

Article

## Mapping Fish Community Variables by Integrating Field and Satellite Data, Object-Based Image Analysis and Modeling in a Traditional Fijian Fisheries Management Area

Anders Knudby <sup>1,\*</sup>, Chris Roelfsema <sup>2</sup>, Mitchell Lyons <sup>2</sup>, Stuart Phinn <sup>2</sup> and Stacy Jupiter <sup>3</sup>

<sup>1</sup> Department of Geography, University of Waterloo, Waterloo, ON K1Z8G6, Canada

<sup>2</sup> School of Geography, Planning and Environmental Management, University of Queensland, Brisbane, QLD 4072, Australia; E-Mails: c.roelfsema@uq.edu.au (C.R.); m.lyons@uq.edu.au (M.L.); s.phinn@uq.edu.au (S.P.)

<sup>3</sup> Wildlife Conservation Society, Fiji Program, Suva, Fiji; E-Mail: sjupiter@wcs.org

\* Author to whom correspondence should be addressed; E-Mail: knudby@gmail.com; Tel.: +1-613-729-5236.

Received: 22 December 2010; in revised form: 1 February 2011 / Accepted: 21 February 2011 / Published: 1 March 2011

---

**Abstract:** The use of marine spatial planning for zoning multi-use areas is growing in both developed and developing countries. Comprehensive maps of marine resources, including those important for local fisheries management and biodiversity conservation, provide a crucial foundation of information for the planning process. Using a combination of field and high spatial resolution satellite data, we use an empirical procedure to create a bathymetric map (RMSE 1.76 m) and object-based image analysis to produce accurate maps of geomorphic and benthic coral reef classes (Kappa values of 0.80 and 0.63; 9 and 33 classes, respectively) covering a large (>260 km<sup>2</sup>) traditional fisheries management area in Fiji. From these maps, we derive per-pixel information on habitat richness, structural complexity, coral cover and the distance from land, and use these variables as input in models to predict fish species richness, diversity and biomass. We show that random forest models outperform five other model types, and that all three fish community variables can be satisfactorily predicted from the high spatial resolution satellite data. We also show geomorphic zone to be the most important predictor on average, with secondary contributions from a range of other variables including benthic class, depth, distance from land, and live coral cover mapped at coarse spatial scales, suggesting that data with lower spatial resolution and lower cost may be sufficient for spatial predictions of the three fish community variables.

**Keywords:** coral reefs; IKONOS; Quickbird; predictive mapping; fish; species richness; species diversity; biomass

---

## 1. Introduction

Many of the half a billion people worldwide who live within 100 km of a coral reef depend on small scale reef fisheries to sustain their livelihoods [1,2]. As nearly 60% of reefs globally are imperiled by human activities [1], protection of reef fish populations requires urgent and active management. However, tools that have long been promoted by fisheries scientists for managing single species fisheries (e.g., maximum sustainable yield) do not capture the ecosystem complexities of multispecies coral reef fisheries [3]. Meanwhile, newer advanced models applied to coral reef fisheries, e.g., Ecopath with Ecosim [4], require large amounts of data that are often resource intensive to acquire and do not include spatially explicit outputs. Marine Protected Areas (MPAs) have emerged as an alternative, widespread tool for coral reef fisheries management that have resulted in increases in abundance and biomass of targeted species where successfully implemented and enforced [5]. To optimally design networks or individual MPAs, coral reef managers could greatly benefit from comprehensive fish resource maps. Maps of fish biomass can be used to guide placement of MPAs towards naturally productive areas, a known driver of MPA success [6]. Further, in the absence of accurate information on single-species distributions, maps of species richness and diversity can be used to ensure coverage of local biodiversity hotspots.

Ecological field research has produced a large body of literature describing functional and statistical relationships between variables quantifying aspects of the reef fish community and the local habitat. For example, the positive effect of structural complexity on reef fish species richness and abundance has been demonstrated experimentally [7], and areas of high structural complexity have been shown to contain more species [8] and greater total biomass [9] compared to areas with less structure. Other factors have shown similar correlations with one or more aspects of reef fish communities, such as water depth [10], disturbance history [11], coral cover [12], location on the reef [9] and shelf [13], presence of seagrass beds and mangroves [14,15], as well as benthic composition and patch structure [16-18]. Such environmental variables, and the patch mosaic they create, interact with and influence the local fish communities at a range of spatial scales.

Methods have been developed for mapping several of these key environmental variables using remote sensing. The spectral reflectance characteristics of an area can be used to characterize the dominant benthos [19], and estimate both live coral cover [20-22] and water depth [23-25]. The location on the reef can both contribute to classification of the benthos and be used to map geomorphic reef zones [26]. Spatial analysis of modeled depth estimates can be used to quantify structural complexity, and maps of the benthos can be used to calculate indices of patch diversity, size and shape, as well as distances to other marine and adjacent terrestrial and freshwater habitats [27].

The ability to map these environmental variables with remote sensing has in turn made it possible to employ statistical models to produce spatially explicit predictions of fish community variables [28]. To date, only a handful of studies have been published using this approach. Some rely solely on seafloor

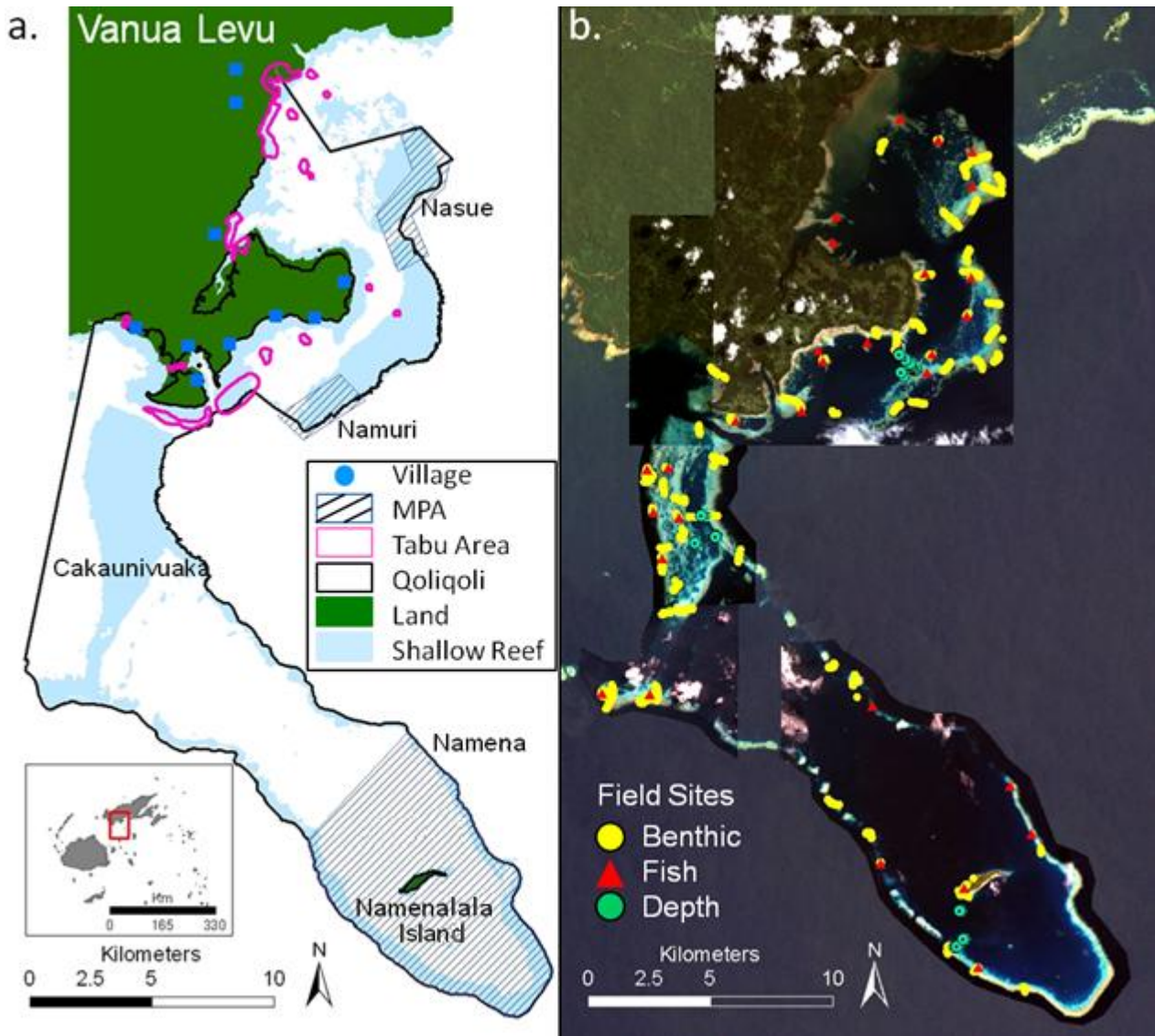
terrain models, typically derived from lidar data, from which metrics of the seafloor's three-dimensional structure are calculated and used to produce predictions of fish community variables [29-31]. These studies cover sufficiently large areas to have relevance in a management context (e.g., ~200 km<sup>2</sup> studied by Pittman *et al.* [30]), but the cost of acquiring lidar data necessary to create seafloor terrain models at sufficient spatial resolution and extent makes the use of these models prohibitive to most institutions. It is therefore important to explore the utility of less expensive data sources from which predictions of fish community variables is also possible. In this study, we build on work by Mellin *et al.* [32] and Knudby *et al.* [28], which successfully tested the application of optical remote sensing data for predicting fish assemblage characteristics over small-scale coral reef sites (<20 km<sup>2</sup>), by scaling the approach up to cover the reef complex in a Fijian fisheries management area (>260 km<sup>2</sup>). We use object-based image analysis to create near-seamless classifications of reef benthos and geomorphology from five high-spatial resolution optical satellite images captured over Kubulau District's traditional fisheries management area in southwest Vanua Levu. We combine these classifications with derived water depth, as well as measures of seafloor structure and substrate diversity at varying spatial scales, to provide spatially explicit predictions of fish species richness, diversity, and biomass. Lastly, we discuss how the results are being used to prioritize alternative configurations for the Kubulau community managed MPA network to improve effectiveness for increasing fisheries yields while reducing current conflict among resource users.

