010), suggesting that the simple MIG metric performs well in capturing this within-reef variation in community structure. Although this percentage of deviance explained is much lower than for some terrestrial ecosystems (e.g., 58% deviance explained in bird species richness in arid landscapes; St-Louis et al. 2006), it agrees with other studies that used

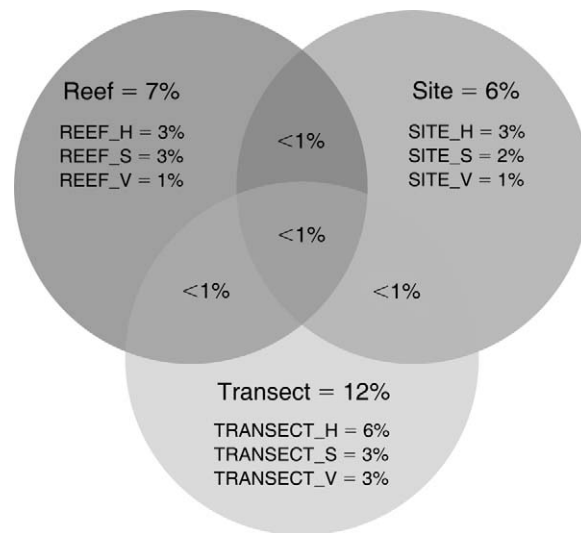


FIG. 3. Venn diagram partitioning the percentage of variation in the fish community matrix as a function of image mean information gain variables (H, hue; S, saturation; V, value) at reef, site, and transect scales.

both remotely sensed and field data for understanding biodiversity patterns (Estes et al. 2008). Indeed, a key advantage of remote sensing is its ability to provide information over a species' entire range, but models incorporating remotely sensed data might be limited by the ability of remote sensing to detect important habitat features (Estes et al. 2008). In coral reefs, and on the Great Barrier Reef in particular, major correlates of fish diversity patterns include distance to the coast and to the barrier reef, sea surface temperature (Mellin et al. 2010a), and reef size and isolation (MacNeil et al. 2009, Mellin et al. 2010b). Therefore, incorporating these factors in models based on image-derived habitat complexity should result in improved predictive power.

Of the total deviance explained in fish species richness, abundance, and assemblage structure, 17%, 30%, and 13% (respectively) were explained by MIG at the reef and/or site scales directly estimated from the Landsat ETM+ mosaic. In previous studies using habitat maps, the best predictive performance was achieved using an 11-class habitat map derived from aerial photography (i.e., 38% of variation in fish species richness; 28% of variation in abundance) and obtained from a series of 12 landscape metrics related to patch configuration in Moreton Bay, Queensland, Australia (Pittman et al. 2004). Although that approach was comparable in its explanatory power to our own, the method required considerable image pre-processing, labeling, and ground-truthing. In contrast, where images are available, MIG can be calculated for the three channels of a regular JPEG format without any pre-processing and in only a few seconds per image on a desktop computer, providing substantial cost savings. The MIG methodology presented here and applied on the Great Barrier Reef now needs to be applied to other regions and ecosystems to enable cost-benefit comparisons among methods for predicting biodiversity patterns.

Both fish biodiversity and MIG of the digital habitat images at the transect scale displayed strong spatial variation reflecting the data set's transect-within-site and site-within-reef hierarchical patterns. Furthermore, we found latitudinal and cross-shelf effects on MIG at the transect scale. These results reflect the gradient of benthic assemblages, which consist of both biotic (e.g., live coral, macroalgae) and abiotic (e.g., dead coral, bare substratum) components. As a result, the similarity of fish and benthic assemblages was higher among sites within the same reef (i.e., 250 m to 1 km apart) than among different reefs (i.e., 3 to 1000 km apart), with the turnover (i.e., β diversity) in benthic communities increasing along latitudinal and cross-shelf gradients. This contrasts with observations of Caribbean reef systems in which differences in fish and coral diversity were observed among geomorphological classes within reefs, but with little difference among reefs along a 400-km latitudinal gradient (Arias-González et al. 2008). On the Great Barrier Reef, conditions such as water quality, temperature, and biological productivity follow latitu-

