Vol. 361: 1–13, 2008
doi: 10.3354/meps07481

MARINE ECOLOGY PROGRESS SERIES
Mar Ecol Prog Ser

Published June 9

OPEN ACCESS

FEATURE ARTICLE

Environmental factors that influence the distribution of coral reef fishes: modeling occurrence data for broad-scale conservation and management

Maria Beger*, Hugh P. Possingham

The Ecology Centre, The School of Integrative Biology, The University of Queensland, St Lucia, Queensland 4072, Australia

ABSTRACT: To manage coral reef species, it is important to identify the factors that determine their distribution inexpensively. We identified the remotely measured environmental factors that are most influential in determining the distributions of coral reef fish species on a regional scale. Logistic regression models for 227 fish species related presence/absence data to 4 remotely determined environmental predictor variables: depth, presence of a land–sea interface, exposure, and the distance to the nearest estuary. We compared modeled Akaike information criterion (AIC) values with AIC values of randomly distributed species with different numbers of occurrences and levels of habitat specificity to evaluate model significance. Efficient species distribution models were identified for 118 predominantly habitat-specific fishes of the 227 species for which we had data. All 4 predictor variables significantly influenced the distributions of at least some fish species. Depth was the most frequently efficient variable for single variable models. For combinations of 2 predictor variables, depth and exposure, as well as depth and distance from the nearest estuary, were the prevalent predictors of fish distributions. Several fish species responded to the combination of variables distance from an estuary and presence of the terrestrial–marine interface, indicating that these species depend on intact coastal reef habitat, which is in decline near the main sediment-laden rivers. Statistically significant models were predominantly developed for habitat-specific species. These habitat-specific species are of greater conservation concern than widespread species because threats affect them more severely if the threats are selectively affecting their habitat. For this reason, species distribution modeling using remotely determined environmental data may be an efficient method to build models for habitat-specific species and inform marine reserve design.

Distribution modelling shows that bigeye jacks *Caranx sexfasciatus* typically school on exposed reefs.

Photo: M. Beger

KEY WORDS: Species distribution modeling · Logistic regression · Environmental parameters · Coral reef fishes · Coral reef conservation · GLM

Resale or republication not permitted without written consent of the publisher

INTRODUCTION

Detailed understanding of the processes determining the spatial distribution of species is of primary importance for resource managers designing effective conservation measures for ecosystems. Focal species for research on species distributions are often confined to a narrow habitat type or set of environmental condi-

*Email: m.beger@uq.edu.au

© Inter-Research 2008 · www.int-res.com

tions by traits such as specialized nutrition, restricted dispersal, or low abundance (Dunk et al. 2004). Such specialized, habitat-specific species are of greater conservation concern than widespread, more generalist species because of their restricted range (Lawler et al. 2003), or because threats that selectively impact their habitat may affect them more severely (Jones et al. 2004). In terrestrial ecosystems, species distribution models commonly serve to elucidate ecological processes affecting distributions (Lehmann et al. 2002), to predict responses of species to changes of environmental variables (Thomas et al. 2004), and to highlight conservation priorities (Peterson et al. 2000). In the present study, we applied these techniques to coral reef fishes, aiming to determine environmental factors that influence the distribution of common and habitat-specific fishes.

Coral reefs are highly diverse ecosystems harboring a significant proportion of the global biodiversity. They support many habitat-specific and restricted-range species (Roberts et al. 2002). Coral reefs are severely threatened by anthropogenic resource use, pollution, and by climate change (Hughes et al. 2003). Resource management to mitigate these threats relies on sound knowledge of ecosystem processes, spatial patterns of species richness and rarity, and evaluation of present and future threats. Comprehensive studies of the distributions of reef taxa, such as crustaceans, echinoderms, and algae, are largely lacking, but coral reef fishes and scleractinian corals are relatively well studied. For example, distributional maps of coral reef species on a global scale exist for corals (Veron 2000), fishes (Allen & Steene 1994), mollusks, and lobsters (Roberts et al. 2002), but they are inaccurate and frequently updated by range extensions and newly described species. Other work lists the species of fish occurring in a region or country, and species checklists exist for many localities (e.g. Myers 1999). Cross-shelf variation patterns of fish communities on the Great Barrier Reef have been widely described for many fish families, genera, or species (Williams 1991). These studies, however, are of limited use to resource managers aiming to represent all species, restricted-range species, or flag-ship species across a project area at a local or sub-regional scale, because they are descriptive in nature and do not usually explain the factors influencing these patterns. A sound knowledge of which environmental parameters are most likely to influence the distribution of species will better enable conservation planners to identify which reefs are most beneficial to manage for these species.

Empirical ecological models describing links between coral reef fishes and their required environment are rare, as the necessary experimental development of such ecological models is time consuming and expensive. Indo-Pacific fish and coral species richness patterns have been attributed to environmental and geometric constraints such as habitat area and the mid-domain effect (Bellwood et al. 2005). Other global scale models relate fish distribution to pelagic larval duration (Mora et al. 2003) and describe patterns of endemism (Robertson 2001). Empirical studies on reef micro-habitat characteristics highlighted environmental factors such as exposure, depth, food availability, topographic complexity, water quality, distance to deep water, distance from consolidated reef, sediment composition and the level of disturbance, and spatial isolation of reefs (e.g. Holbrook et al. 2002) as the main environmental conditions influencing the distribution of reef fishes. Statistical species distribution models for coral reef fauna are lacking on the sub-regional and regional scales relevant to both environmental change and conservation. Identifying valid models may be the first step towards predicting responses of coral reef fauna to changes in the environmental conditions, similar to modeling applications in the terrestrial realm (Thomas et al. 2004).

Species distribution modeling methods have been widely applied in terrestrial ecology to relate presence, presence/absence, or abundance records of species to environmental factors (Elith 2000). Species distribution models thus present a powerful tool to explore a wide range of ecological and biogeographical questions. Such methods were applied to examine spatial biodiversity patterns of species richness (Peterson et al. 2000), beta-diversity (Cayuela et al. 2006), and rarity (Engler et al. 2004), and to hindcast (Hugall et al. 2002) and forecast species distributions (Araujo & New 2007). The application of species distribution modeling to marine species and communities is relatively rare and mostly confined to a single species or a single community. For example, Kelly et al. (2001) predicted seagrass occurrences based on a relative exposure index and water depth with a logistic regression model. Bathymetry data and spectral bands from remote imagery were successfully used to predict the presence of coral communities in Mexico (Garza-Perez et al. 2004).

The increasing availability of spatially explicit, electronic remote sensing environmental data in the marine realm may render the application of species distribution modeling approaches more attractive to marine ecologists. An introduction to these methods can be found in the review by Elith et al. (2006) that compares 16 well-established and novel species distribution modeling methods, providing an overview of the advantages and disadvantages of each approach. Appraisals of species distribution modeling highlight generalized linear models (GLMs) as an efficient

method; it has a strong statistical foundation, yet is straightforward to use and flexible, allowing ecological relationships to be modeled realistically (Austin 2002). Predictions resulting from GLMs can be easily transferred into geographical information systems (GIS) for mapping (Wintle et al. 2005). The underlying regression analysis is explicit, consistent, and repeatable, and creates models that are fully data defined (Lehmann et al. 2002). GLMs can be applied when the statistical distribution function of the data is uncertain (Elith et al. 2006). In particular, GLMs accommodate variance heterogeneity and asymmetric non-normal behavior, such as encountered in binomial data (Venables & Dichmont 2004).

