

An investigation of coral reef fish assemblage modelling with geostatistical methods

J. De Mazières¹, J. Comley²

1) School of Marine Studies, University of the South Pacific, PO Box 1168, Suva, Fiji Islands

2) Institute of Applied Science, University of the South Pacific, PO Box 1168, Suva, Fiji Islands

Abstract. The objective of this study was to determine the spatial distribution of reef fish assemblages according to the habitat types of the Coral Coast, Fiji by using geostatistical methods. The study focused on the reef flat zone covering 45 km² which was divided into 22 geomorphological reef units. Substrate and fish data were obtained from a previous field survey of 311 transects. We worked with five classes of habitats (sand, rubble, macroalgae, seagrass and coral) which were defined according to thresholds of 20% for the biotic substrate and 50% for the abiotic. Nine fish families were selected due to their importance for the fisheries and as reef health indicators. We conducted batches of multivariate statistical tests to identify significant patterns of fish assemblage distribution. Prediction maps of fish abundance were created by using the co-Kriging geostatistical model. Results showed that sand, seagrass and coral habitats hosted significantly different assemblages. The quality of the predictive models varied highly depending on the location and the fish family of which Acanthuridae and Scaridae showed the best accuracy. Reef fish modelling using a geostatistical approach provided encouraging results considering that this investigation was conducted with previously collected data and within a reduced resource context.

Key words: coral reef fish, assemblage, habitat, geostatistical modelling, Fiji

Introduction

Coral reef fish communities are influenced by the distribution and abundance of component species. Their distribution and abundance are in turn influenced by the interaction between their behaviour and the physical and biological environment. Many studies have shown that reef fish abundance and species richness are correlated with the biological nature of the substratum and notably with the extent of live coral cover (Bell and Galzin 1984; Chabanet et al. 1997; Garpe and Ohman 2003; Bouchon-Navaroa et al. 2005). However, habitat composition has not always been found as the main factor explaining fish community structures (Roberts and Ormond 1987; McClanahan and Arthur 2001; Dominici-Arosemena and Wolff 2005).

Although spatial models have been widely developed for terrestrial environments, they have just been recently applied to the marine ecosystems (Pittman et al. 2007). Most of the coral reef studies involving spatial modelling have mainly focused on assessing the effect of marine reserves on single fish species or assemblages (Núñez-Lara et al. 2005) and on improving the ecological knowledge of fish populations to support ecosystem-based management (Mellin et al. 2007; Pittman et al. 2007; Mumby 2006). As collecting data on extensive areas is regarded as too expensive and time costly, ecologists are now focusing on predictive methods and their improvement. Extrapolation is one of the approaches to predict

patterns, where known data are projected, extended or expanded into an area not sampled based on assumptions of continuity and correlation among the variables (Miller et al. 2004).

Coral reefs are of crucial importance for the people of the Pacific Islands Countries since they support numerous economic and cultural activities in addition to providing environmental services. The Fiji Islands have been the location of many marine survey programmes providing a large quantity of data on reef benthic and fish assemblages. However, spatial analysis has been rarely if ever conducted and normal analysis has not been exhaustively carried out.

The main aim of this study is to model the spatial distribution of the reef fish assemblages according to the habitat types of the Coral Coast, Fiji Islands using geostatistical methods. The influence of the benthic composition on family taxa abundance is first assessed and then the feasibility of modelling reef fish distribution with the ArcGIS Geostatistical Analyst extension is investigated. Considering the Fiji context, where scientific resources are limited, this paper illustrates an experimental approach for coral reef research.

Material and Methods

Study Site

This study focused on the Coral Coast region located along the southwest coast of Viti Levu, Fiji's main island (Figure 1a). The Coral Coast is

bordered by the longest fringing reef found within Fiji's waters, stretching along about 80 km with a seaward extension of 500 m to 1 km. The reef system is made of 22 geomorphological reef units separated by channels.