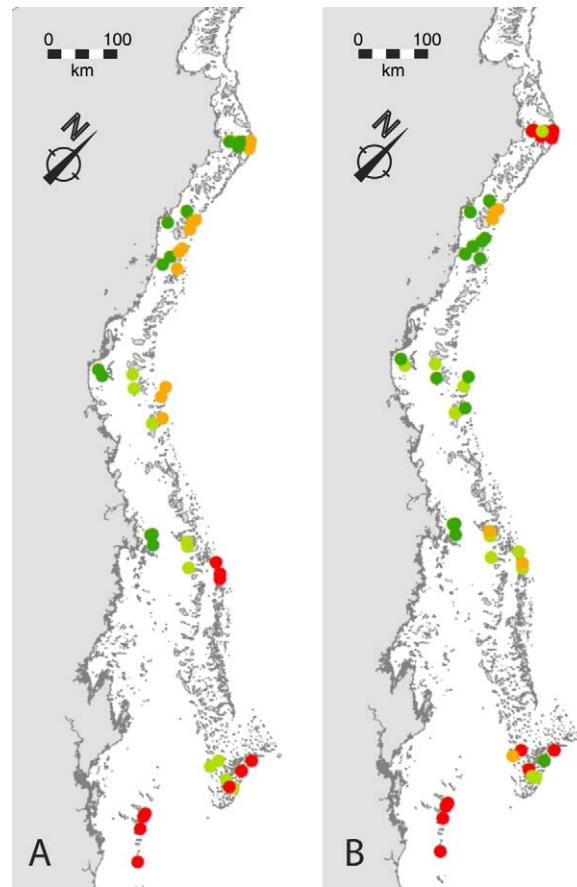


FIG. 4. Clustering based on (A) the fish community matrix and (B) mean information gain variables at the reef, site, and transect scales. For each reef, only the most frequent cluster is shown, with each color indicating a particular cluster. Identity of the clusters is irrelevant; only the grouping is important.

dinal and cross-shelf gradients associated with variation in the structure of fish (Burgess et al. 2010, Mellin et al. 2010a) and benthic communities (Burgess et al. 2010, De'ath and Fabricius 2010, Emslie et al. 2010). The hierarchical structure of fish diversity and habitat complexity as estimated here by MIG thus reflects the combination of environmental and geographical influences on the β diversity of fish and benthic communities on the Great Barrier Reef.

The performance of MIG in predicting fish biodiversity was robust to the choice of spatial scale and resolution, suggesting that this approach could be transferrable to other reef systems and could allow cross-system comparisons. Although the predictive performance of MIG was consistent across image resolutions, image resolution influences the absolute value of MIG, necessitating image pretreatment to standardize resolution, white balance, and color models across sites prior to analyses. Therefore, images combined within a study should be of approximately equal spatial extent to give comparable MIG values and avoid

scale-dependent artifacts. Ideally, uncompressed images in RAW format should be used if feasible. In our study, however, the large number of files used and the normalized mosaic of satellite images compiled for the entire study area, made the compressed (JPEG) format necessary for facilitating image storage and manipulation.

Although the predictive performance of image-derived indices of habitat complexity on their own might still be insufficient to support large-scale conservation planning, they can complement existing methods for studying biodiversity and can contribute to improvements in survey design methods. For example, broad-scale, long-term monitoring programs are commonly stratified with respect to spatial factors based on a priori visual examination of a map of the region of interest. In this case, all sampled sites were similarly positioned on the reefs examined, which were themselves stratified across the shelf and latitudinally based on the premise that such gradients would be reflected in assemblage structure (Sweatman et al. 2008). We suspect that a design based on continuous environmental gradients derived from remotely sensed images (Dalleau et al. 2010) could lead to more discerning process-based modeling with potentially higher predictive performance. For instance, habitat complexity estimated from remotely sensed imagery has been used in the design of the monitoring program for World Heritage Sites on New Caledonian reefs (Andréfouët and Wantiez 2010). Therefore, image-derived indices of habitat complexity could support the design of monitoring programs based on ecological gradients of habitat complexity and the collection of additional (biological and physical) data, eventually to provide inputs for more sophisticated and powerful species distribution models.

Efficient, streamlined, and cost-effective methods for biodiversity mapping are needed, especially in developing countries where the effects of declining natural resources are often acute and where data, facilities, and expertise are often sparse (Fisher et al. 2011). Methods such as those presented here show much promise as they require little processing, no labeling of geomorphology or habitat structure, and no accuracy assessment or ground-truthing. Capacity-building for biodiversity conservation can therefore be promoted through the dissemination of spatial products for research and environmental management in these jurisdictions (Andréfouët 2008). Although the methods presented here could benefit from further refinement and validation across a broader range of ecosystems before they can be applied routinely, they will ultimately be constrained to some extent by the need for consistent image quality, specifications, and normalization among study sites. In spite of these constraints, these methods can now be consistently applied among multiple sites and regions of the world. For example, both remotely sensed and in situ habitat images are readily available worldwide for many ecosystems including coral reefs (e.g., photo

transects generated by the Reef Check program, *available online*).⁷ This wealth of historical records remains under-exploited. When used in conjunction with field-based surveys, spectral signals from such digital images could prove invaluable for assessing the extent of habitat degradation effects on biological communities.