In the present paper we identify the environmental factors most influential in determining distributions of coral reef fish species on a sub-regional scale based on GLM methods. Because of the limited availability and accuracy of environmental data that can serve as predictor variables, this study constitutes a first test of the applicability of species distribution modeling methods to reef fishes as a basis for future, more complex models. We apply a novel method of evaluating model significance for multiple species that compares model Alaike information criterion (AIC) values of real and random distributions for each species, and relates these values to indices for commonness and habitat specificity.

METHODS

Study site. The study was conducted in Kimbe Bay, New Britain (5° 15′ S, 150° 15′ E), Papua New Guinea (PNG) (Fig. 1). This bay is about 100 km wide and supports a large number of discrete patch reefs, barrier reefs, and fringing reefs. All reefs were mapped and assigned to sites, which were either whole patch reefs or fringing reefs divided perpendicular to the coast to create approximately equal-sized reef sections. We assembled an extensive database for Kimbe Bay and the adjacent land areas in a GIS, including physical, topographical, political, geographical, remotely sensed, and biological data (Beger et al. 2006, S. Sheppard pers. comm.). All data were assigned in a way that each site had only a single entry for each type of data.

Species data. The fish species data used in the present study were compiled from 3 sources: (1) a coral and fish biodiversity assessment carried out in 1994 in the western part of the bay (Allen & Munday 1994), (2) a fish biodiversity study conducted in 2001 by M. Beger on the western side of the bay, and (3) a coral and fish biodiversity assessment carried out in 2002 in the eastern part of the bay (Beger 2002). Presence/absence species counts of reef fishes collected by timed-swim visual observations were available for 105 sites (Fig. 1). These data were collected on deep to shallow belts that crossed all reef zones. The timed-

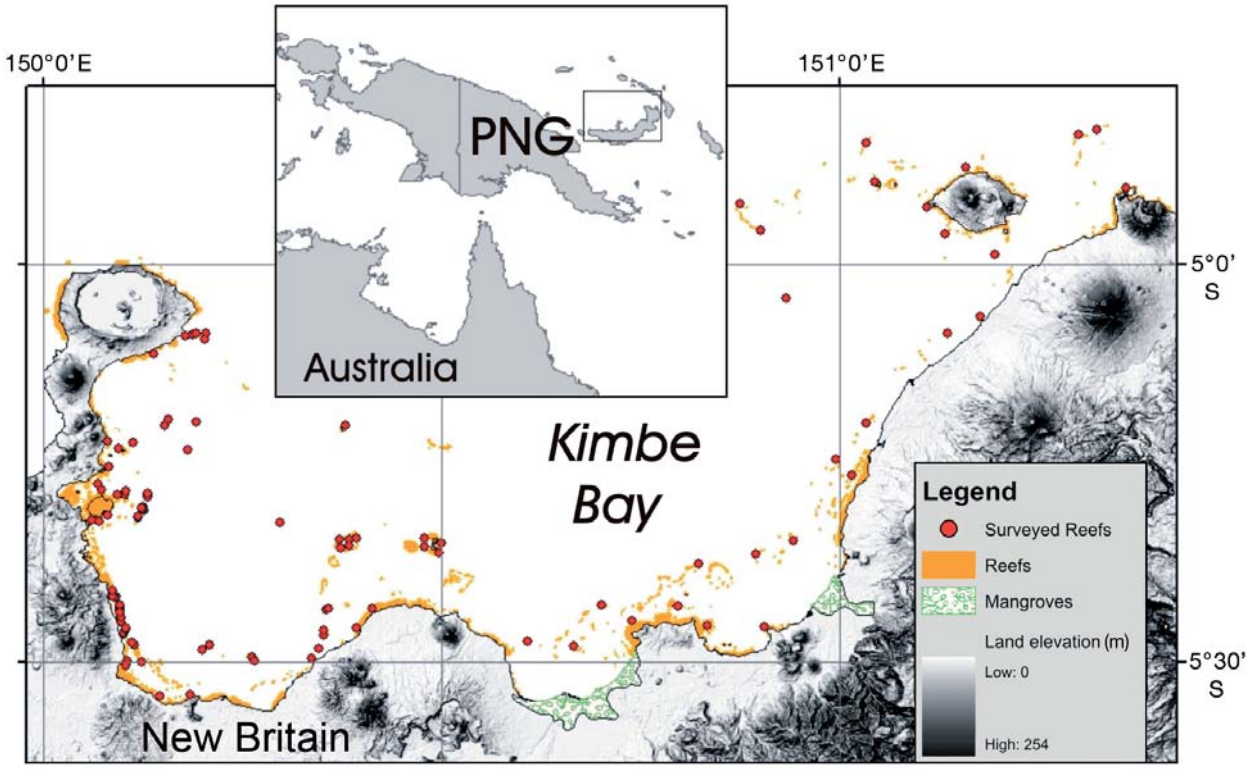

Fig. 1. Map of Kimbe Bay, with surveyed sites as red dots. PNG: Papua New Guinea

swim surveys were standardized to 60–80 min survey time per site; thus, we assumed that the area of each site was reasonably constant (Beger et al. 2003).

Reef characteristics as predictor variables. Data layers of physical characteristics of reefs were used to identify a suite of candidate predictor variables. All parameters were determined remotely from ocean charts, remote sensing, community analysis, and local and expert knowledge (Table 1). Exposure data were defined by local knowledge and prevailing wind directions. The bathymetry data layer was interpolated from a chart of Kimbe Bay (Australian Hydrographic Service, 1967, new edition 1992, AUS 547). Biological variables included 2 different methods of habitat classifications.

Firstly, 2 Landsat 7 images were classified into 5 broad reef classes: fringing reef, patch reef, atoll, outer fringing reef, and barrier reef (S. Sheppard pers. comm. in 2003). Secondly, we mapped 9 bioregions based on a principal component analysis of scleractinian coral diversity data. The data were assessed separately for the eastern and western parts of the bay, in accordance with existing publications and data sources. To account for potential anthropogenic influences on fish distributions, we created data layers for the distance of a reef from the nearest shore and the nearest river estuary, serving as surrogates for the intensity of sedimentation from the common practice of rainforest clearing and subsequent establishment of palm oil plantations. These data layers were calculated by spatial analysis of GIS data layers for rivers derived from the biodiversity assessment and planning (BioRap) study for PNG (Faith et al. 2000).

Co-linearity between variables can reduce the precision in estimating the variable coefficients and their significance. For this reason, we identified a subset of uncorrelated predictors from candidate variables through an iterative process. Because several of the predictors were based on expert opinion, the first step was to also apply expert opinion to identify correlated variables (Wintle et al. 2005). Secondly, we identified relationships between variables with (1) chi-squared cross-tabulation tests for categorical versus categorical variables, (2) Kruskal-Wallis tests for categorical versus ordinal variables, (3) ANOVA for categorical versus numerical variables, (4) Pearson's correlation among numerical variables, and (5) Spearman's rank correlation analysis for ordinal versus ordinal and ordinal versus numerical variables (compare Table 1 for variable types). Lastly, we conducted a multivariate ordination of the variables, based on hierarchical cluster analysis of Bray-Curtis similarities.

