Estuarine, Coastal and Shelf Science 176 (2016) 36–46

Contents lists available at ScienceDirect

Estuarine, Coastal and Shelf Science

journal homepage: www.elsevier.com/locate/ecss

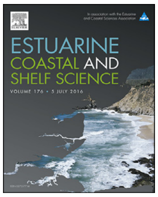

Image subsampling and point scoring approaches for large-scale marine benthic monitoring programs

Nicholas R. Perkins [a,*], Scott D. Foster [b], Nicole A. Hill [a], Neville S. Barrett [a]

[a] *Institute of Marine and Antarctic Studies, University of Tasmania, Australia*
[b] *Commonwealth Scientific and Industrial Research Organisation, Australia*

ARTICLE INFO

Article history:
Received 11 March 2015
Received in revised form
4 March 2016
Accepted 7 April 2016
Available online 9 April 2016

Keywords:
Autonomous underwater vehicle
Image selection
Point count
Precision
Bias
Survey design

ABSTRACT

Benthic imagery is an effective tool for quantitative description of ecologically and economically important benthic habitats and biota. The recent development of autonomous underwater vehicles (AUVs) allows surveying of spatial scales that were previously unfeasible. However, an AUV collects a large number of images, the scoring of which is time and labour intensive. There is a need to optimise the way that subsamples of imagery are chosen and scored to gain meaningful inferences for ecological monitoring studies. We examine the trade-off between the number of images selected within transects and the number of random points scored within images on the percent cover of target biota, the typical output of such monitoring programs. We also investigate the efficacy of various image selection approaches, such as systematic or random, on the bias and precision of cover estimates. We use simulated biotas that have varying size, abundance and distributional patterns. We find that a relatively small sampling effort is required to minimise bias. An increased precision for groups that are likely to be the focus of monitoring programs is best gained through increasing the number of images sampled rather than the number of points scored within images. For rare species, sampling using point count approaches is unlikely to provide sufficient precision, and alternative sampling approaches may need to be employed. The approach by which images are selected (simple random sampling, regularly spaced etc.) had no discernible effect on mean and variance estimates, regardless of the distributional pattern of biota. Field validation of our findings is provided through Monte Carlo resampling analysis of a previously scored benthic survey from temperate waters. We show that point count sampling approaches are capable of providing relatively precise cover estimates for candidate groups that are not overly rare. The amount of sampling required, in terms of both the number of images and number of points, varies with the abundance, size and distributional pattern of target biota. Therefore, we advocate either the incorporation of prior knowledge or the use of baseline surveys to establish key properties of intended target biota in the initial stages of monitoring programs.

1. Introduction

Increasing human impacts on marine ecosystems heighten the need for rapid and cost effective assessment and monitoring methods (Halpern et al., 2008; Brown et al., 2011). Benthic habitats play a vital ecological role and support fisheries with high economic value (Hughes et al., 2005). Quantification of benthic habitats and biota is necessary to better understand spatial patterns, monitor changes and assess the impacts of management strategies (Molloy et al., 2013). In this context, marine imagery collected by unmanned vehicles enables surveys of large areas, access to environments that are difficult to survey such as deeper waters, and creates a permanent record that allows comparison over time (Dumas et al., 2009). Photographic approaches yield high quality quantitative information on benthic communities including species presence-absence, direct counts of individuals or colonies, areal or percent cover estimates, and estimates of size (Hill and Wilkinson, 2004; Dumas et al., 2009; Trygonis and Sini, 2012). Recent technological advances have facilitated the routine collection of increasingly larger amounts of imagery, with some survey platforms able to capture several thousand images over a few hours of deployment (Pizarro et al., 2013). There is now a pressing need to assess the statistical robustness and statistical efficiency of competing

---

* Corresponding author. Institute of Marine and Antarctic Studies, Private Bag 49, University of Tasmania, Hobart, 7001, Australia.
E-mail address: Nicholas.Perkins@utas.edu.au (N.R. Perkins).

http://dx.doi.org/10.1016/j.ecss.2016.04.005
0272-7714/

sampling approaches for extracting ecological data from the large volumes of imagery produced.

One common use of marine imagery in large-scale ecological surveys is the initial inventory and ongoing monitoring of key groups or communities (Brown et al., 2004; Hill and Wilkinson, 2004; Smale et al., 2012). This typically involves the estimation of the areal coverage of “visually conspicuous” (0.01–1.0 m diameter) sessile benthic organisms or assemblages indicative of a particular habitat, and measuring differences in this coverage between regions or over time (Van Rein et al., 2009). A range of platforms have been utilised for this type of work, including diver-swum transects with hand-held cameras, the use of stills from towed camera systems, remotely operated vehicles (ROVs) and autonomous underwater vehicles (AUVs). Techniques using SCUBA divers typically employ multiple short transects (<50 m) and are restricted to shallower depths (Brown et al., 2004; Leujak and Ormond, 2007). Towed systems, ROVs and AUVs are capable of covering much larger areas, can operate at greater depths and return a large amount of data in a relatively short time. For these reasons, technologies such as these are more commonly employed when surveys are to cover meso-scales (Van Rein et al., 2009). For example, an AUV has been deployed as the platform of choice for the benthic surveying and monitoring of a large number of sites around the Australian coastline (Pizarro et al., 2013). Establishing the reliability of coverage estimates in such large-scale projects is of crucial importance from both scientific and economic standpoints.

