Designs for remote sampling: review, discussion, examples of sampling methods and layout and scaling issues

Karen W. Holmes[1,2], Kimberly Van Niel[2], Gary Kendrick[1] and Katrina Baxter[1]

[1]School of Plant Biology
[2]School of Earth and Geographical Sciences
University of Western Australia, 35 Stirling Highway, Crawley, WA 6009

June 2006

CRC for Coastal Zone Estuary & Waterway Management

Designs for remote sampling: Review, discussion, examples of sampling methods and layout and scaling issues

Written by:
Karen W. Holmes, Kimberly Van Niel, Gary Kendrick and Katrina Baxter

Published by the Cooperative Research Centre for Coastal Zone, Estuary and Waterway Management (Coastal CRC)

Indooroopilly Sciences Centre
80 Meiers Road
Indooroopilly Qld 4068
Australia

www.coastal.crc.org.au

National Library of Australia Cataloguing-in-Publication data

Designs for remote sampling: Review, discussion, examples of sampling methods and layout and scaling issues

QNRM06349

ISBN 1 921017 69 4 (print and online)

Contents

Introduction

Marine habitat mapping involves collecting and integrating many different types of data to build realistic models of both the seafloor and the organisms that live there. Data are gathered for determining the natural spatial and temporal heterogeneity of the benthic environment and organisms, evaluating the physical and biological controls on individual and joint habitat distributions, elucidating relationships among habitats and various species, and investigating habitat and organism response to human influences in the coastal zone.

Typical remote data collection for marine habitat research includes electro-optical and acoustic imaging, underwater videography, and still photography to characterise diverse habitats including reef, different geomorphic substrates, and seagrass species assemblages, in order to link structural habitat types with biogeochemical and biological observations. *In situ* sampling in marine environments is required to develop detailed site information and to process the remotely sensed data, yet field work can be time-consuming and expensive. Clearly, all of these datasets will vary in level of detail, scale and extent.

The variety of methods available for collecting information about the marine environment leads to rich datasets incorporating measurements and observations at a range of spatial and temporal scales. Precisely how to integrate these diverse datasets covering different spatial extents and with varying support has not been satisfactorily addressed in the marine mapping literature (Thrush *et al*., 1997).

We review remote sampling methods, standard sampling designs and the general statistical assumptions underlying those designs. Integration of these diverse regional datasets is desirable and necessary, but complicated due to differences in sample spacing, support and general scale issues, which have been addressed extensively in the ecological and geographical literature. We discuss a selection of scaling approaches as examples for modelling based on datasets of differing resolution, and suggest sources for more in depth information. We offer general recommendations to optimise sampling protocols for research projects that involve mixed data types in the coastal zone. Finally, the last two appendices provide examples of optimal sampling in the marine environments. Appendix D demonstrates optimal spatial sampling for increasing the accuracy of models and maps of seagrasses. Appendix E demonstrates optimal sampling design for predictive spatial modelling of marine benthos and substrate using ancillary data at Point Addis Marine National Park in Victoria.

Review of remote sampling methods

The choice of data collection methods depends on (1) the field area size and the level of detail necessary to meet the study objectives, (2) the heterogeneity and distribution of habitat types and (3) the project budget. The most systematic treatment of this topic in the literature is available for electro-optical sensors (Kvitek *et al*., 1999; Mumby *et al*., 1998; Mumby & Harborne, 1999), although there also exists a large number of more anecdotal descriptions of the application of acoustic (Brown *et al*., 2002; Cochrane & Lafferty, 2002; Kenny *et al*., 2003; Kvitek *et al*., 1999; McRea Jr. *et al*., 1999; Penrose & Siwabessy, 2001) and video (Bax *et al*., 1999; Brown *et al*., 2002; Kvitek *et al*., 1999; McDonald, 1998; Norris *et al*., 1997; Parry *et al*., 2003) techniques.

Following is a brief description of the benefits and drawbacks of the main forms of remote sensing for marine habitat research and a summary of relative costs, resolution, and expected positional accuracy (Table 1). It provides an overview of the relative expense and applicability (potential resolution, reliability) for nearshore habitat mapping over hundreds to thousands of square kilometres. The area within the study boundaries is referred to as the study *extent* and the pixel size or measurement unit size is the *grain*. The grain typically increases with increasing extent due to logistical reasons, and the resolution of the study is bounded by extent and grain, as the two extremes. The term 'scale' is used loosely in reference to the combined effect of extent and grain. Temporal scale will not be directly addressed, but is inherently linked to any discussion of spatial scale and issues of sampling over a range of extents (Wiens, 1989).

Electro-optical remote sensing

Satellite imagery provides continuous coverage over large spatial extents, and for some sensors, high temporal coverage as well. Generally it cannot be directly linked with the physical and biological processes inferred to cause the recorded patterns. The grain is typically fairly coarse for multiple channels (20–30 m pixels), which is most useful over very large areas. Higher resolution (smaller grain) data (e.g. IKONOS, with 4m pixels) is now available, but at a greater cost. Manned systems (e.g. SKYLAB and space shuttles) can also provide useful data, but lack temporal repeatability. Highly specialised software and training is necessary for most classification techniques. For marine habitat mapping, utility is limited to areas with deep light penetration, which is generally limited to a narrow swath along the coast (Kenny *et al*., 2003). Also, due to the attenuation of light in water, only the visible wavelengths are applicable for marine applications, with the highest transmission in the blue and green bands.

Airborne sensors are typically better than satellite imagery for coastal habitat assessment, as the areas where light can penetrate to the seafloor are close to the coastline, and require smaller grain to permit identification of boundaries and objects. It can get very expensive depending on the sensor and areal extent needed. Automated classification techniques require specialised software and training, and manual methods require specialist knowledge and skills and strict controls if multiple operators are involved. Satellite and airborne remote sensing for benthic habitats is limited by light attenuation in water both leaving and returning to the sensor. Airborne imagery scale is variable and is controlled by the focal length of the camera and the height of the plane above the target. Sites farther from nadir can be distorted.

Acoustic remote sensing

Sidescan sonar is very useful for mapping seafloor texture, particularly in areas where light penetration is poor. Considered relatively old technology, sidescan sonars are a main tool in seafloor mapping, producing almost photorealistic pictures that assist in visually delineating habitats and their spatial boundaries, as well as detecting patterns in seabed morphology (Kenny *et al*., 2003). Sidescan systems typically consist of two underwater transducers mounted on a sonar or towfish that is pulled behind the survey boat. The transducers emit an acoustic signal, either side of the survey track, which ensonifies a wide swath of the seabed (up to 200 m each side). For each sonar pulse or ping, the received signal is recorded to create a continuous backscatter intensity (textural) image of the seafloor along the swath.

The resolution obtained by sidescan varies depending on factors such as the depth of water, the depth of the towfish, the speed of travel and the width of the survey track. The effective width or range is also limited by the particular frequency of the sonar used. In general, the closer the sonar is to the seafloor the narrower the swath width. In shallow water high-resolution imagery can be obtained at the expense of the area covered, as the resolution is inversely proportional to the swath width. For example, an image with 20 cm resolution at a 50 m range drops to 40 cm at the 100 m range (Kvitek *et al*., 1999).

The quality of the imagery received can also be influenced by the degree of distortion in the outer portions of the sonar swath. Visual delineation of habitats is possible when combined with ground-truthing, but automated methods for classification have not been adequately developed. Distortion in the data increases from near nadir to the edges of the beam. It is not useful for bathymetry measurement (Kenny *et al*., 2003).

Multibeam sonar produces high accuracy bathymetric models but is very expensive to run and time-consuming to cover large areas. Although textural information is recorded (snippets), they are not in a format that can be readily accessed for textural habitat mapping as is the sidescan output (Reson Training class, Curtin University, Perth WA, March 2004). Multibeams acquire depth information across a swath using a collection of acoustic beams, as opposed to a single beam. The swath coverage varies, according to the depth of water and the type of sensor being used. Systems typically provide a swath that is either 3.5 or 7.5 times the water depth. Deeper water applications are more efficient, where large swaths of the seafloor can be mapped. Mapping in shallow water with multibeam can be a very costly exercise.

Multibeam systems can have very wide beam angles, but data from the outer beams are usually of questionable value due to the added effects of vessel roll. However, many of the stability and positioning issues encountered using sidescan sonar can be accounted for with inertial motion sensors and calibration tests that correct for movements in the boat (roll, pitch and yaw) and positioning error, in terms of the latency of the GPS signal. Although declining, purchase and data acquisition costs for multibeam surveys are several orders of magnitude higher than spectral techniques (Kvitek *et al*., 1999), prohibiting most management agencies from obtaining full coverage of actual habitat types and/or seafloor texture using these sensors alone.

Field remote sensing

Underwater videography is an economical method of collecting imagery of the seafloor for ground truthing or mapping. Typically this is only useful under high visibility conditions (Solan *et al*., 2003).

Underwater still photography is useful for documentation of SCUBA findings, and can be used for digital image analysis, depending on the objectives of the study. Typically this is only useful for small areas and for targeting specific species (Solan *et al*., 2003).

Table 1. Summary of relative costs, resolution and expected positional accuracy for the main forms of remote sensing for marine habitat research

Cost estimates are approximate. Expense varies greatly, depending on elements such as start-up costs, minimum purchase requirements and travel costs.

Remote sampling methods	Relative cost[1]	Area mapped (km^2/hr)[2]	Window size/width	Pixel size	Spatial accuracy	Min. recog. unit 2–5 pixels, reg. vs. irregular boundaries[3]	Comments/references
Satellite sensors							
Landsat	\$ A\$0.05/km^2	100+	185 x 185 km	30 m	Georef. 15 m (best)	1800–4500 m^2	(Mumby & Edwards, 2002) Best for large areas (>500 km^2)
SPOT	\$ A\$0.10/km^2	100+	60 x 60 km	20 m	Georef. ~10 m	800–2000 m^2	(Mumby & Edwards, 2002; Mumby *et al.*, 1998) Better than Landsat for <3600 km^2
IKONOS	\$\$ A\$45–69/km^2	100+	1 km^2, min. 100 km^2	4 m, 1 m pan-chromatic	Georef. ~ 3m	32–80 m^2	(Mumby & Edwards, 2002) Good for identifying boundaries, not better than Landsat for habitats
Airborne sensors							
Aerial photography	\$\$	10+	W = depends on survey	0.05–1 m	GPS, ground control	0.2–1 m^2	
CASI	\$\$\$	10+	W = 512 m	1–10 m, typically 1 x 2 m	GPS, 2 m	4–10 m^2	(Kvitek *et al*., 1999; Mumby & Edwards, 2002) Best for areas <100 km^2 Water <30 m deep
LIDAR	\$\$\$\$ US\$3100–3900/km^2	10+	W = 50 m 8–32 km^2 /hr	4 m	3 m horizontal (improved with DGPS) 0.15 m vertical	32–80 m^2	(Kvitek *et al*., 1999) <20 m water depth, to 60 m possible
LLS (Towed, not airborne)	\$\$\$\$\$?	10	W = 4–65 m (depends on water clarity)	0.1–3 cm	DGPS error	<1cm^2	(Kvitek *et al*., 1999; Solan *et al*., 2003) >3 m water, <1500 m Depends on field conditions

Table 1. Summary of relative costs, resolution and expected positional accuracy for the main forms of remote sensing for marine habitat research (continued)

Remote sampling methods	**Relative cost**[1]	**Area mapped (km^2/hr)**[2]	**Window size/width**	**Pixel size**	**Spatial accuracy**	**Min. recog. unit** 2–5 pixels, reg. vs. irregular boundaries[3]	**Comments/references**
Acoustic methods							
Sidescan sonar	$$	5	W = up to 300 m	0.5 m	GPS, layback	2–5 m^2	(Kvitek *et al*., 1999) Distortion across track >2 m water depth, gap under fish Easily distorted by waves, etc. Depends on frequency used and field conditions
Multibeam sonar	$$$$$	3.5	W = 2 x water depth	0.10 m	DGPS error	0.08–0.20 m^2	(Penrose & Siwabessy, 2001) (Kvitek *et al*., 1999) >3 m water depth Depends on equipment and field conditions
Field remote sensing*							
Videography	$$	0.2	0.5–25 m^2	Small	GPS, layback	Depends on FOV, small	(Kvitek *et al*., 1999) <30 m water depth Depends on field conditions
Still photography	$$	<0.01	0.5–25 m^2	Small	GPS	Depends on FOV, small	(McDonald, 1998) <30 m water depth Video less expensive than still, still can be difficult to rectify for accurate analysis

[1] $ = <$5 000; $$ = $5 000–25 000; $$$ = $25 000–75 000; $$$$ = $75 000–150 000; $$$$$ = >$150 000

[2] Modified from (Kenny *et al*., 2003)

[3] 2–5 pixel approximation (from O'Neill *et al*., 1999)

Review of sampling designs

Researchers working with data from a small field area are typically interested not only in understanding how that particular environmental system functions, but also in furthering our knowledge about much larger populations than those directly observed or measured within a single study. Exhaustive sampling of the larger population of interest is not feasible, nor desirable, particularly when sampling is destructive or extremely time-consuming and costly. Instead, a subset or sample of the larger 'true' population is assumed to be representative, and through inference can reveal something meaningful about that part of the larger population not included in the sample (Cochran & Cox, 1957). The validity of this assumption depends on how well the sample captures the spatial and temporal variability of the phenomenon of interest and the level of certainty required for the study. In this section, we explore issues of sampling design for marine habitats. We discuss both random and systematic sampling designs, methods for determining optimal sample spacing, and how sampling methods and outcomes can affect the process and outcome of modelling habitats.

The major issues in sample design or layout are (1) how to select the sample so that it presents an unbiased view of the true population, and (2) how to draw conclusions about the true population from the results of the sample (Cochran & Cox, 1957; Snedecor & Cochran, 1989). Statistical analyses are designed to quantify how well a sample represents the true population in probabilistic terms given several assumptions about the nature of the phenomenon and the sampling procedure. For example, a town council is interested in the overall water quality conditions in the local bay. If only areas around the local sewage treatment plant and the fertiliser factory are sampled, then the average values reported to the town council will be much higher than is true of the bay as a whole. Likewise if only pristine offshore areas are sampled, then the estimation values will be much lower than is truly representative.

In order to produce meaningful estimates for the average bay water quality, bias in sampling must be avoided. The commonly accepted way to avoid bias is to introduce an element of randomness into the sampling procedure, so that every sample has an equal chance of being included in the study. A systematic approach in which measurements are taken at regular intervals, independent of what the collector sees as interesting in the field also guarantees unbiased selection of sample locations. The origin for a systematic sampling plan ideally should be randomly selected.