Fish species distribution models. We developed species distribution models for fishes based on the glm and step functions in R (R Development Core Team 2005). Given the binomial nature of the dependent variable (i.e. presence or absence of fish species), the logit link function and binomial error function were used to find explanatory variables for fish species distributions. This GLM method applied a logistic regression of the type:

$$\ln\left(\frac{p}{1-p}\right) = b_0 + b_1x_1 + b_2x_2 + \ldots + b_kx_k \quad (1)$$

where p was the probability of the species being present and $b_{1,2,..,k}$ were the coefficients for the predictor variables $x_{1,2,..,k}$. In order for the GLM method to work, several properties of the data were assumed to be satisfied here: (1) a dependent variable (presence or absence of a fish species) exists for any possible value of the independent variables (environmental parameters), (2) the relationship between the dependent

Table 1. Candidate predictor variables, their characteristics, and variable type for coral reef fish distribution modeling

Code	Predictor	Proxy for:	Data type
Habitat	Reef class (fringing reef, patch reef, atoll, outer fringing reef, barrier reef)	Reef topography	Categorical, 5 classes
Eco.Name	Bioregions (east inner coastal fringing reefs, east inner exposed coastal reefs, east outer coastal reefs, east outer patch reefs, exposed offshore patch reefs, west inner coastal fringing reefs, west outer coastal reefs, west inner patch reefs, west offshore reefs)	Hard-coral distribution patterns	Categorical, 9 classes
Exposure	Exposure (low, medium, upper medium, high)	Wave energy	Ordinal, 4 classes
Island	Presence of land–water interface	Interface habitat	Ordinal, 2 classes
Depth500	Bathymetry (mean depth at 500 m proximity)	Reef slope angle near reef	Numerical
Depth1000	Bathymetry (mean depth at 1000 m proximity)	Reef slope angle near reef	Numerical
Distance.Est	Distance to nearest estuary	Sedimentation from point source	Numerical
Distance.Land	Distance to nearest land	Sedimentation from diffuse source	Numerical

variable and the independent variables is in the exponential family, (3) the distribution of the dependent variable is known (presence/absence data are binomially distributed), (4) the values of errors are independent for both dependent and independent variables, and (5) the covariance of dependent variables is constant for all reef units (McCullagh & Nelder 1997).

The logistic regressions were fitted with 4 uncorrelated predictor variables derived from a suite of 8 initial candidate variables. We developed models for the 227 fish species that occurred at least 20 times (but not at all sites) throughout the data set of 105 sites, based on the recommendation that 5 to 10 records in the least prevalent class are required for every predictor degree of freedom to obtain accurate models (Harrell 2001). Consequently, models were unlikely to be accurate when all predictors were included, as the maximal possible predictor degree of freedom with our 4 predictors was 10. The remaining 322 fish species reported from Kimbe Bay were omitted from the analysis because they lacked the sufficient number of sightings.

A series of GLMs was fitted for each species and each of the predictor variables, and their combination, but without their interactions, in a step-forward and step-backward fashion. Model choice for each species was accomplished by selecting the best fitting model using AIC (Akaike 1974):

$$\mathrm{AIC} = 2k - 2\log L \quad (2)$$

where L was the likelihood and k was the number of parameters in the fit.

Prevalence and habitat specificity indices. We explored the suitability of models by examining fish occurrence patterns and related AIC values of the models. Fish occurrence patterns were classified by the degree of prevalence and habitat specificity of each species. We calculated a prevalence index P_i as:

$$P_i = \frac{\mathrm{pres}_i}{n_\mathrm{s}} \quad (3)$$

where pres_i was the number of sites that have species i and n_s was the total number of sites surveyed (n_s = 105). P_i ranged from P_i = 0.2 as the lower cut-off for statistical analysis to P_i = 1, which represents a species occurring everywhere; these highly prevalent species were kept in the analysis to investigate the sensitivity of the methods to a high prevalence.

We used a habitat-specificity index S_i to illustrate the degree to which a fish species was observed in particular habitats compared to all available habitats derived from the Eco.Name variable, which was derived from hard-coral data. Because it was used to compare all species that exhibited different habitat preferences, the specificity index was unconcerned with the particular habitat in which the species was observed. The habitat-specificity index was calculated as the variance in the proportion of the number of occurrences of a species i in each of the j habitats in the Eco.Name variable (Table 1), n_j, from the total prevalence of this fish species pres_i, standardized by the maximum variance across all species:

$$S_i = \mathrm{var}\left(\frac{n_j}{\mathrm{pres}_i}\right) / \mathrm{var}_\mathrm{max} \quad (4)$$

The variable S_i ranged from 0 to 1, with 0 being the least habitat specific and 1 being the case where a species was only present in 1 habitat. The analysis of prevalence and habitat specificity allowed us to test models for all species based on their AIC values.

Model checking. Although the distribution of a species can be described as a relationship between its occurrences and the values of a suite of predictor variables in a GLM, these models must be tested against null models to minimize the chance that an apparent relationship occurs by chance alone (Guisan & Zimmermann 2000). For efficient model checking across multiple species, we compared the AIC values of all the best models for each fish species with an AIC value derived from hypothetical fish species that are randomly distributed in space with the same prevalence (Manly 1997). For each prevalence value, we created 1000 permutations of a presence/absence matrix for a hypothetical fish species across all reef units. For each of these permutations, we created GLMs in a stepwise approach, and identified the best model. A reference interval for the distribution of randomly created AIC values was determined as the 2.5 and 97.5 percentiles of the permutation distribution, corresponding to an interval containing 95% of the AIC values obtained by random permutation. We then constructed a polynomial model of the null-model AIC values relative to each prevalence, P_i, with the 95% reference envelope (constructed from the 2.5 and 97.5 percentiles). We defined GLM models for each fish species with AIC values outside this envelope as efficient. Based on these categories of efficient and inefficient models, we evaluated the importance of the 4 predictor variables for predicting distributions of reef fish species.

RESULTS

Analysis of variable correlation

In order to avoid basing the analysis on correlated variables, we identified 4 final predictor variables from 8 candidate environmental variables. The expert opinion analysis was conducted by the lead author based on personal knowledge of the region and coral reef fish ecology and was concerned with an assessment of data

accuracy. The variable Habitat (remotely sensed habitat classes) was derived from an unsupervised classification of a satellite image. It was considered of low accuracy because the classification had not been ground-truthed. The variable Eco.Name (bioregions) was based on species lists of scleractinian corals. Although it was considered of high accuracy, it was not chosen as a final variable because detailed coral data are rarely available in remote locations, and thus this approach could be rarely replicated in other applications.

The statistical tests identified correlated pairs of variables (Table 2). The variables Eco.Name and Habitat were correlated ($p = 0.0006$, $\chi^2 = 122.3$) with 58% similarity (Fig. 2). Eco.Name and Exposure were also colinear ($p < 0.0001$, Kruskal-Wallis $\chi^2 = 41.4$) with 58% similarity. Habitat and Exposure showed 80% similarity. We discarded Habitat and Eco.Name for the further analysis, and retained Exposure for species distribution modeling. Because Eco.Name was developed from coral data, retaining Exposure also meant that we only used predictors that were truly remotely derived.

The 2 depth variables (Depth500 and Depth1000) had 60% similarity, $p < 0.0001$, and $t = 13.1$. We chose the mean surrounding depth at 500 m proximity (Depth500) as the final predictor variable and termed it Depth. The variables Distance.Land (distance to nearest land) and Distance.Est (distance to the nearest estuary) were equally closely related to each other (60% similarity, $p < 0.0001$, $t = 9.12$) (see Fig. 2). The distance to the nearest estuary was chosen as the final predictor. The Island variable (presence of a land–water interface) was also chosen because of its relatively high dissimilarity from all other predictors. Hence, the species distribution modeling was conducted with the variables Exposure, Island, Depth, and Distance.Est.

