7-2008

A Model Framework for Predicting Reef Fish Distributions Across the Seascape Using GIS Topographic Metrics and Benthic Habitat Associations

Brian K. Walker
Nova Southeastern University, walkerb@nova.edu

Proceedings of the 11th International Coral Reef Symposium, Ft. Lauderdale, Florida, 7-11 July 2008
Mini-Symposium 17

A Model Framework for Predicting Reef Fish Distributions Across the Seascape Using GIS Topographic Metrics and Benthic Habitat Associations

B.K. Walker[1]

1) Nova Southeastern University Oceanographic Center, National Coral Reef Institute
8000 North Ocean drive, Dania Beach, FL 33004, walkerb@nova.edu

Abstract. Increased topographic complexity has been linked to increased species diversity and/or abundance in many ecological communities, including coral reefs. Several topographic metrics can be measured remotely in GIS using high resolution bathymetry, including elevation, surface rugosity, and seafloor volume within specified areas. Statistical relationships between these data and organismal distributions within mapped habitats can be used to make predictions across the entire bathymetric dataset. In this study a model framework is presented which utilizes statistically significant relationships between reef fish abundance and species richness and GIS topographic complexity measurements for samples within similar benthic habitats to create GIS-based prediction maps of abundance and species richness for the entire seascape. Reef fish associations with GIS topographic metrics were significant and varied between habitats. Model evaluation showed that patterns in the measured data emerged in the prediction data. The results allow for viewing of data trends throughout the seascape, quantification of assemblages in non-sampled areas, and statistical comparisons of areas within the region to support and guide management related decisions. This model framework can be adapted to other communities (e.g. benthic organisms) and/or parameters (e.g. diversity) that relate to topographic complexity.

Keywords: coral reef, Florida, habitat complexity, prediction, reef fish, rugosity.

Introduction

Studies linking small-scale measurements of abundances and species distributions to broad-scale seascapes are the key to understanding and predicting organismal distributions and their dynamics (Heglund 2002). Reef fish studies are often limited to small spatial scales because of logistical and economic constraints; however, viewing the data at larger spatial scales might elucidate unforeseen relationships and patterns (Sale 1998). Furthermore, the need for large-scale spatial analyses of reef fish is growing due to the over-exploitation of marine resources and the need for management and conservation of large areas (Kendall et al. 2003).

Remote sensing allows the acquisition of large amounts of data quickly and economically, providing the foundation for large-scale resource mapping and modeling. These maps are the basis upon which seascape analyses and modeling efforts are constructed (Pittman et al. 2007; Walker et al. *in press*). Previous research has shown that increased habitat complexity/rugosity positively influence reef fish abundance and/or species richness (Luckhurst and Luckhurst 1978; Gratwicke and Speight 2005). Traditional reef fish rugosity studies used an *in situ* measure of topographic complexity that is not practical on large spatial scales (>km^2) (McCormick 1994); however, this is now possible by analyzing high resolution 3-dimensional topographic surfaces in GIS (Kuffner et al. 2007; Pittman et al. 2007; Wedding et al. 2008). Several topographic metrics can be measured remotely at various scales in GIS using high resolution bathymetry, including elevation, surface rugosity, and seafloor volume within specified areas.

This manuscript presents a model framework that projects the relationships between reef fish assemblage metrics (abundance, richness, etc.) and GIS topographic metrics for multiple habitats in sampled locations across the seascape. Reef fish are used as a case study to show the model design and demonstrate its capabilities. The model framework design, accuracy, strengths, weaknesses, applications and recommended uses are discussed.

Methodology

A subset of 346 stationary daytime visual fish surveys from a larger effort to acquire a baseline census of the coral-reef-associated fishes in Broward County, Florida, USA (Ferro et al. 2005) was used in this study (Figure 1). The subset was chosen on the basis of location accuracy and agreement with independent

GIS data. The fish surveys were conducted using the Bohnsack and Bannerot (1986) method between 2000 and 2002 along 54 east-west transects, each separated by approximately 0.5 km. The surveys assessed fish species, abundance, and length in a 7.5 m radius circle at each location. Each transect consisted of nine fish survey locations that targeted the eastern edge, crest and western edge of each of the three main reef tracts, yet in many cases the nearshore ridge complex (NRC) was mistaken for the Inner and Middle Reefs (Walker et al. *in press*).

