Probability of existence of benthic habitat classes for the Oceanic Shoals CMR

Ben Radford, Marji Puotinen

Introduction

Environment Australia is tasked with managing the networks of Commonwealth Marine Reserves (CMRs) and Key Ecological Features (KEFs). Effective management requires first establishing a baseline of what key biota are located in a given CMR or KEF, and then regularly monitoring their status to ensure all is well. A particular challenge is posed by CMRs and KEFs located in the remote and relatively poorly known N and NW regions, where it is logistically impossible to survey every area in detail. One way to 'fill in the gaps' between field observations is to build spatial predictive habitat models. For the marine environment, such modelling involves collecting and integrating spatial datasets to build realistic representations of both the topography and composition of the seafloor and major biotic groups (Brown et al 2011, Holmes et al 2008). Such models are widely useful to scientists and managers, for example to:

- determine the spatial heterogeneity of the benthic environment and key classes of organisms,
- evaluate the physical and biological controls on individual and joint habitat distributions,
- discover relationships among habitats and various species of interest, and
- investigate how habitats and organisms respond to disturbance from human activities.

The Oceanic Shoals CMR is a classic example of a case where spatial predictive benthic models can be useful. Multiple field campaigns have collected high resolution survey data in seven study areas within or near the Oceanic Shoals, but these studies collectively cover only a small fraction of the total area of the CMR. While this fine-scale data is valuable, managers need a map showing where benthic habitat types exist across the entire Oceanic Shoals CMR. We aimed to address this need by producing a benthic habitat map for the CMR using spatial predictive modelling, along with a guide to its limitations and how it should be used appropriately.

Methods

Benthic spatial predictive habitat models aim to map the spatial distribution of types of bottom-dwelling organisms across an area of interest in as much spatial detail as robustly possible. In producing such models for NESP, we aim to ensure they are:

- Ecologically meaningful on relevant spatial and temporal scales,
- Sufficiently accurate for the intended use, and
- Communicated to stakeholders clearly so that their limits and likely errors are clearly understood.

We build such models following the basic process outlined below (Figure 1).

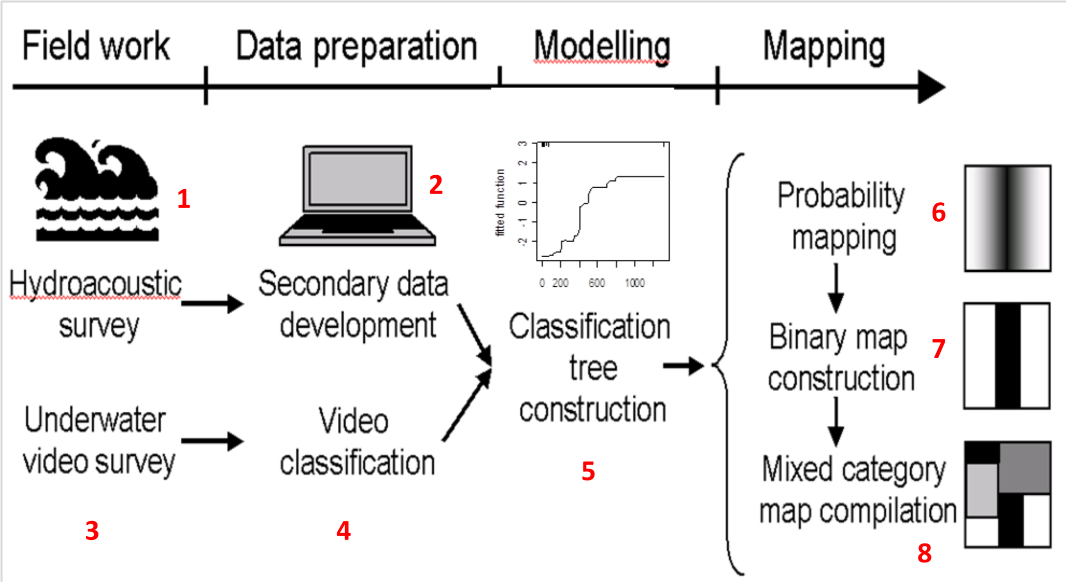

Figure 1. Diagram illustrating the process of producing a spatial benthic prediction model.

Developing predictors

Developing environmental surrogates for the existence and abundance of classes of benthic organisms (step 2) is possible with high resolution bathymetric data (Brown et al 2011). Where such data do not exist in an area of interest, they can be developed from multi-beam sonar data via hydro-acoustic surveys (Holmes et al 2008, Lehmann et al. 2002- step 1 on Figure 1). For the Oceanic Shoals, hydro-acoustic data has been collected at seven locations (Figure 2).

