Using local maps of spatial accuracy for benthic habitat models

J. Miller,[1] M. Puotinen,*[2] and B. Radford[2]

[1]Department of Geography & the Environment, The University of Texas at Austin, Austin, Texas
[2]Australian Institute of Marine Science, Indian Ocean Marine Research Centre, Perth, Western Australia

*Email: M.Puotinen@aims.gov.au

Abstract

Spatial predictive models are used extensively in ecological applications as well as in many other domains. Their effectiveness for a given purpose depends on, among other things, a robust assessment of accuracy, particularly classification accuracy, which quantifies the difference between a prediction/classification and its "true" or observed value. In spite of their limitations, overall accuracy metrics such as percent correctly classified and kappa are still widely used, although more precise metrics that are able to differentiate between errors associated with location versus quantity have been introduced (Pontius and Millones 2011, Pontius and Santacruz 2014). In addition to overall accuracy measures, different aspects of classification accuracy can be measured relative to the observed values (e.g., errors of omission, producer's accuracy), the predicted values (e.g., errors of commission, user's accuracy), and class-specific accuracy. While a variety of accuracy metrics are available, most provide a single assessment value for an entire study area without taking into consideration local variations in spatial accuracy. Global measures of accuracy can be misleading when accuracy varies spatially and may be particularly problematic when the relative importance of predictive accuracy is different throughout the study area. Recent developments in accuracy assessment from remote sensing have illustrated the benefits of providing local measures of accuracy, but these have yet to be applied within the ecological domain. We use a case study of benthic habitat classes in a remote region offshore from NW Australia to demonstrate the value of calculating locally varying diagnostics of classification accuracy for an ecological model by comparing the variation in the resulting maps to the associated global values.

Keywords: spatial accuracy, geographically weighted, species distribution models.

1. Background

Accuracy assessment in thematic or categorical maps quantifies the agreement between 'observed' (i.e. ground truth, often considered to represent 'reality') and 'predicted' classes. The accuracy of such predictive models in general is typically assessed by 'training' a model with in situ (observed) data and 'testing' it with a held-out partition of data not used to build the model (Fielding and Bell 1997). Model predictions are tabulated against observed data in a table called a 'confusion matrix' which not only shows the level of agreement between the two by class, but also the most common types of misclassifications. Standard metrics are used to assess classification accuracy based on the level of agreement between observed (assumed to be correct) and modelled data.

While most spatial predictive models in the ecological literature present classification accuracy estimates for an entire study area as a whole (termed ‘global’), there has been more extensive research on spatial variation in model accuracy in remote sensing applications, such as how classification accuracy varies as a function of land-cover class (Congalton 1988), landscape characteristics (Smith et al 2002, Smith et al 2003, van Oort 2004), and topography (Yu et al 2008). These studies generally use some form of regression to associate these factors with model accuracy and therefore do not generate products analogous to maps of model accuracy. Other studies have focused on pixel level class probability measures and have been used to produce maps of classification quality or certainty/uncertainty (Brown et al 2009, Loosvelt et al 2012, Löw et al 2015, Bogart et al 2016). Probability of misclassification can be used as an indicator of confidence and is generally based on some kind of iterative model process based on “bagging” or bootstrapped aggregation (Steele et al. 1998). This methodology was extended recently to generate “spatially explicit confidence maps” (Mitchell et al. 2018).

As noted above, accuracy and uncertainty are not equivalent and a pixel may be classified accurately but with high uncertainty or inaccurately with low uncertainty (Foody 2005). Applications in which spatially explicit accuracy metrics have been used to produce maps of spatial variation in model accuracy are far less common. In one of the first examples, Foody (2005) used a 50 x 50 pixel constraint to estimate “local” confusion matrices, then interpolated the accuracy measures to generate maps of overall, user’s, and producer’s accuracy. He found that overall accuracy varied locally from 53% to 100%, while it varied for some classes from 0% to 100%. This method was extended to use geographically weighted regression to model spatial variation in model accuracy (Comber et al. 2012, Comber 2013). More recently, Comber et al (2017) provide a framework for calculating geographically weighted confusion matrices, from which local maps of any of the matrix-derived accuracy metrics can be generated.

While the studies above were all based on remote sensing images, here we extend the methods used in Comber et al (2017) to a dataset more commonly used in spatial ecological applications. We use a case study of a benthic habitat model for a remote marine benthic community located offshore from NW Australia (Camden Sound) to demonstrate that valuable ecological insights can be gained from mapping the spatial variation in classification accuracy metrics such as precision, recall and kappa across a study area.