## 2. Study Area

Kubulau District, centered at 16°51'S and 179°0'E in south-west Vanua Levu, is located in one of the least developed regions of the Fiji Islands. Approximately 1,000 residents of predominantly indigenous origin rely heavily on marine resources for subsistence and cash income [33]. Given the relatively low population density (~10 people km<sup>-2</sup>), Kubulau's terrestrial, freshwater and marine ecosystems are still relatively intact, and outer reef areas support very high reef fish biomass and catch rates on par with remote Pacific islands and atolls [6,34].

As market access increased over the past two decades, the chiefs of Kubulau undertook preventative measures to maintain fish stocks for the future. In 2005, with the support of the Wildlife Conservation Society (WCS) and partner NGOs, the chiefs established a network of 20 MPAs, combining 17 small, periodically harvested closures (tabu areas) and three large no-take reserves covering mangrove, seagrass and coral reef habitat within the 261 km<sup>2</sup> traditional fisheries management area [33] (Figure 1). They empowered the Kubulau Resource Management Committee with authority to make recommendations for management of the network, which is regulated under Fiji's first ridge-to-reef management plan [35]. To date, the effectiveness of the protected areas has varied, largely due to challenges related to compliance and enforcement [6,33]. Given the mixed success, as well as more recent conflict involving illegal fishing activity in the Namena Marine Reserve, WCS and research partners are investigating alternative configurations for the MPA network to retain fisheries yields while spreading costs more equitably [36].

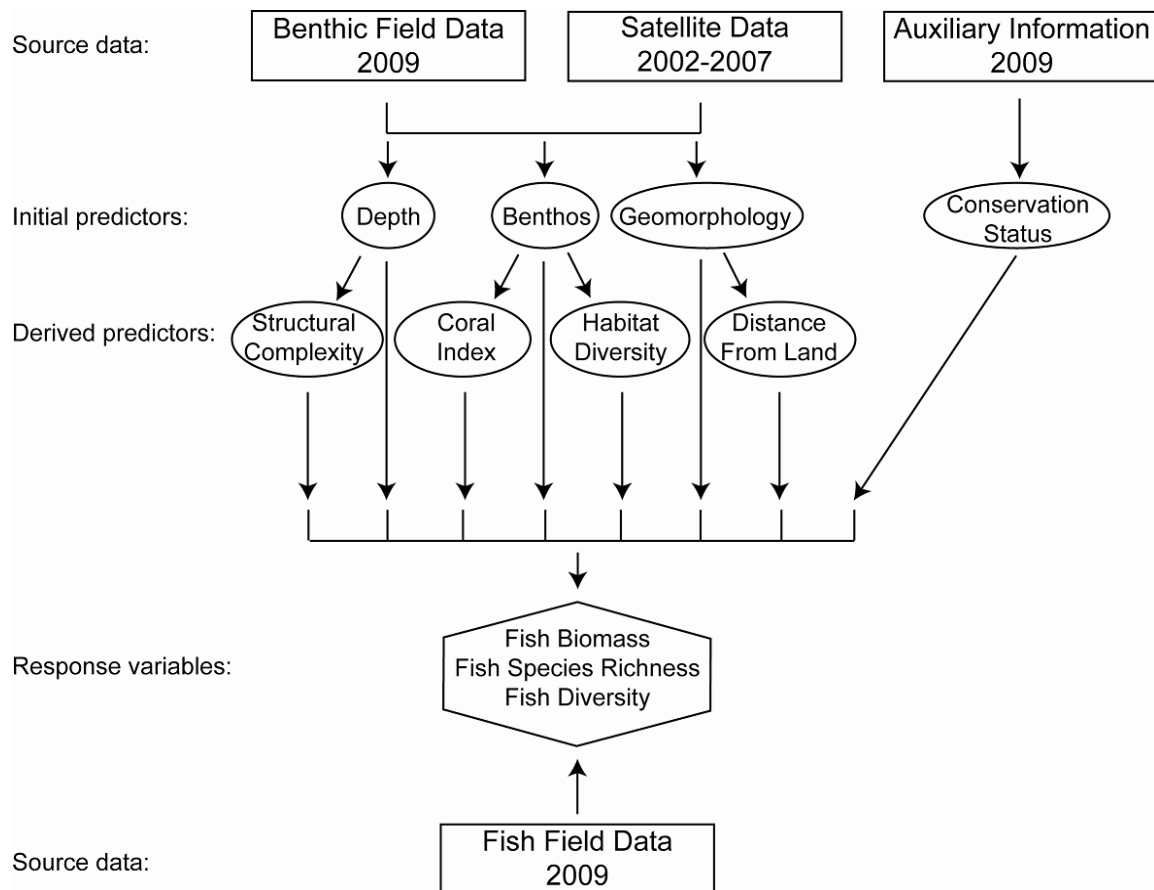
**Figure 1.** Overview of Kubulau fisheries management area (qoliqoli). (a) shows the boundary of the management area, as well as its traditional village-managed tabu areas (pink outline) and district-managed marine protected areas (diagonal lines, names next to areas). Inset illustrates location within Fiji. (b) shows the extent of the Landsat TM data (background) and IKONOS/Quickbird data (foreground), as well as sites where benthic, fish and depth data were collected.



### 3. Methods

Data for this study fall in three categories: benthic field data, fish field data, and satellite data. Collection and processing of each is described below. The flow of data processing and model development is illustrated in Figure 2.

**Figure 2.** Flowchart describing the development of initial and derived predictors from source data, and their combination in predictive model development. Source data in boxes, predictors in circles, response variables in the hexagon.



### 3.1. Field Data Collection and Processing

#### 3.1.1. Benthic Field Data

We acquired benthic cover data using a georeferenced photo transect method [37] for 64 transects throughout the study area (Figure 1). Towing a surface float GPS that logged its position every five seconds and using a standard digital camera in an underwater housing, a snorkeler/diver swam over the bottom taking photos of the benthos every 2–4 m, approximating the size of one pixel in the high-resolution satellite data. At a height of 0.5 m above the benthos, the camera provided a footprint of approximately 1 m × 1 m. The snorkeler/diver would occasionally take overview photos to aid in subsequent interpretation. In order to ensure representative coverage over the major habitat types, photo transect locations, direction and length had been selected prior to surveys by visual assessment of the spatial pattern of benthic structures evident in each satellite image covering the study area, following an approach applied in previous studies [37,38].

An interpreter then manually assigned each of the resulting 9,646 photos to a benthic community cover category using Coral Point Count Excel® software [39]. We used a hierarchical benthic cover scheme containing five first-level categories and 30 second-level categories (Table 1). The categories are similar to those used in previous image-based coral reef mapping [37,40–42]. We then linked each

photo to its geographic coordinates through time synchronization of the GPS and camera using the GPS-Photo Link software. This allowed the photos, the corresponding benthic composition data and benthic mapping category to be viewed at their position in the study area through a GIS interface.

**Table 1.** Hierarchical scheme used for benthic classification of the photo transect data.

First	Description	Second Level	Description
Coral (C)	>10% coral cover	Hard coral dominant	>70% Coral
		Hard coral and Macroalgae	>10% Macroalgae
		Hard coral and seagrass	>10% Seagrass
		Hard coral and sand	>10% Sand
		Hard coral and rubble	>10% Rubble
		Hard coral and reef matrix	>10% Reef Matrix
		Hard coral and less soft coral	>10% Soft Coral
		Hard coral live and dead	>10% Dead Coral
Soft Coral (SC)	>10% soft coral cover	Soft coral dominant	>70% Soft Coral
		Soft coral and macroalgae	>10% Macroalgae
		Soft coral and seagrass	>10% Seagrass
		Soft coral and sand	>10% Sand
		Soft coral and rubble	>10% Rubble
		Soft coral and reef matrix	>10% Reef Matrix
Macroalgae (MA)	<10% live coral cover, macroalgal cover >10% and dominant over seagrass	Macroalgae dominant	>70% Macroalgae
		Macroalgae and Seagrass	>10% Seagrass
		Macroalgae and Sand	>10% Sand
		Macroalgae and Rubble	>10% Rubble
		Macroalgae and Reef Matrix	>10% Reef Matrix
		Macroalgae and dead coral	>10% dead Coral
Seagrass (SG)	<10% live coral cover, seagrass cover >10% and dominant over macroalgae	Seagrass dominant	>70% Seagrass
		Seagrass and sand	>10% Sand
		Seagrass and rubble	>10% Rubble
Non-Living Substratum (BS)	<10% live coral cover, <10% algal cover, <10% seagrass cover, >70% bare substratum	Sand dominant	>70% Sand
		Rubble dominant	>70% Rubble
		Reef matrix dominant	>70% Reef Matrix
		Mud/silt dominant	>70% Mud/Silt
		Sand/rubble dominant	>70% Sand/Rubble
		Dead coral dominant	>70% Dead Coral