Fish species distribution models

The species distribution models for all 227 fish species are in Table S1 available in MEPS supplementary material at http://www.int-res.com/journals/suppl/m361p001_app.pdf. In this paper, we use results for 8 species that represent the variety of fish types with respect to their prevalence and habitat specificity as: (1) fishes that were widespread, (2) fishes that had restricted distributions, (3) fishes that were habitat specific, and (4) fishes that were not habitat specific (Table 3). We used these species to illustrate the kinds of species distribution models generated (Table 4).

The distribution of the 8 representative fish species was explained by different combinations of predictors (Table 4). All example species classified as habitat specific based on a high habitat specificity S_i had efficient models. S_i values ranged from 0.68 to 0.04, with $S_i = 1$ being a species that occurs in only 1 habitat and $S_i = 0$ being a species that occurs in all habitats. *Plectroglyphidodon dickii* is a small sedentary fish that inhabits its small territories on shallow reef flats or offshore reefs, and, correspondingly, their distribution responded to the predictor variables distance to nearest estuary and depth. In contrast, *Caesio lunaris* is a pelagic reef fish that schools on drop-offs of exposed reefs. It was modeled by exposure and depth, indicating its habitat preference. *Cheilinus undulatus*, which is habitat specific with a restricted distribution, likes to loiter on the drop-offs of exposed offshore reefs, as reflected in its model equation (Table 4). *Archamia zosterophora* is an

Table 2. Relationships between pairs of predictor variables given for each pair-wise comparison as probabilities p and test parameters χ^2 (chi-squared test), F (ANOVA), KW χ^2 (Kruskal-Wallis test), t (Pearson correlation test), and rho (ρ) (Spearman rank correlation). **Bold:** $p < 0.01$

Habitat	**p = 0.0006** $\chi^2 = 122.3$						
Depth500	**p < 0.0001** $F = 11.4$	**p < 0.0001** $F = 7.7$					
Depth1000	**p < 0.0001** $F = 20.5$	**p < 0.0001** $F = 16.1$	**p < 0.0001** $t = 13.1$				
Exposure	**p < 0.0001** KW $\chi^2 = 41.4$	**p = 0.004** KW $\chi^2 = 13.0$	**p < 0.0001** ρ = 0.61	**p < 0.0001** ρ = 0.75			
Island	p = 0.02 KW $\chi^2 = 5.1$	p = 0.04 KW $\chi^2 = 4.2$	**p < 0.0001** ρ = –0.42	**p < 0.0001** ρ = –0.43	p = 0.01 ρ = –0.25		
Distance.Land	**p < 0.0001** $F = 10.1$	**p < 0.0001** $F = 11.6$	**p < 0.0001** $t = 12.2$	**p < 0.0001** $t = 10.9$	**p < 0.0001** ρ = 0.39	**p < 0.0001** ρ = –0.73	
Distance.Est	**p < 0.0001** $F = 16.6$	**p < 0.0001** $F = 8.2$	**p < 0.0001** $t = 9.5$	**p < 0.0001** $t = 13.1$	**p < 0.0001** ρ = 0.80	p = 0.29 ρ = –0.10	**p < 0.0001** $t = 9.12$
	Eco.Name	Habitat	Depth500	Depth1000	Exposure	Island	Distance.Land

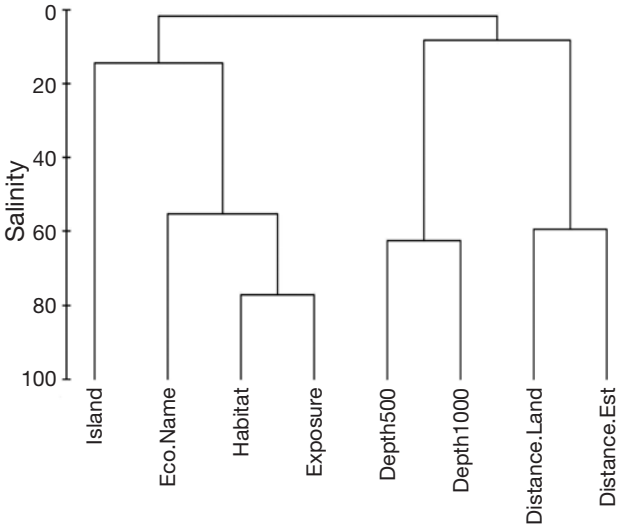

Fig. 2. Dendrogram of the Bray-Curtis similarity between 8 candidate predictors

uncommon small-bodied nocturnal fish that during the day is found in dense aggregations above or among the branches of branching corals of sheltered lagoons and inshore bays. The distribution of this fish was described by the proximity to the nearest island, the distance to the nearest estuary, and depth, reflecting its habitat preference. The models of these species can predict their probability of occurrence in unvisited sites. The resulting species distribution maps are important tools to assist conservation and management decisions in remote locations (Peterson et al. 2000).

All 4 fish species in the 'not habitat specific' category yielded inefficient species distribution models, regardless of their prevalence. Models for species that were not habitat specific and widespread, such as *Balistapus undulatus* and *Chaetodon lunulatus*, reflected this high prevalence by a model equation of $\ln[p/(1-p)] = 1$, indicating that their distributions did not respond to any predictors. While the models for these widespread species had a low AIC value, these AIC values were inside the reference envelope, demonstrating that the distribution of these species was not significantly different from a random distribution. The models for species that had restricted distributions but were not habitat specific, such as *Amphiprion percula* and *Hipposcarus longiceps*, contained predictors, but the high AIC values meant that they lay within the null-model reference envelope, indicating a distribution that is not significantly different from random.

Returning to all the fish species, the comparison of AIC values with those derived from a random distribution null model revealed that we could construct efficient models for 118 of the 227 fish species (Fig. 3). We could not find efficient models for species that occurred at most sites, or at very few. This is logical. Widespread fishes were recorded at almost every site; the resulting models were inefficient because no rela-

Table 3. Examples of 8 fish species that represent groups of prevalence and habitat specificity. Prevalence index (P_i) ranges from 0.19 to 1, and habitat specificity index (S_i) ranges from 0.05 to 0.68 for all 227 fish species. See 'Materials and methods—Prevalence and habitat-specificity indexes' for definitions of indexes

	Habitat specific			Not habitat specific		
	Taxon	P_i	S_i	Taxon	P_i	S_i
Widespread	*Plectroglyphidodon dickii*	0.47	0.37	*Chaetodon lunulatus*	0.98	0.05
	Caesio lunaris	0.41	0.38	*Balistapus undulatus*	0.98	0.06
Restricted	*Cheilinus undulatus*	0.23	0.38	*Amphiprion percula*	0.27	0.19
	Archamia zosterophora	0.27	0.45	*Hipposcarus longiceps*	0.21	0.16

Table 4. Species distribution models for 8 representative fish species. edf: estimated residual degrees of freedom; AIC: Akaike information criterion; Eff.: efficient model compared to null-model reference envelope (E: efficient; nE: inefficient). See Table 1 and 'Results—Analysis of variable correlation' for definitions of model variables