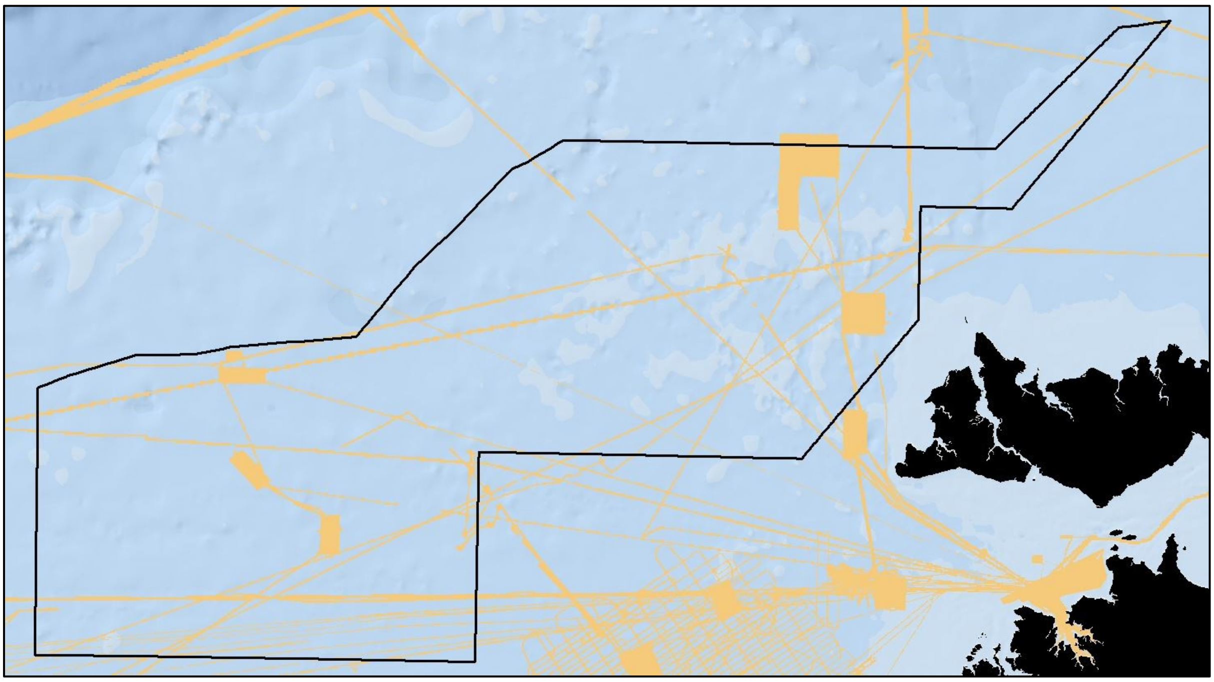

Figure 2. Very high resolution multi-beam sonar coverage of the Oceanic Shoals CMR. The CMR is outlined in black. High resolution multibeam data coverage is shown in orange. Data courtesy of Geoscience Australia.

We used this combined with Geoscience Australia's 250 metre bathymetry coverage of Australian waters to develop the following potential predictors of benthic habitat at 250 m resolution (Figure 1, step 2):

- Depth
- Aspect
- Overall curvature
- Profile curvature
- Plan curvature
- Depth range (5, 10, 25, 50 m windows)
- Standard deviation of depth (5, 10, 25, 50 m windows)
- Mean depth (5, 10, 25, 50 m windows)

Although ocean parameters are also important drivers of benthic community composition and structure (Brown et al 2011), relevant data was not available at spatial and temporal scales sufficient to make it worthwhile to include for the Oceanic Shoals.

<u>Developing training and test data</u>

Building predictive models is not possible without verified field data to document where biota of various types actually exist. The data we use to build and test a given model (Figure 1 – steps 3 & 4) comes from AIMS underwater towed video (http://www.aims.gov.au/docs/research/monitoring/seabed/video-monitoring.html) surveys. It includes both real-time coding of benthic communities from Go-Pro video footage and detailed analysis of downward facing still photos (Figure 3) taken along each of a series of 1.5 km long transects. The former is adept at identifying benthic types like sponge communities that may have a minor profile on a downward facing image while still representing notable biomass. The latter excel at identifying benthic types that do not protrude far into the water column (eg, burrowers, hard corals). For a given field survey, the towed video transects are placed within the study area using a GRTS (Generalized Random Tesselation Stratified) sample design structured to spread transects across a priori classes of habitat complexity while ensuring they are evenly distributed spatially (https://science.nature.nps.gov/im/datamgmt/statistics/r/advanced/grts.cfm).