The degree of randomness required for a study is often left to the discretion of the researcher (Cochran & Cox, 1957), but a major issue for environmental research is that samples taken close together are often more similar than those taken far apart. This spatial autocorrelation violates the basic assumption of sample independence for classical statistical techniques (Burrough & McDonnell, 1998), although options for use within classical statistics are discussed below.

The field of geostatistics was developed to take advantage of spatial autocorrelation to improve and estimation in unsampled areas by incorporating information from neighbouring measurements. Both classical statistics and geostatistics can be used for environmental research depending on the objectives (e.g. comparison of zonal means versus optimal interpolation), but assumptions concerning the independence of samples may affect the calculation of standard error and confidence intervals (Atkinson, 1997; Brus & de Gruijter, 1997), as spatial dependence among measurements effectively reduces the number of degrees of freedom (Legendre, 1993).

Geostatistics tend to be more forgiving of sampling design; however, many data are required for calculating a stable semivariogram [>150 samples (Webster & Oliver, 1992)], and interpolation quality is highest when samples are fairly evenly distributed across the area of interest (Atkinson, 1997; Flatman *et al*., 1988; Isaaks & Srivastava, 1989). Because habitat mapping is fundamentally spatial research, a geostatistical approach to sampling design is highly desirable for many types of analyses, although classical (non-spatial) statistical techniques are also frequently applied for physical modelling and data summaries. Geostatistics can also be applied to the output from nonlinear (generalised) regressive methods, which are often preferred for ecological data due to non-linear relationships (Austin, 2002), and non-parametric geostatistical techniques are frequently used for mapping non-Gaussian datasets including categorical data (Goovaerts, 1997).

‘Random’ versus ‘spatial’ sampling

The sampling designs used for classical and geostatistics are subject to different assumptions required for each approach. Classical random sampling (referred to as random sampling from here on) relies on prior information to select areas for sampling to determine the need for stratification and sample size as it relates to variance. Only the newly collected data is used for analysis while the prior information that drove the sampling process is omitted. The geostatistical data model is based on Random Variable Theory (Goovaerts, 1997), and sampling for geostatistical purposes (referred to as spatial sampling from here on) includes prior information to improve the sampling design, and fuses previously collected data with the new data in analyses. The geostatistical model provides a greater flexibility to incorporate prior information into sampling and statistical analysis (Borgman & Quimby, 1988).

Classical sampling theory assumes that the true population is fixed, and all variance associated with any analyses are a result of sampling design (Borgman & Quimby, 1988). Therefore, if the population were exhaustively sampled, we could calculate the theoretical true mean and standard deviation. There is nothing random about the true population. However, randomness is introduced in the sample design to ensure that every possible measurement location has an equal chance of being included in the sample. In essence, this is the same as taking all of the possible measurement locations, mixing them up, and picking out a sample of size *n*.

In reality randomness in sampling is achieved using random numbers to allocate plots or sampling locations. Many different sets of observations can be selected to represent the true population, or field area. The design itself determines the probability of selection for every possible sample, so conclusions drawn from the sample are considered valid because of the sample design construction, not by assumption, regardless of spatial variability across the study site (Brus & de Gruijter, 1997).

Spatial sampling, in contrast, is based on the idea that the variable under study is a stochastic process, and the reality that is sampled is one realisation of a random field (Borgman & Quimby, 1988; Brus & de Gruijter, 1997). If the same locations were sampled multiple times, multiple values would result, and could be assembled as probability distributions (Goovaerts, 1997). This is known as Regionalised Variable Theory, which further assumes that the spatial variation of any variable can be expressed as the sum of three major components: (1) a trend or constant mean; (2) a random but spatially correlated component (regionalised variable); and (3) spatially uncorrelated random noise, or residual error (Burrough & McDonnell, 1998). Spatial sampling can be defined as those sampling procedures that incorporate the assumption that the variable is stochastic, and rely on estimates of the covariance in previously collected data to drive sampling campaigns (Borgman & Quimby, 1988).

Both the random and spatial approaches can produce satisfactorily ‘independent’ samples for statistical analysis and spatial prediction. Random sampling has benefits in terms of producing strictly valid, unbiased sample data collection, which is sometimes required for legal or regulatory purposes. However, the lack of bias comes at a cost. Truly random surveys ignore all expert opinion in the sampling design, leading to much greater sampling effort, and requiring more samples than necessary in some areas and too few in others. The geostatistical approach rests on several assumptions which are difficult to prove (most notably second order stationarity—that covariance depends only on distance between data points) but offers much more flexibility in terms of sample distribution. This is often desirable as it can simplify field work logistics and permit spatial analysis and it encourages the incorporation of expert knowledge into the analysis process. The primary concern for the spatial approach is that sampling is adequate to estimate the covariance structure of the variable of interest. Common sampling layouts are discussed below, grouped into 'random' and 'systematic' methods, although systematic sampling can also be considered random, as has been discussed.

It is important to note that even 'random' sampling can lead to samples that are spatially autocorrelated, resulting in a well-known ecological problem called pseudoreplication (Allen & Starr, 1982; Levin, 1992). If an inferential approach is preferred to a geostatistical approach, then it is important to ensure that samples are spaced so that they are not spatially autocorrelated, or to remove the spatial autocorrelation from the dataset (called detrending) before proceeding. Another option is to include the level of spatial autocorrelation as an independent variable in the inferential methodology. This method, commonly referred to as autoregressive or autologistic modelling (Klute *et al*., 2002), includes a covariate which allows spatial autocorrelation to influence the prediction. However, in this case, the covariate must theoretically be replacing some known physical function (e.g. seed propagation method).

Random sampling designs

***Simple random sampling*: **Locations for sample collection are selected randomly, using a random number generator (see Van Niel & Laffan, 2003) or a table of random digits to ensure that every member of the population has an equal chance of being selected for the sample (Snedecor & Cochran, 1989). This method ensures allocation of sample locations is not biased, but it cannot incorporate any prior information about the field site, which may be needed to avoid under sampling important populations or excluding difficult to reach or inappropriate areas. The typical problem with simple random sampling is that rarer conditions and therefore habitats are often not sampled or undersampled, while common conditions and habitats are oversampled due to their greater spatial extent.

Stratified random sampling**:** Using expert knowledge, the field area is divided into subpopulations or strata that maximise the variation between units, and minimise the variation within each unit. The one or more strata selected are expected to be major drivers of the system under study, that is, exert some control on the habitat type. A random sample is then drawn from each stratum or unit. When known differences exist between the strata, stratified random sampling with proportionate allocation can sometimes provide improved estimation without introducing bias (Snedecor & Cochran, 1989). For example, a study area is stratified by bottom substrate type to sample for invertebrates, but the number of invertebrate species is known to be higher in reef areas than in fine sand. For stratified random sampling with proportionate allocation, more samples would be taken from reefs than the other categories to permit a more robust calculation of variance.

Multi-level sampling: This is another version of stratified random sampling where only a small number of units in each stratum are sampled. For example, the field area consists of 33 sand patches and 24 seagrass patches, but only five patches of each habitat type are selected for invertebrate surveys.

Cluster sampling: This involves taking a group of samples from a predetermined number of random locations. Clustering can be either spatial or temporal. However, because clustering can bias the overall estimates due to many samples in potentially high or low value locations, the data may need to be weighted before analysis to prevent estimation bias (declustering) (Goovaerts, 1997), or might have to be averaged to produce one value for each cluster location (Greenwood, 1996).

Systematic sampling designs

The various methods of systematic sampling are similar in that once the number and spacing of samples is determined, the distribution of the entire sample or a significant proportion of it is known. Generally the grid origin or the starting point of a systematic sample is drawn randomly. This method has two advantages over random sampling: it is easier to design, since only one random number needs to be chosen, and it guarantees that the measurements are evenly spread over the field area or sample distribution. Systematic sampling often gives more accurate estimates than simple random sampling (Snedecor & Cochran, 1989), except in very large homogeneous regions (Dutilleul, 1993). However there are also disadvantages. The confidence intervals calculated from regularly sampled data (in space or time) for the overall population estimate may be unreliable, and if there is a natural periodic variation in the phenomenon of interest that coincides with the sampling interval, it may go undetected (Atkinson, 1997; Greenwood, 1996). In addition, if the patch scale is much smaller than the sample spacing, the spatial autocorrelation and structure of the patches cannot be determined. The distribution of the samples can be defined by any of the following:

***Grid*:** Regular, square network of sampling points, ideally randomly oriented with a randomly selected origin (Figure 1).

***Equilateral triangular grid*:** Regular, triangular network of sampling points, ideally starting from a randomly selected location. This format is considered ideal for systematic sampling, but is not as convenient for data display in gridded format as the standard square grid.

***Unaligned grid*:** A regular, square grid is defined over the study area, and sample locations are located randomly within each grid cell (Caeiro *et al*., 2003).

***Transects*:** Line transects have been used for many decades in vegetation sampling and other fields. One way of randomising their locations is to randomly select the starting point and/or orientation (Greenwood, 1996). Transects can also be useful for preliminary sampling to develop a variogram (see below), although they only allow variation in one direction to be explored (McBratney & Webster, 1983). They can also be used to assess variability of map units, or boundary spacing (Burgess & Webster, 1984a). Nested transects have proven useful for exploring differences in process and pattern over multiple scales (Oline & Grant, 2002).

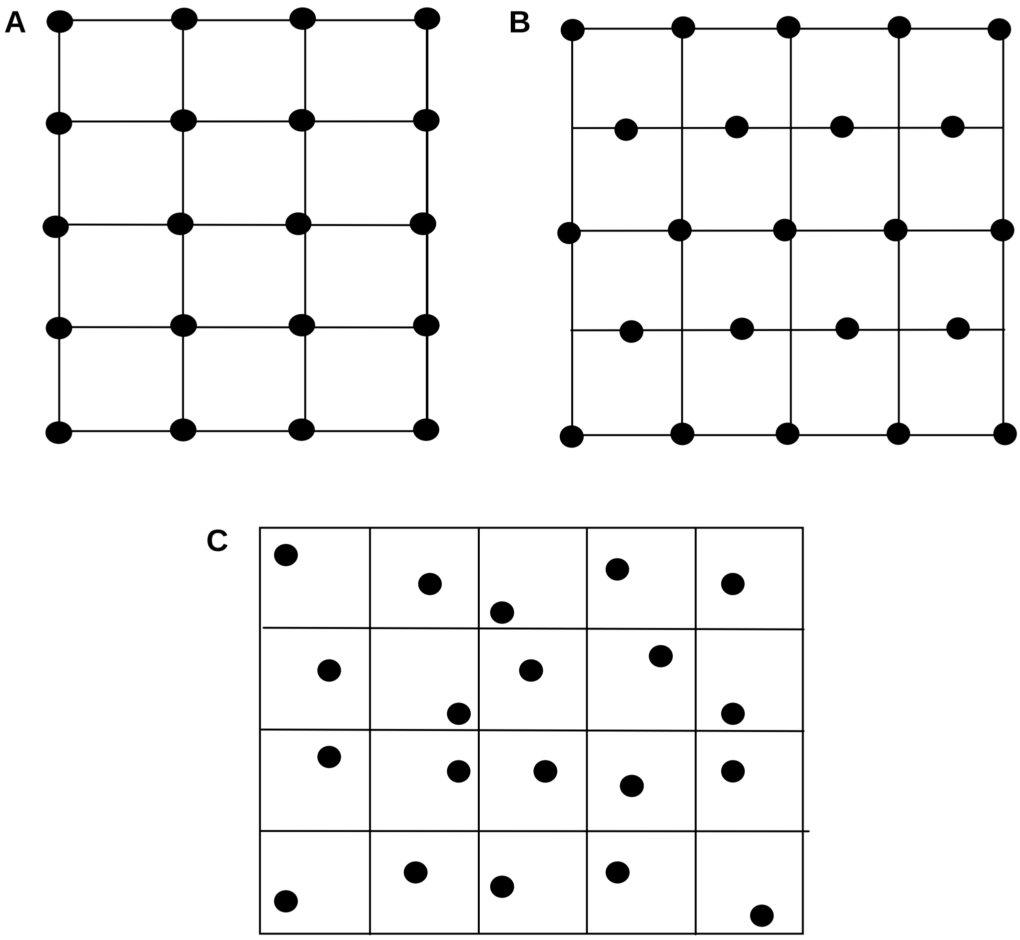

Figure 1. Grid types: A – square grid; B – equilateral triangle grid; and C – unaligned grid

Number of samples

The number of samples necessary to characterise an area depends on the variance of the variable of interest and its distribution, if spatial analysis is planned. Standard equations exist to calculate the number of samples needed to achieve a given level of confidence. Making the prediction requires some information about the standard deviation and the size of the population, and solving for *n* (p. 438 in Snedecor & Cochran, 1989; p. 82 in Greenwood, 1996; Burrough, 1995). If sampling cost can be estimated, it can also be taken into consideration to optimise for efficiency. As a general rule more sample collection should be planned in areas with high variability and those where collection is least expensive.

For studies looking specifically at spatial variation, ideally more than 150 data measurements at different locations are necessary for constructing a reasonable variogram (Webster & Oliver, 1992, 2001), which can be extremely useful for optimising sample spacing and layout (see section on ‘Optimal sample spacing’).

Sample support

Sample support, or volume of sample measurement (i.e. volume of water for analysis per sample, or area over which vegetation cover is estimated), is dependent on what is being studied and the sampling method. For phenomena with high variability over short distances, increasing the size of the support can often help to reveal larger-scale patterns, but this results in losing all information about smaller scale variation. The sample support should remain consistent throughout the study area, otherwise it may be necessary to weight the sample measurements to even out the bias (Flatman *et al*., 1988; Greenwood, 1996). The extent of the field area, divisions for stratified sampling, or quadrat dimensions depend largely on the size and mobility of what's being studied. Power analysis can be a useful technique for optimising sampling efficiency in terms of global estimation or testing for minimum sample support (Heidelbaugh & Nelson, 1996). Sample support can be adjusted through aggregation and disaggregation procedures.

Pre-sampling, or two-phase sampling

No sampling design can be optimised without prior information concerning the variability of the variable of interest and, particularly for spatial sampling, two phases of sampling are recommended (Borgman & Quimby, 1988; Legendre *et al*., 1989). Stratified random nested sampling is probably the best approach for gathering sufficient data to determine spatial variability (Webster & Oliver, 2001), followed by a combination of systematic sampling to provide information about spatial distribution over the full extent of the field area and some version of random sampling at shorter lags to provide enough information to model short distance variability (Flatman *et al*., 1988).