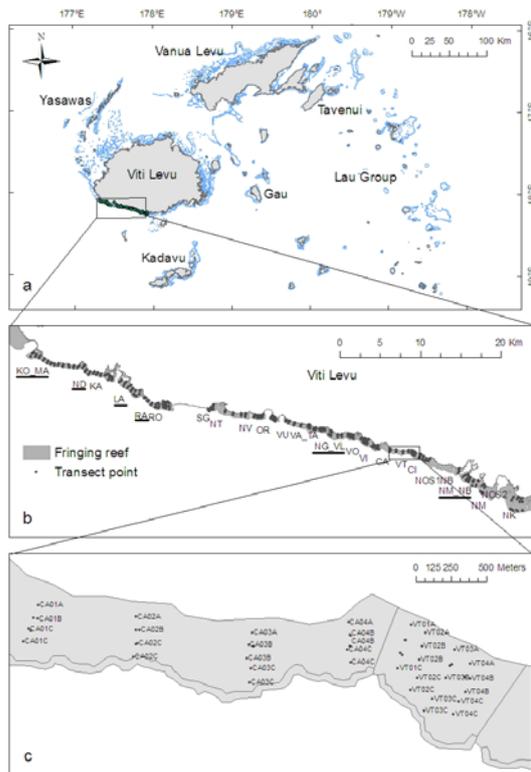


Figure 1: Study site. a) the Fiji Islands; b) the Coral Coast fringing reef and the sampling sites surveyed by CCC. Codes refer to the reef unit names. Underlined reef units were selected for modelling; c) close up showing the sampling design. Two points correspond to the start and end points of one 100 m section (named A, B or C).

The human population along the Coral Coast is distributed into villages of about 100 to 300 inhabitants with the exception of Sigatoka town where about 8,440 people live (Fiji Islands Bureau of Statistics 2007). With development and increasing population, the main threats to the ecosystem and its resources have been identified as being the overexploitation of fish and corals, inefficient waste management leading to coastal water pollution, and erosion and siltation due to inland deforestation (Thaman and Aalbersberg 2004).

Secondary Data and Benthic Habitat Classification

The field data used for this study were provided and collected by Coral Cay Conservation (CCC) between July 2004 and May 2005 (Rowlands et al. 2005) (Figure 1b-c). Data were collected along transects located perpendicular to the shoreline on the reef flat zone. Each transect was divided into three 100 m long sections which were laid down within three different zones of exposure to offshore influence. A total of 311 sections were surveyed by

snorkelling based on a modified Reef Check technique. Assessment of benthic coverage was performed following the point intercept method along a 100 m tape. Substrate type was recorded every 0.5 m, providing a total of 200 data points per 100 m section. Fishes were counted along the same 100 m tape using the belt transect methodology applied within a 5 m wide by 20 m high box.

A hierarchical classification scheme of habitats was established *a priori* based on previous works and on expert knowledge. The classification scheme comprised of five habitat classes: sand, rubble, macroalgae, seagrass and coral. A simple quantitative component was incorporated with arbitrary percentage cover thresholds. Biotic substrates (i.e., coral, seagrass and macroalgae) were considered dominant when covering 20% of a transect, whereas a threshold of 50% was assigned to the abiotic categories (i.e., sand and rubble).

Multivariate Analysis

Multivariate statistical analyses were first conducted to identify significant patterns of fish assemblage distribution. Nine fish families (Chaetodontidae, Lethrinidae, Mullidae, Serranidae, Scaridae, Siganidae, Lutjanidae, Acanthuridae, Labridae) were studied due to their importance in terms of resources for the communities and their ecosystem health indicator status. Analyses were performed using the routines provided by the PRIMER v6 software (Clarke and Warwick 2001). Comparison of fish assemblage structures was based on the Bray-Curtis similarity coefficient. Investigation of the effect of habitat type on fish family abundance was done with the ANOSIM (ANalysis Of SIMilarity) significance test. To identify the characteristic families of each habitat type SIMPER (SIMilarity PERcentage) analysis was applied to the data, allowing identification of typical variables through calculating their average contribution to the intra-group similarity and to the inter-group dissimilarity.