Figure 3. The AIMS towed video system tow body with mounted video camera and down-ward facing camera for stills (lower), with an example of a still photo taken just above the sea floor (above).

Ecologists at AIMS use the CATAMI classification scheme (http://catami.org/classification) to assign benthic categories to what they observe both in real-time via video and by analysing still photos in the lab.

Building models and mapping habitats

We then explore the statistical relationship between the predictors and field data of the presence / absence of benthic classes across the area of interest using a non-parametric statistical method - classification trees (Figure 1, step 5 – Breiman et al 1984). For the Oceanic Shoals CMR model, we used an innovative version of this called random forest (Breiman 2001, Cutler et al 2007). A random forest model first builds hundreds of classification trees that identify all the unique combinations of variable that could predict the distribution of a given class. Those trees that are not useful in predicting that class cancel each other out. This method outperforms standard classification trees that are defined a priori because it ensures that valid relationships in the data are not missed (Cutler et al 2007).

In our process, we withhold a random sample of one-third of the field data to use for model performance estimates (testing set) and use two-thirds of it to establish how various benthic classes of organisms vary with the potential predictors (training set) to enable building a model. When establishing the testing and training sets, we also test of spatial autocorrelation. Where spatial autocorrelation exists, we retain a representative data point for each cluster of auto-correlated points.

We then built a separate model for each class of benthos that predicts the likelihood that class exists (Figure 1, step 6) in each pixel across the study area from 0 (no chance it exists) to 1 (100% certainty that it exists). Below is an example of this for sponges (Figure 4).