ACKNOWLEDGMENTS

We thank all members of the Australian Institute of Marine Science Long Term Monitoring Program who have contributed to data collection. Thanks to K. Johns and A. Dorais for assisting with image acquisition and treatment. The Great Barrier Reef Marine Park Authority (Queensland, Australia) provided the Landsat ETM+ image mosaic (1999–2002).

LITERATURE CITED

- Abdo, D., S. Burgess, G. Coleman, and K. Osborne. 2004. Surveys of benthic reef communities using underwater video. Long-term monitoring of the Great Barrier Reef Standard Operating Procedure Number 2. Australian Institute of Marine Science, Townsville, Australia.
- Anderson, M. J. 2001. A new method for non-parametric multivariate analysis of variance. *Austral Ecology* 26:32–46.
- Andréfouët, S. 2008. Coral reef habitat mapping using remote sensing: A user vs. producer perspective. Implications for research, management and capacity building. *Journal of Spatial Science* 53:113–129.
- Andréfouët, S., C. Payri, M. Kulbicki, J. Scopélitis, M. Dalleau, C. Mellin, M. Scamps, and G. Dirberg. 2010. Mesure, suivi et potentiel économique de la diversité de l'habitat récifo-lagonaire néo-calédonien: inventaire des herbiers, suivi des zones coralliennes et rôle des habitats dans la distribution des ressources en poissons de récifs. Rapport Conventions Sciences de la Mer—Biologie Marine 31. Institut de Recherche pour le Développement Centre de Nouméa/Zone économique de Nouvelle-Calédonie, Nouméa, New Caledonia.
- Andréfouët, S., L. Roux, Y. Chancerelle, and A. Bonneville. 2000. A fuzzy possibilistic scheme of study for objects with indeterminate boundaries: application to French Polynesian reefs. *IEEE Trans Geoscience and Remote Sensing* 38:257–270.
- Andréfouët, S., and L. Wantiez. 2010. Characterizing the diversity of coral reef habitats and fish communities found in a UNESCO World Heritage Site: The strategy developed for lagoons of New Caledonia. *Marine Pollution Bulletin* 61:612–620.
- Andrienko, Y., N. Brilliantov, and J. Kurths. 2000. Complexity of two-dimensional patterns. *European Physical Journal B* 15:539–546.
- Attrill, M. J., J. A. Strong, and A. A. Rowden. 2000. Are macroinvertebrate communities influenced by seagrass structural complexity? *Ecography* 23:114–121.
- Baraldi, A., and F. Parmiggiani. 1995. An investigation of the textural characteristics associated with gray-level cooccurrence matrix statistical parameters. *IEEE Transactions on Geoscience and Remote Sensing* 33:293–304.
- Beger, M., and H. P. Possingham. 2008. Environmental factors that influence the distribution of coral reef fishes: modeling occurrence data for broad-scale conservation and management. *Marine Ecology Progress Series* 361:1–13.
- Bellis, L. M., A. M. Pidgeon, V. C. Radloff, V. St-Louis, J. L. Navarro, and M. B. Martella. 2008. Modeling habitat suitability for greater rheas based on satellite image texture. *Ecological Applications* 18:1956–1966.