Taxon	Model equation: $\ln[p/(1-p)] =$	edf	AIC	Reference AIC 2.5%	Reference AIC 97.5%	Eff.
Plectroglyphidodon dickii	$3.56 + 0.0003 \times$ Distance.Est $+ 0.012 \times$ Depth	3	97.2	140.9	147.6	E
Caesio lunaris	$-1.086 + 2.6 \times$ Exposure $+ 0.003 \times$ Depth	5	114	137.9	144.5	E
Chaetodon lunulatus	3.94	1	21.8	12	22.8	nE
Balistapus undulatus	3.94	1	21.8	12	22.6	nE
Cheilinus undulatus	$-3.2 + 1.11 \times$ Exposure $+ 0.00012 \times$ Distance.Est	5	96.6	107.1	115.3	E
Archamia zosterophora	$2.4 - 1.23 \times$ Island $- 0.0004 \times$ Distance.Est $- 0.005 \times$ Depth	4	92.8	116.7	124.3	E
Amphiprion percula	$-0.13 - 1.28 \times$ Island $- 0.005 \times$ Depth	3	120.4	117.6	124.5	nE
Hipposcarus longiceps	$-1.78 + 0.003 \times$ Depth	2	108.3	102.6	109.8	nE

tionship between the fish distribution and a predictor could be identified. The most frequently observed species that yielded efficient models were *Naso lituratus*, *Thalassoma lunare*, and *Amblyglyphidodon leucogaster* (prevalence p = 0.89, AIC = 68.9, 68.6, and 67.7, respectively); however, these models were only marginally outside the null-model reference envelope. Given that rules of thumb for AIC model comparisons require the difference in AIC values to be 2 to 3 for models to be different from another (Burnham & Anderson 1998), it is unlikely that these models were efficient. Good models for common species were obtained for *Scarus niger* (p = 0.85, AIC = 69), *Pseudanthias tuka* (p = 0.84, AIC = 73.9), and *Melichthys vidua* (p = 0.75, AIC = 54.4).

In summary, not unexpectedly, efficient models were primarily found for fishes that were habitat specific, with both low and high prevalence. An efficient model described a relationship between a species' occurrence patterns and 1 or more predicting variables for those fish species that were distributed non-randomly across the seascape.

Response to predictor variables

Species distributions of coral reef fishes with efficient models were described by either all 4 (1 instance, 0.7% of efficient models), 3 (25 instances, 18.6%), 2 (65 instances, 48.5%), or 1 (27 instances, 20.1%) of the predictor variables. None of the predictors stood out as being relevant to all species (Fig. 4). The most frequent single predictor was Depth, which yielded most of the efficient models, but also most of the inefficient models among all single predictors. Island was the least frequent predictor among efficient models.

The highest number of fish distributions was explained by a combination of Depth with either Exposure or Distance.Est. These 2 combinations of predictor variables yielded the highest number of efficient models, and a relatively low number of inefficient models. Island combined with Distance.Est yielded the next highest number of efficient models. About 11% of fish species responded to combinations of 3 predictor variables. The largest proportion of species (~4.5%) responded to the combination of the 3 predictor variables Island, Depth, and Exposure; this combination of variables yielded only 1 inefficient model. The distribution of only 1 fish species, *Scarus niger*, was linked to all 4 predictor variables in an efficient model.

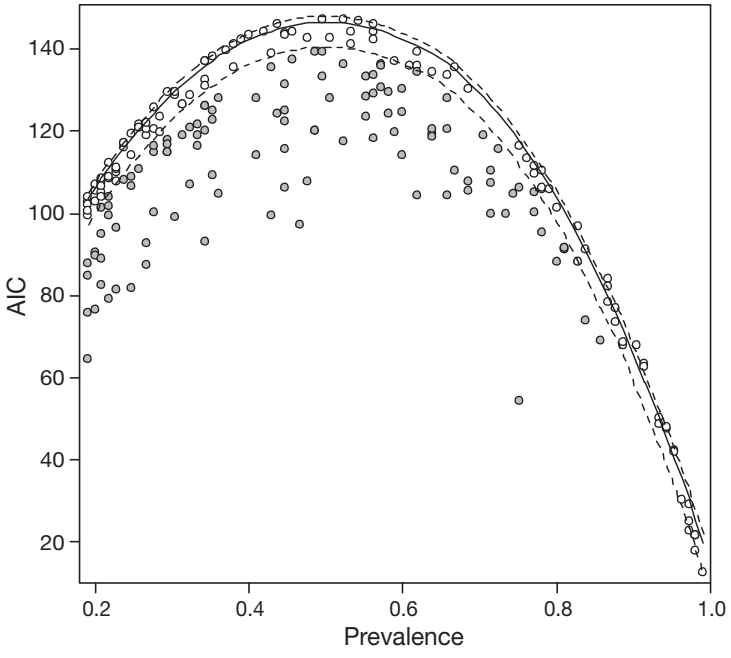

Fig. 3. Akaike information criterion (AIC) values plotted against prevalence P_i for 227 fish species distribution models with the null model (solid line) and the reference envelope (dashed lines). Open dots: inefficient models; filled dots: efficient models

DISCUSSION

This study demonstrates the application of GLM to identifying the relationships between remotely obtained reef characteristics and the distributions of a suite of coral reef fishes. Reef characteristics were described by 4 predictor variables, with 2 variables for physical factors (Exposure and Depth), 1 variable to represent the proximity of main human impact (Distance.Est), and 1 variable signifying the presence or absence of a marine–terrestrial interface (Island). We developed logistic regression models from presence/absence data of 227 fishes, and tested their efficiency by comparing their AIC values to those of hypothetical species with random distributions.

We identified efficient models for 118 species, most of which were highly habitat specific. We established model efficiency by comparing model AIC values of real species with model AIC value of hypothetical species distributions. For this method it was crucial to create random species distribution models of hypothetical species that occupied different fractions of the system. All predictor variables were important for describing fish distributions, but none of them stood out as particularly dominant.

Data for predictor variables

Fish distribution models can be developed based on environmental and habitat data from remote sensing sources. However, the accurate

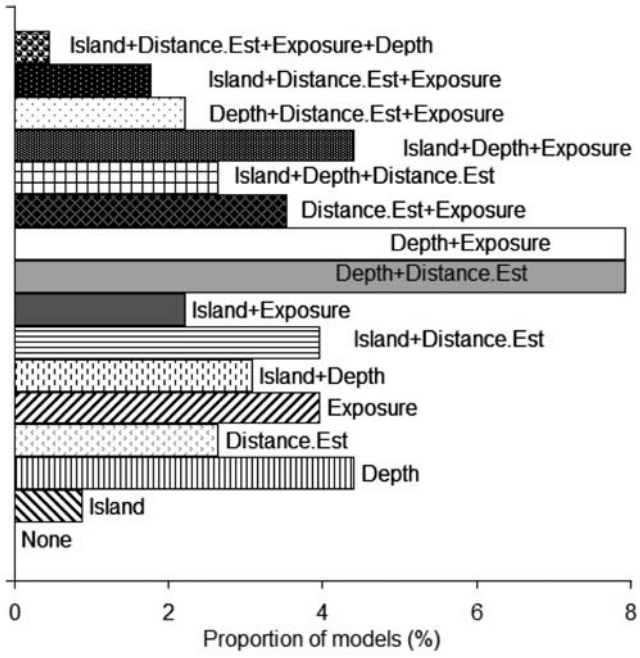

Fig. 4. Proportion of efficient models responding to combinations of predictor variables