### 3.1.2. Fish Field Data

We collected data on the reef fish communities at 110 representative forereef, backreef, patch reef, seagrass and sand sites in Kubulau in April-May and September 2009 (Figure 1) using the underwater visual census methods of Jupiter and Egli [6]. For each site, georeferenced by its central coordinate, we

collected data from three to five replicate 5 m × 50 m belt transects at depths varying from 1 m to 13.5 m. We included the following families in the census: Acanthuridae, Balistidae, Carangidae, Carcharhinidae, Chaetodontidae, Chanidae, Ephippidae, Haemulidae, Kyphosidae, Labridae, Lethrinidae, Lutjanidae, Mullidae, Nemipteridae, Pomacanthidae, Scaridae, Scombridae, Serranidae (groupers only), Siganidae, Sphraenidae, and Zanclidae. These families cover the large majority of fishes in the region, and include those considered priorities for fisheries and conservation. For each transect, we recorded the number of observed fish in each species in 5 cm length classes for fish less than 40cm, and exact length for each fish over 40 cm. We then calculated the biomass of each fish from the mean length of its size class and existing published values from Fishbase [43] used in the standard length-weight (L-W) expression  $W = a \cdot L^b$ , with  $a$  and  $b$  parameter values preferentially selected from sites closest to Fiji. If no L-W parameters were available for the species, we used the factors for the species with the most similar morphology in the same genus. If a suitable similar species could not be determined, we used average L-W values for the genus. As many of the L-W conversions required fork length (FL), we converted our observations from total length (TL) to FL when necessary using a length-length (L-L) conversion factor obtained from Fishbase. For each site, we then quantified three aspects of the fish community, all expressed as averages per transect to account for the variable number of transects at each site. ‘Species richness’ was calculated as the number of species, ‘species diversity’ was calculated as Shannon’s diversity index [44] based on each species’ biomass, and ‘biomass’ was calculated as the total biomass of all fish, expressed as kg/ha.

### 3.2. Satellite Data

We combined three types of multi-spectral satellite data in this study: three Quickbird images, two Ikonos images, and one Landsat Thematic Mapper (TM) image (Table 2). Prior to other processing, we corrected all images for radiometric and atmospheric distortions to at-surface reflectance [37] and created two image mosaics, one from the two overlapping Ikonos images, and one from two overlapping Quickbird images.

**Table 2.** Attributes of satellite data sets used.

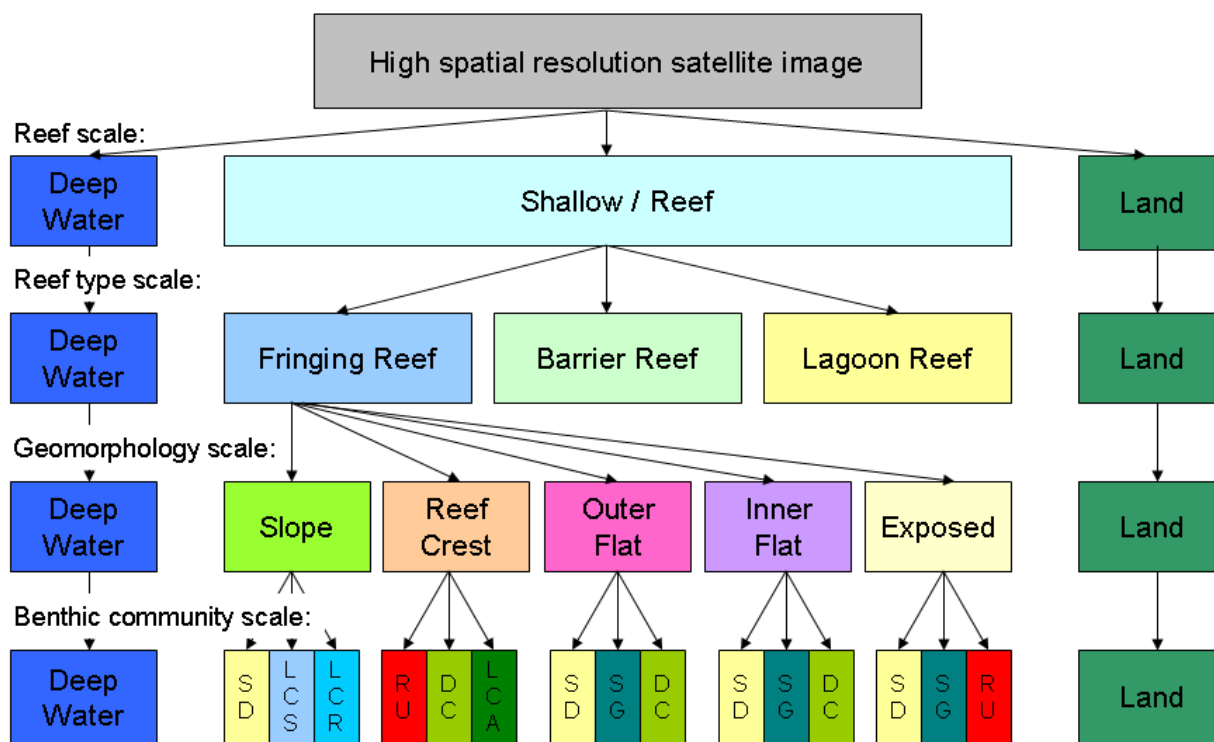
Attribute	Quickbird	Ikonos	Landsat 5 TM
Data provider	Digitalglobe	Geoeye	NASA
Spatial resolution	2.4 m	4 m	30 m
Maximum extent	12 km × 12 km	12 km × 12 km	185 km × 185 km
Number of scenes	3	2	1
Acquisition year	2006	2007	2002
Mapping approach	Object based analysis	Object based analysis	Manual delineation

#### 3.2.1. Object-Based Classification of Benthos and Geomorphic Zonation

For the areas covered by high spatial resolution data, we used object-based image analysis in Definiens Developer 7.0 to create habitat maps at four hierarchical spatial scales: reef, reef type, geomorphic zone, and benthic community (Figure 3). Object-based image analysis consists of two

steps repeated across each mapping scale: (1) image segmentation, and (2) segment classification [45]. Starting at the coarsest spatial scale, segmentation rules are applied to divide the satellite image into segments, and class membership rules are then applied to assign each segment to a reef class. A second set of segmentation rules is then used to sub-divide each reef segment into smaller ‘sub-segments’, and a second set of membership rules is used to assign each of these to a reef type class. This process is then repeated for the third and fourth spatial scales. Throughout the segmentation steps we also homogenized the spatial resolution of the data to 4 m.

**Figure 3.** Example of a hierarchical class structure diagram of mapping categories at four different spatial scales: reef, reef type, geomorphology zone, and benthic community. For clarity, only a limited range of example classes at the geomorphology and benthic community scales are shown, and only for the fringing reef type. Abbreviations at the benthic community scale are: SD = sand, LCS = live coral and sand, LCR = live coral and rock, RU = rubble, DC = dead coral, LCA = live coral and algae, SG = seagrass.



We developed both the segmentation and class membership rule sets from interpretation of the satellite imagery. For the first three spatial scales, we relied on expert knowledge for image interpretation, whereas we also used the benthic field data to aid interpretation for the benthic community scale. After development of both rule sets for all spatial scales, we applied them to the full extent of the satellite image data to produce the four habitat maps. We based segmentation rules on the color and shape of groups of pixels as well as the spatial scale of the features to be mapped, and we based membership rules on the color, shape, texture and position of each segment as well as the biophysical properties of the individual mapping categories. We initially based our rule sets on those developed for the same hierarchical scales in previous studies [37,46]. Although we left the basis of many of the membership rules for individual classes unchanged, we modified thresholds used in the



rules to account for differences in illumination conditions, water clarity, water surface and atmospheric conditions. In some cases we created new membership rules to account for new benthic classes present in Kubulau, e.g., the “Algae/Cyanobacteria on Sand” that covered large areas of outer reef flat on Cakaunivuaka Reef (Figure 1). Following application of the rule sets, we manually checked for errors and assigned new class labels to erroneously classified and unclassified segments, which together consisted of less than 1% of all segments. The manual labeling was primarily based on image interpretation and expert knowledge for the reef, reef type and geomorphic zone scales, and assisted by field data for the benthic community scale. For the 3% of the shallow reefs in the area not covered by high resolution satellite data, we manually segmented the Landsat TM image and assigned reef, reef type and geomorphic zone classes, but were unable to assign benthic community classes.

### 3.2.2. Accuracy Assessment of Benthic and Geomorphic Maps

To assess the accuracy of the habitat maps at the geomorphic scale (nine classes) and benthic community scale (33 classes), we developed error matrices and calculated overall accuracies and Kappa coefficients based on reference data for the individual map scales [47]. Based on previous research [37,46], we developed the reference data sets for the geomorphic zone and the benthic community scales using different approaches. For the geomorphic zones, we generated 15 randomly distributed points within each class, for a total of 135 points, and manually assigned them to a reference geomorphic zone class. For the benthic community classes, we selected all the 1,649 segments that contained at least one benthic photo. These segments were then manually assigned a benthic community class by assessing the photo, its benthic community label, and the underlying satellite image. We note that photos from ~30% of the 1,649 validation segments had also been used to aid the image interpretation that we used to develop the class membership rules for the benthic community scale, potentially resulting in a slightly optimistic assessment of accuracy for that scale.

## 3.3. Mapping of Additional Variables

### 3.3.1. Depth

We mapped bathymetry using the empirical approach developed by Lyzenga [23], calibrating the model parameters with 24 depth data points from reef zones with homogenous high-albedo benthic cover. Of the 24 depth points, we measured 12 in the field with a Vexilar® hand held digital depth finder, and derived another 12 from bathymetric charts and first-hand knowledge. We then combined the depth estimates from the five images into one layer, where depth values in areas covered by more than one image were calculated with an inverse distance weighting function, calculating distance from the border between overlap and non-overlap areas.