If sampling for classical statistical analysis is required, the first-phase sampling can be used to optimise the number of samples necessary for a given level of confidence, and a random placement of a systematic design permits full coverage of the field area. Remotely sensed imagery and ancillary environmental datasets can also be used to help estimate sample variance and potential stratification across the field area during the sampling design phase.

Optimal sample spacing

Geostatistics were designed to take advantage of spatial dependence to make more realistic estimates of values at unmeasured locations. A model is built that represents the spatial dependence in the data (a semivariogram), which can be used to assess and map covariance between sample locations, and among different variables over space. The semivariogram is modelled to produce estimates of covariance between data separated by any distance. The model type plus three parameters are typically sufficient to describe spatial dependence. Semivariograms are calculated as the average squared semi-difference between every possible pair of data in the study area (Goovaerts, 1997):

$$\gamma(h) = \frac{1}{2N(h)} \sum_{\alpha=1}^{N(h)} [z(u_\alpha) - z(u_\alpha + h)]^2 \qquad \text{(Eqn 1)}$$

where *N(h)* is the number of data pairs with the class of distance and direction (vector *h*), *z* is the random variable of interest at a location with coordinates of vector u_α.

The variogram is plotted with semi-variance (γ) on the Y-axis, and distance between the data compared within each pair (lag, *h*) on the X-axis (Figure 2). The model can be determined through an iterative weighted least squares fitting procedure, but requires user judgment to find the appropriate fit and parameter estimates. The three main parameters of interest are the nugget, sill and range. The *nugget* is the point at which the model crosses the Y-axis, and represents sub-measurement scale variation in the data, caused by natural variation at smaller scales than captured through sampling, laboratory and measurement error and locational error. Variance typically increases from the nugget (at zero distance) up to a plateau in the graph, after which increasing separation between the measurements no longer causes an increase in semivariance. The semivariance value of the plateau is the *sill*, and the distance at which it is reached is the *range*.

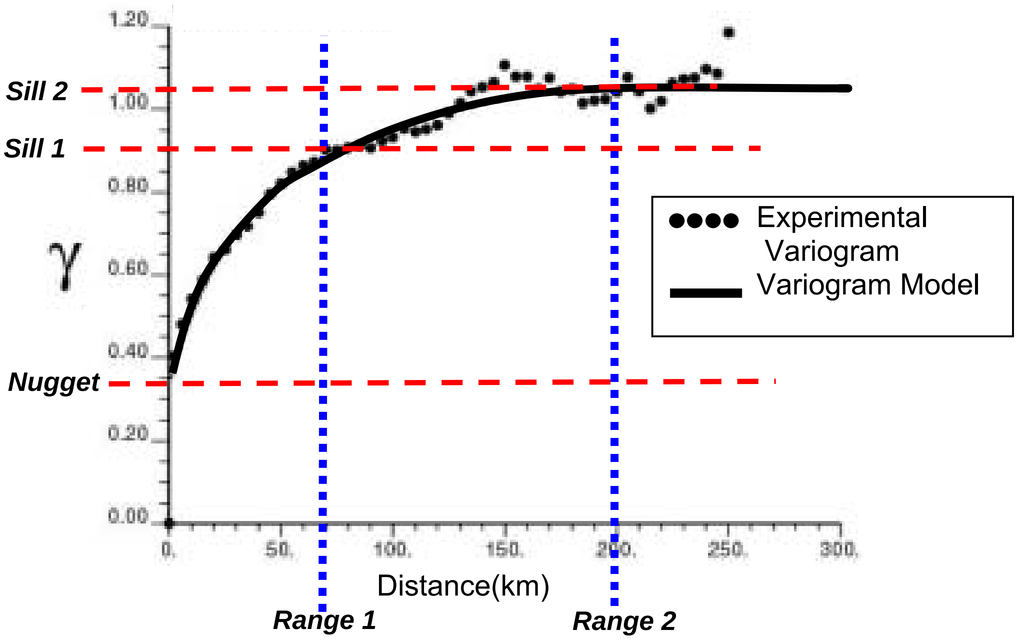

Figure 2. Example of an experimental variogram (black dots) which has been fitted using a spherical model (black line)

In this case, two model structures were fitted, resulting in a nested variogram with two sills and two ranges, as shown by the red and blue lines, respectively.

Spatial correlation can be visualised and measured (modelled) using semivariograms. This example is a nested variogram model composed of two variogram structures, which suggests the spatial patterns are the result of two independent processes operating at different spatial scales (0–70 km and 0–200 km). The variable (Figure 2) shows a fairly large nugget (γ = 0.34) which accounts for more than a third of the overall variance (Sill 2, γ = 1.02). Variance increases to a range of 70 km for the first model structure, and continues to increase to 200 km for the second model structure, beyond which variance levels off with increasing distance.

There are methods for analysing these patterns separately (Dobermann *et al*., 1995; Goovaerts, 1992; Holmes *et al*., in review), but for sampling design the most important feature is the range. Because the major sill for this variable is Sill 2, the discussion will focus around Sill 2 and Range 2 as if it were a simple (one structure) variogram. The range, or distance at which the sill is reached, is the range of correlation and is important to the sampling design, estimation of variable values at unsampled locations and interpolation error. Samples that are separated by a distance greater than the range are spatially uncorrelated, while those closer than the range are spatially correlated. If the samples are for true classical statistical analysis and complete independence of data values is necessary, then the samples must be collected at least the range distance apart. If any kind of mapping or spatial analysis is planned, then samples must be closer together than the range (Flatman *et al.*, 1988).

In two-phase sampling the first phase is primarily meant to evaluate heterogeneity across the field area by permitting the analysis of the semivariogram. An adequate number of samples must be measured over a range of short distances to determine the nugget, and broad coverage is needed to test for the distribution of the variable of interest and any gradients across the study area. If a physical relationship exists between the variable of interest and a more readily available measured variable (e.g. bathymetry or sea surface temperature) then the secondary variables can be used to stratify the field area and ensure sampling of the full range of variable values. Some samples should be taken quite close together to determine how much microvariation there is, and evaluate the sample support.

Many different approaches have been used to sample for the variogram, and among the most useful are transects (McBratney & Webster, 1983), nested transects (Oline & Grant, 2002), and stratified nested sampling (Burrough & McDonnell, 1998; Webster & Oliver, 2001). Caution should be taken when using transects, however, as they limit the ability to explore directional variation. Variograms can be calculated over all directions (isotropic or omnidirectional variograms) and over specific directions (anisotropic or directional variograms). If information about directionality in the field area is known, it can be incorporated to direct preliminary sampling for the variogram, and should be tested in exploratory analysis prior to finalising the sample design. For instance, in coastal areas, strong gradients of increasing water depth and changes in exposure or swell tend to be oriented perpendicular to the coastline, and thus may strongly affect patterns in sediment grain size, vegetation or pelagic properties. At the minimum, the data should be tested for directionality perpendicular and parallel to the coast.

The second phase of two-phase sampling typically is designed to optimise maps of the variable of interest within an acceptable range of interpolation error. Grids are almost always the most efficient sampling design for interpolation because they provide even sample coverage across the field area. The grid spacing may be rectangular, to accommodate different ranges in different directions, which can also be considered a series of parallel transects (Webster & Oliver, 2001). Sample spacing over the study area is determined from the directional semivariograms calculated from the first phase of sampling. If the study area is much larger than the area actually sampled, then the data must be extrapolated, which leads to large estimation errors. Extent of the study should be defined to accommodate the scale over which the processes or variables of interest are operating. The resolution of the final maps must be adequately detailed to capture the spatial patterns of interest, but not so detailed that it causes computational problems.

The range from the semivariogram determines the distance between samples, or the grid spacing, for the second phase. Sampling at distances closer than two-thirds of the range results in diminishing returns if there is a large nugget. If the nugget is very small (very little micro-scale variability), then diminishing returns set in at about half the range (Flatman *et al*., 1988). Studies have shown that using the spatial approach and interpolation to calculate the relationship between standard error and sample size results in either more (interpolation) precision from a given number of samples, or fewer samples needed to achieve a given level of precision compared with the classical statistical approach (Burrough, 1995).

A number of examples of using semivariograms to optimise sampling design are found in the literature, including: characterising spatial heterogeneity in coastal ocean characteristics for locating appropriate positions for monitoring stations (Kitsiou *et al*., 2001; Legendre *et al*., 1989), defining sampling grid orientation and spacing (750 x 500 m) for mapping estuarine sediments (Caeiro *et al*., 2003), determining optimal transect and sample spacing for benthic habitat mapping (Pinn & Robertson, 2003), optimising transect sample spacing for assessing estuarine water quality (Jassby *et al*., 1997), and determining sampling intervals based on variograms of map unit boundary spacings (Burgess & Webster, 1984b).

Linking datasets with varied resolutions and extents (scaling)

Data for benthic habitat mapping are collected using a range of methods by collaborative research teams. This necessitates integrating data with various spatial extents and grains (e.g. Table 1). Changing the scale of observation can alter the conclusions drawn from the analysis (Turner *et al*., 1989). Scale has been identified not only as a critical issue in ecology but as a phenomenon that is central to ecology (Levin 1992). Measurements of the same variable or process at different scales often reveal different patterns or behaviour resulting from nonlinear scaling effects due to strong feedback effects of physical and biological processes (Jarvis, 1995; Rastetter *et al*., 1992; Wessman, 1992). How can we deal with scaling the extent of the study area and the grain and resolution of habitat mapping?

The dynamics of different organisms and environmental variables follow different trajectories in both space and time. What is pattern at one scale can be considered noise at another, and as the extent of the study increases, so does the heterogeneity of the system (Oline & Grant, 2002; Wiens, 1989). As the sample support increases in volume, the extent of the study area also tends to increase. While coarsening sample support tends to reduce the short distance variance, increasing the study extent typically introduces new sources of variability. To successfully compare datasets gathered at different scales, both the support and the extent of the small-scale data must be increased to match that of the larger scale data (Atkinson, 1997; King, 1990). If we study a system at an inappropriate scale, we may end by identifying patterns which are artefacts of the scale of analysis rather than detecting the actual dynamics and patterns of the system (Wiens, 1989).

The researchers define the desired extent for data collection when they delineate the study area boundary. Datasets larger than the area can generally be trimmed without repercussions. However for datasets smaller than the desired extent, generalising beyond the smaller area where data were actually collected can difficult. Direct extrapolation requires accepting the assumption that patterns and processes are uniform over all scales, which we know is not true (Wiens, 1989). Extremely detailed information is necessary to assess spatially variable ecological systems, but it is not logistically possible to collect such detailed information over large regions.

In addition to matching the extent of the data coverage for a study area, the sample support must also be scaled to equivalent volumes or areas to detect meaningful relationships among variables (Atkinson, 1997). Adjusting all the datasets to a similar grain and extent involves choosing the appropriate scale for analysis according to the variable being studied and, in the case of large-area habitat mapping, typically upscaling data to match the largest grain variable included in the study. Various methods have been suggested for upscaling information in terms of extent (extrapolation and interpolation) and grain (aggregation), and are described in more detail in the appendixes. In Appendix A we provide a brief introduction to scaling extent and grain using arithmetic, geostatistical, process modelling and hierarchical methods and the related literature. Readers interested in learning more about ecological issues in scaling and methods for specific scenarios are referred to the extensive ecological literature on the subject.

References

Allen, T.F.H. & Starr, T.B. (1982) *Hierarchy: Perspectives for ecological complexity.* University of Chicago Press, Chicago.

Atkinson, P.M. (1997) Scale and spatial dependence. In: P.P. van Gardingen, G.M. Foody & P.J. Curran (Editors), *Scaling up from cell to landscape*. Society for Experimental Biology, seminar series 63. Cambridge University Press, Cambridge.

Austin, M.P. (2002) Spatial prediction of species distribution: An interface between ecological theory and statistical modelling. *Ecological Modelling,* 157: 101–118.

Bax, N., Kloser, R., Williams, A., Gowlett-Holmes, K. & Ryan, T. (1999) Seafloor habitat definition for spatial management in fisheries: A case study on the continental shelf of southeast Australia. *Oceanologica Acta,* 22(6): 705–720.

Borgman, L.E. & Quimby, W.F. (1988) Sampling for tests of hypothesis when data are correlated in space and time. In: L.H. Keith (Editor), *Principles of environmental sampling*. American Chemical Society, Washington DC, pp. 25-43.

Brown, C.J., Cooper, K.M., Meadows, W.J., Limpenny, D.S. & Rees, H.L. (2002) Small-scale mapping of sea-bed assemblages in the eastern English Channel using sidescan sonar and remote sampling techniques. *Estuarine, Coastal and Shelf Science*, 54: 263–278.

Brus, D.J. & de Gruijter, J.J. (1997) Random sampling or geostatistical modelling? Choosing between design-based and model-based sampling strategies for soil (with discussion). *Geoderma*, 80: 1–59.

Burgess, T.M. & Webster, R. (1984a) Optimal sampling strategies for mapping soil types. I. Distribution of boundary spacings. J*ournal of Soil Science*, 35: 641–654.

Burgess, T.M. & Webster, R. (1984b) Optimal sampling strategies for mapping soil types. II. Risk functions and sampling intervals. J*ournal of Soil Science,* 35: 655–665.

Burrough, P.A. (1995) Spatial aspects of ecological data. In: R.H.G. Jongman, C.J.F. Ter Braak & O.F.R. Van Tongeren (Editors), *Data analysis in community and landscape ecology*. Cambridge University Press, Cambridge, pp. 213–251.

Burrough, P.A. & McDonnell, R.A. (1998) *Principles of geographical information systems*. Oxford University Press, New York, 333 pp.

Caeiro, S., Painho, M., Goovaerts, P., Costa, H. & Sousa, S. (2003) Spatial sampling design for sediment quality assessment in estuaries. *Environmental Modelling & Software*, 18: 853–859.

Cochran, W.G. & Cox, G.M. (1957) *Experimental designs.* 2nd edition. John Wiley & Sons Inc., New York.

Cochrane, G.R. & Lafferty, K.D. (2002) Use of acoustic classification of sidescan sonar data for mapping benthic habitat in the Northern Channel Islands, California. *Continental Shelf Research*, 22: 683–690.

Dobermann, A., Goovaerts, P. & George, T. (1995) Sources of soil variation in an acid Ultisol of the Philippines. *Geoderma*, 68: 173–191.

Dutilleul, P. (1993) Spatial heterogeneity and the design of ecological field experiments. *Ecology,* 74(6): 1646–1658.