representation of environmental and biological entities by such data requires consideration of uncertainty, colinearity, and spatial scale, which affect the suitability of species distribution maps to end-users. Large amounts of data can be gleaned from internet-based free databases of biophysical information at large-scale resolution. Such databases can yield exposure data, e.g. tidal data (NOAA 2007) combined with wave height data (e.g. Directorate of Oceanography and Meteorology Australia 2007), and wind data (e.g. International Research Institute for Climate and Society 2007), bathymetry (e.g. British Oceanographic Data Centre 2007), coastlines (National Geophysical Data Center NOAA 2007), sea surface temperature (NASA 2007), and ocean currents (CSIRO 2006), to name only a few. However, these data sets are often too coarse to be informative at scales relevant to distribution modeling. For this reason, local-scale data layers for Kimbe Bay were collated from printed marine charts, expert consultation, GIS analysis, and multivariate bioregionalization of coral data. These data were used in our models without explicitly considering the uncertainty associated with them. However, these data layers are likely to have varying levels of accuracy. For example, our exposure data classified up to 4 exposure regimes and were determined by expert opinion. A more accurate approach would be to calculate exposure by measuring fetch and applying linear wave theory to convert fetch to wave energy (Harborne et al. 2006). Similarly, several methods of determining bathymetry from remote imagery exist (e.g. Stumpf et al. 2003). Such remote sensing methods, however, depend on the pixel size of remote measurements. The coarse resolution of many of these measurements (e.g. the 30 m pixel size of the most readily available imagery such as Landsat) excluded their use in our study. In future regional models attempting to spatially generalize reef-scale patterns, *in situ* measurements may be required to obtain accurate exposure and bathymetry data. Alternatively, high resolution images such as those from the QuickBird satellite exist, but are prohibitively expensive.

Applications of species distribution modeling in conservation

Conservation planning in remote locations is hampered by incomplete data that bias reserve site selection towards surveyed sites if used directly (Rondinini et al. 2006). Therefore, one of the most compelling applications of species distribution modeling is to enable reef managers and conservation planners to incorporate modeled distributions in their decisions about the location and size of reserve sites. Species distribution modeling is a statistically sound and repeatable method to transform point data into the spatially explicit data required by decision support software for systematic conservation planning (Rondinini et al. 2006). Often, a conservation aim is to represent as many species as possible, therefore fulfilling the conservation planning principle of comprehensiveness, requiring species modeling rather than community level modeling approaches. In addition, modeled probabilities of species occurrence can easily be adjusted to projected change in environmental conditions on a species-by-species basis (e.g. changes in distributions of fishes with increased sediment levels owing to major coastal deforestation). While this would also be possible with multivariate analysis, the ability to distinguish small community changes will be low given the high number of species within a reef assemblage.

Habitat-specific fish species are the focus for many conservation efforts, because they efficiently inform a complementarity-based reserve selection process. This means that habitat-specific species are likely to be protected if a complete range of reef habitats is incorporated in a reserve system. Modeled distributions of these species that relate species records to remotely collated environmental information may aid this process and expand point data to the spatially explicit complete coverage required by conservation decision support tools (Possingham et al. 2006). Fish surveys and other biodiversity assessments are expensive, and expert taxonomists are required. Identifying a suite of

fish species with distributions that have a robust association with available predictor variables, which are easily recognisable by non-experts, may be a way to reduce the cost of data collection.

The classification of fishes by their prevalence and habitat specificity allowed us to identify fish species that represented different habitat associations, and therefore have varying degrees of importance in conservation planning approaches where species distribution modeling is performed. For example, a species that is widespread and has low habitat specificity will always be automatically included in any reserve system, and is likely to experience a low level of threat. Distributions of widespread species could rarely be described by efficient models as we had no predictor variables that, over this extent, helped to explain the distribution of the widespread species and because they are flexible in their ecological requirements. On the contrary, species with restricted distributions and high habitat specificity are more likely to be under higher threat. Most of these species yielded an efficient distribution model.

A sound knowledge of how environmental conditions affect the distributions of coral reef fishes is necessary to manage populations for persistence while considering known and future threats, species vulnerability, and potential sustainable harvest. In the case study of Kimbe Bay, we found that depth was the most frequently included environmental parameter in our models, both in single and multiple predictor models. This finding is consistent with previous studies showing depth to be important in influencing fish communities at local (e.g. Syms & Jones 2004) and regional scales (Leathwick et al. 2006a). Depth is also a major predictor for sponge and coral diversity on a subregional scale (Cleary et al. 2005). The depth of a reef is a proxy for environmental conditions such as turbidity and intensity of sunlight (Williams 1991). In our study, the depth variable also represented the steepness of the reef slope, as the depth parameter signified the depth at approximately 500 m distance from the reef crest. Thus, for steep slopes, the depth at 500 m distance was greater than for shallow slopes.

The highest number of fish distributions was explained by combinations of 2 predictors: either depth and exposure or depth and the distance from the nearest estuary. Exposure to wave action and the associated water motion strongly affects the distributions of reef organisms (Tunnicliffe 1982, Madin et al. 2006). For example, some hard corals reproduce asexually through the dispersal and re-attachment of coral fragments broken by water motion (Smith & Hughes 1999). In contrast, benthic organisms can be dislodged or damaged by high wave impacts (Tunnicliffe 1982), but may be smothered by sediment in very sheltered environments (Victor et al. 2006). The association of reef fishes with reef zones (experiencing a gradient of exposure) depends strongly on their swimming ability (Fulton et al. 2005), because a failure to control their position at high levels of wave action may lead to injury or mortality (Madin et al. 2006).

The distance to the nearest estuary parameter represents a proxy for the exposure of a reef to levels of suspended terrigenous sediment (causing turbidity) and pollution that may be anthropogenically enhanced. In Kimbe Bay, deforestation and subsequent oil palm plantations are spreading at a rapid rate (M. Beger pers. obs.); and resulting sediment input into rivers creates large plumes of suspended solids near estuaries. Our results highlight the combination of depth and distance to the nearest estuary as an important predictor for the distributions of some fish species, pointing to the importance of the effects of terrestrial run-off on coral communities (Fabricius 2005) that support coral reef fishes. The combination of the distance to the nearest estuary and the presence of an island yielded a number of efficient models, highlighting that several fish species were affected by the level of disturbance to shallow reef habitats on the terrestrial–marine interface. These results for fishes are consistent with other taxonomic groups on reefs. For example, a large proportion of the variation in sponge and coral diversity on the west-coast of Sulawesi was explained by models using distance offshore, depth, and exposure (Cleary et al. 2005).

Statistical considerations

The present study employed linear models to predict responses of fish species to environmental gradients. However, these gradients and the responses are unlikely to be linear, and future models should explore if the relationships described here hold when non-linear modeling methods such as generalized additive models (Wintle et al. 2005) or multivariate adaptive regression splines (Leathwick et al. 2006b) are tested. For example, we expect the degree of suspended solids from land erosion represented by the Distance.Est variable will decrease from the estuary to outer reefs in a non-linear fashion. Near the estuary, the impacts of these sediments will be high, and then drop off quickly towards an ambient background sediment level. As a result, fish species that live on coastal reefs will experience high levels of sediment, whereas these impacts could be negligible at reefs furthest away from the pollution sources.