### 3.3.2. Structural Complexity

Based on the combined depth layer, we quantified the structural complexity of the seafloor using five metrics: standard deviation of water depth, surface rugosity, the slope of slope measure used by Pittman *et al.* [30], curvature, and absolute curvature. We calculated standard deviation of water depth using the Neighborhood-Focal Statistics-STD function in ArcGIS's Spatial Analyst, and calculated

surface rugosity as the ratio of surface area to planar area [30] with a customized script. We calculated both metrics for circular areas around each pixel center using 10 different radii: 4, 10, 30, 50, 75, 100, 200, 300, 500 and 1,000 m. We calculated slope of slope with two applications of the Surface-Slope function, and curvature using the Surface-Curvature function, both in ArcGIS's Spatial Analyst. We also determined absolute curvature as the absolute value of the curvature metric in order to eliminate the differentiation between convex and concave surfaces. In order to derive the slope and curvature metrics at spatial scales comparable to those used for standard deviation of water depth and surface rugosity, we resampled (nearest neighbor) the combined depth layer nine times to produce a total of 10 depth layers with cell sizes equal to the above list of radii. We then applied the slope and curvature functions, calculated on the basis of  $3 \times 3$  cell kernels, to each of the depth layers. We chose the list of 10 radii/cell sizes to cover a range spatial scales from the smallest possible given the spatial resolution of the satellite data to very large radii when calculations for some metrics become computationally demanding; the list of radii/cell sizes does not reflect an *a priori* judgment about the relevant ecological scales. We masked out land prior to all calculations of structural complexity. Except for curvature, all these measures were significantly collinear. In addition to curvature we therefore retained only absolute curvature, which, of the four remaining structural complexity metrics, had the highest individual correlations with the response variables. We eliminated all other variables from the subsequent analysis.

### 3.3.3. Live Coral Cover Index

As a proxy for live coral cover, we generated a live coral cover index by assigning pixels with coral as the dominant benthic class a value of 2, pixels with coral as a secondary benthic class a value of 1, and remaining pixels a value of 0. In addition to the per-pixel values of this index, we calculated average values using the same list of radii as was used for curvature, resulting in a total of 10 live coral cover index layers.

### 3.3.4. Habitat Richness

We used two metrics to quantify the spatial heterogeneity of habitat types, both calculated on the basis of the benthic community classes present within a radius around each pixel. We used the Neighborhood—Focal Statistics—VARIETY function in ArcGIS's Spatial Analyst to calculate habitat richness, quantified as the number of individual classes present within the radius, and a customized script to calculate the Shannon diversity of classes based on the frequency of occurrence of each class. We used the same list of 10 radii as for structural complexity and live coral cover index calculations. Given high collinearity between the two measures, we retained the one with the highest individual correlations with the response variables, habitat richness, while we eliminated the other, Shannon diversity, from the subsequent analysis.

### 3.3.5. Distance from Land

We calculated the distance from land for each pixel using ArcGIS's Euclidian distance tool. We did not include Namenalala Island as "land" for this calculation; distances were calculated from Vanua

Levu only. Namenalala Island has very limited freshwater input to the surrounding reefs because of its small size and has no resident population besides a tourist resort. Not including the island as “land” in this calculation therefore allowed the “distance from land” variable to also function as a proxy both for the distance from freshwater inputs and for anthropogenic impacts (mainly fishing pressure).

### 3.3.6. Conservation Status

Finally, we generated a map layer describing the conservation status of each pixel based on information from Kubulau District’s management plan [35]. We used three classes in this layer: unprotected, inside a village-managed tabu area (typically small, located on fringing reefs and/or reef flats, and harvested between one and three times per year), and inside a district-managed no-take marine protected area.

### 3.4. Modeling Methods

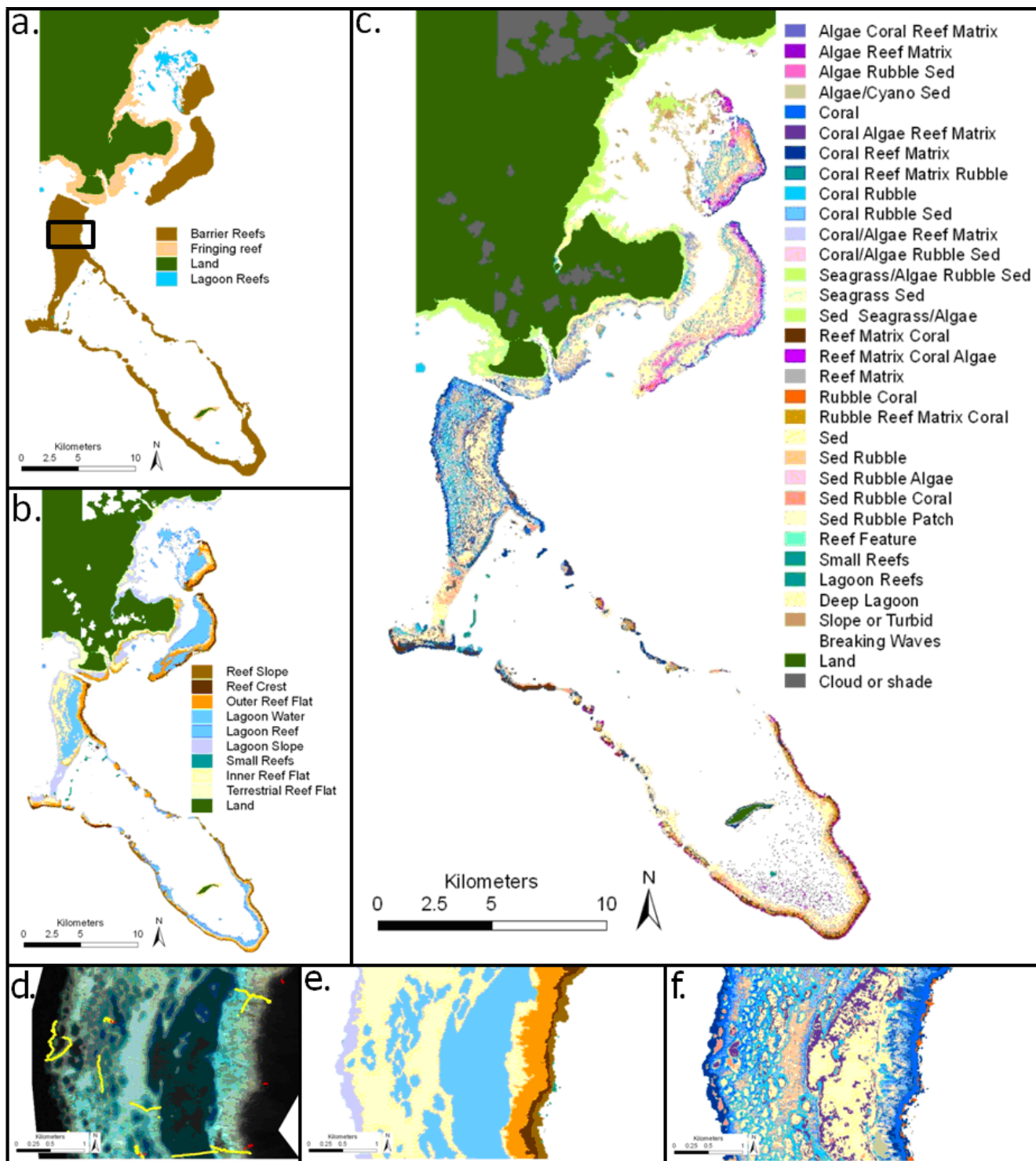
Using the suite of remote sensing-derived habitat variables as predictors and the fish species richness, diversity and biomass data as response variables, we developed predictive models for each response variable using the procedure described in Knudby *et al.* [28,48]. We developed six types of predictive models, two parametric models (linear and generalized additive) and four non-parametric models (bagging, random forest, boosted regression trees and support vector machine), all implemented using R and its contributed packages [49,50]. We quantified the predictive performance of each model type as the RMSE of its predictions, calculated using 10-fold cross-validation repeated 100 times. We considered the model type with lowest average RMSE to perform best, and used it to produce spatially explicit predictions for each response variable. Prior to use in the models, we log/power-transformed variables when necessary to approach normality. Semi-variograms were used to check the level of spatial autocorrelation in model residuals; all were negligible. We also calculated a measure of each predictor’s importance for predictions, using a permutation-based approach where importance is quantified as the increase in RMSE value caused by permutation of a given variable in the test set [48]. This measure is considered an initial pointer to variables that were important for the specific predictions made in our study, keeping in mind the known difficulties in objectively assessing the importance of individual variables in multi-variable models [51-53].

## 4. Results

### 4.1. Geomorphic, Benthic and Depth Maps

The object-based image analysis produced maps of the Kubulau reef complex at four hierarchical mapping levels: reef, reef type, geomorphic, and benthic community (Figure 4). The map of geomorphic zones with nine classes had an overall accuracy of 82.1% and a kappa accuracy of 0.80, while the benthic community map had overall and kappa accuracies of 66.6 % and 0.63 with 33 classes. The depth map had an RMSE of 1.76 m. The accuracy of the derived variables (coral cover index, habitat richness and curvature) could not be assessed as no reference data exist at corresponding spatial scales.