2. Methods

We use an example of a spatial predictive model - in this case, a benthic habitat model - to demonstrate the value to ecologists of not only standard classification accuracy metrics, but also their local (or spatially varying) versions. Well established ecological theory (Elith et al., 2008, Elith et al., 2009, Holmes et al., 2006, Leathwick et al, 2008, Pittman et al, 2009) details how seafloor physical properties act as both direct and indirect drivers of landscape scale ecological processes on the benthos. Thus it is possible to identify and map the potential drivers of the spatial distribution of habitats (or related surrogates), and use field observations of the spatial distribution of those biota to predict where biota are likely located in areas not surveyed (Brown et al 2011, Holmes et al 2008). This is particularly important for remote marine communities that are both poorly known and expensive and difficult to survey.

We built a spatial predictive model fusing the following process (Figure 1) for Camden Sound - a remote and poorly studied region located offshore from NW Australia (Figure 2). In short, we used hydroacoustic survey data to develop predictor maps (step 2) based on fine-scale bathymetry and combined these with in situ data collected with towed video (step 3) to predict the probability of existence of a set of mixed benthic classes (step 5) using a Random Forest model (steps 6 and 7). We then mapped the most likely mixed class for each pixel using a maximum likelihood model (MaxEnt - step 8). We tested for spatial auto-correlation in the towed video data using local and global variograms but the semi-variance indicated virtually no spatial autocorrelation (the nugget/intersect was basically the same as the slope asymptote point). The priority was checking the test split but we checked both. The complete dataset and all the variograms were almost flat with very small amounts of semi-variance (over ~ 1% - 2% of total variance in the highest levels). We suspect the reason for this is that the ecological gradients are very gradual between these sites and the habitats in nearly all cases are inherently very patchy and diffuse – that is, they are very stochastic at the finer spatial scales at which you'd expect to find spatial autocorrelation in other ecological communities.