The reliability of estimates is intrinsically linked to their precision (how variable one sample is from the next) and bias (the difference between the expected value of an estimator and the true value of the parameter being estimated). Results of surveys, and estimates of variability in particular, will not only be dependent upon the natural variability in the biota but also dependent upon: (i) the method used and effort spent on scoring each image, (ii) the number of images selected from each transect and their method of selection, (iii) the number of replicate transects used in an area, and (iv) the length and design of the transects (Houk and Van Woesik, 2006). Survey approaches which provide greater precision minimise the variation attributable to the sampling design and maximise the information about the biota. Issues surrounding survey and transect design are dealt with by Foster et al. (2014). Here we focus on the issues of subsampling imagery from transects and the intensity with which individual images are scored.

When scoring individual images, percent cover is the most common metric used to quantify benthic organisms (Van Rein et al., 2011; Deter et al., 2012). Images are usually selected along a transect or within sites and arguably the most common approach is the use of point count methodology, whereby the proportion of randomly overlain points intersecting an organism or substratum is used to calculate its coverage (Pielou, 1974). This approach has been facilitated by the development of dedicated point-count software, such as Coral Point Counts (CPC) (Kohler and Gill, 2006). The optimum number of points for scoring individual images (e.g. Dumas et al., 2009; Deter et al., 2012), and the interplay between the number of points per image and the number of images has been well studied (e.g. Brown et al., 2004). However, attributes of biological organisms, such as size, abundance and spatial pattern may also affect this trade-off, but have not been considered. Further, different image selection approaches (e.g. random or systematic sampling) within a transect will result in a different spatial spread of samples, which may affect efficiency (or bias and variance), particularly when the distribution of biota is clustered, yet a formal assessment of these potential affects is lacking. We aim to take a holistic approach by simultaneously examining the effect of both attributes of the biota and sampling approaches on the accuracy and precision of estimates of the precent cover of biota. Due to the likelihood of a strong interplay between these factors, an approach is required that allows an examination of the various trade-offs simultaneously. We do this through the use of Monte Carlo simulation and resampling approaches, which offer a flexible way to vary these parameters, whilst comparing the outputs to known percent covers, a quantity typically not known in real surveys (Bros and Cowell, 1987).

We report findings on the interplay between: (i) the number of images selected, (ii) the number of points used, and (iii) image selection approach on estimates of the cover of biota of varying spatial distribution and size classes. We use AUVs as a case study as they have a number of features that make them an ideal platform for benthic surveys in cross-shelf waters although other platforms share many common issues with AUVs. In order to “ground truth” the findings of our simulation we use Monte Carlo resampling techniques on actual scored AUV imagery to analyse the implications of reducing the number of points scored per image. Through analysis of simulated and actual benthic imagery we provide general recommendations for the level of subsampling and scoring that may be appropriate for scoring of marine benthic imagery.

2. Methods

2.1. Simulation methodology

As the basis for our simulation study, we used spatial scales of deployment and transect designs that have been used in AUV deployments under the government funded Integrated Marine Observing System (IMOS) monitoring program in Australian shelf waters (Barrett et al., 2010; Pizarro et al., 2013). All surveys for the IMOS program were conducted using the AUV Sirius, an AUV sampling platform designed by the Australian Centre for Field Robotics (ACFR) at the University of Sydney (http://www.acfr.usyd.edu.au/research/projects/subsea/auvSIRIUS.shtml).

Our simulation involved two major components: (i) the creation of biological distributions within a virtual seascape, and (ii) the sampling and subsampling of this distribution with various image and point scoring approaches along an AUV transect.

2.2. Simulating the distribution of biota

To create the biological distributions, a sample frame of 400 × 800 m, that encompasses the approximate survey area that has been used for AUV transects in the IMOS program, was established. Within the sample frame a biological distribution was created using either random or clustered spatial point patterns (see Diggle, 1983), with points assigned one of three size classes and abundances (Table 1).

Random biological distributions were simulated using a Poisson process. Clustered distributions of biota are commonly observed in ecological studies due to fine-scale biological and environmental factors such as dispersal and habitat structure (Dormann et al., 2007). Clustered distributions were simulated using a Neyman-Scott clustering process (see Diggle, 1983). The intensity of the processes were given values that resulted in percent covers that were similar to those of organisms of interest in previously scored data (see below). For the Neyman-Scott clustering processes, a homogenous Poisson process was used to establish the ‘parent points’. Each parent point spawned daughter points whose number follows a Poisson distribution. These daughter points were spatially distributed at a random angle and distance from the parent point. We examined two different clustered processes: (i) Mildly clustered: 10 daughter points, with the distance daughter points fell from the parent being drawn from a random normal distribution with a mean of zero and standard deviation of 5 and maximum