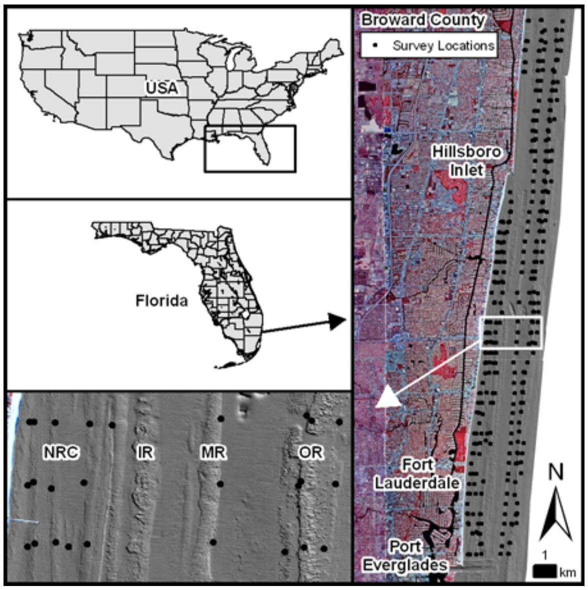

Figure 1. Aerial photo-LIDAR mosaic with the 346 point-count fish assessment sites in northern Broward County, FL, USA. NRC = Nearshore ridge complex; IR=Inner reef; MR=Middle reef; and OR=Outer reef.

GIS topographic analyses of the fish survey locations were performed in ArcGIS 9.2. Triangulated irregular networks (TIN) were created using LIDAR bathymetry for a 7.5 m radius area around each fish survey. This allowed over 12 bathymetric points per area for topographic analyses. The individual TINs were analyzed in 3D Analyst for Z min, Z max, 2D area, 3D surface area, and volume. Elevation was the positive difference between the min and max Z value. The surface rugosity index was the surface area of the TIN divided by its planar area. Volume was calculated as the space between the 3D surface and a horizontal plane at Z min.

Reef fish surveys were categorized by their location in relation to the benthic habitat characterization of Walker et al. (2008). Some habitats were excluded in the prediction model due to low fish survey sample sizes. The benthic habitats used herein were Ridge-Shallow, Colonized Pavement (CP)-Shallow, Linear Reef (LR)-Middle Shallow, LR-Middle Deep, CP-Deep, LR-Outer, and Aggregated Patch Reefs.

Analysis of variance (ANOVA) was used to analyze the data for differences in abundance and number of species per count (i.e., species richness). Abundance data (x) were log transformed using the formula $\log_{10}(x+1)$ to homogenize variance. Tukey HSD *post-hoc* tests were used to determine significance when more than two categories were examined. Linear regression was performed in Statistica 6.0 (Stat Soft Inc.) and an r^2, r, and p-value were reported for a best-fit linear regression line.

Predictions of reef fish abundance and species richness were made based on the linear regression equation of the GIS topographic measurements within each habitat. The model was created at the same scale as the fish surveys. A grid of 15 m square polygons was projected over the entire survey area. Depth, elevation, volume, and surface rugosity index were calculated for each polygon in GIS resulting in each grid polygon having individual topographic statistics and habitat characterization based on its location to the seafloor. The grid polygon topographic data values were then input into the appropriate regression equation based on the GIS metric predictor and its habitat. This generated six columns of prediction data for each grid polygon: a predicted abundance and richness for each of the three GIS metrics.

Results

A comprehensive analysis on how the fish data relate to topographic complexity is presented in Walker et al. *in press*. In summary, both abundance and richness increased with increasing topographic complexity and these relationships changed across the seascape. Richness related to topographic complexity stronger in the shallow habitats, whereas, abundance exhibited a stronger relationship in offshore habitats. *In situ* rugosity measurement yielded the best explanation of fish assemblage structure parameters, but the weaker GIS metric correlations followed similar trends. Since linear regression results varied between habitats and between GIS metrics, a separate regression equation was determined for each. Several of the relationships were not statistically significant but were included in the model for completeness.