**Figure 4.** Classification results from areas within Kubulau’s traditional fisheries management area: (a) reef type, (b) geomorphic zones, (c) benthic community, (d) subset showing IKONOS true-color composite and transects used to collect benthic field data, (e) subset of (b), (f) subset of (c). Subset areas outlined by black box in (a). Sed = Sediment.



#### 4.2. Predictive Models

Cross-validated RMSE values for predictions of the three response variables are listed in Table 3, which shows that the Random Forest model type performed best regardless of response variable. The predictive maps were therefore all produced using this model type. The Support Vector Machine model performed almost as well for predictions of Shannon diversity, and the Bagging models also performed well for all response variables.

**Table 3.** Average cross-validated RMSE values for each response variable and model type (biomass and diversity variables were re-transformed prior to RMSE calculations).

Model type	Species richness	Biomass	Shannon diversity
Bagging	6.395	613.6	0.563
Boosted Regression Trees	6.713	654.9	0.573
Generalized Additive Model	7.283	703.8	0.572
Linear Model	7.545	689.5	0.568
Random Forest	6.371	601.2	0.558
Support Vector Machine	7.434	637.8	0.558

The three most important predictors for each response variable and model type are shown in Table 4, which shows that the importance of individual predictors varied substantially between response variables as well as between model types. For the random forest models used to produce the spatial predictions geomorphic zone was the most or second-most important predictor for all response variables, with other substantial contributions from a range of variables including benthos, which was the most important predictor of fish diversity. For the support vector machine models, depth was the most important predictor regardless of response variable, while curvature calculated at 4 m radius was important for predictions of species richness and biomass in both the linear and generalized additive models.

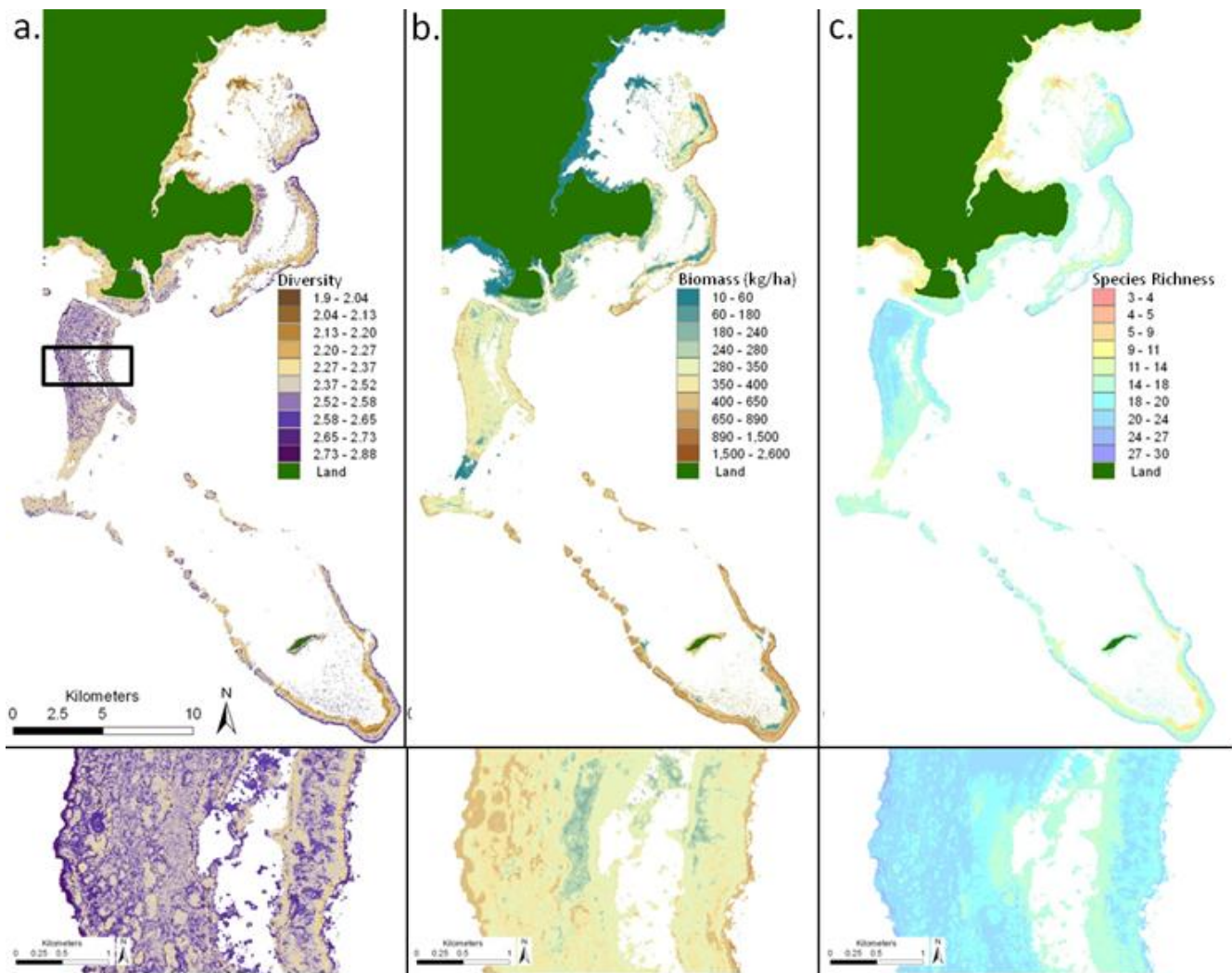
#### 4.3. Predictive Maps

The predictive maps of Shannon diversity and fish species richness indicate highest scores along the outer barrier reefs of Kubulau, as well as notably high diversity within the patch mosaic of the Cakaunivuaka reef complex. The predictions of fish biomass show a marked increase with distance from land, with the greatest biomass found on the outer barrier reefs, particularly within the Nasue and Namuri MPAs and the steep reef slopes within and adjacent to the Namena Marine Reserve (Figure 5).

**Table 4.** The three most important predictors for each response variable and their importance value, listed in order of importance and by model type. Cor = Coral, Geo = Geomorphology, Dist = Distance from land, Cur = curvature, A Cur = Absolute curvature. Numbers following predictor name indicate radius of calculation.

Linear model	Generalized additive model	Bagging	Random forest	Boosted regression trees	Support vector machine						
<b>Species richness (<math>\Delta</math>RMSE)</b>											
Cor 4	1.791	Benthos	1.801	Benthos	1.424	Geo	0.656	Geo	0.639	Depth	0.772
Benthos	1.374	Cor 4	1.678	Cor 500	0.178	Depth	0.183	Benthos	0.183	Cur 4	0.173
Geo	1.215	Geo	1.239	Depth	0.127	Benthos	0.135	Depth	0.155	A Cur 4	0.132
<b>Biomass (<math>\Delta</math>RMSE)</b>											
Cor 4	139.4	Cor 4	98.8	Geo	46.6	Geo	44.3	Geo	44.2	Depth	67.1
Benthos	112.8	Benthos	94.7	Cor 1000	22.3	Dist	25.2	Dist	39.0	Dist	45.2
Dist	42.1	Geo	32.6	Dist	17.0	Depth	15.1	Cur 300	24.0	A Cur 4	10.3
<b>Shannon diversity (<math>\Delta</math>RMSE <math>\times 10^{-3}</math>)</b>											
Geo	16.9	Geo	23.1	Geo	12.4	Benthos	8.46	Geo	10.3	Depth	6.55
Cor 200	10.7	Cor 1000	4.38	Benthos	10.7	Geo	7.38	Benthos	5.99	Cur 4	3.22
Cor 30	7.89	Dist	0.92	A Cur 10	5.39	Cor 500	5.64	Depth	4.20	A Cur 4	1.73

**Figure 5.** Predictions of the three fish response variables: (a) Shannon diversity, (b) biomass, and (c) species richness. Subset areas are outlined by the black box in (a), and correspond to those shown in Figure 4.



## 5. Discussion

Predicted spatial inventories of coral reef fisheries resources have multiple advantages for fisheries management and conservation [54], including improved information for prioritizing locations for conservation and management through marine spatial planning and the ability to assess potential changes to fish communities following large disturbances that alter reef habitat structure. However, their utility depends on the accuracies of the model predictions, which themselves depend on the accuracy of mapped predictor variables as well as the statistical relationships between predictors and response variables.

The use of hierarchical spatial scales and object-based image analysis to map coral reef habitats in this study is unique, and allowed a much greater level of thematic detail than is normally produced with pixel-based approaches (33 benthic and nine geomorphic classes covering the entire Kubulau fisheries management area). Our classification accuracies also compare favorably with pixel-based classifiers given the number of classes [26,55,56]. As classification errors will propagate to the

relevant derived predictor variables, in our case coral cover and habitat richness, this high accuracy lends confidence that the derived predictors are also good reflections of the actual spatial distribution of coral cover and habitat richness in the area. The depth map, on the other hand, was less accurate than what has been achieved in other comparable studies (e.g., [25,57-59]). This was primarily due to the limited number of in situ depth measurements available, particularly over high-albedo areas used to optimize the depth model parameters. Some of this error will have propagated to the curvature and absolute curvature variables, although spatial autocorrelation of errors as reported by Su *et al.* [57], if also present in our data, will have limited the impact of imprecise depth estimates on these variables. Nevertheless some reduction in prediction error would likely result from further improvements in the derivation of habitat variables in general, and an improved depth map in particular.

Of the six model types developed, the random forest model produced predictions with the lowest errors for all three response variables (Table 3). This differs from the result of Knudby *et al.* [28], where the bagging model outperformed the other five model types. However, bagging and random forest models are similar as both rely on ensembles of individual decision trees, the difference being that the individual trees are decorrelated in random forests by randomly selecting a subset of predictor variables to split each node in the tree [60]. Whether this randomization produces higher or lower prediction errors thus depends on the structure of the particular dataset in question. A tentative conclusion from the two studies is that an ensemble of decision trees is superior to the other tested model types for predictions of fish community variables, although a functional explanation for this superiority would be largely speculative without further investigation.