Flatman, G.T., Englund, E.J. & Yfantis, A.A. (1988) Geostatistical approaches to the design of sampling regimes. In: L.H. Keith (Editor), *Principles of environmental sampling*. American Chemical Society, Washington DC, pp. 73–84.

Goovaerts, P. (1992) Factorial kriging analysis: A useful tool for exploring the structure of multivariate spatial soil information. *Journal of Soil Science,* 43: 597–619.

Goovaerts, P. (1997) *Geostatistics for natural resources evaluation*. Oxford University Press Inc., New York, 483 pp.

Greenwood, J.J.D. (1996) Basic techniques. In: W.J. Sutherland (Editor), *Ecological census techniques*: A handbook. Cambridge University Press, Cambridge, pp. 11–110.

Hay, G.J., Niemann, K.O. & Goodenough, D.G. (1997) Spatial thresholds, image-objects, and upscaling: A multiscale evaluation. *Remote Sensing of the Environment*, 62: 1–19.

Heidelbaugh, W.S. & Nelson, W.G. (1996) A power analysis of methods for assessment of change in seagrass cover. *Aquatic Botany*, 53: 227–233.

Heuvelink, G.B.M. (1998) *Error propagation in environmental modelling with GIS*. Research monographs in geographic information systems. Taylor & Francis, Bristol, PA, 127 pp.

Holmes, K.W., Kyriakidis, P.C., Chadwick, O.A., Soares, J.V. & Roberts, D.A. (in review) Tropical land-cover change impacts on multi-scale soil nutrient variability. *Biogeochemistry.*

Isaaks, E.H. & Srivastava, R.M. (1989) *An introduction to applied geostatistics*. Oxford University Press Inc., New York, 561 pp.

Jarvis, P.G. (1995) Scaling processes and problems. *Plant, Cell and Environment*, 18: 1079–1089.

Jassby, A.D., Cole, B.E. & Cloern, J.E. (1997) The design of sampling transects for characterising water quality in estuaries. *Estuarine, Coastal and Shelf Science*, 45: 285–302.

Kenny, A.J., Cato I., Desprez M., Fader G., Schuttenhelm R.T.E. & Side J. (2003) An overview of seabed-mapping technologies in the context of marine habitat classification. *Journal of Marine Science*, 60: 411–418.

King, A.W. (1990) Translating models across scales in the landscape. In: M.G. Turner & R.H. Gardner (Editors), *Quantitative methods in landscape ecology*. Ecological Studies 82. Springer-Verlag, New York, pp. 479–517.

Kitsiou, D., Tsirtsis, G. & Karydis, M. (2001) Developing an optimal sampling design: A case study in a coastal marine ecosystem. *Environmental Monitoring and Assessment*, 71: 1–12.

Klute, D.S., Lovallo, M.J. & Tzilkowski, W.M. (2002) Autologistic regression modelling of American woodcock habitat use with spatially dependent data. In: M.J. Scott, P.J. Heglund, M.L. Morrison, J.B. Haufler, M.G. Raphael, W.A. Wall & F.B. Samson (Editors), *Predicting species occurrences: Issues of accuracy and scale*. Island Press, Washington DC, pp. 335–344.

Kvitek, R., Iampietro, P.J. & Thomas K. (1999) *Final report: Early implementation of nearshore ecosystem database project*. Prepared for California Department of Fish and Game and Nearshore Ecosystem Database Project (NEDP).

Legendre, P. (1993) Spatial autocorrelation: Trouble or new paradigm? *Ecology,* 74(6): 1659–1673.

Legendre, P., Troussellier, M., Jarry, V. & Fortin, M.-J. (1989) Design for simultaneous sampling of ecological variables: from concepts to numerical solutions. *Oikos*, 55(1): 30–42.

Levin, S.A. (1992) The problem of pattern and scale in ecology. *Ecology,* 73: 1943–1967.

McBratney, A.B. & Webster, R. (1983) How big a sample is needed for regional estimation of soil properties? *Soil Science*, 135: 177–183.

McDonald, J.I. (1998) *An evaluation of video photography as a tool for monitoring seagrass systems*. University of Western Australia, Perth, 155 pp.

McRea Jr., J.E., Greene, H.G., O'Connell, V.M. & Wakefield, W.W. (1999) Mapping marine habitats with high resolution sidescan sonar. *Oceanologica Acta*, 22(6): 679–686.

Mumby, P.J. & Edwards, A.J. (2002) Mapping marine environments with IKONOS imagery: enhanced spatial resolution can deliver greater thematic accuracy. *Remote Sensing of the Environment*, 82: 248–257.

Mumby, P.J., Green, E., Edwards, A.J. & Clark, C. (1998) Cost-effectiveness of remote sensing for coastal management. In: A.J. Edwards (Editor), *Remote sensing handbook for tropical coastal environments.* UNESCO, [http://www.unesco.org/csi/pub/source/rs.htm].

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

Norris, J.G., Wyllie-Echeverria, S., Mumford, T., Bailey, A. & Turner, T. (1997) Estimating basal area coverage of subtidal seagrass beds using underwater videography. *Aquatic Botany,* 58: 269–287.

Oline, D.K. & Grant, M.C. (2002) Scaling patterns of biomass and soil properties: an empirical analysis. *Landscape Ecology*, 17(1): 13–26.

O'Neill, R.V., Riitters, K.H., Wickham, J.D. & Jones, K.B. (1999) Landscape pattern metrics and regional assessment. *Ecosystem Health*, 5(4): 225–233.

Parry, D.M., Kendall, M.A., Pilgrim, D.A. & Jones, M.B. (2003) Identification of patch structure within marine benthic landscapes using a remotely operated vehicle. *Journal of Experimental Marine Biology and Ecology*, 285–286: 497–511.

Penrose, J.D. & Siwabessy, P.J.W. (2001) *Acoustic techniques for seabed classification.* Curtin University of Technology, Perth, prepared for the Marine Conservation Branch of the Western Australian Department of Conservation and Land Management (CALM).

Phinn, S.R., Menges, C., Hill, G.J.E. & Stanford, M. (2000) Optimising remotely sensed solutions for monitoring, modelling, and managing coastal environments. *Remote Sensing of the Environment*, 73: 117–132.

Pinn, E.H. & Robertson, M.R. (2003) Effect of track spacing and data interpolation on the interpretation of benthic community distributions derived from RoxAnn(TM) acoustic surveys. *Journal of Marine Science,* 60: 1288–1297.

Raffaeli, D. Bell, E., Weithoff, G., Matsumoto, A., Cruz-Motta, J.J., Kershaw, P., Parker, R., Parry, D. & Jones, M. (2003) The ups and downs of benthic ecology: Considerations of scale, heterogeneity and surveillance for benthic-pelagic coupling. J*ournal of Experimental Marine Biology and Ecology*, 285–286: 191–203.

Rastetter, E.B., King, A.W., Cosby, B.J., Hornberger, G.M., O'Neill, R.V. & Hobbie, J.E. (1992) Aggregating fine-scale ecological knowledge to model coarser-scale attributes of ecosystems. *Ecological Applications*, 2(1): 55–70.

Snedecor, G.W. & Cochran, W.G. (1989) *Statistical methods*. Iowa State University Press, Ames.

Solan, M., Germano, J.D., Rhoads, D.C, Smith, C., Michaude, E., Parry, D., Wenzhofer, F., Kennedy, R., Henriques, C., Battlea, E., Carey, D., Iocco, L., Valente, R., Watson, J. & Rosenberg, R. (2003) Towards a greater understanding of pattern, scale, and process in marine benthic systems: a picture is worth a thousand worms. *Journal of Experimental Marine Biology and Ecology,* 285–286: 313–338.

Thrush, S.F., Schneider, D.C., Legendre, P., Whitlatch, R.B., Dayton, P.K., Hewitt, J.E., Hines, A.H., Cummings, V.J., Lawrie, S.M., Grant, J., Pridmore, R.D., Turner, S.J., & McArdle, B.H. (1997) Scaling-up from experiments to complex ecological systems: Where to next? *Journal of Experimental Marine Biology and Ecology,* 216: 243–254.

Turner, M.G., O'Neill, R.V., Gardner, R.H. & Milne, B.T. (1989) Effects of changing spatial scale on the analysis of landscape pattern. *Landscape Ecology,* 3(3/4): 153–162.

Van Niel, K. & Laffan, S.W. (2003) Gambling with randomness: The use of pseudo-random number generators in GIS. *International Journal of Geographical Information Science*, 17(1): 49–68.

Webster, R. & Oliver, M.A. (1992) Sample adequately to estimate variograms of soil properties. *Journal of Soil Science*, 43: 177–192.

Webster, R. & Oliver, M.A. (2001) *Geostatistics for environmental scientists: Statistics in practice*. John Wiley & Sons Ltd, New York, 271 pp.

Wessman, C.A. (1992) Spatial scales and global change: Bridging the gap from plots to GCM cells. *Annual Review of Ecology and Systematics*, 23: 175–200.

Wiens, J.A. (1989) Spatial scaling in ecology. *Functional Ecology*, 3: 385–397.

Wu, J. & David, J.L. (2002) A spatially explicit hierarchical approach to modelling complex ecological systems: Theory and applications. *Ecological Modelling,* 153: 7–26.

Appendix A: Scaling data with different resolutions

I. Arithmetic scaling

The landscape can be viewed as a nested hierarchy of process and pattern. Moving between levels in the hierarchy corresponds with changing the study grain and extent.

Aggregation of detailed information to integrate variables at a larger scale can be done in a variety of ways. Three general methods for scaling detailed information in the form of a mathematical model for the purpose of prediction of large-area expression of the detailed phenomenon are discussed, namely lumping, extrapolation and explicit integration (King, 1990).

'*Lumping*' refers to averaging variability in the landscape by calculating mean values over larger areas, effectively increasing sample support. This can be done for model parameters and for estimates, resulting in increasing both the grain and the extent of the study area. The values change, but the model structure remains the same. The assumption that the same model adequately represents both small- and large-scale phenomena is often not applicable, and holds only if the system behaves linearly; otherwise, bias is introduced into the modelling process. The simplicity of the lumping approach makes it easy to implement but errors generated through scaling can be significant.

Extrapolation, in contrast to lumping, does not change data values but rather applies the model for small-scale phenomena over a much larger extent, basically simulating small-scale patterns over areas where no data measurements are available. This would typically be applied in a GIS format, either as grid-or polygon-based modelling, and is probably the most commonly applied method in environmental research for scaling information from detailed study sites to larger areas.

Extrapolation can be taken a step further, and treated as a stochastic small-scale model applied over a larger area. King (1990) refers to this method as 'extrapolation by expected value'. The parameters of the small-scale model are treated as stochastic variables, and their joint probability distribution defines the spatial variability of the larger extent of the study area. The output of the small-scale model is also a random variable, and the expectation of the output variable determines the small-scale behaviour over the landscape or larger extent of interest. This approach is compatible with the theory of random variables, but the joint outcome can be extremely complex. The basic requirements for using extrapolation by expected value are the following: (1) a model simulating small-scale system behaviour; (2) the larger extent over which the model is to be extrapolated; (3) frequency distributions for all the variables describing the large area spatial heterogeneity; and (4) the calculation of the expected value. If all the data are continuous and readily available, this can be calculated analytically. Alternatively, if a solution for the equation does not exist, it can be solved numerically or through stochastic simulation. The main source of error for this approach is in estimating the probability distributions of the small-scale model variables.

In contrast with the previous methods discussed, extrapolation by *explicit integration* transforms or rescales the original small-scale model to create a new, larger-scale model. Both the grain and the extent are changed to represent the larger area of interest. This approach explicitly integrates the small-scale model with space as variable of integration. Consider a model where *l* is some locally measured quantity and *z* is a spatially distributed variable:

$$l = m(z) \qquad \text{(Eqn 2)}$$

If *z* can be described as a function of space *g(x,y)* where *(x,y)* represents a location, then:

$$l = m(g(x,y)) = f(x,y) \qquad \text{(Eqn 3)}$$

Given a rectangular study area bounded by *a* and *b* in the easting direction and *c* and *d* in the northing direction, then the aggregate landscape-scale representation *(L)* of the small-scale model is:

$$L = \int_c^d \int_a^b f(x,y)dxdy = \int_c^d \left[\int_a^b f(x,y)dx \right] dy \qquad \text{(Eqn 4)}$$

This method assumes that landscape heterogeneity can be described as a function of space, and that this integral exists and can be solved. Very complex functions can make it impossible to find such a solution (King, 1990; Rastetter *et al*., 1992).

II. Geostatistical scaling

The semivariogram, used to describe spatial dependence among measurements, can also be used to describe the effect of changing the size of sample support. The following discussion draws heavily on Atkinson (1997).

Measurement values represent the spatial average or integral of the property of interest *Z(x)* over a support v centred on location x_0:

$$Z_v(x_0) = \frac{1}{v} \int_{v(x_0)} Z(y)dy \qquad \text{(Eqn 5)}$$

where *Z(y)* is the property Z defined on point support. All measurements taken by any method are made using this type of integration; by definition there is no such thing as a sample with a support of zero. The semivariogram will change shape depending on the measurement support, implying different scales of variation depending on the way the phenomenon was sampled. The following equation defines the relationship between semivariance defined on point support and rescaled semivariance at a given lag (*h*):

$$\gamma_v(h) = \bar{\gamma}(v, v_h) - \bar{\gamma}(v, v) \qquad \text{(Eqn 6)}$$

where $\gamma(v, v_h)$ is the integral of the point semivariance between two supports of size /v/ with centroids separated by *h*. The average point semivariance is $\gamma(v,v)$ within an observation of size /v/ (within block variance). The within block variance is completely overwhelmed by integration over the support (Equation 5), and the only detectable variation from the sample values is described by the left-hand side of Equation 6. The effect of changing sample support is to obscure small-scale variability, and emphasise larger scale differences within the extent of the study area. In the case where most of the variability occurs over short distances, most of the detail will be lost. In terms of the experimental variogram this equates to depressing the sill and decreasing the nugget effect, generally smoothing out short-distance variability.

Scaling sample support using variograms makes it possible to predict spatial dependence based on different sample support without ever physically measuring at that support. The method supplies an estimate of changing variance with distance to which a variogram model can be fitted and used to model spatial dependence at all lags. Atkinson (1997) suggests that for coarse resolution remote sensing data, it may be necessary to first downscale the data to point support, then calculate the variogram based on intermediate levels of support. To calculate the point support, a plausible variogram model is chosen for the point-based variogram, and its coefficients are determined by iteratively adjusting Equation 6 until the fit between the estimated (scaled) variogram and the measured variogram is acceptable. Equation 5 can also be used to calculate summary statistics such as variance and to produce simulations from the new support again without physical measurements.