Spatial Modelling

Modelling was based on the results of the multivariate analysis. Prediction maps of fish abundance were created at the reef unit scale, using the co-Kriging geostatistical model. The extrapolation is based on a trend identified from the data and the spatial autocorrelation theory assuming that features close to each other are likely to be more similar than the ones apart from each other. Co-Kriging is a stochastic model which provides an estimation of the variance for any point predicted. It is thus possible to quantify the quality of the surface models by measuring the statistical error of predicted surfaces. In addition, the co-Kriging model allows multivariate analysis where the predicted values are estimated according to the correlation with an influencing variable.

Several extrapolation models were tested by setting different parameters, such as the lag number and size, or the searching neighbourhood shape and methods (Johnson et al. 2001). They were compared using the cross-validation results. This provided information on the statistical performance of the model in predicting fish abundance values. A model was qualified as good if the mean prediction error is close to zero, the root-mean-square standardized prediction error is close to one and the root-mean-square prediction error is small. These measurements indicate, respectively, unbiased predictions, accurate standard errors and predicted values similar to the measured values (Johnson et al. 2001). The models with the lowest statistical errors obtained for each family and reef unit were then selected and the surface maps were converted into polygon layers.

Results

Fish-Habitat Associations

Exploratory analyses and statistical tests showed that fish assemblages were structured differently among different habitat types (Table 1). Differences were highly significant for macroalgae vs. sand ($p=0.001$), seagrass vs. coral ($p=0.001$) and rubble vs. sand ($p=0.003$). In addition, fish assemblages found within coral and within rubble habitats had significantly different structures compared to those associated with sand and seagrass dominated areas respectively.

Table 1: Results of the ANOSIM analysis. Global $R=0.116$ and $p=0.001$. Confidence intervals are represented by asterisks as follow: * 95%; ** 99.9%. Only significant results are shown here.

Pair-wise tests	R value	p value
macroalgae vs. sand	0.46	0.001**
seagrass vs. coral	0.558	0.001**
rubble vs. sand	0.291	0.003**
coral vs. sand	0.182	0.026*
rubble vs. seagrass	0.169	0.044*

According to the SIMPER analysis, the average contribution of the family to the intra-group similarity showed two patterns of association family-habitat (Figure 2). Labridae contributed at 95% to the similarity within the seagrass group and was thus considered as a discriminating family. In

addition, given their high percentage contributions, Scaridae and Acanthuridae can be considered as discriminating families of the groups seagrass vs. coral and coral vs. sand.

Labridae had the highest contribution to the inter-group dissimilarity and the order of percentage contribution of the families was the same in all of the pair-wise results (Table 2). Dissimilarity of fish assemblages between macroalgae and sand habitats was due to the high abundance of Labridae in sand areas. Three pair-wise tests (macroalgae vs. sand; rubble vs. sand; rubble vs. seagrass) showed that the contribution to the dissimilarity was dominated by Labridae while in the two others (seagrass vs. coral; coral vs. sand), the contribution were more evenly distributed among Labridae, Scaridae and Acanthuridae families. Given their high percentage contributions, Scaridae and Acanthuridae can be considered as discriminating families of the groups seagrass vs. coral and coral vs. sand.

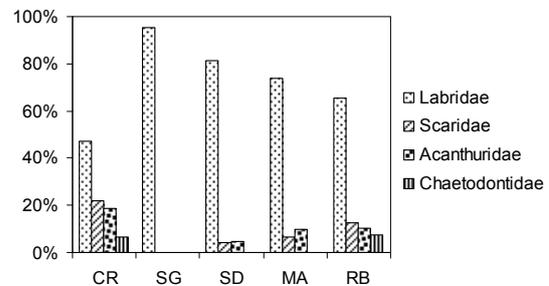


Figure 2: Average contribution of the fish families to the intra-group (habitat type) similarity. CR: coral; SG: seagrass; SD: sand; MA: macroalgae; RB: rubble