The prediction model yielded 134,704 square polygons, each with a value for predicted fish abundance and richness using the elevation, volume, and surface rugosity values generated from the regression equations in their respective habitats, resulting in six separate prediction maps.

Linear regressions of the total measured fish abundance and richness versus the predicted values for all metrics showed statistically significant relationships ($p<0.0001$). Elevation had the highest r^2 values in both abundance and richness of the three GIS metrics, $r^2=0.27$ and 0.39 respectively; surface rugosity had slightly lower r² values than elevation for abundance (r^2=0.25) and richness (r^2=0.38); and volume had the lowest r² values for both abundance (r^2=0.19) and richness (r^2=0.31).

ANOVA comparisons of reef fish abundance between the surveys and predictions within habitats showed only one statistical difference where volume abundance was significantly higher than the measured abundance on the LR-Middle Shallow ($p<0.05$) (Fig. 2, upper). In every other case, the predicted means were not significantly different from the measured means for each habitat. This resulted in the data trends of the empirical values emerging in most of the predictions. For example, both measured and predicted reef fish abundance were significantly lower ($p<0.05$) in the CP-Shallow than the LR-Middle Shallow, the LR-Middle Deep, the LR-Outer Reef, and the Aggregated Patch Reef.

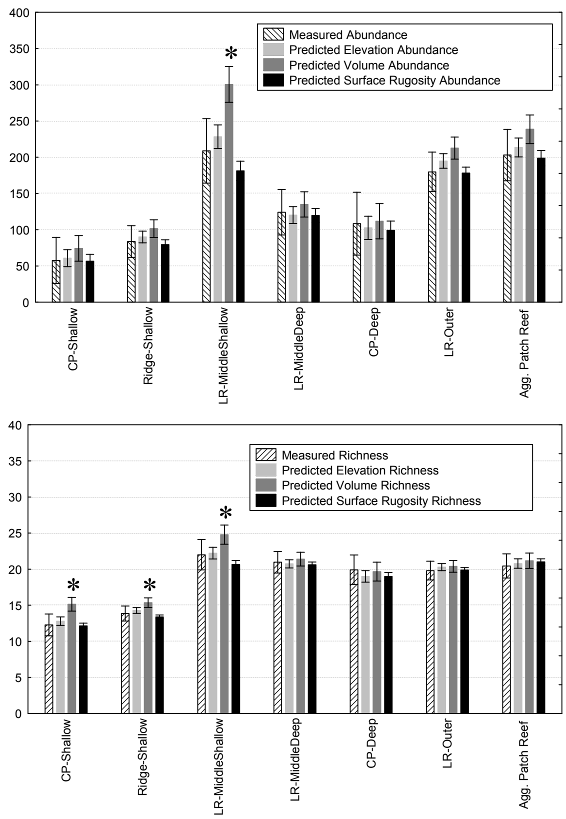

Figure 2. Measured abundance (upper) and species richness (lower) (hashed) and predicted values of reef fish by GIS calculated elevation (light grey), volume (medium grey), and surface rugosity (black) by benthic habitat. Error bars show one standard deviation about the mean. * indicates significant difference from measured abundance ($p<0.05$).

Comparisons of species richness by ANOVA between the surveys and predictions within habitats showed that the predictions did not significantly differ from the measured data with the exception of three significantly higher volume predictions ($p<0.05$) (Fig. 2, lower). In every other case the modeled data showed the same trends between habitats. For example, the CP-Shallow and Ridge-Shallow had significantly lower richness than the other habitats which did not significantly differ from one another for the measured and predicted data ($p<0.05$).

Discussion

Previous analyses of reef fish and LIDAR topography have either not attempted modeling (Kuffner et al. 2007) or focused modeling efforts on species richness (Pittman et al. 2007). The model presented herein adopts an approach to predicting reef fish distribution not previously reported. By using new technologies to project the relationship of both species richness and abundance to large-scale topographic complexity across the seascape, it provides the ability to view, quantify, and relate these predicted data.

