Identifying Fish Habitats: the use of spatially explicit habitat modeling and prediction in marine research

Katrina Baxter and Dr Mark Shortis

Department of Geomatics
University of Melbourne, Victoria, Australia
Phone: +61 3 8344 6806   Fax: +61 3 9347 2916
Email: k.baxter1@pgrad.unimelb.edu.au

Presented at SIRC 2002 – The 14th Annual Colloquium of the Spatial Information Research Centre
University of Otago, Dunedin, New Zealand
December 3-5th 2002

ABSTRACT

New methods of optimally identifying and predicting marine habitat occurrence are needed to help best address management issues such as marine reserve designation, fisheries stock assessment and aquaculture planning across large areas. A combination of video sampling, acoustic remote sensing and learning-based classification methods are proposed as a means of optimally identifying marine habitats. More commonly used in the identification of terrestrial landscape patterns, learning based classifiers, such as decision trees and artificial neural networks have the advantage of being able to accurately and rapidly identify patterns in complex systems. Opportunities exist to use similar techniques to both classify and predict the distribution of marine habitats. Both decision trees and neural networks are assessed as a means of deciding upon the most appropriate method for developing a spatially explicit marine habitat model, with reference to a case study within the Recherché Archipelago, Western Australia. Classifying habitats, predicting habitats and also understanding the importance of particular variables that define habitats are feasible outcomes through the adoption of a learning-based classification approach.

Keywords and phrases**:** marine habitats, spatial analysis, acoustics, decision trees, neural networks

1.0 INTRODUCTION

Marine researchers and managers are often required to understand patterns and processes operating across large areas and at a variety of scales to best address management issues such as marine reserve designation, fisheries stock assessment and aquaculture planning. Depending on the management question the area of interest may vary from the scale of an individual reef or habitat to very large geographic areas with significant variation in marine habitat types. Often presented with the difficultly that data a) do not exist for the areas of interest or b) are scarce and expensive to collect, researchers and managers require optimal mapping techniques to intelligently, accurately and cost-effectively survey and classify large areas of marine habitat.

Australia's Ocean Policy recommends the use of acoustic remote sensing devices along with visual methods, such as video, as a means of broadly sampling large areas of the marine environment (National Oceans Office, 1998). Recent advances in high-resolution acoustic sensors, such as side scan sonar and multibeam swath mappers, hold great promise for the development of optimal habitat mapping methodologies, enabling broad scale imagery of seafloor texture and depth to be acquired in both shallow and deep water habitats (Hughes-Clarke, 1998; Kvitek *et al*, 1999; Mayer *et al,* 2000; Kloser *et al*, 2001a; Kloser *et al*, 2001b; Ferns and Hough, 2002). It is expected that this 3 dimensional (or rather 2.5d) data of the seafloor will significantly improve our ability to describe marine habitats, but as with any new technology underpinning research is needed to develop appropriate methods of interpreting and classifying the data (Kloser *et al*, 2001b). In this sense, many lessons can be learnt and applied from the classification of terrestrial landscapes, which has been the focus of a number of geographic disciplines, including remote sensing, physical geography, landscape ecology and spatial analysis. Research directions in these fields have shown a shift from classification of landscape features using single variables alone (e.g. pixel value) to those classification algorithms that allow the use of multiple variables to describe and also predict pattern.

Learning based classification approaches such as Artificial Neural Networks (ANN's) and Decision Tree classifiers are known for their abilities to learn and also generalise patterns in complex non-linear datasets using multiple variables. Such techniques have been increasingly applied within terrestrial remote sensing, yielding significant improvements in accuracy over traditional regression and maximum likelihood techniques (Benediktsson *et al*, 1990, Foody and Arora, 1997). More recently, learning based approaches have been applied in ecological modelling to classify and also predict the occurrence of environmental patterns. A key focus of this paper is to review the use of learning based classifiers to quantitatively classify landscape pattern and illustrate the potential for similar methods to be applied to the marine environment.

Future research directions in marine habitat mapping seek to not only improve the accuracy of habitat classification, but also develop the ability to *predict* the occurrence of habitats from physical and biological parameters that can be remotely sampled (Kvitek *et al*,. 1999; Kloser and Penrose, 2000; Kloser *et al,* 2001b). Learning based classification approaches offer a potential means of achieving these objectives and providing new insights into understanding the relationships contributing to those patterns, in not only 2 but also 3 dimensions.

2.0 SPATIAL APPROACHES TO MARINE HABITAT MAPPING

Mapping of marine habitats is often undertaken by research and management agencies for a variety of purposes including:

- Assessment of habitat change due to natural or human impacts (e.g. climate change, oil spills, trawl disturbance)
- Monitoring and protecting important habitats (e.g. marine reserves, spawning areas, harvest closure areas)
- Design and location of marine reserves and aquaculture projects
- Species distributions and stock assessment

(Kvitek *et al*, 1999)

There are many parameters that can contribute to the distribution and abundance of species and habitats including: depth, exposure, substrate type, surface roughness, relief, sediment type, temperature, current and the presence of other species. Methods used to either directly or remotely sample these habitat parameters can include: diver transects; diver video; stereo video; 'drop' or towed video; aerial photography, satellite and acoustic imagery, such as sidescan sonar and multibeam swath mappers. The goal of the habitat mapping exercise will dictate the appropriate resolution required and the corresponding methods to be used, although increasingly large area mapping is best undertaken using a combination of remotely sensed data and direct observations, as a means of validating habitats and reducing overall costs associated with direct sampling methods such as diving (Kvitek *et al*, 1999; Kracker, 1999; Kloser *et al*, 2001a).