Spatial Prediction of Fish Abundance

Surface predictions of abundance were created for the three families which displayed the strongest patterns in their relationship with habitat types: Labridae, Acanthuridae and Scaridae. Labridae distribution was predicted at two reef units, NM_NB and RA (Figure 1), where seagrass habitats have been mapped. Modelling of Acanthuridae and Scaridae abundance was performed based on coral cover and thus at four reef units (KO_MA, LA, ND and NG_VL). Models including data on the three biotic habitats (seagrass, macroalgae and coral) were also developed in order to investigate whether considering additional habitat characteristics improved the results.

Table 2: Percentage contribution of family to assemblage dissimilarity between habitat types.

Family	rubble vs. seagrass	seagrass vs. coral	macroalgae vs. sand	rubble vs. sand	coral vs. Sand
Labridae	45.21	30.96	52.57	48.07	37.75
Scaridae	19.96	26.54	14.93	19.55	24.32
Acanthuridae	14.07	23.73	13.5	12.94	20.41
Chaetodontidae	7.49	6.84	7.42	7.94	6.29
Siganidae	6.15	4.33	5.6	4.91	3.98

Table 3: Results of the cross-validations of the fish abundance models which showed the lowest statistical errors. Models based on “3 habitats” included seagrass, macroalgae and coral covers.

Model	Reef unit	Mean error	Root-mean-square	Average standard error	Mean standardized	Root-mean-square standardized
Labridae + Seagrass	RA	0.236	63.39	62.76	0.006	1.018
Acanthuridae + 3 habitats	ND	0.008	22.18	23.08	0.001	0.962
Scaridae + 3 habitats	ND	0.055	17.65	17.96	0.002	0.987

Comparison between the models using cross-validation showed that the quality of the models was highly variable among families and reef units (Table 3). Surface prediction of Labridae abundance appeared to be more difficult to create with high accuracy since the models were not satisfactory with mean prediction errors ranging from -1.546 to 0.236. However, prediction was best fitted at RA when only seagrass was included in the model.

Models of Acanthuridae and Scaridae seemed to provide better results at some locations. The best mean prediction errors were for both families found at ND. Considering the models of Acanthuridae, the best model predicted abundance with an average standard error of 22 at ND. This uncertainty appeared to be overestimated since the root-mean-square standardized error was less than one (0.962). It seemed that the model quality was slightly improved with the integration of the three biotic habitat types. Mean prediction error of Scaridae abundance was lowest at ND with the coral based model. However, Scaridae abundance was estimated with an average standard error of 18 at ND based on the three habitat covers. The root-mean-square standardized error was 0.987 indicating a better assessment of the variability though there remained an apparent overestimation of abundance. Cross-validation results showed that Acanthuridae and Scaridae predictions were of better quality when they were based on the three-habitat model while Labridae model was best fitted with the one-habitat model.

Discussion

The modelling of selected fish families using geostatistical methods provided contrasting results depending on the site and the families. It appeared that models were likely to create accurate surface predictions of abundance at one specific reef unit (ND) for one family (Acanthuridae). It is possible that the level of spatial autocorrelation in reef fish distribution was too low to provide accurate models. Reef fish distribution is influenced by spatial, temporal and behavioural factors which might have also affected the level of accuracy. In addition, the number of samples was a constraint preventing the modelling of fish abundance at several reef units. It suggests that additional field

data may improve the model accuracy but it compromises the intent of the method which is to obtain results with already available data. Prediction maps of Scaridae and Acanthuridae showed a narrow range of abundance variation across the reef unit (Figure 3). Such patterns are unlikely to occur in reality and when the predicted values were compared with the measured ones, the variation was not reflected within the former. However, the results need to be compared with the reality of the field.

It has been argued that the tendency of the fish species to move and migrate influences the results of spatial variation models (Ault and Johnson 1998). They found that the best predictive models of fish density occurred for site-attached species such as the members of the Pomacentridae (Damsel fish) family.