In remote sensing applications, the issue of support is more problematic than extent. If intermediate data resolutions are available, these can be used to bridge the gap between detailed field data with point support and coarse resolution imagery. The more likely scenario, however, is that only irregularly spaced ground-collected data are available for comparison with coarser-scale remote sensing imagery. The data can be directly compared, but correlations tend to be very low, and an adjustment of the smaller support to match the imagery is recommended (Atkinson, 1997). If few to no data points are located within each pixel of the imagery, then it is not possible to estimate a mean pixel value. Rather, block kriging and conditional simulation are the recommended methods for scaling sample support.

Block kriging is optimal estimation of data values based on the variogram, or model of spatial dependence. Neighbouring values are used to estimate aggregated area (larger support, or block) mean values in areas where few or no measurements exist (Goovaerts, 1997). While this approach optimises estimation in areas with few measurements, it also leads to oversmoothing of the data, because neighbouring values are incorporated in estimation, effectively extending the support even further. This means that the bivariate relationship between the kriged variable and the satellite imagery is also affected by oversmoothing, and may not represent the true correlation. The preferable method for extending sample support is conditional simulation. Conditional simulation utilises kriging to calculate the optimal estimate and the amount of variation lost through smoothing. Simulated values, drawn from a probability distribution of possible values at each estimate location, are conditional on the original data measurements, the variance and the variogram, and the data are not smoothed. However, there are still unresolved issues surrounding the effect of the stochastic variation added into the realisations on bivariate relationships.

The problem of extending the extent of coverage for the small-scale data still remains a problem for most marine habitat mapping programs owing to the expense and difficulty of full data coverage with video or sidescan. This can be addressed by gathering regular sample coverage across the field area, particularly around the borders, and using sound interpolation techniques for estimation. If no data is collected around the perimeter, values in these areas can only be estimated through extrapolation, which is not constrained to the measured data and produces less reliable estimates.

III. Process model scaling

In some cases, biogeochemical or biological fluxes may be incorporated in habitat modelling. The ideal situation for process upscaling is to collect flux or other process information at multiple sample supports (e.g. the plant, patch and meadow) to model system changes with scale and characterise nonlinear behaviour (Wessman, 1992). Several common ecological process upscaling procedures include summation, averaging and aggregation, as described by Jarvis (1995).

If all fluxes in the system are measured reliably, then summation at the scale of measurement gives the flux at the next larger scale. For instance, summation of CO_2 fluxes from all seagrass plants is equal to the patch flux, and fluxes over a number of patches gives the meadow gas flux. However, not all plants in a seascape can be directly measured, reducing the problem to a classic sampling problem of defining how many samples are necessary to characterise the population as a whole. In spatially heterogeneous areas with mixed species cover, the number of samples and timing constraints may be prohibitive. Another issue is that as larger and larger areas are considered, more and more parameters enter into the equation, requiring new methodologies for measurement and more and more complex models. The data collection requirements of this method are prohibitive, and it can be applied in few real-world situations.

A second approach involves averaging relevant parameters at the scale of measurement, and using these averages to estimate fluxes at larger scales. The requirement is that the fluxes that are most important at the larger scales are given appropriate weighting in the averaging process, which is often not straightforward. As for direct summation, sampling to adequately estimate mean fluxes for all important processes involved can be a difficult problem. Stratified sampling has been shown to reduce the measurement effort, but a relatively large number of flux measurements are typically required (Jarvis, 1995).

Finally, the aggregation approach is much more integrative than the previous two approaches discussed in that it takes into account feedbacks in the system, such as temporal changes that affect flux rates or between-system flows. This method requires iterative modelling of both the variable of interest (e.g. seagrass gas fluxes) and other processes affecting these rates (e.g. hydrodynamics) at multiple scales, to account for the interplay between the systems, rather than calculating the flux from one set of field measurements alone (Jarvis, 1995). Although this approach promises more transparency in the scaling procedure from the intermediate products, it also is much more time- and data-intensive, and regularising data support for all levels may be difficult.

IV. Nested hierarchical scaling

The landscape (and seascape) has commonly been conceptualised as a hierarchy which contains different object sizes and patterns at particular scales resulting from different processes acting over defined ranges of scale (Allen & Starr, 1982; Hay *et al*., 1997). The concept of levels of hierarchy provides a structure for decomposing complex systems based on the assumption that the levels are fundamentally different: the upper levels impose constraints on lower level dynamics (boundary conditions) and the lower levels determine the starting conditions for higher level processes (Wu & David, 2002).

Extrapolating from small-scale information over several hierarchical levels can be done using aggregation methods such as those described above, but there is always a risk of wrongly assuming the dominant processes, adequately quantifying heterogeneity or detecting and modelling system nonlinearities. Wu & David (2002) propose a spatially explicit hierarchical patch dynamics paradigm, which couples spatial patterns and processes (patch dynamics) in a 'scaling ladder' strategy to model changing dynamics with changes in scale. The approach is composed of three steps: (1) identifying the appropriate scales to study; (2) making observations and modelling at focal levels; and (3) extrapolating information across the domains of scale.

The approach accommodates all types of modelling and extrapolation methods, but also provides a nice conceptual framework in which to consider datasets of various grains in the habitat classification context. For a classification hierarchy, such as for benthic habitat mapping, data is necessary at every level to categorise both pattern and process for classification. One possibility is to avoid the problem of data aggregation by aligning the sampling methods with the corresponding hierarchical levels for classification (Raffaeli *et al*., 2003). For example, satellite and airborne imagery might be used for the coarsest scale (Mumby & Edwards, 2002; Mumby & Harborne, 1999), sidescan and automated classification at intermediate scales (Brown *et al*., 2002; Cochrane & Lafferty, 2002; Kenny *et al*., 2003; Pinn & Robertson, 2003), and video and ground sampling (Norris *et al*., 1997; Parry *et al*., 2003) at the finest scale. Rather than directly integrating the data, this means developing classification models specific to each spatial scale, and linking them through hierarchical association. This amounts to a nested analysis in which the extent is not fully covered at each scale, but rather detailed windows are analysed at each scale.

It is critical to remember that manipulating data extent and grain has the potential for introducing large amounts of error into the measurement values, and changes the nature of data from directly measurable values to theoretical area-averaged values which can no longer be directly validated by field collection nor correspond with the original measurement units (Rastetter *et al*., 1992). Error in the data models will propagate through all analyses, and as more parameters are included, the estimates become less and less precise (Heuvelink, 1998). By understanding the margin of error of the original data, the transformed data and any additional data manipulations, limitations of the resulting datasets and habitat models can be assessed quantitatively.

Appendix B: Recommendations for sampling protocols

All sampling plans must be customised for a given field area, research objectives and method of data collection, but there are some general field sampling principles and theoretical guidelines that can help structure sampling design. The purpose of the survey and existing data availability typically determine whether one or more sampling campaigns are necessary as well as the appropriate sample layout for the planned analyses. Unfortunately sampling costs, in terms of both money (equipment rental, field and lab personnel and sample analysis) and time, are often the main factor that determines the number of samples to be collected.

Some recommendations for sampling protocols are outlined below.

- Explicitly define the analytical objectives of the study before sampling, the assumptions associated with those methods, and the sample resolution and support necessary to meet the objectives. Are the variables considered stochastic, or must the sample design meet rigorous standards for unbiased selection?
- Determine the data resolution necessary to detect spatial or temporal variation. Phinn *et al*. (2000) is a good example for satellite imagery, but this needs to be done at all scales of interest. Evaluate all data sources in terms of accuracy (measurement accuracy, positional accuracy and magnitude of short-scale variation) to ensure that the sampling methodology chosen will adequately capture the patterns or process of interest. An equivalent sample support in all datasets is desirable for statistical analysis, so the full suite of variables in the study area should be considered prior to sampling both to determine appropriate sample support and possible field area stratification.
- Find all existing data sources, do some preliminary sampling if no data exists. This provides a basis for evaluation of sampling design and statistical analysis options.
- Aim for unbiased sampling designs, but use expert knowledge (first-phase sampling) to ensure adequate coverage in areas with high variability. For any kind of mapping, short distance relationships must be assessed, as well as obtaining full study area coverage.
- Verify that the methodology chosen for sampling will operate satisfactorily over the full field area and all environmental conditions which may occur.

- Focus sampling in areas of high variability once general trends in field area are understood. Stratify the sampling effort on the basis of variability to minimise sample collection and laboratory analysis.
- If using multiple sampling techniques, overlap the data collection in each stratified zone (if identified) to allow full characterisation of the field area, calibration of the different data types, and comparison of variance.
- Allow for time after the first sampling campaign to analyse the data and define the second sampling phase based on gaps in data or high variance.
- Use semivariogram analysis to determine sample spacing, or optimise selection of few locations for intensive temporal monitoring (Flatman *et al*., 1988; Webster & Oliver, 2001).
- Track the measurement and aggregation error in the data, and consider stochastic simulation to assess error propagation through modelling efforts. At the very least, calculate positional accuracy to estimate boundary location accuracy regardless of classification uncertainty.

Appendix C. Notes on other methods

(References cited below are only examples of these techniques)

(1) Bayesian hierarchical aggregation/disaggregation

Kolaczyk, E.D. (1999) Bayesian multi-scale models for Poisson processes. *Journal of the American Statistics Association*. 94, 920–933.

(2) Multiscale integration of categorical and continuous data

Cardille, J.A., Foley, J.A. & Costa, M.H. (2002) Characterising patterns of agricultural land use across Amazonia by merging satellite images and census data. *Global Biogeochemical Cycles*, 16(3): 1045.

(3) Multi-scale mapping given three datasets at different resolutions (same variable)

Zhu, J., Morgan, C.L.S., Norman, J.M., Yue, W. & Lowery, B. (2004) Combined mapping of soil properties using a multi-scale tree-structured spatial model. *Geoderma,* 118: 321–334.

Huang, H.C., Cressie, N. & Gabrosek, J. (2002) Fast, resolution-consistent spatial prediction of global processes from satellite data. *Journal of Computers, Graphings, and Statistics*, 11, 1–26.

(4) Wavelet analysis

Epinat, V., Stein, A., DeJong, S.M. & Bouma, J. (2001) A wavelet characterisation of high-resolution NDVI patterns for precision agriculture. *International Journal of Applied Earth Observation and Geoinformation*, 3, 121–132.

Appendix D. Spatial sampling designs for mapping the benthos: A case study for seagrass species mapping, Owen Anchorage, Western Australia

Introduction

Benthic habitat maps are increasingly in demand for a variety of applications, including: marine park placement and zoning (Friedlander *et al*., 2003); marine resource management (Bax *et al*., 1999); environmental monitoring in areas with human activities (Kendrick *et al*., 2002; Plummer *et al*., 2003); holistic catchment management; reducing ecological impacts of economic exploration; and integrative ecological research (Durako *et al*., 2002; Fonseca *et al*., 2002). Effort has been focussed on developing efficient marine survey equipment and data acquisition systems (Kenny *et al*., 2003; McRea Jr. *et al*., 1999; Solan *et al*., 2003), and sophisticated statistical modelling methods are increasingly being applied to benthic datasets to analyse, classify and map a wide range of abiotic and biotic features on the seafloor [e.g. multivariate analysis (Freitas *et al*., 2003); generalised additive models (Garza-Perez *et al*., 2004); and geostatistics (Bulit *et al*., 2004; Rueda, 2001)].

Benthic habitat maps vary greatly in scope from simple polygon maps of substrate texture to the use of fuzzy logic and stochastic simulations for mapping habitat probabilities (Urbanski & Szymelfenig, 2003) and to correlative (Valavanis *et al*., 2004) and ultimately physically-based (Fonseca *et al*., 2002) predictive models of ecological communities or species assemblages (Diaz *et al*., 2004). Maps of classified benthic habitats can be very useful for gaining general knowledge about what is likely to be found on the seafloor, but they typically represent a highly simplified, discrete view of the seafloor, represented as polygon maps. However, biological phenomena typically form gradients in response to light penetration, exposure to wave action, water biogeochemistry and other ecological factors.

The definition of seafloor maps based on a combination of geophysical and biological features as benthic *habitat* maps implies that they assist in understanding the distribution of plant and animal species, often highly mobile species that use different habitats during different stages of their life cycles. Unfortunately, most methods used to develop continuous spatial information on marine habitats (e.g. multibeam and sidescan technologies) generally provide information on seafloor morphology, which includes mainly the abiotic features of the environment. Yet, researchers have shown that in many cases, the abiotic data used to classify habitats are not highly correlated with the biota they intended to study (Stevens & Connolly, 2004), the correlations may be scale-dependent (Zajac *et al*., 2000), and evaluation of the quality of habitat types is difficult (Diaz *et al*., 2004).

There has been a movement in terrestrial ecology to focus directly on and map the species or processes of interest, rather than relying on generalised habitat maps. This movement is likely due to the shift from a community to an individualistic view of species distributions (Gleason, 1926; Gleason, 1939). Although this theory was slow to be accepted by other ecologists (McIntosh, 1967), it is the basis of the now widely accepted continuum concept (Austin, 1985; Collins *et al*., 1993; McIntosh, 1967), which describes the biotic landscape as a continuously changing set of species along environmental gradients (Austin & Heyligers, 1989). It is only recently that there has been a call for individual species prediction and mapping, but for many reasons, such as the practical needs of cartography, characterisation, and land management, generalised community level analyses have continued (Franklin, 1995).

The development of generalised habitat maps, as with all mapping, requires that certain aspects of a habitat must be deemed as important and thus preserved while others are ignored. Poor selection of these aspects can lead to misrepresentation or loss of information about the important relationships within habitats (James, 1967). Generalisation of habitat data for classification is a difficult process since environmental data is complex and exhibits emergent features at different scales. In addition, selection of the important aspects of a habitat may be good for every application (e.g. extent of reef structures) and poor for another (e.g. monitoring fish biodiversity). Thus there is often a requirement for a separate classification for each application. Franklin (1995) points out that, theoretically, species distribution predictions are simpler than community distribution predictions since they have fewer ambiguities and abstractions, but communities are still mapped due to lack of sufficient data or other constraints.