Initially, satellite methods were used to broadly quantify reef habitats, using Landsat TM and SPOT imagery (for a review see Green *et al*, 1996 and Mumby, 2000). However, the coarse pixel resolution ($30m^2$ for Landsat TM, $10m^2$ for SPOT) and subsequent lack of habitat detail have prevented satellite methods from being widely used by management agencies. Aerial photography has been one of the most common methods of delineating marine habitats across large areas. Ground truth data collected by diver or towed video has been used to classify aerial imagery of these shallow water habitats to quantify biodiversity (Mumby *et al*, 1998), detect reef health (Thamrongnawasawat and Catt, 1994; Baxter, 1998) and changes in the extent of seagrass beds (Mulhearn, 2001; Kendrick *et al*, 2002a; Kendrick *et al*, 2002b). Classified habitats can then be incorporated into a Geographic Information System (GIS) and used by management agencies for reserve planning, biodiversity assessment and monitoring (Sotheran *et al*, 1997; Kvitek *et al*, 1999; Ferns and Hough, 2000; Donoghue & Mironet, 2002). The finer resolution of data acquired from aerial photography (0.5-10metres) represents a significant advantage over satellite imagers and is a useful method of habitat mapping in shallow water environments. However, due to the limits of light penetration, only habitats up to maximum depths of 15 to 20 metres can be delineated.

Advances in acoustic remote sensing have helped overcome the limitation of light penetration. Technologies, such as sidescan sonar and multibeam swath mapping, have delivered opportunities to remotely sample seafloor bathymetry and texture in waters as deep as 200 metres (Hughes-Clarke *et al*, 1996; Blondel and Murton, 1997; Mayer *et al*, 2000; Kloser *et al*, 2001a; Brown *et al*, 2002). Acoustic sensors have been used for broad scale purposes such as pipeline planning across the seafloor (Mayer et al**,** 2000), fisheries stock assessment (Kloser *et al*, 1996; Diachok et al, 2001) and the detection of mid ocean ridges (Wright, 1999). More recently, research has focussed on detection of seafloor habitat characteristics at depths and scales appropriate to management of marine reserves and the assessment of human impacts (McRea *et al*, 1999; Brown *et al*, 2002; Ferns and Hough, 2002). Investigations into the detection of seafloor targets using multibeam technology illustrate features less than $1m^2$ may now be discerned (Hughes-Clarke, 1998; Kenny et al, 2000; Galway, 2000). It is expected improved classification of seafloor texture will result, allowing large areas to be classified and features of interest to be accurately identified. Direct sampling locations can be intelligently chosen from images of seafloor texture, ultimately resulting in significant time and cost savings. Similarly, such high-resolution depth data will be a welcome addition to the data needs of marine researchers, enabling terrain models of the seafloor to be developed that will allow powerful visualisation of the seafloor and derivation of additional factors such as slope, relief and surface roughness.

Together with direct sampling techniques such as video, acoustic devices provide a very powerful tool-kit for optimal habitat mapping and biodiversity assessment (Kloser *et al*, 2001b). Although multibeam swath mapping tools promise markedly superior improvements in habitat mapping compared to other acoustic methods, much work still needs to be done to improve the accuracy and interpretation of the data collected (Kloser *et al*,. 2001b). For this reason, the technology should not be widely lauded as a panacea. The future of habitat mapping should focus on utilising these new technologies as tools for obtaining data as inputs to broader, spatially explicit classification and modelling approaches. Access to multibeam technology is currently expensive and it is unlikely the immediate benefits will be widely available to state management agencies, who at best have meagre budgets for habitat mapping. Concurrent research investigating methods that will benefit not just acoustic image classification but habitat mapping as a whole are needed.

As marine habitat mapping is still in its infancy, opportunities exist to draw from the experiences of other fields of research. Significant parallels exist between marine and terrestrial landscape classification techniques (Kracker, 1999). In this sense, there is rationale in reviewing the development of landscape classification and modelling methodologies that have successfully identified techniques for classifying complex systems. Optimal marine habitat mapping will benefit from the development and application of methodologies that utilise multiple variables and allow not only classification but also the spatial prediction of habitat characteristics.

3.0 LEARNING BASED CLASSIFICATION APPROACHES – POTENTIAL APPLICATION IN THE MARINE ENVIRONMENT

Learning based classification methods, such as neural networks and decision trees, provide an attractive approach to modelling ecological systems, due to their ability to learn and predict patterns of a non-parametric nature (Dowla and Rogers, 1995; German *et al*, 1997; Özesmi and Özesmi, 1999; Drumm *et al*, 2000). Multiple variables can be included in the analysis, using either an unsupervised or supervised approach. As new example or training data are added, output feature classes are adjusted to account for the occurrence of different patterns or groupings of features (White, 1989; German *et al*, 1997; Dzeroski, 2001). This improved understanding of what makes one pattern structurally different from another can then be used to predict the occurrence of a feature class or clusters of classes given input variables from new observations or test data. In this sense, learning based classification approaches have a number of advantages over traditional data classification techniques, particularly those that rely on a single data source such as remotely sensed images.

This ability to learn and 'generalise' makes neural networks and decision trees highly suited to classifying complex environments where there are likely to be many variables contributing to the formation of pattern (Dowla and Rogers, 1995; Dzeroski, 2001). However, differences exist in the way decision trees and ANN's classify patterns and express relationships between variables defining the pattern. The advantages and limitations of both decision trees and ANN's are reviewed below with respect to applications in terrestrial classification, remote sensing and ecological modelling. Opportunities to apply similar learning based classification approaches within marine research to identify habitat classes and the variables that best define them are outlined.