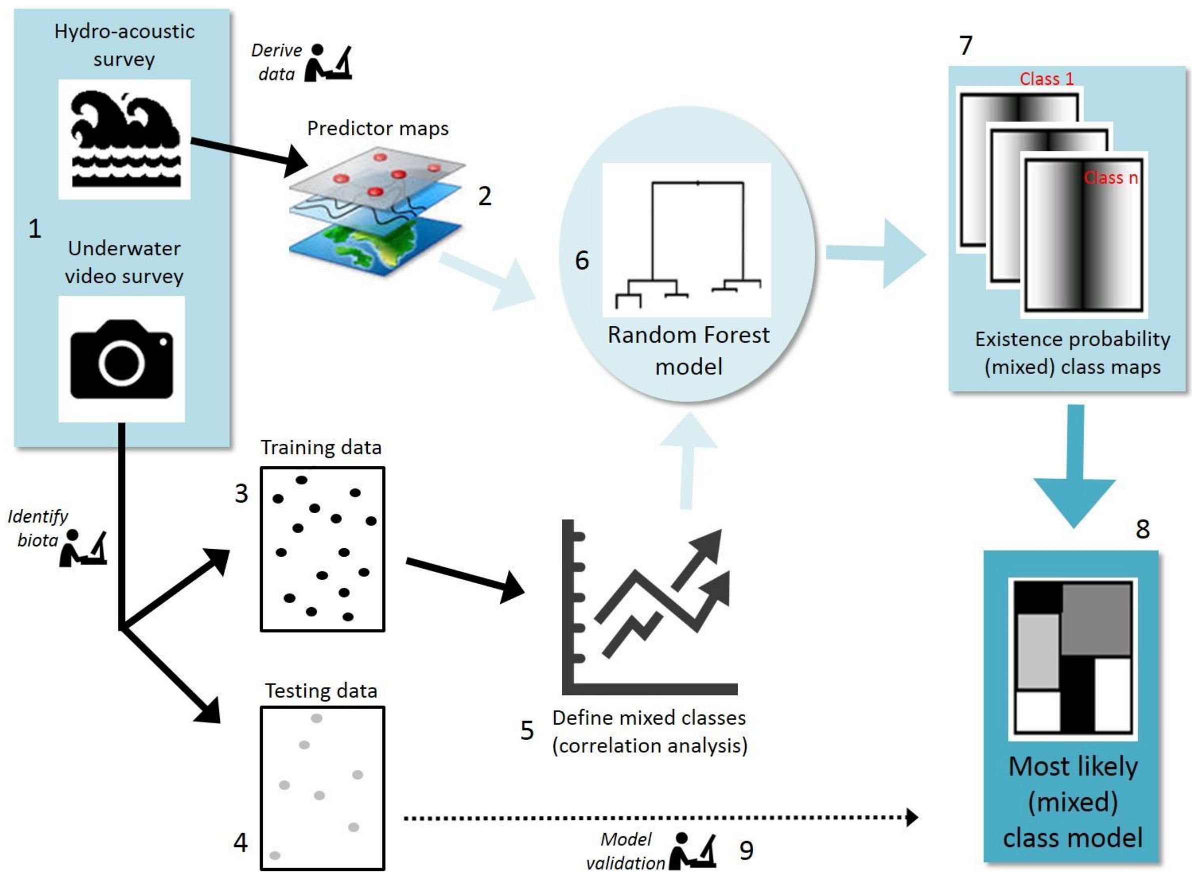


Figure 1: Work flow for building spatial predictive models of benthic communities from field data collected in Camden Sound, offshore from NW Australia (see Figure 2).

We held out one-third of the towed video data to use to calculate classification accuracy using a range of standard and metrics well established within either or both the remote sensing and data mining / signal processing literature, including: confusion matrix (Congalton 1991), Cohen's kappa (Cohen 1960), precision and recall (Fawcett 2006), and allocation difference and quantity difference (Pontius and Millones 2011). We calculated these metrics first for the study area as whole (global) for all classes, and separately by class. We then mapped the variation of each metric within each study area (Comber et al 2017; extending methods from Comber et al 2012, Comber 2013 and Tsutsumida & Comber 2015). In addition to more widely used accuracy metrics such as kappa and overall accuracy, we also use recently introduced accuracy metrics such as quantity, allocation, shift, and exchange (Pontius and Santacruz, 2014). We used geographically weighted crosstabs (Comber et al., 2016) with 5 km kernels to estimate local confusion matrices from which accuracy metrics were derived. Accuracy surfaces were interpolated for a variety of metrics (see Fig. 3).