⁷ www.reefcheck.org

- Burgess, S. C., K. Osborne, and M. J. Caley. 2010. Similar regional effects among local habitats on the structure of tropical reef fish and coral communities. *Global Ecology and Biogeography* 19:363–375.
- Caley, M. J., and J. St John. 1996. Refuge availability structures assemblages of tropical reef fishes. *Journal of Animal Ecology* 65:414–428.
- Chittaro, P. M. 2004. Fish–habitat associations across multiple spatial scales. *Coral Reefs* 23:235–244.
- Cornell, H. V., R. H. Karlson, and T. P. Hughes. 2007. Scale-dependent variation in coral community similarity across sites, islands, and island groups. *Ecology* 88:1707–1715.
- Dalleau, M., S. Andréfouët, C. C. C. Wabnitz, C. Payri, L. Wantiez, M. Pichon, K. Friedman, L. Vigliola, and F. Benzoni. 2010. Use of habitats as surrogates of biodiversity for efficient coral reef conservation planning in Pacific Ocean islands. *Conservation Biology* 24:541–552.
- Davison, A. C., and D. V. Hinkley. 1997. *Bootstrap methods and their application*. Cambridge University Press, Cambridge, UK.
- De'ath, G., and K. Fabricius. 2010. Water quality as a regional driver of coral biodiversity and macroalgae on the Great Barrier Reef. *Ecological Applications* 20:840–850.
- Diggle, P. J., and P. J. Ribeiro, Jr. 2007. *Model-based geostatistics*. Springer, New York, New York, USA.
- Dumas, P., A. Bertaud, C. Peignon, M. Leopold, and D. Pelletier. 2009. A “quick and clean” photographic method for the description of coral reef habitats. *Journal of Experimental Marine Biology and Ecology* 368:161–168.
- Emslie, M. J., M. S. Pratchett, A. J. Cheal, and K. Osborne. 2010. Great Barrier Reef butterflyfish community structure: the role of shelf position and benthic community type. *Coral Reefs* 29:705–715.
- Estes, L. D., G. S. Okin, A. G. Mwangi, and H. H. Shugart. 2008. Habitat selection by a rare forest antelope: A multi-scale approach combining field data and imagery from three sensors. *Remote Sensing of Environment* 112:2033–2050.
- Fisher, R., B. Radford, N. Knowlton, R. E. Brainard, and M. J. Caley. 2011. Global mismatch between research effort and conservation needs on coral reefs. *Conservation Letters* 4:64–72.
- Friedlander, A. M., and J. Parrish. 1998. Habitat characteristics affecting fish assemblages on a Hawaiian coral reef. *Journal of Experimental Marine Biology and Ecology* 224:1–30.
- Frost, N. J., M. T. Burrows, M. P. Johnson, M. E. Hanley, and S. J. Hawkins. 2005. Measuring surface complexity in ecological studies. *Limnology and Oceanography: Methods* 3:203–210.
- Garza-Pérez, J. L., A. Lehmann, and J. E. Arias-González. 2004. Spatial prediction of coral reef habitats: integrating ecology with spatial modeling and remote sensing. *Marine Ecology Progress Series* 269:141–152.
- Gell-Mann, M., and S. Lloyd. 1996. Information measures, effective complexity and total information. *Complexity* 2:44–52.
- Gelman, A., and J. Hill. 2007. *Data analysis using regression and multilevel hierarchical models*. Cambridge University Press, New York, New York, USA.
- Grassberger, P. 1988. Complexity and forecasting in dynamical systems. Pages 1–22 *in* L. Peliti and A. Vulpiani, editors. *Measures of complexity*. Springer-Verlag, Berlin, Germany.
- Halford, A. R., and M. J. Caley. 2009. Towards an understanding of resilience in isolated coral reefs. *Global Change Biology* 15:3031–3045.
- Halford, A. R., and A. A. Thompson. 1996. Visual census surveys of reef fish. Long term monitoring of the Great Barrier Reef. Standard Operational Procedure Number 3. Australian Institute of Marine Science, Townsville, Australia.
- Huston, M. 1979. General hypothesis of species diversity. *American Naturalist* 113:81–101.
- Kuffner, I. B., J. C. Brock, R. Grober-Dunsmore, V. E. Bonito, T. D. Hickey, and C. W. Wright. 2007. Relationships between reef fish communities and remotely sensed rugosity measurements in Biscayne National Park, Florida, USA. *Environmental Biology of Fishes* 78:71–82.
- Legendre, P., and M. J. Anderson. 1999. Distance-based redundancy analysis: testing multispecies responses in multifactorial ecological experiments. *Ecological Monographs* 69:1–24.
- Legendre, P., and L. Legendre. 1998. *Numerical ecology*. Second edition. Elsevier, Amsterdam, The Netherlands.
- Levin, S. 1999. *Fragile dominion*. Perseus Publishing, Cambridge, UK.
- Lloyd, S. 2001. Measures of complexity: a non exhaustive list. *IEEE Control Systems* 21:7–8.
- Luckhurst, B. E., and K. Luckhurst. 1978. Analysis of influence of substrate variables on coral-reef fish communities. *Marine Biology* 49:317–323.
- MacNeil, M. A., N. A. J. Graham, J. E. Cinner, N. K. Dulvy, P. A. Loring, S. Jennings, N. V. C. Polunin, A. T. Fisk, and T. R. McClanahan. 2010. Transitional states in marine fisheries: adapting to cope with global change. *Philosophical Transactions of the Royal Society B* 365:3753–3763.
- MacNeil, M. A., N. A. J. Graham, N. V. C. Polunin, M. Kulbicki, R. Galzin, M. Harmelin-Vivien, and S. P. Rushton. 2009. Hierarchical drivers of reef-fish metacommunity structure. *Ecology* 90:252–264.
- Mattio, L., G. Dirberg, C. Payri, and S. Andréfouët. 2008. Diversity, biomass and distribution pattern of Sargassum beds in the South West lagoon of New Caledonia (South Pacific). *Journal of Applied Phycology* 20:811–823.
- Mellin, C., S. Andréfouët, M. Kulbicki, M. Dalleau, and L. Vigliola. 2009. Remote sensing and fish–habitat relationships in coral reef ecosystems: Review and pathways for systematic multi-scale hierarchical research. *Marine Pollution Bulletin* 58:11–19.
- Mellin, C., S. Andréfouët, and D. Ponton. 2007. Spatial predictability of juvenile fish species richness and abundance in a coral reef environment. *Coral Reefs* 26:895–907.
- Mellin, C., C. J. A. Bradshaw, M. G. Meekan, and M. J. Caley. 2010a. Environmental and spatial predictors of species richness and abundance in coral reef fishes. *Global Ecology and Biogeography* 19:212–222.
- Mellin, C., C. Huchery, M. J. Caley, M. G. Meekan, and C. J. A. Bradshaw. 2010b. Reef size and isolation determine the temporal stability of coral reef fish populations. *Ecology* 91:3138–3145.
- Oldeland, J., D. Wesuls, D. Rocchini, M. Schmidt, and N. Jurgens. 2010. Does using species abundance data improve estimates of species diversity from remotely sensed spectral heterogeneity? *Ecological Indicators* 10:390–396.
- Parrott, L. 2010. Measuring ecological complexity. *Ecological Indicators* 10:1069–1076.
- Pittman, S. J., C. A. McAlpine, and K. M. Pittman. 2004. Linking fish and prawns to their environment: a hierarchical landscape approach. *Marine Ecology Progress Series* 283:233–254.
- Proulx, R., and L. Parrott. 2008. Measures of structural complexity in digital images for monitoring the ecological signature of an old-growth forest ecosystem. *Ecological Indicators* 8:270–284.
- Proulx, R., and L. Parrott. 2009. Structural complexity in digital images as an ecological indicator for monitoring forest dynamics across scale, space and time. *Ecological Indicators* 9:1248–1256.
- Sale, P. F., and W. A. Douglas. 1984. Temporal variability in the community structure of fish on coral patch reefs and the relation of community structure to reef structure. *Ecology* 65:409–422.
- St-Louis, V., A. M. Pidgeon, V. C. Radeloff, T. J. Hawbaker, and M. K. Clayton. 2006. High-resolution image texture as a