3.1 Artificial Neural Networks

Neural networks have been used extensively in the field of remote sensing, with accuracies achieved often being higher than those achieved by traditional data classification techniques, such as maximum likelihood methods (Bendiktsson *et al*, 1990; Lek *et al*, 1996; Foody and Arora, 1997, Evans, 1998; Berberoglu *et al*, 2000). Like traditional classification techniques, some neural networks require training data to guide the classification process, others can adopt an unsupervised approach that requires no coaching or guidance of what defines a particular class. ANN's *learn* patterns and trends in the data, adjusting weightings between input variables and the relationships they form. This is significant in that the relationships amongst the data are always being evaluated and the modeling algorithm updated after each iteration (White, 1989). In this respect ANN's are time saving, especially when classifying complex systems whereby the types of information needed to separate classes are likely to be numerous. Atkinson and Tatnall (1997) cite one of the advantages of using neural networks with remote sensing data is the ability to be able to "incorporate different types of data into the analysis" although initially very few studies utilised remote sensing data with inputs other than spectral data (although see Fitzgerald and Lees, 1992). Increasingly, digital elevation models and their derived outputs, such as slope, relief and aspect have been used with remote sensing data as inputs to neural network models to predict terrestrial landscape features such as soil erosion, salinity and vegetation cover type (Ellis, 1997; Evans, 1998; Blackard & Dean, 1999). More recently, a number of ecological studies have used a mix of physical and biological descriptors (including satellite remote sensing data) to understand associations between: macroinvertebrates and water quality (Chon *et al*, 1996); trout density and stream habitat variables (Lek *et al*, 1996); and water quality parameters and algal blooms (Recknagel *et al*, 1997; Karul *et al*, 2000).

ANN's have been applied to estimate and *predict* the spatial distributions of certain species or habitats as functions of environmental parameters. Recknagel *et al*,. (1997) used ANN's with input variables of nutrient levels, light and temperatures, depth and water retention time to successfully predict the locations of algal blooms in four freshwater systems. GIS data of reef habitat at the Cook Islands was included within a neural network to predict sea cucumber habitat preferences (Drumm *et al*,. 1999). The presence of rubble was found to be the most important factor influencing the occurrence and density of sea cucumbers (Drumm *et al*,. 1999).

Spatial models for habitat selection have been developed using neural networks to predict the probable nesting locations of blackbirds and marsh wrens (Özesmi and Özesmi, 1999) and also to predict the suitability of coastline segments for colonisation and breeding by New Zealand fur seals (Bradshaw *et al*, 2002). Pup condition (over a 3 year period) was used as a surrogate indicator of relative prey availability in surrounding waters. Similarly, the proximity of breeding sites to deep bathymetry offshore was used as a factor indicating the degree of upwelling and hence food availability. When predicted breeding ground preferences were compared to actual breeding colonies, both coastal substrate type and food availability factors improved the classification and prediction accuracy, indicating the importance of the factors in influencing seal colonisation (Bradshaw *et al*, 2002). The ability to extract inferences or rules from ANN's to derive an understanding of the importance of different variables is discussed below in section 3.3.

3.2 Decision Trees

Decision tree classifiers are also able to extract information about the importance of particular variables in defining a particular pattern or relationship. Through an iterative process input data is split into 'branches' or sub-classes based on common attributes. These commonalities define *rules* by which splits in the classification are made, resulting in a tree-like decision structure. The decision rules of the tree can either reinforce current knowledge about relationships between variables in the system or define new understandings about how the observed pattern is formed (Dzerovski, 2001). Decision trees can be formulated using two general approaches. Input features can be assigned to different classes or subclasses according to decision rules defined either by (1) 'expert' knowledge or (2) machine learning techniques. Expert decision trees, as the name suggests, enable knowledge to be added to the model about how classes may be split, using some previous understanding of how the phenomena are to be classified (e.g. from literature, experiments, expert or local knowledge). Expert knowledge has not been used widely in ecology and natural resource management modelling (see Dzeroski *et al*, (1997) for a review of limitations).

Many of the common decision tree algorithms (eg. Classification and Regression Tree or CART® and ID3) allow decision rules to be induced directly from training data, through a process of machine learning (Quinlan, 1986; Lees and Ritman, 1991; Guissan and Zimmermann, 2000). Machine learning decision models have been developed for: the identification of dominant algal species and the prediction of algal blooms (Kompare and Dzeroski, 1995); predicting salinity risk (Evans *et al*, 1996); modeling areas at risk of soil erosion (Ellis, 1997); predicting antelope habitat (Bell, 1999); classifying water quality indicators (Dzeroski *et al*, 1997); mapping wetland habitat (Huang and Jensen, 1997); determining relationships between soil habitat characteristics and insect populations (Kampichler *et al*, 2000); distinguishing forested and non-forested areas to develop habitat suitability models for brown bears (Kobler & Adamic, 2001); and identifying the influence of deer population size on vegetation quality (Debeljak *et al*, 2001). Further use of machine learning within decision tree classifiers is reviewed in Guissan and Zimmermann (2000).

Decision trees that develop rule based classifications using machine learning are often likened to ANN's that identify and learn patterns in the data using sample data to train on, although significant differences exist in the way the knowledge within each of the models is represented. A disadvantage of using traditional neural networks is that it is difficult to impose prior knowledge about relationships within the classification structure (Guissan and Zimmermann, 2000). In contrast, decision trees (including both expert and machine learning systems) include a knowledge base that contains the rules in the form of if –then statements that split data into information classes. In many instances in ecological modelling it can be highly beneficial to include some prior knowledge about the system under study. For example, if it were known that a particular marine habitat is found only on leeward sides of an island, between a certain depth range, it would be useful and time saving to input this information within the model to reduce mis-classifications and overall error.