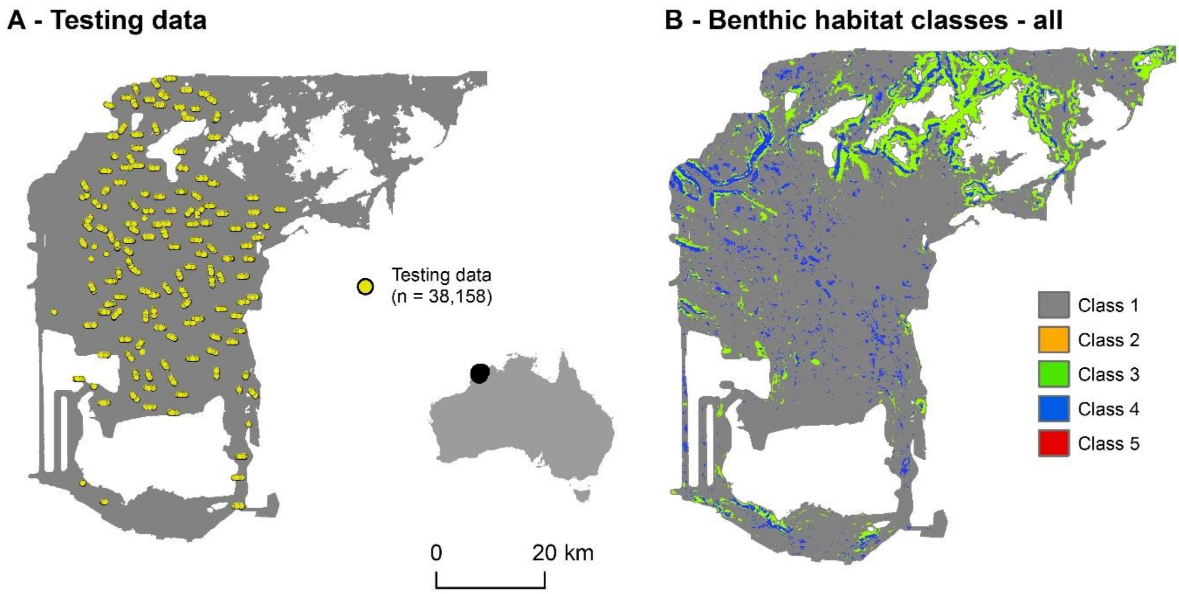
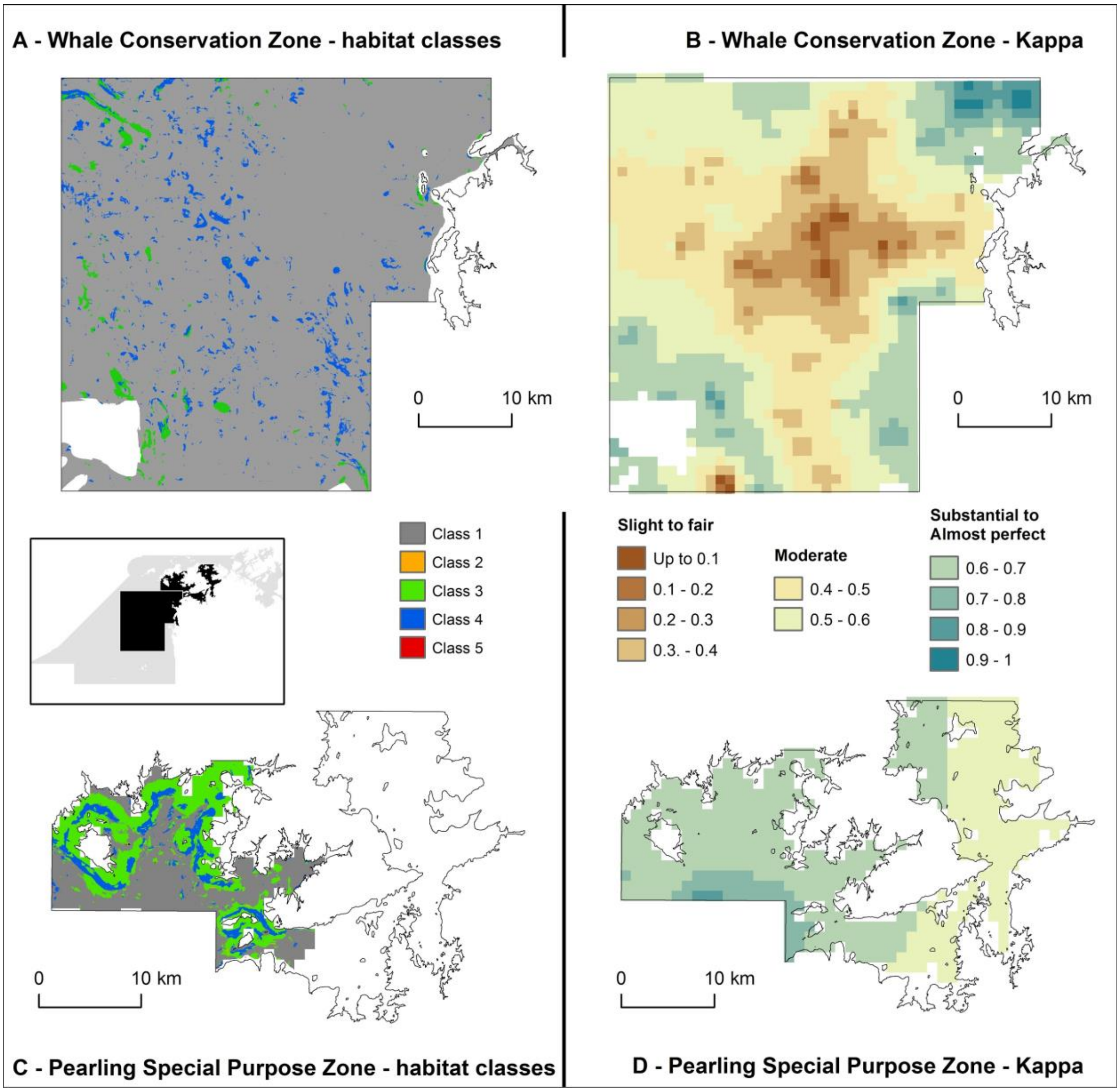


Figure 2: Spatial distribution of mixed benthic habitat classes predicted for part of Camden Sound, NW Australia and their classification accuracy. Panel A shows the study area extent (grey) and the location of validation data points (yellow dots, n=115,630). Panel B shows the predicted spatial distribution of five mixed benthic classes: 1 - No biota detected; 2 - Dense Gorgonian & Sponge; 3 - Medium Gorgonian & Sponge; 4 - Sparse Gorgonian, Sponge & Whips, 5 - Sparse to Medium Hard Coral & Bryozoans.