Biological modeling involves less certainty than models based on physics or chemistry, which are derived from fundamental laws (Mitasova and Mitas 2002). The accuracy of the model presented herein relies heavily on the observed data. Although statistically relevant, the regressions showed a relatively low agreement between the predicted and measured data (r^2 = 0.27 for abundance and 0.39 for richness). This relationship was expected to be very high (r^2 > 0.80) since the model was developed using the same data. The output weaknesses were likely caused by the weak measured relationships between the measured reef fish variables and GIS metrics (Walker et al. *in press*). Because the initial relationship is weak, the output did not yield a high degree of accuracy. However, the comparisons between mean abundance and richness values of the fish surveys (measured data) and the predicted values among benthic habitats showed high agreement. In most of the habitats neither mean predicted abundance nor richness significantly differed from the mean measured values with the exception of volume. Hence, the empirical data patterns between habitats emerged in both predicted abundance and richness exhibiting the same trends in the data within each habitat. This suggests the model is more powerful as a comparative tool than a tool to predict absolute values in an area.

As a comparative tool, the model can provide very useful information for decisions on Marine Protected Area (MPA) placement. An MPA's location is of key importance to optimize its potential (Baker 2000). MPAs representing a full range of habitats are most effective (Carr et al. 2003) and they should contain essential fish habitat (Rieser 2000) and highly rugose areas (Friedlander et al. 2007). This model provides the information necessary to statistically compare different areas based on the organism's relationship to topography throughout the seascape. For example, a comparison of model data between two 1 km stretches of Middle Reef shows clear quantifiable differences (Figure 3). A T-test comparison showed predicted mean abundance in area A (253.9 ± 4.5 SEM) was significantly higher than area B (178.8 ± 2.7 SEM) and area A contained significantly higher species richness (23.8 ± 0.16 SEM) than area B (21.6 ± 0.09 SEM). In this example, area A would be a better conservation area based on predicted fish data and because these data are in GIS, they can be analyzed in relation to other data relevant to MPA design and implementation.

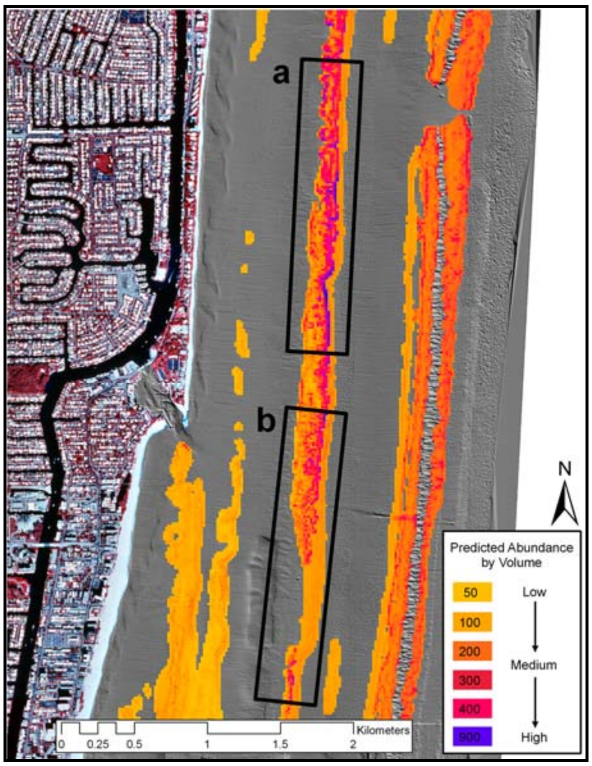

Figure 3. A map of the predicted fish abundance by volume showing two identically-sized areas used for statistical comparison. Box a contains significantly higher mean predicted abundance and richness than box b.

Bathymetry

High resolution bathymetry is some of the most valuable data to acquire in mapping submerged lands. These data, which have many uses beyond the scope of this study, were essential to mapping the benthic habitats and obtaining topographic measurements of discrete areas over the seascape. The 4 m resolution bathymetry was sufficient to map the habitats; however, it was not ideal for measuring the topographic variables at a sufficient operational scale to the fish assemblage (Walker et al. *in press*). Differences in bathymetric resolution have implications on the topographic measurements calculated in the GIS (Wolock and McCabe 2000). It is recommended that future bathymetric surveys be taken at a higher density to obtain more accurate topographic information.