3.3 Transparency and Rule Extraction

Decision tree based models have in the past offered a more transparent means of analyzing multi-dimensional datasets than traditional neural networks. Decision trees have provided a relatively easy to understand model that extracts rules as to how classes are separated, offering an improved understanding of the importance of particular variables in defining pattern (Breiman *et al*, 1984; Dzeroski *et al*, 1997; Kampichler *et al*, 2000; Guissan and Zimmermann, 2000; Dzeroski, 2001). In many instances, ANN's have been a black box, from which it is difficult to extract rules and subsequently inferences about the system (Caudill, 1991; Guissan and Zimmermann, 2000; Dzeroski, 2001). In recent years, efforts have been made to improve the transparency of ANN's, with the formation of rule extraction methods (Fu, 1994; Setiono and Liu, 1997; Purvis *et al*, 1997; Drumm *et al*, 1999). Ozesmi and Ozesmi (1999) successfully predicted the nesting locations of blackbirds and marsh wrens using a neural network 'perception' model designed to give an increased understanding of the relationships between the different input variables. Drumm *et al* (1999) generated rule sets from ANN's to successfully identify the influence of rubble and sand variables on the habitat preferences of sea cucumbers. Bradshaw *et al* (2002) generated extraction rules to provide some additional insight in understanding fur seal colonisation. The results of the network classification and rule extraction successfully corresponded to existing understanding of the system, however complex associations that were difficult to understand were also derived from rule extraction, suggesting additional information may assist in future understanding of the system.

In many instances, the relationships between factors that define a class are not always clear-cut but rather can be quite fuzzy, particular in environmental models. Fuzzy neural networks have been developed to analyse vague or uncertain relationships in a systematic fashion, incorporating human-like reasoning to derive decision rules (Purvis *et al*, 1997). Combining the attributes of decision trees, fuzzy logic and neural networks, fuzzy neural networks would appear ideally suited to understanding complex environmental systems. Applied previously to determine land use suitability using topography and climate variables (Purvis *et al*, 1997), fuzzy neural networks offer both inference and predictive qualities. Kampichler *et al* (2000) recommended that if the goal of a learning based classification approach is to predict the occurrence of pattern, ANN's are the are the proper tools for achieving that goal. If, however, the understanding of abundance and diversity patterns is also desired, then tree-based models may be more appropriate (Kampichler *et al*, 2000). Fuzzy neural networks offer a combined approach that warrants further comparison with both decision trees and ANN's. The successful application of learning based classification techniques in terrestrial and ecological modeling suggests similar techniques could be adopted to classify and predict marine habitats. A suggested approach is outlined below with reference to the Recherché Archipelago, Western Australia.

4.0 SPATIALLY EXPLICIT MARINE HABITAT MODELING AND PREDICTION

Developing new methods of classification and utilising innovative technologies to better characterise marine habitats is of key relevance to management agencies. Managers require knowledge of the distribution of habitats with verifiable accuracy and reliability. Learning based classifiers can associate measures of accuracy to different classification outcomes, allowing reasoned information to be provided to decision makers. Combined remote sensing and learning based classification approaches can offer a simple modeling environment to not only classify pattern across large areas but also provide reliable measures of classification accuracy from which predictions of habitat occurrence can be made. As marine data collection is inherently costly, predicting the occurrence of habitats allows resources to be spent in areas of importance. For instance, areas that can be distinguished with high accuracy will not need to be sampled, however fuzzy areas where habitats are not easily distinguished from one another may need to be surveyed in more detail. On this basis, the application of learning based classifiers within marine environments would appear highly suitable to both marine researchers and management agencies.

A learning based classification and prediction approach is proposed for delineating fish habitats within the Recherché Archipelago, Western Australia. Located on the south coast of Western Australia near the town of Esperance (Figure 1), the Archipelago is a chain of approximately 105 islands and 1500 islets extending over 470 km of coastline (230 km linear distance) (Lee & Bancroft, 2001). The Archipelago is valued for its marine resources and is an important habitat for numerous commercial fisheries, including Abalone, Pilchard, Shark and the Southern Rock Lobster. The area has also been identified as suitable for a number of aquaculture programs, including the rearing of Southern Blue Fin tuna. However, substantial gaps exist in our knowledge of the marine habitats of the Recherché Archipelago (Kendrick et al, 2002a). Bathymetric data from this area is poor with approximately 33% of the Recherché Archipelago region having inadequate, or no bathymetric information (Kendrick et al, 2002a). The oceanography of the region is yet to be studied in any great detail, either through field, analytical or numerical modelling methods (van Hazel *et al*, 2001). Regional habitat data has only been interpreted at coarse scales from Landsat TM data (Kirkman, 1997) and a broad scale benthic habitat survey conducted using towed underwater video (Fisheries WA & Everall, 1999). The towed video identified eight categories of sea bottom including: Dense seagrass; Medium seagrass; Sparse seagrass; Patchy seagrass; Bare sand; Flat platform or low profile reef; Heavy limestone reef; and Granite reef. Despite these general observations, Fisheries WA and Everall (1999) note that the video record of the surveys contains much more information that could be analysed at a more detailed level to accurately define habitats.

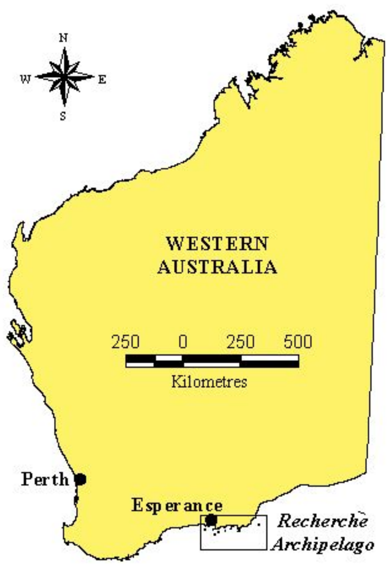

Figure 1: Location of the Recherché Archipelago

This current lack of knowledge presents significant difficulty for both state and local agencies required to make informed planning decisions about the use of marine areas in the Archipelago. The Recherché is a priority area identified by the West Australian Marine Parks and Reserves Authority to undergo the process of marine reserve planning (CALM, 1994). It is expected that this will occur in 2004-2005 given the extensive commercial fisheries within the Archipelago and the popularity of the region for tourism and recreational fishing. A research program, funded by the Fisheries Research Development Corporation (FRDC), has begun to characterise fish habitats of the Archipelago. A primary objective of the overall project is to collect baseline data to enable government agencies and the community to make decisions about the use and management of marine areas and reserves.