Apart from comparisons with the results from other model types, interpretation of the RMSE values obtained by the random forest models is not straight-forward as there is little to compare them with. They can be judged in the context of the range of values each response variable has in the field data, where species richness ranges from 1.0 to 38.2 species per transect, biomass from 0 to 5,374 kg/ha, and diversity from 0 to 3.09 (unitless). The RMSE values thus correspond to 17% of the range of species richness values, 11% of the range of biomass values, and 18% of the range of diversity values. Although this measure is sensitive to extreme values in the field data (e.g., the maximum biomass observed), it gives a quantitative indication that all models performed well. In addition, a subjective evaluation of the predictive maps, given our knowledge of the area, shows a very reasonable spatial distribution of predicted values for all response variables. However, these evaluations do not answer the important question, in the management context, of whether the maps are sufficiently accurate to be used in conservation planning. The answer to this question will depend on numerous factors not related to the maps themselves, including what alternative sources of comparable information are available and the acceptance of the modeling process by stakeholders.

As shown in Table 4, variable importance differed substantially depending on model type. Although the higher predictive power of the random forest models does not guarantee that this model type better reflects the complex functional relationships between predictors and response variables, it is more likely to do so than a model type with lower predictive power. However, the differences between the RMSE values of most model types are small, and none of the other models perform so poorly as to be discounted completely. This situation, which reflects the multi-collinearity of our set of predictors, makes it difficult to draw ecologically relevant conclusions from the reported variable importances.



With this in mind, the following discussion is based on variable importance results from the random forest model.

In the random forest models, the high importance of geomorphic zone for predictions of all response variables is interesting because this predictor has been investigated by very few studies [13], including several that have explored the use of a large variety of metrics for quantifying the seascape (e.g., [16-18,32,61]). The importance of geomorphic zone in our study may be a function of the spatial extent of our study area, a function of the high accuracy with which this variable was mapped, a function of its collinearity with another predictor, and/or a true description of the importance of reef geomorphology for structuring the spatial distribution of the three response variables in our study area. More detailed research would be needed to draw firmer conclusions. Other predictors of importance in the random forest models include depth and benthos, which echoed findings from Friedlander *et al.* [62] and Knudby *et al.* [28]. Of note, ‘distance from land’ was found important for predictions of fish biomass. The importance of ‘distance from land’ may partially come from its function as a proxy variable for fishing pressure, known to decrease with distance from local villages in Kubulau [36]. Meanwhile, the variable ‘conservation status’ was striking for its lack of importance in all three models of fish assemblage characteristics. This may be attributed to the mixed effectiveness of the MPAs surveyed during 2009, which is partly due to illegal fishing from both traditional fishing rights owners and poachers coming from adjacent districts [6]. In cases where protection has been strongly enforced, such as the Chumbe Island MPA in Tanzania, the variable has had much greater importance in model outputs of fish biomass [28]. Habitat richness was also of low importance, regardless of model type and spatial scale. Although this corresponds to findings from other studies (e.g., [16,17]), the question remains whether other measures of the spatial composition of habitats, e.g., taking into account functional differences or complementarities between habitat types [63], would contribute more to predictions of the fish community variables. The success of landscape composition metrics in terrestrial spatial ecology [64] suggests that this is an important avenue for future research.

None of the three most important predictors in the random forest models (Table 4) were directly calculated at fine spatial scale, which suggests that high spatial resolution satellite data may not be crucial for accurate predictions of our response variables. Although predictors calculated at a fine spatial scale had high importance in several other model types, these models produced relatively poorer predictions of the response variables, and their important variables can therefore not be regarded as essential. Any substantial gain from high spatial resolution satellite data would thus have to come primarily from its influence on the accuracy of the geomorphic and benthic classifications [56] and the depth estimation, and the propagated influence on predictors derived from these three variables. This suggestion stands in contrast to results reported elsewhere. Pittman *et al.* [30] found that structural complexity calculated at 15–25 m radii were the most important predictors in models of a variety of fish community metrics, Purkis *et al.* [65] found rugosity calculated at an 8 m radius to have the highest correlations with fish species richness and abundance, and Knudby *et al.* [28] found rugosity calculated at a 6 m radius to have the highest importance for predictions of fish species richness, diversity, and biomass. Meanwhile, Wedding *et al.* [66] found that slightly higher correlations with fish species richness, abundance and biomass were obtained with rugosity calculated at a 25 m radius, compared to similar rugosity calculations at 4 m, 10 m and 15 m radii. However, comparisons between

these studies are complicated by several factors. Study sites vary between the Caribbean and the Indo-Pacific, with associated differences in environment and species composition. The species composition can impact the predictability of fish community variables, as illustrated by the differences in predictability between functional groups shown by Pittman *et al.* [30]. The specific fish survey method may also play a role. Both Purkis *et al.* [65] and Knudby *et al.* [28] used 5 m radius point counts while Pittman *et al.* [30] and Wedding *et al.* [66], as well as this study, used belt transects. Point counts cover a smaller area, but give the observer more time to search for shy and cryptic species, a greater proportion of which are therefore recorded. The impact of survey method on the predictability of fish community variables is likely to be greater for species richness and diversity than for biomass, but has not been tested. In addition, the metric used to quantify structural complexity differs between studies, as does the way in which the spatial scale of its calculation is varied; while Purkis *et al.* [65] and Pittman *et al.* [30] used a fixed terrain model with 4 m cell size and varied the window size used to calculate their structure metrics (as for standard deviation of water depth and surface rugosity in this study), Wedding *et al.* [66] used terrain models of varying cell size and a fixed  $3 \times 3$  cell window to calculate rugosity (as for the slope and curvature metrics in this study). Another reason why variables at fine spatial scales have limited importance in our results may be the spatial coverage represented by each data point in the fish survey. As opposed to the 3–5 transects each covering  $250 \text{ m}^2$  used in our study, the other studies surveyed their fish community data at a single  $100 \text{ m}^2$  transect [30], four replicate  $79 \text{ m}^2$  point counts [65], a single  $79 \text{ m}^2$  radius point count [28], and a single  $125 \text{ m}^2$  transect [66]. Given the ambiguity of results from individual studies, focused investigation of the predictions that could be produced on the basis of coarser scale satellite data is important because its use would enable predictive models to cover much larger areas at a reduced cost, using e.g., free Landsat Thematic Mapper data or maps of areas already classified and available from the Millennium Coral Reef Landsat Archive [67]. Landsat data have already been used extensively to map reef geomorphology at a near global scale [68] and predict biomass of giant clams [69], another important reef resource for livelihoods.

We speculate that rather than the spatial resolution of the satellite data, a more significant obstacle to improved predictions is the uncertainty surrounding how well fish data from one-off underwater visual censuses represent the true means of the fish community variables at each site. The presence/absence of large individuals or schools of transient species during the time of a particular survey introduces substantial noise in the data, particularly influencing the fish biomass variable. Although we used 3–5 replicate transects per site to reduce this source of error, natural near-instantaneous differences in fish biomass has been shown elsewhere be greater than between-observer variation and similar to between-site variation [70]. However, given a fixed time and budget for field work, it will always be necessary to strike a balance between the number of sites surveyed versus replication at the site and transect levels to optimize habitat representation while minimizing noise.

In addition to the substantial difference between the importances of individual predictors depending on model type, there are several other caveats related to the interpretation of the variable importance results. Highly correlated predictor variables, most clearly exemplified by the same variable calculated at neighboring spatial scales, can substitute for one another in the random forest models, and therefore both be given lower importance than if only one had been present in the data set. In addition, variable

importance depends on the spatial extent of the study, e.g., the “distance from land” variable would have been less important if the study site had been limited to areas close to the mainland of Vanua Levu. Similarly, variable importance depends on the data coverage of the parameter space, e.g., if only a narrow range of coral cover values were sampled, the importance of coral cover in the resulting model would be reduced. Furthermore, as suggested above, variable importances depend on the detail and error with which each variable is mapped. A poorly estimated variable, or a variable that has been mapped at a coarse thematic resolution, will, regardless of its theoretical importance, be unable to contribute to predictions of the response variable and will therefore have low importance. This is also the case if an important ecological factor (e.g., structural complexity) has been quantified using an inappropriate metric (e.g., standard deviation of depth). These issues may contribute to the high importance of geomorphic zone, an accurately mapped variable in our study, as well as the low importance of curvature and absolute curvature at fine spatial scales, variables that were expected from other studies to have substantial importance for all response variables [28,65,66] but in our study were calculated on the basis of a depth map with relatively high associated errors. In addition, variable selection in the individual decision trees that comprise the random forest is known to be biased, e.g., for categorical variables those with a greater number of categories (e.g., benthos, geomorphology) are chosen more frequently than those with fewer categories (e.g., conservation status) regardless of information content [52]. Given these numerous caveats, conclusions about variable importance from multi-variable models should be cautious, and variables with higher importance should primarily be seen as relevant variables to target for further investigation in future studies.