- predictor of bird species richness. *Remote Sensing of Environment* 105:299–312.
- Sweatman, H., S. Burgess, A. Cheal, G. Coleman, S. Delean, M. Emslie, A. McDonald, I. Miller, K. Osborne, and A. Thompson. 2005. Long-term monitoring of the Great Barrier Reef. Status Report 7. Australian Institute of Marine Science, Townsville, Australia.
- Sweatman, H., A. Cheal, G. Coleman, M. Jonker, K. Johns, M. Emslie, I. Miller, and K. Osborne. 2008. Long-term monitoring of the Great Barrier Reef. Status Report 8. Australian Institute of Marine Science, Townsville, Australia.
- Wackerbauer, R., A. Witt, H. Altmanspacher, J. Kurths, and H. Scheingraber. 1994. A comparative classification of complexity measures. *Chaos, Solitons & Fractals* 4:133–173.

SUPPLEMENTAL MATERIAL

Appendix A

Variation partitioning in image complexity variables calculated at each spatial scale (*Ecological Archives* XXXXXXXX).

Appendix B

Summary of the hierarchical linear models of fish species richness as a function of image complexity variables (*Ecological Archives* XXXXXXXX).

Appendix C

Summary of the hierarchical linear models of fish abundance as a function of image complexity variables (*Ecological Archives* XXXXXXXX).

Appendix D

Observed values of fish species richness and abundance against predictions resulting from the 10-fold cross-validation (*Ecological Archives* XXXXXXXX).

Appendix E

Spatial correlograms of Moran's I for observations and model residuals of fish species richness and abundance (*Ecological Archives* XXXXXXXX).