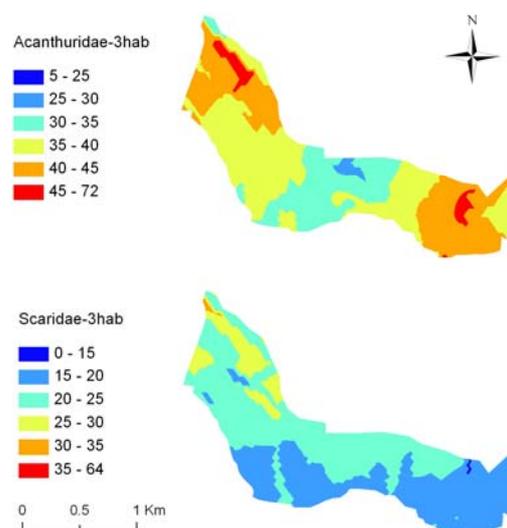


Figure 3: Surface predictions produced by the two statistically best co-Kriging models of the Scaridae and Acanthuridae abundances at the ND reef unit.

Reef ecologists first conducted their research with the notion that fish communities were organized in an equilibrium system to later shift toward an approach which comprehends reef fish communities as an open non-equilibrium system (Sale 1991). Owing to the difficulty in drawing general patterns in reef fish community organization, ecologists have for some time accepted that reef fish assemblages were driven by

stochastic processes. Recent research has increasingly shown that the habitat utilization by reef fishes was not a random process (Ault and Johnson 1998; Friedlander et al. 2007; Garpe and Ohman 2007).

If reef fish assemblages are non-equilibrium systems which develop within a dynamic environment, their modelling can be done through simulation models (Guisan and Zimmermann 2000). However, this needs a high level of understanding of the relationship between assemblage and habitat. For instance, spatial extrapolation methods encompass the determination of the scale effect on the response variable, in this case on the fish assemblage structures (Miller et al. 2004). It thus appears necessary to gain a good understanding of the scale-dependence of the reef fish assemblages in order to reach accurate predictions. Yet again it implies a design of sampling methodologies allowing the estimation of the scale effect on the results.

Modelling of reef fish assemblages using geostatistics showed some satisfactory results in selected taxa and defined geographic areas. However, the outcomes of this investigation suggested that geostatistics applied to reef fish ecology could provide better results if used to define sampling design for a non-parametric spatial approach. In addition, there is a need to deepen the investigation of the Coral Coast reef fish assemblage patterns and processes. By examining the impacts of other influencing factors such as the habitat complexity and human pressure, the model results could be improved. Integrating human pressure could be based on the fishing pressure gradients documented by previous studies done in Fiji (Jennings and Polunin 1995; Dulvy et al. 2004). Study on the temporal variation and the recruitment effects would be of great interest but would be more costly in terms of time and funds. Similarly, the model may be improved by basing it at the species level which would allow the examination of other indicators such as species richness, biomass and more clearly defined feeding, ecological and behavioural guilds.

This investigation provided encouraging results considering that it was conducted with previously collected data and considering the Fiji context where resources are limited. Moreover, it demonstrated to local scientists and practitioners the potential of Geographic Information System and spatial analysis to coral reef science and management.

Acknowledgements

This study was funded by the Initiative for the Coral Reefs in the South Pacific (CRISP) and the University of the South Pacific Research Fund. We thank the Coral Cay Conservation organization for sharing their data on the Coral Coast. We are grateful to the people who provided help for the analysis,

mapping and field work: Dr. Nick Rollings, Chris Roelfsema, Akuila Cakacaka, Rusiate Ratuniata and Ron Simpson.