The research intends to develop intuitive data collection and classification methods that allow optimal distinction and prediction of marine habitats across the Archipelago. Given the sheer size of the study area, a combination of video sampling, acoustic remote sensing and learning-based classification methods are proposed as a means of optimally identifying habitats. A 'drop camera' video survey has been undertaken to broadly characterise representative habitats of the region. Variables suspected of defining the habitat type, such as exposure, depth, relief, substrate and dominant species types have been characterised at 3000 spatially referenced locations during an eight-week survey (figure 2). A relational marine database management structure has been established within a GIS to act as a framework for streamlining classification of the data, and has been designed to be adaptable to either a GIS or a learning based classification environment. The architecture of both the GIS database and the marine classification system that it is based upon is such that it is simple enough to be integrated with other systems and applications (Huber and Schneider, 1999) yet practical and easily integrated with the existing state marine classification system (Bancroft, 2002).

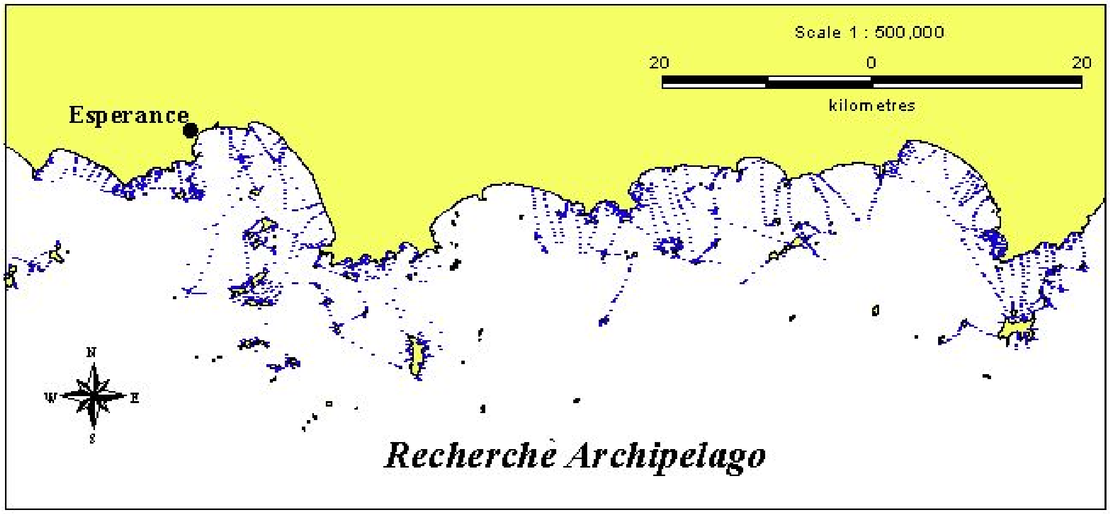

Figure 2: Drop camera video habitat survey locations, Recherché Archipelago.

Derived variables, such as distance to island, slope and surface roughness can also be included in the habitat model. Remotely sensed imagery, such as aerial photography, sidescan sonar and multibeam swath mapper, will be collected to broadly characterise the spatial extent of habitats, through either spectral or textural classification. Processed data can then be used as an additional input to the learning classifier to refine overall habitat classification. Validation of habitat classes, through further ground truth, will allow not only evaluation of the accuracy of the classification approach but will also provide a means for optimising further field sampling. Researchers can focus on discriminating or quantifying the diversity of habitats that are not classified as accurately as others or are of specific interest. From a practical viewpoint, this limits the need for extensive and expensive field surveys, allowing researchers to concentrate their efforts according to their research goals.

The predictive powers of learning based classifiers will also assist in optimising marine field research across large areas. Classifiers, such as neural networks, can be utilised to recognise relationships amongst input features (eg. depth, exposure) to predict the occurrence of output classes (e.g. habitat type). Through a process of weighting and error minimisation the ability of the model to separate habitats and predict habitat occurrence can be explained in terms of a probability or percentage error (Özesmi and Özesmi, 1999). The probability of a habitat being accurately classified can also be represented spatially (see Özesmi and Özesmi, 1999). Assignment of this degree of error or reliability is of key relevance to management agencies, particularly when faced with making difficult planning decisions.

Training learning based classifiers to predict the occurrence of a particular feature in the marine environment need not be limited to habitats. The potential of learning based classifiers to predict habitat occurrence based on relationships between variables contributing to pattern will also help the definition of surrogate habitats. Surrogates or 'indicator species' represent a suite of species commonly found in or associated with a similar habitat. Identifying surrogates offers a means of optimally characterising habitats and habitat diversity across large areas in a cost effective manner. Similarly, prediction of the habitat preferences of particular species can be derived using learning based classification methods. Within 300 of the 'drop camera' locations surveyed within the Recherché Archipelago, fish populations have also been surveyed, using stereo video. Stereo video techniques, developed to accurately and precisely determine the size and age class of marine fauna such as reef fish (see Harvey and Shortis, 1996; Harvey *et al*, 2001) will be used to obtain age-length characteristics of a range of species associated with a particular habitat type. The inclusion of such data in a learning based classifier can identify the habitat preferences of particular species. These fish habitat associations can then be used as a basis for reserve planning that protects not only habitat but also a variety of fish age classes.