The predictive maps of fish biomass, species richness and diversity represent an improvement over previous spatial modeling of fish abundance and biomass in Kubulau, which only considered three habitat classes (fringing reef, patch reef, barrier reef) as well as other binary predictor variables (exposure to tides, exposure to waves, depth <10 m) and distance from land, within Poisson, negative-binomial and zero-inflated count models [36]. In order to translate the results into management implementation, biomass and species diversity targets for the Kubulau fisheries management area will be developed and weighed against prior calculations of opportunity costs to fishers [36], enabling revision of recommendations for MPA network configuration in Kubulau before these are presented to stakeholders for consultation. The communities of Kubulau will then have the option to adapt their MPA boundaries to reduce internal conflict among resource users [33] and foregone revenue [36] while maximizing diversity and fish biomass within protected areas to promote ecosystem function, resilience and fisheries benefits from spillover. Incorporating knowledge of each variable’s associated prediction errors in the planning process will be crucial in order to balance model predictions optimally with other information sources; this is an area of ongoing research.

## 6. Conclusion

The work described in this paper has demonstrated the feasibility of producing spatially explicit predictions of fish species richness, biomass and diversity, at the spatial extent of the fisheries management area in Kubulau, Fiji. The object-based approach to classification of the IKONOS and Quickbird data produced accurate maps at hierarchical spatial scales of the geomorphology and benthos, both of which were important for predictions of the response variables. Of the six predictive

model types investigated, random forests produced lowest cross-validated errors for all response variables. Prediction errors are reasonable considering the range of values for each response variable in the data set, and the resulting maps show a spatial distribution of the response variables that is in accord with our knowledge of the area. Although this approach can theoretically be applied to any reef system in the world, major obstacles to its replication include lack of technical expertise and funding for data and software in relevant organizations. However, free open source software packages are available for both image segmentation and object-based classification (e.g., GRASS GIS) and traditional pixel-based image processing and classification (e.g., Opticks, ILWIS), as well as for statistical modeling (e.g., R). In addition, there is an increasing number of options available for medium/high resolution satellite data, some free (e.g., Landsat), and many commercial (e.g., Alos, GeoEye-1, WorldView-2, RapidEye). Future efforts should attempt to quantify the influence on prediction errors of both the spatial resolution of the satellite data and the classification approach—if comparable results could be achieved with free data and free software the development and application of maps such as those produced here would have great potential to improve marine spatial planning in other managed coral reef areas.

### Acknowledgements

WCS acknowledges financial support from the David and Lucile Packard Foundation (#2007-31847, #2009-34839), the Gordon and Betty Moore Foundation (540.01), and the US National Oceanic and Atmospheric Administration (NA07NOS4630035). The authors gratefully acknowledge the support of the chiefs and communities of Kubulau and the WCS staff and volunteers who assisted with field data collection and photo analysis: A. Cakacaka, S. Dulunaqio, U. Mara, W. Moy, W. Naisilisili, I. Qauqau, T. Tui, and P. Veileqe. We additionally thank the Centre for Spatial Environmental Research at the University Of Queensland for providing software, hardware and assistance with the development of the object-based analysis approaches.

### References

1. Bryant, D.; Burke, L.; McManus, J.; Spalding, M. *Reefs at Risk: A Map-Based Indicator of Potential Threats to the World's Coral Reefs*; World Resources Institute: Washington, DC, USA, 1998; p. 56.
2. Andrew, N.L.; Bene, C.; Hall, S.J.; Allison, E.H.; Heck, S.; Ratner, B.D. Diagnosis and management of small-scale fisheries in developing countries. *Fish Fish.* **2007**, *8*, 227-240.
3. Gaichas, S.K. A context for ecosystem-based fishery management: Developing concepts of ecosystems and sustainability. *Marine Policy* **2008**, *32*, 393-401.
4. Ainsworth, C.H.; Varkey, D.A.; Pitcher, T.J. Ecosystem simulations supporting ecosystem-based fisheries management in the Coral Triangle, Indonesia. *Ecol. Model.* **2008**, *214*, 361-374.
5. Lester, S.E.; Halpern, B.S.; Grorud-Colvert, K.; Lubchenco, J.; Ruttenberg, B.I.; Gaines, S.D.; Airame, S.; Warner, R.R. Biological effects within no-take marine reserves: A global synthesis. *Marine Ecol. Progr. Ser.* **2009**, *384*, 33-46.
6. Jupiter, S.D.; Egli, D.P. Ecosystem-based management in Fiji: Successes and challenges after five years of implementation. *J. Marine Biology* **2011**, doi:10.1155/2011/940765.

7. Syms, C.; Jones, G.P. Disturbance, habitat structure, and the dynamics of a coral-reef fish community. *Ecology* **2000**, *81*, 2714-2729.
8. Gratwicke, B.; Speight, M.R. The relationship between fish species richness, abundance and habitat complexity in a range of shallow tropical marine habitats. *J. Fish Biol.* **2005**, *66*, 650-667.
9. Friedlander, A.M.; Parrish, J.D. Habitat characteristics affecting fish assemblages on a Hawaiian coral reef. *J. Exp. Mar. Biol. Ecol.* **1998**, *224*, 1-30.
10. Lara, E.N.; González, E.A. The relationship between reef fish community structure and environmental variables in the southern Mexican Caribbean. *J. Fish Biol.* **1998**, *53*, 209-221.
11. Connell, J.H. Diversity in tropical rain forests and coral reefs—High diversity of trees and corals is maintained only in a non-equilibrium state. *Science* **1978**, *199*, 1302-1310.
12. Jones, G.P.; McCormick, M.I.; Srinivasan, M.; Eagle, J.V. Coral decline threatens fish biodiversity in marine reserves. *Proc. Nat. Acad. Sci. USA* **2004**, *101*, 8251-8253.
13. Christensen, J.D.; Jeffrey, C.F.; Caldow, C.; Monaco, M.E.; Kendall, M.S.; Appeldoorn, R.S. Cross shelf habitat utilization patterns of reef fishes in southwestern Puerto Rico. *Gulf Caribbean Res.* **2003**, *14*, 9-27.
14. Dorenbosch, M.; Grol, M.G.G.; Christianen, M.J.A.; Nagelkerken, I.; van der Velde, G. Indo-Pacific seagrass beds and mangroves contribute to fish density coral and diversity on adjacent reefs. *Marine Ecol. Progr. Ser.* **2005**, *302*, 63-76.
15. Mumby, P.J.; Edwards, A.J.; Arias-Gonzalez, J.E.; Lindeman, K.C.; Blackwell, P.G.; Gall, A.; Gorczynska, M.I.; Harborne, A.R.; Pescod, C.L.; Renken, H.; Wabnitz, C.C.C.; Llewellyn, G. Mangroves enhance the biomass of coral reef fish communities in the Caribbean. *Nature* **2004**, *427*, 533-536.
16. Pittman, S.J.; McAlpine, C.A.; Pittman, K.M. Linking fish and prawns to their environment: a hierarchical landscape approach. *Marine Ecol. Progr. Ser.* **2004**, *283*, 233-254.
17. Pittman, S.J.; Christensen, J.D.; Caldow, C.; Menza, C.; Monaco, M.E. Predictive mapping of fish species richness across shallow-water seascapes in the Caribbean. *Ecol. Model.* **2007**, *204*, 9-21.
18. Grober-Dunsmore, R.; Frazer, T.K.; Beets, J.P.; Lindberg, W.J.; Zwick, P.; Funicelli, N.A. Influence of landscape structure, on reef fish assemblages. *Landscape Ecol.* **2008**, *23*, 37-53.
19. Mumby, P.J.; Green, E.P.; Edwards, A.J.; Clark, C.D. Coral reef habitat-mapping: How much detail can remote sensing provide? *Marine Biol.* **1997**, *130*, 193-202.
20. Mumby, P.J.; Hedley, J.D.; Chisholm, J.R.M.; Clark, C.D.; Ripley, H.; Jaubert, J. The cover of living and dead corals from airborne remote sensing. *Coral Reefs* **2004**, *23*, 171-183.
21. Newman, C.; Knudby, A.; LeDrew, E. Assessing the effect of management zonation on live coral cover using multi-date IKONOS satellite imagery. *J. Appl. Remote Sens.* **2007**, *1*, doi:10.1117/1.2822612.
22. Isoun, E.; Fletcher, C.; Frazer, N.; Gradie, J. Multi-spectral mapping of reef bathymetry and coral cover; Kailua Bay, Hawaii. *Coral Reefs* **2003**, *22*, 68-82.
23. Lyzenga, D.R. Passive remote-sensing techniques for mapping water depth and bottom features. *Appl. Opt.* **1978**, *17*, 379-383.
24. Maritorena, S.; Morel, A.; Gentili, B. Diffuse-reflectance of oceanic shallow waters—Influence of water depth and bottom albedo. *Limnol. Oceanogr.* **1994**, *39*, 1689-1703.