3. Results

Using local accuracy metrics, we are able to see where in the study area accuracy is low or high, as well as where different types of accuracy vary. We illustrate the use of these maps for management in Figure 3: a. shows a whale sanctuary that is predicted to contain scattered patches of medium Gorgonian & Sponge (class 3 - green) and sparse Gorgonian, sponge and whips (class 4 - blue) interspersed within a seascape largely dominated by bare ground (class 1 - grey). Because whales do not depend on benthic habitats for food, the widespread low classification accuracy (77.6% of pixels with classification accuracy below 0.61 - Figure 3B) is of minimal consequence for the utility of the zoning. In contrast, the



relatively widespread higher classification accuracy (only 7.5% of pixels with classification accuracy below 0.61) evident across the Pearling Special Purpose Zone (Figure 3D) is

Figure 3: Predicted benthic habitat classes and their estimated accuracy (as given by kappa) for two management zones of the Camden Sound Marine Park: a whale conservation zone (panels A and B) and a pearling special management zone (panels C and D). Predicted habitat classes are given in Panels A and C and estimated kappa is shown in Panels B and D.

important because pearl oyster abundance is correlated to benthic habitat type. Pearl oysters are much more likely to be found in areas of habitat classes 3 and 4 rather than 1 (Figure 3C), for example. The examples above demonstrate that knowledge of the spatial distribution of classification accuracy is valuable to managers that want to use habitat models for decision-making. Such maps can also be used as a basis for spatial prioritisation of future data collection when designing field surveys.

3. Acknowledgements

Data used in this study was collected with funding from the Western Australian Marine Science Association (WAMSI). JM was supported by funds from the Australian Institute of Marine Science Capability Development Fund.

4. References

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.

Cohen, J. 1960, A Coefficient of Agreement for Nominal Scales, *Educational and Psychological Measurement*, **20**, 37-46.

Comber, A., Brunsdon, C., Charlton, M. and Harris, P., 2017. Geographically weighted correspondence matrices for local error reporting and change analyses: mapping the spatial distribution of errors and change. *Remote Sensing Letters*, **8**(3), 234-243.

Comber, A. J. 2013. Geographically weighted methods for estimating local surfaces of overall, user and producer accuracies. *Remote Sensing Letters*, **4**(4), 373-380.

Comber, A., Fisher, P., Brunsdon, C., & Khmag, A. 2012. Spatial analysis of remote sensing image classification accuracy. *Remote Sensing of Environment*, **127**, 237-246.

Congalton, R. G. 1991. A review of assessing the accuracy of classifications of remotely sensed data. *Remote sensing of environment*, **37**(1), 35-46.

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.

Fielding, A. H., & Bell, J. F. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental conservation*, **24**(1), 38-49.

Foody, G.M., 2002. Status of land cover classification accuracy assessment. *Remote sensing of Environment*, **80**(1), 185-201.

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.

Mitchell, P.J., Downie, A.L. and Diesing, M., 2018. How good is my map? A tool for semi-automated thematic mapping and spatially explicit confidence assessment. *Environmental Modelling & Software*, **108**, 111-122.

Monk, J., Ierodiaconou, D., Bellgrove, A., Harvey, E., & Laurenson, L. 2011. Remotely sensed hydroacoustics and observation data for predicting fish habitat suitability. *Continental Shelf Research*, **31**(2), S17-S27.

Pontius R.G., Millones, M., 2011. Death to Kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment. *International Journal of Remote Sensing* **32**, 4407–4429. https://doi.org/10.1080/01431161.2011.552923

Pontius, R.G.P., Santacruz, A., 2014. Quantity, exchange, and shift components of difference in a square contingency table. *International Journal of Remote Sensing* **35**, 7543–7554. https://doi.org/10.1080/2150704X.2014.969814

Smith, J.H., Stehman, S.V., Wickham, J.D., Yang, L., 2003. Effects of landscape characteristics on land-cover class accuracy. *Remote Sensing of Environment* **84,** 342–349. https://doi.org/10.1016/S0034-4257(02)00126-8

Steele, B.M., Winne, J.C. and Redmond, R.L., 1998. Estimation and mapping of misclassification probabilities for thematic land cover maps. *Remote Sensing of Environment*, **66**(2), 92-202.

Stehman, S.V. and Wickham, J.D., 2011. Pixels, blocks of pixels, and polygons: Choosing a spatial unit for thematic accuracy assessment. *Remote Sensing of Environment*, **115**(12), 3044-3055.

Tsutsumida, N., & Comber, A. J. 2015. Measures of spatio-temporal accuracy for time series land cover data. *International Journal of Applied Earth Observation and Geoinformation*, **41,** 46-55.

Williams, A., F. Althaus, P.K. Dunstan, G.C.B. Poore, N.J. Bax, R.J. Kloser and F.R. McEnnulty. 2010. Scales of habitat heterogeneity and megabenthos biodiversity on an extensive Australian continental margin (100-1,000m depths). *Marine Ecology* **31**: 222-236.