1 Processing spatial data for predictive mapping of benthic substrate and biota: Point Addis Marine National Park

The following is an overview of methods and data development for mapping:

(1) Data processing:
   a. Developed secondary datasets from bathymetry;
   b. Designed and executed video interpretation plan to use as primary field collected data identifying the substrate and biological features on the seafloor;

(2) Modelling: Modelled all substrate and biological categories as a function of bathymetry and derived datasets to provide a framework for predicting values in unsampled locations;

(3) Mapping: Applied the models to the rest of the Park to predict the most likely substrate and biological features present where no field data was available.

In this document, the processing of field data (point 1, above) for mapping is described. The modelling (2) and mapping (3) of the benthic substrate and biota are covered in Milestone 6.03 Classification under Task 3.6: Revision and review of designs and methodologies.

(1) Data Processing

Predictor dataset development: The hydroacoustic surveys for Point Addis include both bathymetry and backscatter imagery, collected in March, 2005, using an 8101 Reson Multibeam. Both datasets potentially will be useful for classifying the seafloor, however, at the time of map production there were a number of missing segments and artifacts in the backscatter data which heavily impacted the mapping results. This data was excluded from the current modeling, but should be revisited in the future when the backscatter processing has been improved.

A variety of secondary (textural) datasets which may correlate with seafloor properties were developed from the bathymetry using terrain analysis techniques. These techniques are applied to elevation data, quantifying the relationships among elevation values in small neighborhoods to reveal textural differences. Using a gridded elevation (or bathymetry) dataset, calculations are run on a small number of cells surrounding each pixel. A 3 x 3-cell analysis window for two pixels, A and B, is shown in **Figure 1** for illustration. In this case, all neighborhood calculations (such as the mean, mode, or slope) are run on the central cell plus the 8 surrounding cells, and the value assigned to the central cell in the output, thus creating a derivative dataset. The full suite of datasets used for modeling at Point Addis are described in **Table 1**. Gray-level co-occurrence matrix methods developed for textural analysis of radar images were also explored, but did not significantly improve the models, and hence were not included in the final product.

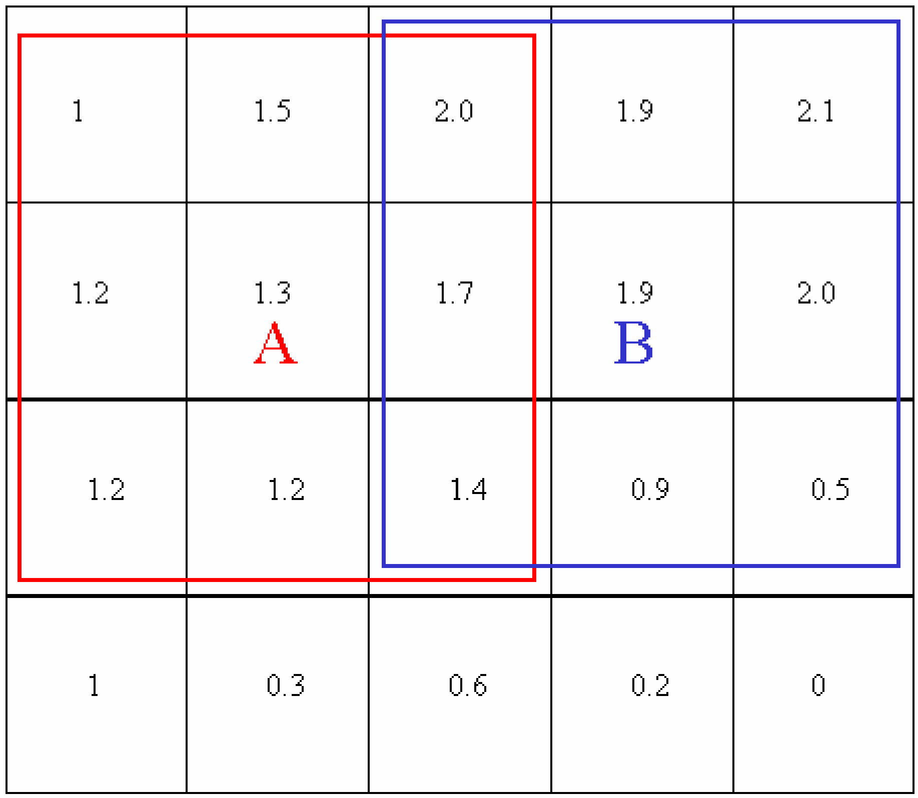

Figure 1-1 Example of the definition of local neighborhoods for terrain analysis. Red and blue squares highlight the analysis neighborhood for cells A and B.

Table 1: Datasets derived from bathymetry that were tested for use as predictors of seafloor substrate and biota at Point Addis.