5.0 CONCLUSION

Improving the accuracy and cost-effectiveness of mapping marine habitats across large areas requires a new approach. In order to reliably detect and also predict the occurrence of habitats future work needs to overcome the current limitations of traditional classification techniques. This paper proposes the use of decision trees, neural networks and fuzzy neural networks as a means of developing a spatially explicit habitat model to accurately delineate and predict marine habitats. Learning based classifiers and the inferences that may be produced from them can better assist in understanding the processes contributing to habitat formation and the habitat preferences of particular species. The development of simple rules by which habitat relationships can be easily understood will be of practical use to community and management agencies involved in reserve planning and marine management. Decision tree classifiers are ideally suited to this task, however algorithms are increasingly being developed to improve the transparency of neural networks (see Purvis *et al*, 1997; Drumm *et al*, 1999; Özesmi and Özesmi, 1999; Bradshaw *et al*, 2002). Improved understanding of the relationships within neural networks, combined with their predictive abilities makes them an attractive tool for habitat mapping and reserve planning in marine environments.

The success of any model is its interaction with real world problems and decision-making processes (Huber & Schneider, 1999). ANN's, decision trees and fuzzy neural networks will be evaluated within the Recherché Archipelago given the multiple objectives of the habitat mapping exercise. A learning based classification approach, evaluating each method, will allow habitats to be classified and predicted in a spatial explicit modeling environment. The importance of this work will be to provide an efficient means to extract rules from which the community can further understand the processes of habitat formation and researchers can evaluate optimal habitat mapping methodologies. It is hoped that advances in the application of marine remote sensing can be made whilst at the same time highlighting the application of learning based classification approaches in the marine environment (Wright and Goodchild, 1997; Kracker, 1999).

ACKNOWLEDGEMENTS

This research is supported by grant FRDC2001/060 "Characterising the fish habitats in the Recherché Archipelago, Western Australia", awarded by the Fisheries Research and Development Corporation over the period 2001-2004. The research program is jointly carried out by the University of Western Australia, the Australian Institute of Marine Science, the University of Melbourne, the Western Australia Museum and the people of the Esperance region.

REFERENCES

Atkinson, P.M. & A.R.L. Tatnall (1997) Introduction: neural networks in remote sensing. International Journal of Remote Sensing 18, pp699-709.

Bancroft, K.P. (2002). Developing a marine ecological community classification scheme. Report MCB-05/2002. Marine Conservation Branch, Department of Conservation and Land Management, 47 Henry St Fremantle, Western Australia, 6160. Unpublished report, January 2002.

Baxter, K. (1998). Monitoring Coral Reefs at a High Resolution: a Study of Pattern Using Aerial Photography and Image Processing Techniques, Great Barrier Reef, Australia. Unpublished Honours thesis, School of Tropical Environment Studies and Geography (TESAG)**,** James Cook University, Townsville, Australia.

Bell, J.F. (1999). Tree based methods. *In:* Fielding, A.H. (Ed.), Machine Learning Methods for Ecological Applications. Kluwer, pp. 89–105.

Berberoglu S., C.D. Lloyd, P.M. Atkinson & P.J. Curran (2000). The integration of spectral and textural information using neural networks for land cover mapping in the Mediterranean. *Computers & Geosciences* 26, pp385-396.

Benediktsson, J. A., P. H. Swain & O.K. Ersoy (1990). Neural network approaches versus statistical methods in classification of multisource remote sensing data. *IEEE transactions on Geoscience and Remote Sensing* 28(4), pp540-551.

Blackard, J.A. & D.J. Dean (1999). Comparative accuracies of artificial neural networks and discriminant analysis in predicting forest cover types from cartographic variables. *Computers and Electronics in Agriculture* 24, pp131-151.

Blondel, P. & B.J. Murton (1997). Handbook of seafloor sonar imagery. Chichester, England. Wiley-Praxis Publishing.

Bradshaw, C.J.A., L.S. Davis, M. Purvis, Q. Zhou & G.E. Benwell (2002). Using artificial neural networks to model the suitability of coastline for breeding by New Zealand fur seals *(Arctocephalus forsteri*). *Ecological Modelling* 148:111-131

Breiman, L., J.H. Friedman, R.A. Olshen & C.J. Stone (1984). Classification and Regression Trees. Chapman and Hall, New York.

Brown, C.J., K.M. Cooper, W.J. Meadows, D.S. Limpenny, & H.L. Rees (2002). Small-scale mapping of seabed assemblages in the Eastern English Channel using sidescan sonar and remote sampling techniques. *Estuarine, Coastal and Shelf Science* 54, pp263-278.

CALM (1994). A representative marine reserve system for Western Australia. Report of the marine parks and reserves selection working group. Marine Conservation Branch, Department of Conservation and Land Management, 47 Henry St Fremantle, Western Australia, 6160.

Caudill, M. (1991). Neural network training tips and techniques. *AI Expert,* January 1991, pp56–61.

Chon, T.S., Y.S. Park, K.H. Moon, & E.Y. Cha (1996). Patternising communities by using an artificial neural network. *Ecological Modelling* 90, pp69–78.

Diachok, B. Liorzou, and C. Scalabrin. (2001). Estimation of the number density of fish from resonance absorptivity and echo sounder data. *ICES Journal of Marine Science* 58(1), pp137-153

Debeljak, M., S. Dzeroski, K. Jerina, A. Kobler & M. Adamic¡ (2001). Habitat suitability modelling of red deer (*Cervus elaphus, L*.) in South-Central Slovenia with classification trees. *Ecological Modelling* 138(1–3), pp321–330.

Donoghue, D.N.M. & N. Mironnet (2002) Development of an integrated geographical information system prototype for coastal habitat monitoring. *Computers &Geosciences* 28, pp129 –141

Dowla, F.U. & L.L. Rogers (1995). Solving Problems in Environmental Engineering and Geosciences with Artificial Neural Networks. MIT Press, Cambridge, Massachusetts. pp1-14.

Drumm, D., M. Purvis & Q. Zhou (2000). Spatial ecology and artificial neural networks: modeling the habitat preference of the sea cucumber *(Holothuria leucospilota)* on Rarotonga, Cook Islands. The 11th Annual Colloquium of the Spatial Information Research Centre, University of Otago, Dunedin, New Zealand, December 13-15th 1999

Dzeroski, S., J. Grbovic, W.J. Walley, & B. Kompare (1997). Using machine learning techniques in the construction of models: II data analysis with rule induction. *Ecological Modelling* 95, pp95–111.