Benthic Habitat Mapping

Benthic habitat mapping is an essential tool for effective management of submerged resources (Friedlander et al. 2007). Mapping the resources not only aids resource managers in the determination of mitigation for impacts, the designation for marine protected areas, and the identification of essential fish habitat, it also can elucidate previously unforeseen relationships in data brought on by the proper classification of the sample sites (Walker et al. *in press*). For Example, on a patch reef system in Biscayne National Park, FL, Kuffner et al. (2007) did not find significant differences between abundance and richness with rugosity in pooled data, but found significance when the data were split by individual patch reef. Hence, measuring changes in relationships between habitats is essential to the accuracy of prediction models.

The scale of habitat mapping can also affect the model and it is likely that a map at a finer scale would produce better results. In the current map, the area within each polygon is homogenous as described by each classifier (Walker et al. 2008). The absence of within-polygon variation might significantly underestimate the total variance of the polygonal data (Bian 1997). The variation of benthic cover within habitats could introduce significant variation in the data, obscuring other relationships (Aaby et al. 2004). Since variations within habitats (patchiness) were acoustically detected (Walker et al. 2008), it is possible that this confounded the reef fish-topographic complexity relationship.

It is recommended that benthic habitat mapping be created at the finest scale possible to include variations of patchiness within major habitat categories. This can be accomplished through high density acoustic surveys or LIDAR backscatter habitat classification (Foster et al. *in press*).

Model Adaptation

This empirical static model has been developed based on the statistical analyses of observed data enabling views of the relationship between reef fishes and their habitats on a large scale (>100 km²), allowing for statistically comparable analyses between areas based on empirical data, and thus giving statistical support to resource management decisions. Its simple design makes it highly adaptable to other uses. The framework can be used to predict any biological/ecological relationship to topographic complexity provided the bathymetry and mapping data are of the appropriate scale. For example, it could be used to predict coral reef biodiversity via topographic complexity. The grid polygon size could be adjusted to change the scale of the model and the benthic habitat resolution could be tuned accordingly.

This system could also be taken to the next level as a spatial decision support system- a computer-based system designed to assist decision making (Corbett et al. 2002). The framework could be assembled in a user-friendly program with more automated processes and the ability to obtain instant viewable results in a GIS. Once the grid has been created and the topographic statistics calculated, fine tuning the ecological process relationship is a statistical procedure that could be self-contained in a program that would allow a user to specify the relationship (i.e. input the regression equations) and quickly view the results. This could be extremely useful to scientists studying different ecological processes and resource managers in making decisions on resource use and/or mitigation.

Future research can greatly increase this model's accuracy. Increasing the resolution of bathymetric data and habitat mapping units would eliminate several possible error sources; however, research is still needed to better understand the dynamics of how reef fish relate to topographic complexity and the other ecological factors influencing their distributions. Better understandings of the appropriate measurement scale and the scales at which reef fish operate would help to model their distributions more accurately. As these relationships are uncovered, modeling efforts using topographic complexity as a proxy for organism distribution may become more accurate.

Acknowledgements

This was a product of my PhD dissertation research at Nova Southeastern University. I thank my committee, Bernhard Riegl, Richard Dodge, Richard Spieler, and Mark Monaco for their guidance. This manuscript is a result of research partially funded by the National Oceanic and Atmospheric Administration Coastal Ocean Program under award NA03NOS4260046 to Nova Southeastern University for the National Coral Reef Institute (NCRI).This is NCRI contribution #101.

References

Aaby A, Wright D, Tissot B (2004) Case study of the ArcGIS marine data model: examining habitat utilization patterns of reef fish along the west coast of Hawaii. Proceedings of the 24th Annual ESRI User Conference, San Diego, CA, 1458:1-28.

Baker JL (2000) Guide to marine protected areas. Prepared for the Office for Coast and Marine, Department for Environment and Heritage, South Australia. ISBN 1 876562 24 2.

Bohnsack JA, Bannerot SP (1986) A stationary visual technique for quantitatively assessing community structure of coral reef fishes. NOAA techincal report. NMFS 41:1-15.