Both benthic habitat classification mapping and benthic species mapping are wholly dependent on the quality of the input datasets, both abiotic and biotic, and how well they represent the features and region of interest. In addition, how representative data are of the underlying ‘true’ population impacts all statistically based interpretations and mapping efforts. Nearly all environmental phenomena (e.g. terrain morphology, plant distributions or geochemical and biogeochemical properties) are spatially autocorrelated, meaning they form spatial patterns such that locations close together in geographic space tend to be more similar than locations separated by larger distances (Burrough, 1995; Legendre, 1993). For quantitative environmental mapping the ideal spatial sample configuration depends on the purpose of the investigation, the intensity of the spatial patterning and distance over which it occurs, and the planned statistical analysis methods (Kitsiou *et al*., 2001; Legendre *et al*., 1989). The analysis methods must be chosen to suit the intended research purpose, plus be compatible with the type of data collected (e.g. categorical, continuous, counts and area averaged), the number of samples and number of sampling locations, sampling unit (area represented by each sample or sample support), the spatial layout of the samples in relation to patterns in the phenomenon under study and the nature of the phenomenon itself. Otherwise, the model or map may present a biased or distorted view of the true seafloor characteristics.

In this appendix we discuss the application of classic (terrestrial) sampling approaches to benthic habitat mapping, and the development of a sampling design for generating maps of seagrass species over a 100 km^2 area in Owen Anchorage, near Perth, Western Australia. We suggest this is a flexible approach for planning sampling for mapping that can be readily adapted for characterising benthic features or organisms elsewhere.

Spatial sampling designs

Exhaustive sampling of every occurrence of a feature of interest is neither feasible nor desirable, particularly when collection is destructive, time-consuming or costly. Instead, a subset or sample is taken, which is presumed to be representative of the larger population, and thus through statistical inference can reveal something meaningful about the population as a whole (Cochran & Cox, 1957). The validity of this assumption depends on how well the sample characterises the underlying distribution (magnitude and distribution of values) and the confidence level or accuracy required for the study.

For nearly all environmental data, spatial autocorrelation exists at some spatial scale because the processes affecting environmental phenomena tend to be distributed over spatially continuous gradients. Spatial autocorrelation must be considered in sampling design and analysis, as it may affect interpretations depending on the distance over which data are autocorrelated (that is, pseudoreplicated, *sensu* Hurlbert 1984), the study extent (field area size) and resolution (map scale, and/or pixel size) of the final maps. Studies have addressed optimisation of spatial sampling designs for habitat mapping in terrestrial environments (e.g. (Hirzel & Guisan, 2002; Plotkin & Muller-Landau, 2002), for estuarine transect surveys (Jassby *et al*., 1997), general benthic mapping and ecological process measurements (Caeiro *et al*., 2003; Fonseca *et al*., 2002), and choosing marine monitoring sites (Kitsiou *et al*., 2001; Legendre *et al*., 1989). The type of sample layout needed varies with the application and logistical considerations.

Limitations on sampling designs for benthic mapping

Sampling underwater takes significant resources in terms of funding and time, plus makes it impossible to use many common sampling approaches and technologies such as optical remote sensing, some *in situ* experiments or high-frequency data collection. For map production, samples must be collected over the full range of environmental conditions in the study area at a high enough density to meet the map scale requirements for the project. For underwater sampling plans, this is extremely difficult to do if the study area is large and information about the area is lacking. Most vessel-towed sensors permit data collection at points (drop cameras), along transects (video, ROV), or across areas when full coverage is possible (acoustic imaging: sidescan or multibeam). Diver data collection is common to 35 m in water depth, but health concerns and the need for specialised equipment make it difficult at greater depths, plus equipment for determining diver position is often prohibitively expensive.

In terms of sampling design, the methods available for marine sample collection (e.g. boat-towed sensors or diving) impact the potential spatial configuration of samples. Tow video, for example, results in data gathered along transects. This spatial structure consequently affects the appropriateness and applicability of many analysis methods. For instance, for many of the classical statistical (aspatial) techniques applied for modelling and mapping, such as calculating ANOVAs to determine differences between group means, or classification algorithms such as principle components analysis, require that samples are independent, a condition which is not met by strongly spatially autocorrelated data.

The impact of using spatially autocorrelated data in classical statistical analyses can be extensive, as it affects the estimate of model errors. This is caused by both potential bias in the data (i.e. the sample is not representative of the general population) and autocorrelation of error. When the errors are not independent, interpretation of the significance of error is not possible (Hurlbert, 1984). In Hurlbert's (1984) seminal paper on pseudoreplication, his survey of published papers indicated that statistical analysis of marine benthos had the highest incidence of pseudoreplication (62% of all studies in this field).

For benthic sampling, random sampling designs are difficult to achieve over large areas, although it has been successfully used for dive sampling. However, while the spatial autocorrelation found in marine data confounds classic aspatial statistical techniques, it can be exploited by geostatistical techniques to improve spatial predictions based on sparse data. Due to the logistical problems associated with marine benthos data collection, these techniques should be more extensively explored for their application for marine benthos prediction and mapping. Therefore, systematic sampling designs are often a better option for these applications.

Systematic designs for benthic mapping

Systematic sampling designs often result in more accurate map production than unsystematic sampling (Flatman *et al*., 1988; Snedecor & Cochran, 1989), except in very large homogeneous regions (Dutilleul, 1993). Numerous systematic sampling schemes have been developed, but all are similar in that once the number and spacing of samples is determined, the distribution of all samples is known. The spacing of samples is usually determined by the spatial autocorrelation inherent in the phenomenon and whether the goal is to optimise this attribute and use geostatistics in the calculation or to remove it and use classical statistics.

Generally, systematic sample designs are grids, for which the grid origin or the starting point is chosen randomly to allow any location an equal chance of being sampled, thus reducing bias (Greenwood, 1996). However, the confidence intervals calculated from regularly spaced data for the underlying ‘true’ population estimate may be unreliable, and if there is natural periodic variation in the mapped feature that coincides with the sampling interval, it may go undetected (Atkinson, 1997; Greenwood, 1996). In addition, if any patterns exist that are smaller than the distance between samples, they will not be detected, nor can be predicted from nearby samples. Standard systematic sample distributions include grids [square, equilateral triangles and unaligned (samples located randomly within square cells)] and transects.

Systematic sampling designs can include stratification, which involves subdividing the field area to maximise between-unit variation, and minimise within-unit variation. The stratification can be systematic (e.g. five units with equal areas) or chosen based on a physical understanding of the system (e.g. different hydrodynamic regimes). Once stratified, each unit can be sampled systematically and proportional allocation of samples is sometimes used to increase or decrease sample density in different units according to area or prior knowledge of within unit variability (Snedecor & Cochran, 1989). For example, a study area is stratified by substrate type to sample for invertebrates, but the number of invertebrate species is known to be higher in reef areas than in fine sand.

For stratified random sampling with proportionate allocation, more samples would be taken from reefs than the other categories to permit a more robust calculation of variance. For terrestrial species modelling, equal stratification (same number of samples per unit, regardless of unit area) has been found to be more effective for accurate and robust model results (Hirzel & Guisan, 2002). Stratification can also be used to reduce sampling redundancy if high efficiency is necessary, using a multi-level approach where only a small number of units in each strata are sampled. For example, a field area consists of 33 sand patches and 24 seagrass patches, but only five patches of each habitat type are selected for invertebrate surveys.

Nesting sampling is widely used to investigate changes in spatial distributions with changes in scale (Hewitt *et al*., 1998; Oline & Grant, 2002; Wiens, 1989). The use of nested systematic sampling is generally recommended for collecting a sufficient number of data to investigate patterns at multiple scales. The field collection can be nested as well, using an early phase of sampling to determine the spatial variability of the phenomenon of interest, and thus the necessary sampling density or 'representative' monitoring locations (Kitsiou *et al*., 2001; Legendre *et al*., 1989).

Unfortunately, transect data collection provides dense information in the direction of travel and no samples are collected between transects. Anisotropic (directional) effects can be missed or misrepresented, due to the lack of continuous data in all directions, making it difficult to estimate the most appropriate scale (or interspersion) of data samples for collection and analysis. Another approach would be to collect sparse samples over the full extent of the study area, then choose (randomly or systematically) samples around which to cluster new sampling points (Hewitt *et al*., 1998; Oline & Grant, 2002). However, because spatial clustering of autocorrelated data can bias analyses due to many samples in potentially high or low value locations, the data may need to be weighted before analysis (declustering) (Goovaerts, 1997) or averaged to produce one value for each cluster location (Greenwood, 1996).

Seagrass species mapping

Project overview

The coastal region near Perth, Western Australia (Figure D1), has undergone major environmental changes over the last 30 years, including a massive loss of seagrass as a result of changes in light penetration from algal blooms related to nutrient loading from industries discharging waste water along the coastline (Cambridge *et al*., 1986). Since that time, the areas covered by seagrass and species composition have been studied as potential indicators of destructive human impacts on the coastal ecosystem in this commercially important area. Numerous detailed field studies have focussed on high spatial resolution issues such as seagrass physiology, and broadscale analyses of seagrass distribution and historical changes based on classification of time-series aerial photographs (Kendrick *et al*., 2002; Kendrick *et al*., 1999; Kendrick *et al*., 2000).

Commercial dredging in Owen Anchorage has been ongoing since the 1970s, which has resulted in the removal of some seagrass habitats, and increased sedimentation and turbidity in the surrounding areas. Quantifying the magnitude and distribution of the effects of dredging on the total area occupied by seagrass and on species composition is important for understanding the long-term impact of economic activity in relation to the natural dynamics of the seagrass-dominated benthic ecosystem. This has led to various research initiatives to develop objective and repeatable methods of sampling and analysis to quantify the distribution of seagrass species in the area and generate a baseline map and mapping standards for future comparisons. Here we discuss the development of a sampling design for collecting underwater video footage which supplies the raw data of georeferenced species presence or absence used for species mapping.

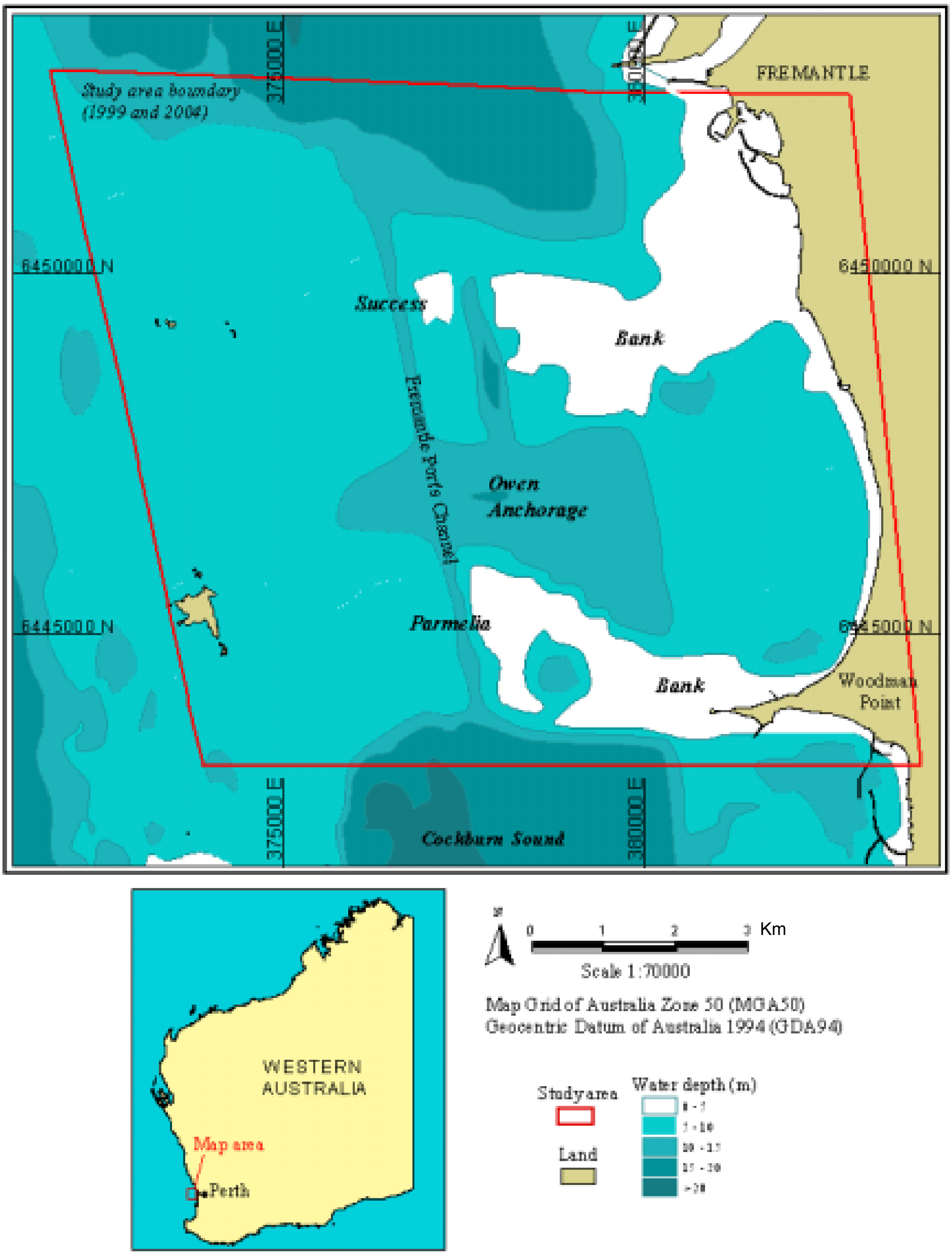

Figure D1. Location map for Owen Anchorage and the seagrass species study

A large number of spatially referenced datasets were available for incorporation into the sampling plan, including: bathymetry, a time series of historical photographs, and interpreted underwater video (Figure D2**)** from 1999 which provided identification to the species level. In this relatively shallow coastal region, classification of airphotos is effective for large-area mapping of dense canopy perennial seagrass (DAL *et al*., 2000), but it does not provide information about seagrass species. Underwater tow video is the most practical method of gathering species-level information over large areas, although subtle differences between species sometimes cannot be detected.

The analysis planned for producing semi-quantitative maps of species distributions was the geostatistical method of block indicator kriging. This technique requires binary or indicator data (1 for species presence, 0 for species absence), ideally full systematic coverage of the study area, some information concerning species distribution over short distances, typically gathered through some form of nested sampling, and a sufficient number of presence data to model spatial dependence [typically >150 (Webster & Oliver, 1992)].