Predictor datasets	Definition	Source
Bathymetry	Elevation relative to the Australian Height Datum (AHD)	ArcGIS 9.0
Std Dev *	Standard deviation of elevation (local neighborhood analysis)	ArcGIS 9.0
Slope	First derivative of elevation: Average change in elevation / distance)	ArcGIS 9.0
Aspect	Azimuthal direction of the steepest slope.	ArcGIS 9.0
Profile curvature	Second derivative of elevation: concavity/convexity parallel to the slope	ArcGIS 9.0
Plan curvature	Second derivative of elevation: concavity/convexity perpendicular to the slope	ArcGIS 9.0
Curvature	Combined index of profile and plan curvature.	ArcGIS 9.0
Hypsometric index*	Indicator of whether a cell is a high or low point within the local neighborhood	ArcGIS 9.0
Moran's I*, r = 5	A weighted correlation coefficient used to detect spatial dependence. Calculated on the residuals from a linear trend surface.	Laffan, 1998
Local relief (Range)*	Maximum minus the minimum elevation in a local neighborhood	ArcGIS 9.0
Rugosity (surface area)	Surface area of the local neighborhood	Jenness, 2002
Rugosity (surface ratio)	The ratio of the surface area divided by the planimetric area of the neighborhood.	Jenness, 2002

*Local neighborhood analysis: run on circles of radius 10m, 20m, 50m, and 100m.

Video data interpretation: All the video observations had to be carefully georeferenced in order to directly compare information from the video, bathymetry, and the secondary datasets. Locations of over 80,000 frames were recorded. Before interpretation of the video could begin, a small subset had to be chosen. To this purpose, the video data were interpreted in two phases. First a 'coarse classification' was done to identify the major categories that were visible on the video, and to delineate major transitions between bottom types. The second phase of video interpretation, 'point interpretation', involved much more detailed observations on individual frames (i.e. point locations). The coarse classification was used to target the locations of the point samples. Areas identified as having high variability were allocated higher density point samples, and more homogeneous areas were more sparsely sampled. Out of 58 km of video footage, 34 km (57%) were categorized as high variability based on mixed substrate categories, and 24 km (41%) as low variability. The high variability sections were regularly sampled at approximately one frame per 20 meters, and the low variability sections at one frame per 60 meters. For information spanning a range of spatial scales, additional locations (57 in high and 43 in low variability areas) were selected for clustered sampling. At each location, five additional video frames were interpreted at 4-m intervals. This resulted in a total of 2862 georeferenced data points for which all aspects of substrate and sessile biota were described. The major substrate categories that were identifiable from video included type (reef, sediment, or cobble) and several variations of texture and structure. The biota identified were at the coarsest level vegetation and sessile invertebrates, and included a wide range of subcategories (**Table 2**). All attributes that were reasonably clearly identified from the video were classified as present (1) or absent (0), to prepare for modeling. Several types of sessile invertebrates were identifiable, but too sparse to model. The distributions of these plus maps of poor video quality or interpretation uncertainty are shown in **Figure 2**.

Table 2: Frequency of occurrence of substrate and biota classes in the video footage. ( ) indicate number of presences and percentage out of 2862 total frames interpreted.

Substrate (2862, 100%):	**Vegetation** (1136, 40%):	**Sessile invertebrates** (1450, 51%):
Sediment (2457; 86%)	Seagrass (0, 0%)*	Sponges (888, 31%)
Coarse sediment (856; 30%)	Algae (1136, 40%)	Ascidians (122, 4%)
Fine ripples (609; 21%)	Mixed Green (1, 0%)*	Bryzoa (0, 0%)*
Coarse ripples (623; 22%)	Mixed Red (967, 34%)	Soft corals (15, 1%)*
Reef (1206; 42%)	Rhodoliths (480, 17	Gorgonians (15, 1%)*
Vertical reef (38,1%)	Mixed Brown (662, 23%)	
Solid reef (840; 29%)	Kelp (597, 21%)	
Cobbles (480; 17%)	Ecklonia (586, 20%)	

*Insufficient data for modelling.

Figure 1-2 Location of Gorgonians (A) and soft corals (B); both classes were too uncommon at Point Addis for modelling

References

Hosmer, D.W. and Lemeshow, S. 2000. Applied logistic regression (2nd Edition). New York: Wiley.

Jenness, J. 2002. Surface Areas and Ratios from Elevation Grid (surfgrids.avx) extension for ArcView 3.x, v. 1.2. Jenness Enterprises. Available at: http://www.jennessent.com/arcview/surface_areas.htm .

Laffan, S. 1998. Program for calculating a local Moran's I statistic. Arc Macro Language (AML).