The uncertainty associated with our models was defined by the accuracy and extent of the underlying

data and the correct execution of modeling methods (Guisan et al. 2006). For example, the coral reef biodiversity surveys providing the baseline for species distribution models were limited to 105 sites. Model comparisons with the stepwise addition or removal of variables could be improved by using the small sample AIC_c (Burnham & Anderson 1998). If more sites in a greater variety of habitats could be added, these models, their accuracy, and their predictions are likely to change for many species. Several authors found that logistic regression methods of distribution modeling are less powerful for fewer occurrences of a species within a data set, and are sensitive to predictor-variable accuracy (Manel et al. 1999). This drawback of GLMs for low numbers of sightings excludes the modeling of rare species distributions. However, rare species are often specifically targeted by conservation planners because of their distinct and relatively unknown contribution to biodiversity patterns (Lennon et al. 2004) and because they are often not captured by other methods (Lawler et al. 2003). A future research challenge is to accurately define and model the distributions of rare coral reef species.

Spatial autocorrelation in the species occurrence among sites can violate the assumption of independence in many statistical tests (Dormann et al. 2007). For example, we could expect reefs in close proximity to be similar because of similar larval supply and environmental conditions. In this study, we considered spatial autocorrelation in the context of our study site and data to be low, as the larvae of most species will be well mixed at the scale of Kimbe Bay (Steinberg et al. 2006). With such relatively uniform mixing, we consider that recruitment and population maintenance of fishes is relatively independent of proximity to other reefs. We believe this is a valid approach because the model improvement achieved by correcting for spatial autocorrelation in the models would only be moderate (Dormann et al. 2007). The study reefs were spread relatively evenly across the bay, and thus the low resolution of our predictor variables is unlikely to allow us to distinguish small differences in recruitment based on the vicinity of sources. However, future applications should employ and evaluate statistical modeling techniques that take spatial autocorrelation into account (reviewed in Dormann et al. 2007), particularly when probabilities of occurrences are created for adjacent reefs from models.

CONCLUSIONS

Conservation and management requires an explicit understanding of species' responses to environmental parameters. We discovered that, while all studied parameters affected the distributions of some fish species, depth was the most frequent independent variable in efficient single variable models. For models with combinations of 2 predictor variables, depth and exposure as well as depth and distance from the nearest estuary were the prevalent predictors of fish distributions, underpinning known ecological relationships of fish assemblages on coral reefs. Several fish species responded to the combination of the distance from an estuary and the presence of the terrestrial–marine interface, indicating that these species depend on intact coastal reef habitat. Coastal reefs are in decline in many locations, and their conservation and management depend on considering processes linking them to upstream terrestrial landscapes.

In the marine realm, comprehensive species data are extremely rare, data sets are small, and environmental variables are not readily available at local spatial scales. However, modeling species distributions across a seascape can greatly improve the spatial information available to management applications (Polasky et al. 2000), as long as the uncertainties associated with the data and propagated through to predictions of species distributions are taken into account. Further research is required to quantify such model uncertainties specific to the marine realm and to test how species distribution modeling could enhance the efficacy of reef management and conservation applications in remote locations.

Acknowledgements. We acknowledge The Nature Conservancy for assistance in providing data and GIS advice. G. Jones, James Cook University, assisted with part of the data collection. Many people at The Ecology Centre, University of Queensland, and J. Elith, University of Melbourne, provided essential advice in the modeling process. Thanks go to C. Elkin, D. Ceccarelli, and P. Baxter for comments on an earlier version of the manuscript. M.B. acknowledges a postgraduate scholarship from the University of Queensland, and a PADI Foundation grant. H.P.P. was supported by grants from the Australian Research Council and The Department of The Environment and Water Resources through a Commonwealth Environment Resource Facility.

LITERATURE CITED

Akaike H (1974) A new look at the statistical model identification. IEEE Transactions on Automatic Control 19:716–723

Allen GR, Munday P (1994) Fish diversity of Kimbe Bay. The Nature Conservancy, Auckland

Allen GR, Steene R (1994) Indo-Pacific coral reef field guide. Tropical Reef Research, Singapore

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

Austin MP (2002) Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. Ecol Model 157:101–118

Beger M (2002) The diversity and status of coral reef fishes of eastern Kimbe Bay. The Nature Conservancy, Brisbane

Beger M, Jones GP, Munday PL (2003) Conservation of coral reef biodiversity: a comparison of reserve selection procedures for corals and fishes. Biol Conserv 111:53–62

Beger M, Jones GP, Possingham HP (2006) Models of coral reef fish biodiversity and species habitat association in conservation planning. In: Suzuki Y, Nakamori T, Hidaka M, Kayanne H and others (eds) Proceedings of the 10th international coral reef symposium. Japanese Coral Reef Society, Okinawa, p 1445–1456

Bellwood DR, Hughes TP, Connolly SR, Tanner J (2005) Environmental and geometric constraints on Indo-Pacific coral reef biodiversity. Ecol Lett 8:643–651

British Oceanographic Data Centre (2007) Gridded bathymetry data. Available at: www.bodc.ac.uk/data/online_delivery/gebco

Burnham KP, Anderson DR (1998) Model selection and inference: a practical information-theoretic approach. Springer-Verlag, New York

Cayuela L, Benayas JM, Justel A, Salas-Rey J (2006) Modelling tree diversity in a highly fragmented tropical montane landscape. Glob Ecol Biogeogr 15:602–613

Cleary DFR, Becking LE, de Voogd NJ, Renema W, de Beer M, van Soest RWM, Hoeksema BW (2005) Variation in the diversity and composition of benthic taxa as a function of distance offshore, depth and exposure in the Spermonde Archipelago, Indonesia. Estuar Coast Shelf Sci 65:557–570

CSIRO (Australian Commonwealth Scientific and Research Organization) (2006) Access to Bluelink ReANalysis (BRAN) output. Available at: www.marine.csiro.au/ofam1/

Directorate of Oceanography and Meteorology Australia (2007) Online data services. Available at: www.metoc.gov.au/products/data.html

Dormann CF, McPherson JM, Araújo MB, Bivand R and others (2007) Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. Ecography 30:609–628

Dunk JR, Zielinski WJ, Preisler HK (2004) Predicting the occurrence of rare mollusks in northern California forests. Ecol Appl 14:713–729

Elith J (2000) Quantitative methods for modeling species habitat: comparative performance and application to Australian plants. In: Ferson S, Burgman M (eds) Quantitative methods for conservation biology. Springer-Verlag, New York, p 33–58

Elith J, Graham CH, Anderson RP, Dudik M and others (2006) Novel methods improve prediction of species' distributions from occurrence data. Ecography 29:129–151

Engler R, Guisan A, Rechsteiner L (2004) An improved approach for predicting the distribution of rare and endangered species from occurrence and pseudo-absence data. J Appl Ecol 41:263–274

Fabricius KE (2005) Effects of terrestrial runoff on the ecology of corals and coral reefs: review and synthesis. Mar Pollut Bull 50:125–146

Faith DP, Nix HA, Hutchinson MF, Stein JL and others (2000) The BioRap biodiversity assessment and planning study for Papua New Guinea. Pac Conserv Biol 6:279–288

Fulton CJ, Bellwood DR, Wainwright PC (2005) Wave energy and swimming performance shape coral reef fish assemblages. Proc R Soc Lond B 272:827–832

Garza-Perez JR, Lehmann A, Arias-Gonzalez JE (2004) Spatial prediction of coral reef habitats: integrating ecology with spatial modeling and remote sensing. Mar Ecol Prog Ser 269:141–152