Dzeroski, S. (2001). Applications of symbolic machine learning to ecological modeling. *Ecological Modelling* 146, pp263–273

Ellis, F.G. (1997). Evaluating techniques for soil erosion modeling: a role for Artificial Intelligence? PhD thesis, Australian National University.

Evans, F. H., P. C. Caccetta, & R. Ferdowsian (1996). Integrating remotely sensed data with other spatial data sets to predict areas at risk from salinity. *Proceedings of the 8th Australasian Remote Sensing Conference*

Evans, F. (1998). An investigation into the use of maximum likelihood classifiers, decision trees, neural networks and conditional probabilistic networks for mapping and predicting salinity. MSc thesis. University of Western Australia.

Ferns. L.W. and D. Hough (Eds.) (2000). *Environmental Inventory of Victoria's Marine Ecosystems Stage 3 (2nd Edition) - Understanding biodiversity representativeness of Victoria's rocky reef*s. Parks, Flora and Fauna Division, Department of Natural Resources and Environment, East Melbourne, Australia.

Ferns, L.W. & D. Hough (Eds.) (2002). *High Resolution Marine Habitat Mapping of the Bunurong Coast (Victoria) – Including the Bunurong Marine and Coastal Par*k. Parks, Flora and Fauna Division, Department of Natural Resources and Environment, East Melbourne, Australia.

Fisheries WA & Everall Consulting Biologists (1999). Draft aquaculture plan for the Recherché Archipelago, WA: Benthic habitat survey of the Remark, Mart, Mondrain, Tory and York Island groups in the Recherché Archipelago.

Fitzgerald, R.W. & B.G. Lees (1992). The application of neural networks to the floristic classification of remote sensing and GIS data in complex terrain. In: American Society of Photogrammetry and Remote Sensing (Eds.), Proceedings of the XVII Congress ASPRS, Bethesda, MD, pp. 570–573.

Foody, G.M. and M.K. Arora (1997). An evaluation of some factors affecting the accuracy of classification by an artificial neural network. *International Journal of Remote Sensing* 18, pp799-810.

Fu, M. (1994) Rule Generation from Neural Networks. *IEEE Transactions on Systems, Machines and Cybernetics* 28(8), pp1114-1124.

Galway, S. (2000). Comparision of Target Detection Capabilities of the Reson Seabat 8101 and Reson Seabat 9001 Multibeam Sonars. Accessed 10th August 2002. http://www.omg.unb.ca/omg/papers/MBSS_TermPaper.pdf

German, G., M. Gahegan, & G. West (1997). Predictive assessment of neural network classifiers for applications in GIS. *Proceedings of the 2nd International Conference on Geocomputation*, pp41-50.

Green E.P., P.J. Mumby, A.J. Edwards & C.D. Clark (1996) A review of remote sensing for the assessment and management of tropical coastal resources. *Coastal Management* 24, pp1–40

Guissan, A. & N.E. Zimmermann (2000) Predictive habitat distribution models in ecology. *Ecological Modelling* 135, pp147–186

Harvey, E. & M. R. Shortis (1996). A system for stereo-video measurement of sub-tidal organisms. *Marine Technology Society Journal* 29(4), pp10-22.

Harvey, E., D. Fletcher & M.Shortis (2001). A comparison of the precision and accuracy of estimates of reef fish lengths determined visually by divers with estimates produced by a stereo-video system, Fishery Bulletin 99(1), pp63-71

Huang, X. & J.R. Jensen (1997) A machine-learning approach to automated knowledge-base building for remote sensing image analysis with GIS data. *Photogrammetry, Engineering and Remote Sensing* 63(10), pp1185–1194.

Huber, M. & D. Schneider (1999) Spatial data standards in view of models of space and the functions operating on them. *Computers and Geosciences* 25, pp25-38.

Hughes Clarke, J.E., L.A. Mayer, & D.E. Wells (1996) Shallow-water imaging multibeam sonars: A new tool for investigating seafloor processes in the coastal zone and on the continental shelf. *Marine Geophysical Research* 18, pp607-629.

Hughes-Clarke, J.E. (1998) Detecting small seabed targets using high frequency multibeam sonar. *Sea Technology*, June 1998

Kampichler, C., S. Dz¡eroski & R. Wieland (2000). The application of machine learning techniques to the analysis of soil ecological databases: relationships between habitat features and Collembola community characteristics. *Soil Biology & Biochemistry* 32, pp197–209.

Karul, C., S. Soyupak, A.F. Cilesiz, N. Akbay & E. Germen (2000) Case studies on the use of neural networks in eutrophication modeling. *Ecological Modelling* 134, pp145–152

Kendrick, G.A., E. Harvey, J. Hill, J.I. McDonald & S. Grove. (2002a) Characterising the fish habitats of the Recherché Archipelago: Review of existing information, Fisheries Research Development Corporation Project 2001/060. Report to the FRDC.

Kendrick, G.A., M.J. Aylward, B.J. Hegge, M.L. Cambridge, K. Hillman, A. Wyllie & D.A. Lord (2002b). Changes in seagrass coverage in Cockburn Sound, Western Australia between 1967 and 1999. *Aquatic Botany* 73, pp75-87.

Kenny A. J., E. Andrulewicz, H. Bokuniewicz, S. E. Boyd & J. Breslin (2000). An overview of seabed mapping technologies in the context of marine habitat classification. ICES ASC September 2000: Theme session on classification and mapping of marine habitats.

Kirkman, H. (1997). Mapping Australia's underwater features. CSIRO Division of Marine Research, Perth, Western Australia.