The model estimates represent the probability of species presence in unsampled locations, based on the values of nearby samples, and are presented as a continuous (gridded) probability map. The number and distribution of neighbouring samples has a major influence on the accuracy of estimates in the final map (Goovaerts, 1997). The major sample design constraints were the need to gather a large number of samples over a large area (6 species x 300 samples for modelling, 10 km x 10 km study area), non-destructive and rapid sample collection, and use a maximum of 6 boat days. We chose underwater tow video, as it is nondestructive and can cover a fairly large area more efficiently than using diver surveys, plus the data are easily archived and can be interpreted multiple times for different analysis purposes. However, one major drawback is that the samples are aligned along transects, providing dense information in the direction of travel, and no samples collected between transects.

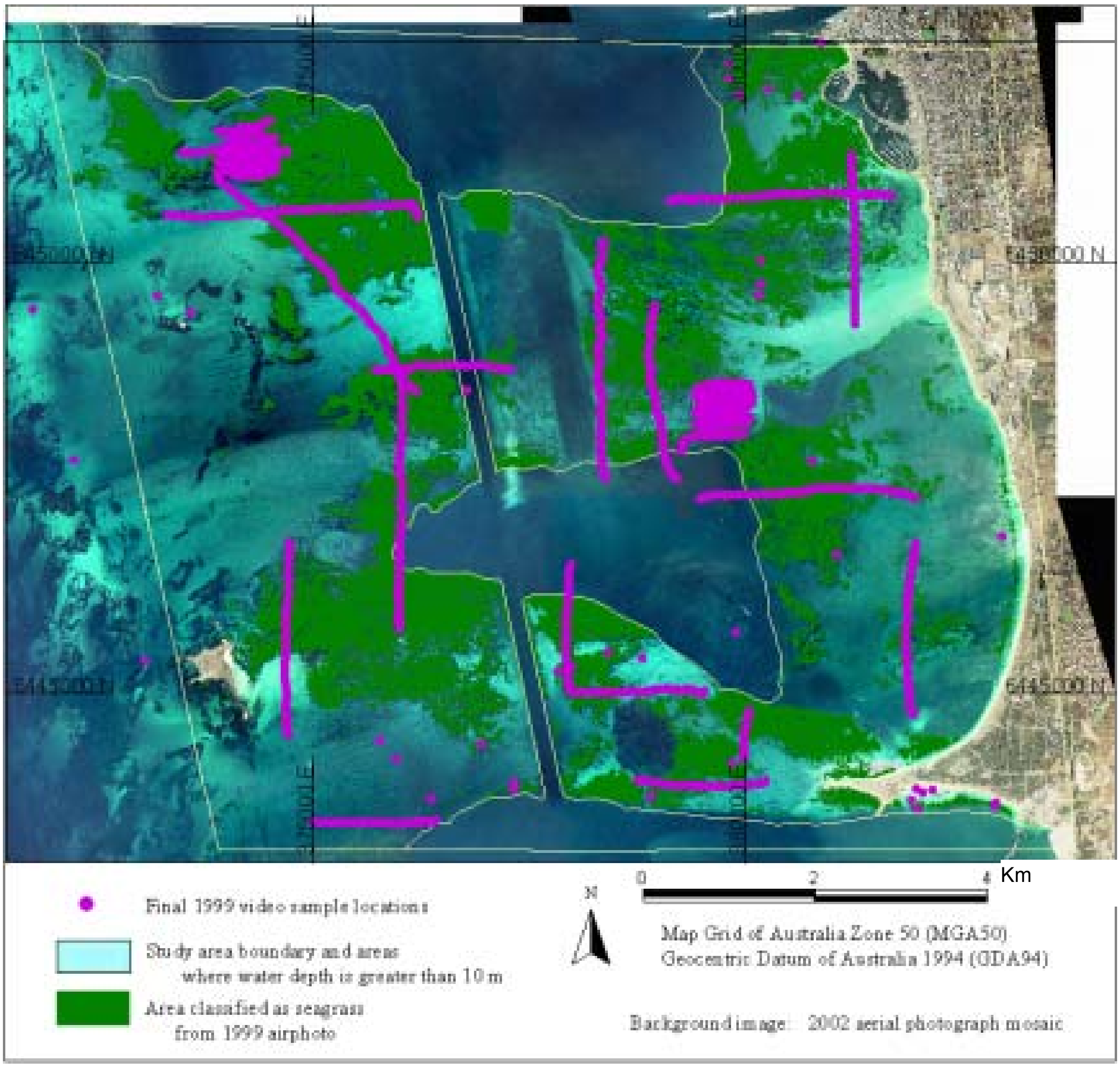

Figure D2. Most recent airphoto (2002) and map of seagrass coverage for Owen Anchorage derived from airphoto classification (1999)

Data collection

A video camera in an underwater housing was towed behind a vessel travelling at 1–3 knots. The video was viewed live on the boat during the survey, and recorded on a Sony miniDV recorder. Acquisition software developed at the University of Western Australia running on a laptop computer controlled the system and generated a unique reference point (GPS time stamp) at five-second intervals along each transect. The position, heading and speed were recorded from the boat's differential GPS for each of these points, along with depths from the vessel's echo sounder. All the information, including the frame reference, was referenced to the video frames. The camera position above the seabed was controlled manually to maintain an optimal distance from the seabed (<1 m).

Each frame of the video footage captured between 0.25 m^2 and 9 m^2 of the seafloor, depending on the height of the camera above the seabed. The area occupied by bare sand, limestone reef, wrack and the occurrence of the following seagrass species of seagrass was noted: *Amphibolis griffithii, Posidonia australis, Posidonia coriacea, Posidonia sinuosa, Holophila ovalis,* and *Heterozostera tasmanica*.

Sampling design

The sampling design for the 2004 videography relied on information on environmental gradients across the study area, previously collected species information and an iterative design in which data from the first few days was processed and analysed before finalising data collection locations for the final phase of sampling. The total coverage of video footage possible was determined by the number of field days available, and the average number of kilometres it was possible to cover per day, as estimated by the boat operators (20 km/day).

All sampling plans have to be flexible, as unforeseen events such as weather, equipment failure or personnel availability can affect the timing and extent of fieldwork. However, for the species mapping methods planned, samples spanning the full extent of the study area were needed, plus higher density samples in areas with more complex patterns of species distributions. Also, some coverage of the same areas observed in 1999 was needed for a temporal comparison at higher resolution.

Incorporation of existing information

The bathymetry map for the area (gridded 100 m postings) provided a general overview of the seafloor geomorphology. Two large sandbanks oriented roughly east–west, perpendicular to the coastline, are the dominant features, and seagrass distribution is presumed to be largely controlled by hydrodynamic stresses partially controlled by the local bathymetry. The strongest ecological gradients then would be expected to be related to water depth (affects light availability for photosynthesis) and hydrodynamics (determines stability of substrate).

Water depth increases east–west and north–south, as related to the positions of the sand banks, while the hydrodynamic regime is quite complex, but can be simplified to a measure of protection from wave activity, related to local geomorphic features and overall water depth. The preliminary video transects needed to be oriented roughly east–west (perpendicular to the shore line) and north–south (parallel to the shoreline, perpendicular to the sandbanks) to capture the full range of light and energy conditions in the study area which were expected to control seagrass species distribution.

The 1999 species data were used to target areas with a high likelihood of spatial variability and high species diversity, which require higher density sampling. Semivariogram analysis of the 1999 species data was used to define the desirable spacing between video transects and the selection of video frames for interpretation. The semivariogram of seagrass presence or absence showed strong patterning as expected from observations in the field and from the aerial photography.

The semivariogram for the most commonly identified species in 1999, *Posidonia coriacea*, shows strong spatial dependence, likely related to average meadow size, over a distance of approximately 2 km (Figure D3a), and its steep slope at short distances (Figure D3b) supports the observation that it is commonly found in fairly small patches. Different species are distributed differently across the field area. In particular, *Posidonia sinuosa* tends to form large homogeneous meadows, and *Amphibolis griffithii* forms smaller patches and is commonly associated with *Posidonia coriacea*. Not all of the species identified were found in sufficient numbers to construct a semivariogram, but from the more common species semivariograms, the spacing between transects was calculated as ideally 0.5 km (half the semivariogram range, or distance where the plot plateaus), but acceptable up to approximately 1 km.

Proposed sampling transects were oriented both north–south and east–west to give broad coverage of all areas known to have seagrass (Figure D4). A higher density of transects were located in areas mapped as having high species variability during the 1999 surveys. On average, transects were spaced less than 1 km apart. Transects, although not an ideal layout for mapping exercises, can be useful for preliminary sampling to determine the spatial extent of typical patches or patterns in the field area, although they only allow variation along the transect orientation to be explored (McBratney & Webster, 1983). They can also be used to assess variability of map units, or boundary spacing (Burgess & Webster, 1984).

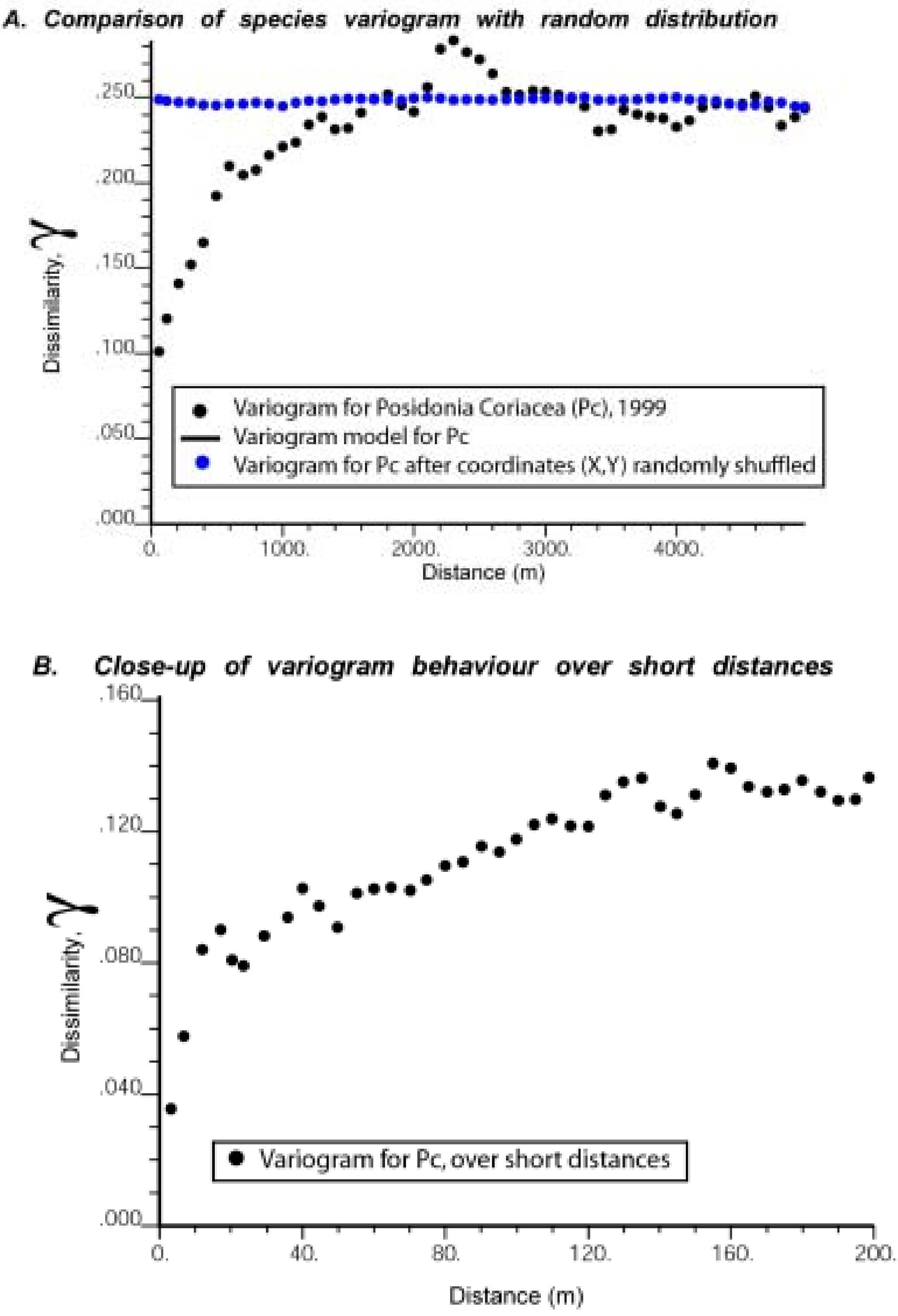

Figure D3. Semivariograms and models for species estimation shown for the most common species identified in 1999, *Posidonia coriacea*