Bian L (1997) Multiscale nature of spatial data in scaling up environmental models. In: Quattrochhi DA, Goodchild MF (eds) Scale in remote sensing and GIS. CRC press, Boca Raton, Florida, pp 13-26.

Carr M, Neigel J, Estes J, Andelman S, Warner R, Largier J (2003) Comparing marine and terrestrial ecosystems: implications for the design of coastal marine reserves. Ecol Appls 13: S90-S107.

Corbett JD, White JW, Collins SN (2002) Spatial decision support systems and environmental modeling: an application approach. In: Clarke KC, Parks BO, Crane MP (eds) Geographic information systems and environmental modeling. Prentice Hall, New Jersey, pp 51-66.

Ferro F, Jordan LKB, Spieler RE (2005) The marine fishes of Broward County, FL: final report of 1998-2002 survey results. NOAA Technical Memorandum, NMFS-SEFSC-532, p 73.

Foster G, Walker BK, Riegl B. (in press) Interpretation of single-beam acoustic backscatter using LIDAR-derived topographic complexity and benthic habitat classifications in a coral reef environment. J Coastal Res, 'in press'.

Friedlander AM, Brown E, Monaco ME (2007b) Defining reef fish habitat utilization patterns in Hawaii: comparisons between marine protected areas and areas open to fishing. Mar Ecol Prog Ser 351:221-233.

Gratwicke B, Speight MR (2005) The relationship between fish species richness, abundance and habitat complexity in a range of shallow tropical marine habitats. J Fish Biol 66:650-667.

Heglund PJ (2002) Foundations of species-environment relations. In: Scott MJ, Heglund PJ, Morrison ML (eds) Predicting species occurrences: issues of accuracy and scale. Island Press, Washington, pp 35-41.

Kendall MS, Christensen JD, Hillis-Starr Z (2003) Multi-scale data used to analyze the spatial distribution of french grunts, *Haemulon flavolineatum*, relative to hard and soft bottom in a benthic landscape. Environ Biol Fish 66:19-26.

Kuffner IB, Brock JC, Grober-Dunsmore R, Bonito VE, Hickey TD, Wright CW (2007) Relationships between reef fish communities and remotely sensed rugosity measurements in Biscayne National Park, Florida, USA. Environ Biol Fish 78:71-82.

Luckhurst BE, Luckhurst K (1978) Analysis of substrate variables on coral reef fish communities. Marine Biology 49:317-323.

McCormick MI (1994) Comparison of field methods for measuring surface topography and their associations with a tropical reef fish assemblage. Mar Ecol Prog Ser 112:87-96.

Mitasova H, Mitas L (2002) Modeling physical systems. In: Clarke KC, Parks BO, and Crane MP (eds) Geographic information systems and environmental modeling. Prentice Hall, New Jersey, pp 189-210.

Pittman SJ, Christensen JD, Caldow C, Menza C, Monaco ME (2007) Predictive mapping of fish species richness across shallow-water seascapes in the Caribbean. Ecol Model 204: 9-21.

Rieser A (2000) Essential fish habitat as a basis for marine protected areas in the U.S. Exclusive Economic Zone. Bull Mar Sci 66:889–900.

Sale PF (1998) Appropriate spatial scales for studies of reef fish ecology. Australian J Ecol 23:202-208.

Walker BK, Riegl BM, Dodge RE (2008) Mapping coral reef habitats in southeast Florida using a combined technique approach. Journal of Coastal Research 24:1138–1150.

Walker BK, Jordan LKB, Spieler RS (in press) Relationship of the reef fish assemblage and topographic complexity on southeastern Florida coral reef habitats. J Coastal Res 'in press'.

Wedding L, Friedlander A, McGranaghan M, Yost R, Monaco M (in press) Using bathymetric LIDAR to define nearshore benthic habitat complexity: implications for management of reef fish assemblages in Hawaii. Remote Sens Environ 'in press'.

Wolock DM, McCabe GJ (2000) Differences in topographic characteristics computed from 100- and 1000-m resolution digital elevation model data. Hydrol Process 14:987-10.