25. Stumpf, R.P.; Holderied, K.; Sinclair, M. Determination of water depth with high-resolution satellite imagery over variable bottom types. *Limnol. Oceanogr.* **2003**, *48*, 547-556.
26. Roelfsema, C.; Phinn, S.; Jupiter, S.; Comley, M.; Beger, M.; Peterson, E. Object Based Analysis of High Spatial Resolution Imagery for Mapping Large Coral Reef Systems in The West Pacific at Geomorphic and Benthic Community Scales. In *Proceedings of 2010 IEEE International Geoscience and Remote Sensing Symposium*, Honolulu, HI, USA, 25–30 July 2010.
27. Mellin, C.; Bradshaw, C.J.A.; Meekan, M.G.; Caley, M.J. Environmental and spatial predictors of species richness and abundance in coral reef fishes. *Glob. Ecol. Biogeogr.* **2010**, *19*, 212-222.
28. Knudby, A.; LeDrew, E.; Brenning, A. Predictive mapping of reef fish species richness, diversity and biomass in Zanzibar using IKONOS imagery and machine-learning techniques. *Remote Sens. Environ.* **2010**, *114*, 1230-1241.
29. Wedding, L.M.; Friedlander, A.M. Determining the Influence of Seascape Structure on Coral Reef Fishes in Hawaii Using a Geospatial Approach. *Marine Geodesy* **2008**, *31*, 246-266.
30. Pittman, S.J.; Costa, M.B.; Battista, T.A. Using lidar bathymetry and boosted regression trees to predict the diversity and abundance of fish and corals. *J. Coastal Res.* **2009**, *SI*, 27-38.
31. Kuffner, I.B.; Brock, J.C.; Grober-Dunsmore, R.; Bonito, V.E.; Hickey, T.D.; Wright, C.W. Relationships between reef fish communities and remotely sensed rugosity measurements in Biscayne National Park, Florida, USA. *Environ. Biol. Fish.* **2007**, *78*, 71-82.
32. Mellin, C.; Andréfouët, S.; Ponton, D. Spatial predictability of juvenile fish species richness and abundance in a coral reef environment. *Coral Reefs* **2007**, *26*, 895-907.
33. Clarke, P.; Jupiter, S.D. Law, custom and community-based natural resource management in Kubulau District (Fiji). *Environ. Conserv.* **2010**, *37*, 98-106.
34. Cakacaka, A.; Jupiter, S.D.; Egli, D.P.; Moy, W. *Status of Fin Fisheries in a Fijian Traditional Fishing Ground, Kubulau District, Vanua Levu*; Technical Report No. 06/10; Wildlife Conservation Society: Suva, Fiji, 2010; p. 21.
35. WCS. *Ecosystem-Based Management Plan: Kubulau District, Vanua Levu, Fiji*; Wildlife Conservation Society: Suva, Fiji, 2009; p. 121.
36. Adams, V.M.; Mills, M.; Jupiter, S.D.; Pressey, R.L. Improving social acceptability of marine protected area networks: A method for estimating opportunity costs to multiple gear types in both fished and currently unfished areas. *Biol. Conserv.* **2010**, *144*, 350-361.
37. Roelfsema, C.; Phinn, S. Integrating field data with high spatial resolution multispectral satellite imagery for calibration and validation of coral reef benthic community maps. *J. Appl. Remote Sens.* **2010**, *4*, 043527, doi:10.1117/1.3430107.
38. Andréfouët, S.; Guzman, H.M. Coral reef distribution, status and geomorphology-biodiversity relationship in Kuna Yala (San Blas) archipelago, Caribbean Panama. *Coral Reefs* **2005**, *24*, 31-42.
39. Kohler, K.E.; Gill, S.M. Coral Point Count with Excel extensions (CPCe): A Visual Basic program for the determination of coral and substrate coverage using random point count methodology. *Computer. Geosci.* **2006**, *32*, 1259-1269.
40. English, S.; Wilkinson, C.; Baker, V. *Survey Manual for Tropical Marine Resources*; Australian Institute of Marine Science: Townsville, QLD, Australia, 1997; p. 390.

41. Mumby, P.J.; Harborne, A.R. Development of a systematic classification scheme of marine habitats to facilitate regional management and mapping of Caribbean coral reefs. *Biol. Conserv.* **1999**, *88*, 155-163.
42. Hill, J.; Wilkinson, C. *Methods for Ecological Monitoring of Coral Reefs. Version 1: A Resource for Managers*; Australian Institute of Marine Science and Reef Check: Townsville, QLD, Australia, 2004; p. 117.
43. Froese, R.; Pauly, D. Fishbase. Available online: [www.fishbase.org](http://www.fishbase.org). (accessed on 12 January 2009).
44. Shannon, C.E.; Weaver, W. *The Mathematical Theory of Communication*; University of Illinois Press: Urbana, IL, USA, 1963; p. 144.
45. Blaschke, T. Object based image analysis for remote sensing. *ISPRS J. Photogramm. Remote Sens.* **2010**, *65*, 2-16.
46. Phinn, S.; Roelfsema, C.; Mumby, P. Multi-scale image segmentation for mapping coral reef geomorphic and benthic community zone. *Int. J. Remote Sens.* **2011**, in press.
47. Congalton, R.; Green, K. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*; Lewis Publishers: Boca Raton, FL, USA, 1999; p. 137.
48. Knudby, A.; Brenning, A.; LeDrew, E. New approaches to modelling fish-habitat relationship. *Ecol. Model.* **2010**, *221*, 503-511.
49. R Core Development Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2008. Available online: <http://www.R-project.org/> (accessed on 28 February 2011).
50. Hastie, T.; Tibshirani, R.; Friedman, J.H. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*; Springer-Verlag: New York, NY, USA, 2001; p. 763.
51. Murray, K.; Conner, M.M. Methods to quantify variable importance: Implications for the analysis of noisy ecological data. *Ecology* **2009**, *90*, 348-355.
52. Strobl, C.; Boulesteix, A.-L.; Zeileis, A.; Hothorn, T. Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC Bioinf.* **2007**, *8*, doi:10.1186/1471-2105-8-25.
53. Strobl, C.; Boulesteix, A.L.; Kneib, T.; Augustin, T.; Zeileis, A. Conditional variable importance for random forests. *BMC Bioinf.* **2008**, *9*, doi:10.1186/1471-2105-9-307.
54. Hamel, M.A.; Andr  ou  , S. Using very high resolution remote sensing for the management of coral reef fisheries: Review and perspectives. *Mar. Pollut. Bull.* **2010**, *60*, 1397-1405.
55. Andr  ou  , S.; Kramer, P.; Torres-Pulliza, D.; Joyce, K.E.; Hochberg, E.J.; Garza-Perez, R.; Mumby, P.J.; Riegl, B.; Yamano, H.; White, W.H.; Zubia, M.; Brock, J.C.; Phinn, S.R.; Naseer, A.; Hatcher, B.G.; Muller-Karger, F.E. Multi-site evaluation of IKONOS data for classification of tropical coral reef environments. *Remote Sens. Environ.* **2003**, *88*, 128-143.
56. Mumby, P.J.; Edwards, A.J. Mapping marine environments with IKONOS imagery: Enhanced spatial resolution can deliver greater thematic accuracy. *Remote Sens. Environ.* **2002**, *82*, 248-257.
57. Su, H.B.; Liu, H.X.; Heyman, W.D. Automated derivation of bathymetric information from multi-spectral satellite imagery using a Non-Linear Inversion Model. *Marine Geodesy* **2008**, *31*, 281-298.

58. Muslim, A.M.; Foody, G.M. DEM and bathymetry estimation for mapping a tide-coordinated shoreline from fine spatial resolution satellite sensor imagery. *Int. J. Remote Sens.* **2008**, *29*, 4515-4536.
59. Lyzenga, D.R.; Malinas, N.R.; Tanis, F.J. Multispectral bathymetry using a simple physically based algorithm. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 2251-2259.
60. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5-32.
61. Grober-Dunsmore, R.; Frazer, T.K.; Lindberg, W.J.; Beets, J.P. Reef fish and habitat relationships in a Caribbean seascape: The importance of reef context. *Coral Reefs* **2007**, *26*, 201-216.
62. Friedlander, A.M.; Brown, E.K.; Monaco, M.E. Coupling ecology and GIS to evaluate efficacy of marine protected areas in Hawaii. *Ecol. Appl.* **2007**, *17*, 715-730.
63. Mumby, P.J. Beta and habitat diversity in marine systems: A new approach to measurement, scaling and interpretation. *Oecologia* **2001**, *128*, 274-280.
64. Turner, M.G.; Gardner, R.H.; O'Neill, R.V. *Landscape Ecology in Theory and Practice: Pattern and Process*; Springer Science + Business Media: New York, NY, USA, 2001; p. 404.
65. Purkis, S.J.; Graham, N.A.J.; Riegl, B.M. Predictability of reef fish diversity and abundance using remote sensing data in Diego Garcia (Chagos Archipelago). *Coral Reefs* **2008**, *27*, 167-178.
66. Wedding, L.; Friedlander, A.M.; McGranaghan, M.; Yost, R.; Monaco, M. Using bathymetric LIDAR to define nearshore benthic habitat complexity: Implications for management of reef fish assemblages in Hawaii. *Remote Sens. Environ.* **2008**, *112*, 4159-4165.
67. Androuët, S.; Muller-Krager, F.E.; Robinson, J.A. Millennium Coral Reef Landsat Archive. Available online: <http://oceancolor.gsfc.nasa.gov/cgi/landsat.pl> (accessed on 13 October 2010).
68. Androuët, S.; Muller-Krager, F.E.; Robinson, J.A.; Kranenburg, C.J.; Torres-Pulliza, D.; Spraggins, S.A.; Murch, B. Global Assessment of Modern Coral Reef Extent and Diversity for Regional Science and Management Applications: A View from Space. In *Proceedings of 10th International Coral Reef Symposium*, Okinawa, Japan, 28 June–2 July 2006; pp. 1732-1745.
69. Androuët, S.; Gilbert, A.; Yan, L.; Remoissenet, G.; Payri, C.; Chancerelle, Y. The remarkable population size of the endangered clam *Tridacna maxima* assessed in Fangatau Atoll (Eastern Tuamotu, French Polynesia) using in situ and remote sensing data. *Ices J. Mar. Sci.* **2005**, *62*, 1037-1048.
70. McClanahan, T.R.; Graham, N.A.J.; Maina, J.; Chabanet, P.; Bruggemann, J.H.; Polunin, N.V.C. Influence of instantaneous variation on estimates of coral reef fish populations and communities. *Marine Ecol. Progr. Ser.* **2007**, *340*, 221-234.