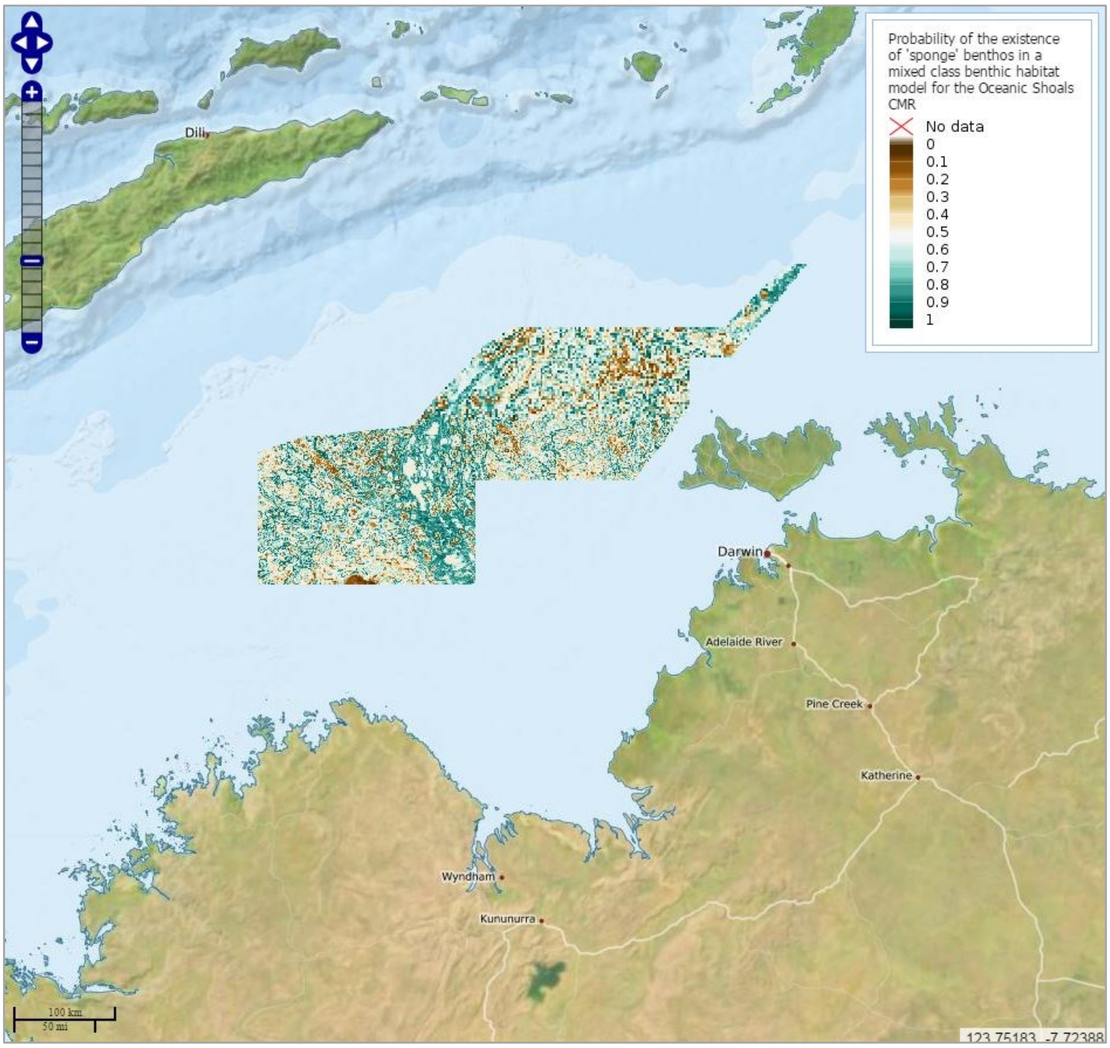

Figure 4. An example of a 'probability map' (Figure 1, step 6) of sponges for the Oceanic Shoals CMR at a 250 metre spatial resolution. Sponge is most likely to exist in pixels shaded dark turquoise, and least likely to exist in pixels shaded dark brown.

After each model is built with the training data, we use the testing data to assess its performance (based on the AUC – 'area under curve' parameter in ROC analysis – Faucet 2006). Models whose AUC values are less than 0.7 are discarded, as is standard. Table 1 shows the benthic habitat classes for which this standard was met for the Oceanic Shoals CMR model:

Table 1: Habitat classes successfully modelled for the Oceanic Shoals CMR.

Habitat class value	Habitat class name	AUC value
0	Indeterminant	0.980327908
1	Alcyons	0.997724852
2	Burrowers and crinoids	0.980157023
3	Filterer dominated (filterers and hard corals)	0.987782706
4	Gorgonians	0.991297317
5	Halimeda	0.976685003
6	Hard coral dominated (hard corals; halimeda; filterers)	0.99732923
7	Macro-algae dominated (macro-algae; filterers)	0.986979452
8	Soft coral dominated (soft corals; halimeda, sponges)	0.990430216
9	Seagrasses	0.9670035
10	Sponges	0.966650245
11	Whips	0.988385689

References

Andrefouet, S. et al (2003), 'Multi-site evaluation of IKONOS data for classification of tropical coral reef environments', Remote Sensing of Environment, 88, 128-143;

Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). *Classification and regression trees*. CRC press.

Breiman, L. (2001). Random forests. *Machine learning*, *45*(1), 5-32.

Brown, C. J., Smith, S. J., Lawton, P., & Anderson, J. T. (2011). Benthic habitat mapping: A review of progress towards improved understanding of the spatial ecology of the seafloor using acoustic techniques. *Estuarine, Coastal and Shelf Science*, *92*(3), 502-520.

Cutler, D. R., Edwards, T. C., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J., & Lawler, J. J. (2007). Random forests for classification in ecology. *Ecology*, *88*(11), 2783-2792.

Fawcett, T. (2006). An introduction to ROC analysis. Pattern Recognition Letters, 27, 861-374.

Gray, J. S. (1997). "Marine biodiversity: patterns, threats and conservation needs." Biodiversity and Conservation **6**: 153-175.

Gray, J. S. (2001). "Antarctic marine benthic biodiversity in a world-wide latitudinal context." Polar Biology **24**(9): 1432-2056.

Holmes, K. W., et al. (2008). "Modelling distribution of marine benthos from hydroacoustics and underwater video." Continental Shelf Research **28**(14): 1800-1810.

Lehmann, A., et al. (2002). "Regression models for spatial prediction: their role for biodiversity and conservation." Biodiversity and Conservation **11**: 2085-2092.

Mumby, P. J., et al. (1997). "Coral reef habitat mapping: how much detail can remote sensing provide?" Marine Biology **130**: 193-202.

Mumby, P. J., et al. (1999). "The cost-effectiveness of remote sensing for tropical coastal resources assessment and management." Journal of Environmental Management **3**: 157-166.

Mumby, P. J. and A. R. Harborne (1999). "Development of a systematic classification scheme of marine habitats to facilitate regional management and mapping of Caribbean coral reefs." Biological Conservation **88**(2): 155-163.