References

- Ault TR, Johnson CR (1998) Spatially and temporally predictable fish communities on coral reefs. *Ecol Monogr* 68:25–50
- Bell JD, Galzin R (1984) Influence of live coral cover on coral reef fish communities. *Mar Ecol Prog Ser* 15:265–274
- Bouchon-Navaroa Y, Bouchona C, Louisa M, Legendre P (2005) Biogeographic patterns of coastal fish assemblages in the West Indies. *J Exp Mar Biol Ecol* 315:31–47
- Chabanet P, Ralambondrainy H, Amanieu M, Faure G, Galzin R (1997) Relationships between coral reef substrata and fish. *Coral Reefs* 16:93–102
- Clarke K, Warwick R (2001) Change in marine communities: an approach to statistical analysis and interpretation, 2nd edition. Primer-E Ltd., Plymouth
- Dominici-Arosemena A, Wolff M (2005) Reef fish community structure in Bocas del Toro (Caribbean, Panama): gradients in habitat complexity and exposure. *Caribb J Sci* 41:613–637
- Dulvy NK, Polunin NVC, Mill AC, Graham NAJ (2004) Size structural change in lightly exploited coral reef fish communities: evidence for weak indirect effects. *Canadian J Fish Aquat Sci* 61:466–475
- Fiji Islands Bureau of Statistics (2007) Provisional Results of the 2007 Population and Housing Census, Press Release no. 52 (31/10/07)
- Friedlander AM, Brown EK, Monaco ME (2007) Coupling ecology and GIS to evaluate efficacy of marine protected areas in Hawaii. *Ecol Appl* 17:715–730
- Garpe KC, Öhman MC (2003) Coral and fish distribution patterns in Mafia Island Marine Park, Tanzania: fish–habitat interactions. *Hydrobiologia* 498:191–211
- Garpe KC, Öhman MC (2007) Non-random habitat use by coral reef fish recruits in Mafia Island Marine Park, Tanzania. *Afr J Mar Sci* 29:187–199
- Guisan A, Zimmermann NE (2000) Predictive habitat distribution models in ecology. *Ecol Model* 135:147–186
- Jennings S, Polunin NVC (1995) Comparative size and composition of yield from six Fijian reef fisheries. *J Fish Biol* 46:28–46
- Johnston K, Ver Heof JM, Krivoruchko K, Lucas, N (2001) Using ArcGIS™ Geostatistical Analyst. ESRI, Redlands
- McClanahan TR, Arthur R (2001) The effect of marine reserves and habitat on population of East African coral reef fishes. *Ecol App* 11:559–569
- Mellin C, Andréfouët S, Ponton D (2007) Spatial predictability of juvenile fish species richness and abundance in a coral reef environment. *Coral Reefs* 26:895–907
- Miller JR, Turner MG, Smithwick EAH, Dent CL, Stanley EH (2004) Spatial extrapolation: the science of predicting ecological patterns and processes. *BioScience* 54:310–320
- Mummy PJ (2006) The impact of exploiting grazers (scaridae) on the dynamics of Caribbean coral reefs. *Ecol Appl* 16:747–769
- Núñez-Lara E, Arias-González J, Legendre P (2005) Spatial patterns of Yucatan reef fish communities: Testing models using a multi-scale survey design. *J Exp Mar Biol Ecol* 324:157–169
- Pittman SJ, Christensen JD, Caldwell C, Menza C, Monaco ME (2007) Predictive mapping of fish species richness across shallow-water seascapes in the Caribbean. *Ecol Model* 204:9–21
- Roberts C, Ormond RFG (1987) Habitat complexity and coral reef fish diversity and abundance on Red Sea fringing reefs. *Mar Ecol Prog Ser* 41:1–8
- Rowlands G, Comley J, Raines P (2005) The Coral Coast, Viti Levu, Fiji: A Marine Resource Assessment. Coral Cay Conservation Report
- Sale PF (1991) Reef fish communities: open nonequilibrium systems. In: Sale PF (ed) *The ecology of fishes on coral reefs*. Academic, London, pp 564–598
- Thaman B, Aalbersberg B (2004) Initiating Integrated Coastal Management (ICM) in the Fiji Islands. Coastal Zone Asia-Pacific Symposium 2004, Brisbane