Kloser, R. J., J. A. Koslow & A. Williams (1996). Acoustic assessment of the biomass of a spawning aggregation of orange roughy (*Hoplostethus atlanticus*, Collett) off south-eastern Australia, 1990-93. *Marine and Freshwater Research*, 47 (8), pp1015-1024.

Kloser, R.J. *&* J.D. Penrose (2000). Optimal seabed habitat mapping using multibeam acoustics with associated physical and visual sampling devices – at sea trials. Australian Acoustical Society Conference, Joondalup, Australia. 15-17 November 2000.

Kloser, R.J., N.J. Bax, T. Ryan, A. Williams & B.A. Barker (2001a). Remote sensing of seabed types in the Australian South East Fishery; development and application of normal incident acoustic techniques and associated 'ground truthing ' *Marine & Freshwater Research* 52 , pp475 –89

Kloser, R.J., A. Williams, & A. Butler (2001b). Acoustic, biological and physical data for seabed characterisation. Final Report to the National Oceans Office. Project OP2000-SE02.

Kobler, A. & M. Adamic (2000). Identifying brown bear habitat by a combined GIS and machine learning method. *Ecological Modelling* 135, pp291–300

Kompare, B. & S. Dzeroski (1995). Getting more out of data: automated modelling of algal growth with machine learn-ing. In: Proc. International Conference on Coastal Ocean Space Utilization. University of Hawaii, Hawaii, pp. 209-220.

Kracker, L. (1999). The Geography of Fish: the use of remote sensing and spatial analysis tools in fisheries research. *The Professional Geographer* 51(3), pp440-450

Kvitek, R., P. Iampietro, E. Sandoval, M. Castleton, C. Bretz, T. Manouki, & A. Green (1999). Final Report - Early Implementation Of Nearshore Ecosystem Database Project. Report for California Department of Fish and Game Nearshore Ecosystem Database Project

Lee, S. & K.P. Bancroft (2001). Review of the existing ecological information for the proposed Recherché Archipelago marine conservation reserve. Literature review. MRI/WSA, EUC/SIN, RAR-51/2001 (Marine Conservation Branch, CALM).

Lees, B.G. & K. Ritman (1991). Decision-tree and rule-induction approach to integration of remotely sensed and GIS data in mapping vegetation in disturbed or hilly environment. *Environmental Management* 15, pp823–831.

Lek, S., M. Delacoste, P. Baran, I. Dimopoulos, J. Lauga & S. Aulagnier (1996). Application of neural networks to modelling nonlinear relationships in ecology. *Ecological Modelling* 90, pp39–52.

Mayer, L.. M. Paton, L. Gee, J. Gardner and C. Ware (2000). Interactive 3-D Visualization: A Tool for Seafloor Navigation, Exploration and Engineering. Accessed 14th April 2002. http://www.ccom.unh.edu/vislab/PDFs/3DGIS.pdf

McRea, J.E., H.G. Greene, V.M. O'Connell & W.W. Wakefield (1999). Mapping marine habitats with high resolution sonar. *Oceanologica Acta* 22(6), pp679-686.

Mulhearn, P.J. (2001). Mapping seabed vegetation with sidescan sonar. Defence Science and Technology Organisation, Report TN-0381, Sydney Maritime Operations Division.

Mumby, P.J. (2000). Remote sensing of tropical coastal resources: progress and fresh challenges for the new millennium. In: Sheppard C.R.C. (ed) Seas at the millennium, vol 3. Elsevier, London, pp 283–291

Mumby, P.J., E.P. Green, C.D. Clark, & A.J. Edwards (1998). Digital analysis of multispectral airborne imagery of coral reefs. *Coral Reefs* 17(1): pp59-69

National Oceans Office (1998). Australia's Ocean Policy. Report to the Commonwealth Government.

Özesmi, S.L., & U. Özesmi (1999) An artificial neural network approach to spatial habitat modelling with interspecific interaction. *Ecological Modelling* 116, pp15–31.

Purvis, M., Kasabov, N., Benwell, G., Zhou, Q., and Zhang, F. (1997). Neuro-fuzzy methods for Environmental Modelling. In: *Proceedings of the Second International Symposium on Environmental Software Systems.* Whistler, Canada. pp30 - 37

Quinlan, J.R. (1986). Induction of Decision Trees. *Machine Learning* (1), pp81-106.

Recknagel, F., M. French, P. Harkonen & K. Yabunaka (1997). Artificial neural network approach for modelling and prediction of algal blooms. *Ecological Modelling* 96, pp.11-28.

Setiono, R. and H. Liu (1997). NeuroLinear: From Neural Networks to Oblique Decision Rules. *Neurocomputing* 17(1), pp.1-25.

Sotheran I.S., R.L. Foster-Smith & J. Davies (1997). Mapping marine benthic habitats using image processing techniques with a raster based geographic information system. *Estuarine & Coastal Shelf Science* 44(suppl A), pp25–31

Thamrongnawasawat T., and Catt P., 1994, High resolution remote sensing of reef biology: the application of digitised air photography to coral mapping. *Proceedings of the 7th Australasian Remote Sensing Conference, Melbourne.* pp690–697.

Van Hazel, J.H., C. Pattiaratchi & N. D'Adamo (2001). Review of the climate and physical oceanography of the Recherche Archipelago and adjacent waters. A report prepared for the Department of Conservation and Land Management of Western Australia by Centre for Water Research, The University of Western Australia.

White H. (1989) Learning in Artificial Neural Networks: A statistical Perspective. *Neural Computation* 1(2), pp425-464.

Wright, D.J. (1999) Getting to the bottom of it: Tools, techniques, and discoveries of deep ocean geography. *The Professional Geographer* 51(3), pp 426-439.

Wright, D. J. and Goodchild, M. F. (1997) Data from the deep: implications for the GIS community. *International Journal of Geographical Information Systems* 11(6), pp523-528.