Guisan A, Zimmermann NE (2000) Predictive habitat distribution models in ecology. Ecol Model 135:147–186

Guisan A, Lehmann A, Ferrier S, Austin M, Overton JMC, Aspinall R, Hastie T (2006) Making better biogeographical predictions of species' distributions. J Appl Ecol 43:386–392

Harborne AR, Mumby PJ, Zychaluk K, Hedley JD, Blackwell PG (2006) Modeling the beta diversity of coral reefs. Ecology 87:2871–2881

Harrell FE (2001) Regression modeling strategies with applications to linear models, logistic regression and survival analysis. Springer-Verlag, New York

Holbrook SJ, Brooks AJ, Schmitt RJ (2002) Predictability of fish assemblages on coral patch reefs. Mar Freshw Res 53:181–188

Hugall A, Moritz C, Moussalli A, Stanisic J (2002) Reconciling paleodistribution models and comparative phylogeography in the Wet Tropics rainforest land snail *Gnarosophia bellendenkerensis* (Brazier, 1875). Proc Natl Acad Sci USA 99:6112–6117

Hughes TP, Baird AH, Bellwood DR, Card M and others (2003) Climate change, human impacts, and the resilience of coral reefs. Science 301:929–933

International Research Institute for Climate and Society (2007) Data library—wind vectors and sea surface temperatures. Available at: ingrid.Idgo.columbia.edu/SOURCES/.IGOSS/.TOGA-TAO.cdf/

Jones GP, McCormick MI, Srinivasan M, Eagle JV (2004) Coral decline threatens fish biodiversity in marine reserves. Proc Natl Acad Sci USA 101:8251–8253

Kelly NM, Fonseca M, Whitfield P (2001) Predictive mapping for management and conservation of seagrass beds in North Carolina. Aquat Conserv Mar Freshw Ecosyst 11:437–451

Lawler JJ, White D, Sifneos JC, Master LL (2003) Rare species and the use of indicator groups for conservation planning. Conserv Biol 17:875–882

Leathwick JR, Elith J, Francis MP, Hastie T, Taylor P (2006a) Variation in demersal fish species richness in the oceans surrounding New Zealand: an analysis using boosted regression trees. Mar Ecol Prog Ser 321:267–281

Leathwick JR, Elith J, Hastie T (2006b) Comparative performance of generalized additive models and multivariate adaptive regression splines for statistical modelling of species distributions. Ecol Model 199:188–196

Lehmann A, Overton JM, Austin MP (2002) Regression models for spatial prediction: their role for biodiversity and conservation. Biodivers Conserv 11:2085–2092

Lennon JJ, Koleff P, Greenwood JJD, Gaston KJ (2004) Contribution of rarity and commonness to patterns of species richness. Ecol Lett 7:81–87

Madin JS, Black KP, Connolly SR (2006) Scaling water motion on coral reefs: from regional to organismal scales. Coral Reefs 25:635–644

Manel S, Dias JM, Ormerod SJ (1999) Comparing discriminant analysis, neural networks and logistic regression for predicting species distributions: a case study with a Himalayan river bird. Ecol Model 120:337–347

Manly BFJ (1997) Randomization, bootstrap and Monte Carlo methods in biology. Chapman & Hall, London

McCullagh P, Nelder JA (1997) Generalised linear models. Chapman & Hall, London

Mora C, Chittaro PM, Sale PF, Kritzer JP, Ludsin SA (2003) Patterns and processes in reef fish diversity. Nature 421: 933–936

Myers RF (1999) Micronesian reef fishes. Coral Graphics, Barrigada

NASA (National Aeronautics and Space Administration) (2007) SeaWiFS project. Available at: oceancolor.gfsc.nasa.gov/SeaWiFS

National Geophysical Data Center NOAA (2007) Shoreline/

Coastline data. Available at: www.ngdc.noaa.gov/mgg/shorelines/shorelines.html

NOAA (National Oceanic and Atmospheric Administration) (2007) Tides and currents. Available at: tidesandcurrents.noaa.gov

Peterson AT, Egbert SL, Sanchez-Cordero V, Price KP (2000) Geographic analysis of conservation priority: endemic birds and mammals in Veracruz, Mexico. Biol Conserv 93:85–94

Polasky S, Camm JD, Solow AR, Csuti B, White D, Ding RG (2000) Choosing reserve networks with incomplete species information. Biol Conserv 94:1–10

Possingham HP, Wilson KA, Andelman SJ, Vynne CH (2006) Protected areas: goals, limitations, and design. In: Groom MJ, Meffe GK, Carroll CR (eds) Principles of conservation biology. Sinauer Associates, Sunderland, MA, p 509–533

R Development Core Team (2005) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna

Roberts CM, McClean CJ, Veron JEN, Hawkins JP and others (2002) Marine biodiversity hotspots and conservation priorities for tropical reefs. Science 295:1280–1284

Robertson DR (2001) Population maintenance among tropical reef fishes: inferences from small-island endemics. Proc Natl Acad Sci USA 98:5667–5670

Rondinini C, Wilson KA, Boitani L, Grantham H, Possingham HP (2006) Tradeoffs of different types of species occurrence data for use in systematic conservation planning. Ecol Lett 9:1136–1145

Smith LD, Hughes TP (1999) An experimental assessment of survival, re-attachment and fecundity of coral fragments. J Exp Mar Biol Ecol 235:147–164

Steinberg CR, Choukroun SM, Slivkoff MM, Mahoney MV, Brinkman RM (2006) Currents in the Bismarck Sea and Kimbe Bay, Papua New Guinea. TNC Pacific Island Countries Report No. 6/06, Australian Institute of Marine Science and The Nature Conservancy, Brisbane

Stumpf RP, Holderied K, Sinclair M (2003) Determination of water depth with high-resolution satellite imagery over variable bottom types. Limnol Oceanogr 48:547–556

Syms C, Jones GP (2004) Habitat structure, disturbance and the composition of sand-dwelling goby assemblages in a coral reef lagoon. Mar Ecol Prog Ser 268:221–230

Thomas CD, Cameron A, Green RE, Bakkenes M and others (2004) Extinction risk from climate change. Nature 427: 145–148

Tunnicliffe V (1982) The effects of wave-induced flow on a reef coral. J Exp Mar Biol Ecol 64:1–10

Venables WN, Dichmont CM (2004) GLMs, GAMs and GLMMs: an overview of theory for applications in fisheries research. Fish Res 70:319–337

Veron JEN (2000) Corals of the world. Australian Institute of Marine Science, Townsville

Victor S, Neth L, Gobuu Y, Wolanski E, Richmond RH (2006) Sedimentation in mangroves and coral reefs in a wet tropical island, Pohnpei, Micronesia. Estuar Coast Shelf Sci 66:409–416

Williams DM (1991) Patterns and processes in the distribution of coral reef fishes. In: Sale PF (ed) The ecology of fishes on coral reefs. Academic Press, San Diego, CA, p 437–474

Wintle BA, Elith J, Potts JM (2005) Fauna habitat modelling and mapping: a review and case study in the Lower Hunter Central Coast region of NSW. Austral Ecol 30: 719–738

Editorial responsibility: Hans-Heinrich Janssen, Oldendorf/Luhe, Germany

Submitted: June 14, 2007; Accepted: March 6, 2008
Proofs received from author(s): May 9, 2008