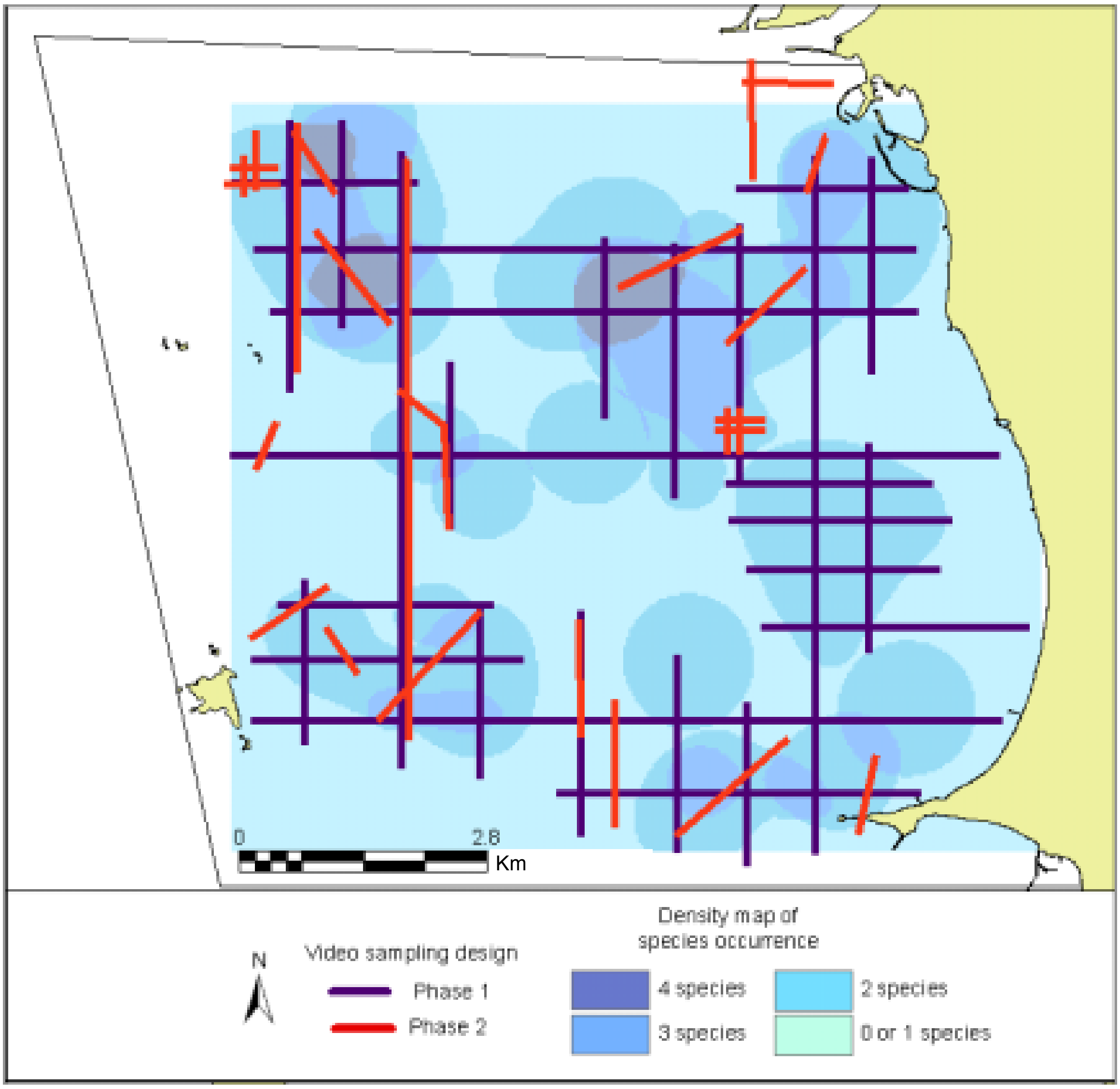

Figure D4. Sampling plan for 2004

The video acquisition software used for the survey recorded GPS locations approximately every 5 seconds. To permit multiple interpretations of the same video frames, the frame in which the new GPS recording occurred was used for identifying species presence and percent coverage of seagrass. Given the large amount of video to be collected (6 days x 20 km/day = 120 km of video footage, 1 video frame every 5 seconds, average speed 3.5 km/hr = ~25 000 video frames with GPS referencing), the number of frames for interpretation were prohibitively high. We had one week allocated for video interpretation, which would allow for interpretation of approximately 2000–3000 video frames. The plan was to interpret 1 in every 10 GPS video frames to get systematic coverage over all video transects as a first cut, then to stratify the interpretations by species, and randomly select 20 frames within each of 7 species around which to interpret an additional 9 frames to document short distance variability, resulting in a total of 3250–4250 interpreted frames. This resulted in interpreted video frames approximately every 50 m across all transects, and stratified, unsystematic placement of clusters of frames to be identified at a later date.

Iterative sampling

In areas where little is known about the seafloor, it may be useful to plan several phases of sampling to allow the sampling design to be adapted according to what is identified in the early phases. By breaking the Owen Anchorage field time into two phases, we were able to roughly analyse the first set of samples collected, and adjust the design. First, as much of the systematic, broad coverage sampling was done as time permitted (Figure D4). The preliminary video footage was interpreted for the percent cover of seagrass species, reef, sand and wrack, plus presence of algae and epiphytes, and examined for video image quality. A spatially referenced dataset was constructed, and the number of samples of each species and its distribution assessed.

Plots of species count density (number of occurrences of a species divided by area) were constructed to assist in identifying areas with high variability in species, which therefore required a higher sampling density for mapping. The first standard deviation for species count density was chosen as the cut-off for presence or absence of the species, and made into binary maps. The binary density maps for all of the species were combined using map algebra to produce a final map showing the number of species estimated (based on Phase 1 sampling) across Owen Anchorage (Figure D4).

This map and further variogram analyses were used in determining the layout of the second phase of video sampling. Our overall purpose was to gather information fairly evenly across the study area, collect sufficient information to characterise areas with high spatial variability based on former mapping and knowledge of the region, and to collect information at some of the same locations observed in 1999 to permit a temporal comparison.

The final video dataset consists of approximately 110 km of video footage, out of which 2160 frames were categorised in terms of percent cover of seagrass species and substrate (Figure D5). The interpretation of another 1250 frames is planned to assess short-distance relationships in species distributions.

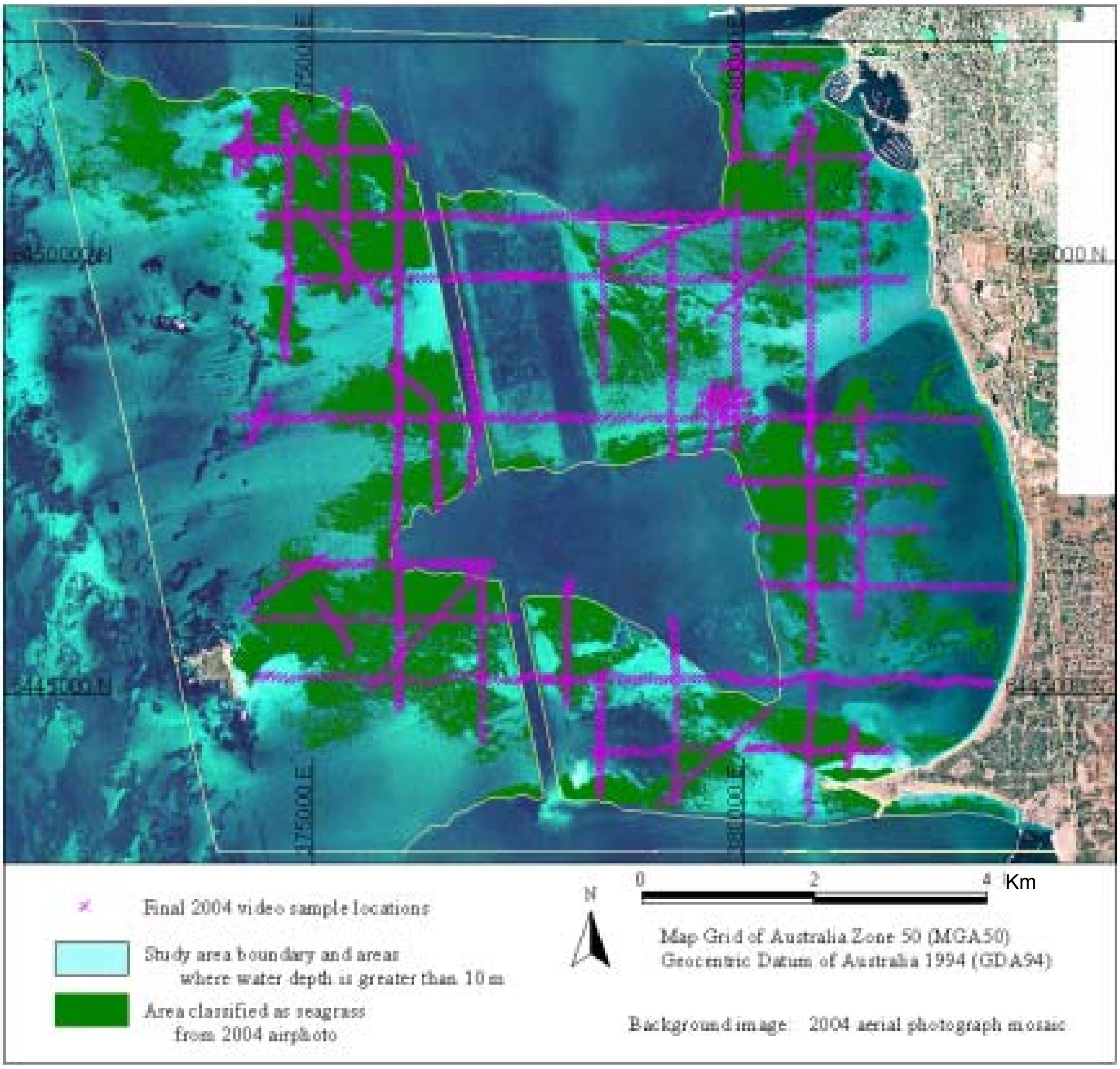

Figure D5. Final 2004 samples interpreted from the video footage

Impact of sampling design on map accuracy

Once the 2004 video footage was captured and interpreted, the presence/ absence data for seagrass species were interpolated as planned, and maps for both 1999 and 2004 were produced of the interpolation variance, which is roughly equivalent to map uncertainty arising from the interpolation methods used (Figure D6).

Figure D6. Map errors (kriging variance) from (a) 1999 and (b) 2004 sample distributions

High variance equates to high uncertainty in the model, meaning potentially large errors in the mapped distributions of species. The variance is tightly linked with the spatial configuration of the samples across the study area, and the number and distribution of samples in each 'neighbourhood' for interpolation. Therefore, areas with very few samples, particularly along the edges, have much larger estimation errors. A comparison of the 1999 and 2004 interpolation variance maps clearly shows that an even distribution of samples produces a higher accuracy map.

All the samples for 2004 were taken along transects, but because the interpreted video frames were well distributed and covered the major portion of the study area, the final result was an improvement over the map of the 1999 data. In fact, in 1999, about 2800 video identifications of seagrass species were made from the video footage for mapping, and in 2004 only 1915 were made. Although fewer video frames were interpreted in 2004, the overall interpolation error was lower because of the improved sample layout. Of course, it must be noted that the 1999 sampling was not designed specifically for kriging, but it is a good illustration of the difference in resulting map accuracy from the sampling design.

Sampling recommendations

Some sampling recommendations arising from the Owen Anchorage study include:

- Sampling designs must be tailored to the specific study objectives and nature of phenomenon of interest.
- By careful planning and leveraging off previously collected information, fieldwork expenses can be reduced and, at the same time, our ability to interpret the results can be improved. Use all prior information collected to help optimise the locations for sampling. Use any known relationships between the phenomenon of interest and bathymetry, or potential identification from imagery, hydrodynamic data, etc. to assist in stratifying the study area, thereby reducing redundant sampling.
- Use 'back of the envelope' uncertainty analysis on the planned methods for surveying and mapping, and concentrate efforts and/or resources on the main sources of error.
- The research objectives must drive the sampling design, not the other way around (Solan *et al*, 2003).
- A broad regional view is usually very instructive before focussing in on specific areas, either for more intensive research, choosing monitoring sites, etc. This requires appropriate broad scale mapping techniques, including sampling design.

Appendix D references

Atkinson, P.M. (1997) Scale and spatial dependence. In: P.P. van Gardingen, G.M. Foody & P.J. Curran (Editors), *Scaling up from cell to landscape.* Society for Experimental Biology, seminar series 63. Cambridge University Press, Cambridge.

Austin, M. (1985) Continuum concept, ordination methods and niche theory. *Annual Review of Ecology and Systematics*, 16: 39–61.

Austin, M. & Heyligers, P. (1989) Vegetation survey design for conservation: Gradsect sampling of forests in north-eastern New South Wales. *Biological Conservation*, 50: 13–32.

Bax, N., Kloser, R., Williams, A., Gowlett-Holmes, K. & Ryan, T. (1999) Seafloor habitat definition for spatial management in fisheries: A case study on the continental shelf of southeast Australia. *Oceanologica Acta,* 22(6): 705–720.

Bulit, C., Diaz-Avalos, C. & Montagnes, D.J.S. (2004) Assessing spatial and temporal patchiness of the autotrophic ciliate *Myrionecta rubra*: A case study in a coastal lagoon. *Marine Ecology Progress Series*, 268: 55–67.

Burgess, T.M. & Webster, R. (1984) Optimal sampling strategies for mapping soil types. I. Distribution of boundary spacings. *Journal of Soil Science,* 35: 641–654.

Burrough, P.A. (1995) Spatial aspects of ecological data. In: R.H.G. Jongman, C.J.F. Ter Braak & O.F.R. Van Tongeren (Editors), *Data analysis in community and landscape ecology.* Cambridge University Press, Cambridge, pp. 213–251.

Caeiro, S., Painho, M., Goovaerts, P., Costa, H. & Sousa, S. (2003) Spatial sampling design for sediment quality assessment in estuaries. *Environmental Modelling & Software*, 18: 853–859.

Cambridge, M.L., Chiffings, A.W., Brittain, C., Moore, L. & McComb, A.J. (1986) The loss of seagrass in Cockburn Sound, Western Australia. II. Possible causes of seagrass decline. *Aquatic Botany,* 24: 269–285.

Cochran, W.G. & Cox, G.M. (1957) *Experimental designs*. 2nd edition. John Wiley & Sons Inc., New York.

Collins, S.L., Glenn, S.M. & Roberts, D.W. (1993) The hierarchical continuum concept. *Journal of Vegetation Science*, 4: 149–156.

DAL (DA Lord & Associates Pty Ltd), Botany Department (University of Western Australia), Alex Wyllie & Associates Pty Ltd, NGIS Australia & Kevron Aerial Surveys Pty Ltd (2000) *Seagrass mapping, Owen Anchorage and Cockburn Sound, 1999.* Report No. 94/026/S3/2, prepared for Cockburn Cement Ltd.

Diaz, R.J., Solan, M. & Valente, R.M. (2004) A review of approaches for classifying benthic habitats and evaluating habitat quality. *Journal of Environmental Management*, 73: 165–181.

Durako, M.J., Hall, M.O. & Merello, M. (2002) Patterns of change in the seagrass-dominated Florida Bay hydroscape. In: J.W. Porter & K.G. Porter (Editors), *The Everglades, Florida Bay, and coral reefs of the Florida Keys: An ecosystem sourcebook*. CRC Press, Washington DC, pp. 523–537.

Dutilleul, P. (1993) Spatial heterogeneity and the design of ecological field experiments. *Ecology*, 74(6): 1646–1658.

Flatman, G.T., Englund, E.J. & Yfantis, A.A. (1988) Geostatistical approaches to the design of sampling regimes. In: L.H. Keith (Editor), *Principles of environmental sampling*. American Chemical Society, Washington DC, pp. 73–84.

Fonseca, M., Whitfield, P.E., Kelly, N.M. & Bell, S.S. (2002) Modelling seagrass landscape pattern and associated ecological attributes. *Ecological Applications*, 12(1): 218–237.

Franklin, J. (1995) Predictive vegetation mapping: Geographic modelling of biospatial patterns in relation to environmental gradients. *Progress in Physical Geography,* 19(4): 474–499.

Freitas, R., Rodrigues, A.M. & Quintino, V. (2003) Benthic biotopes remote sensing using acoustics. *Journal of Experimental Marine Biology and Ecology*, 285–286: 339–353.

Friedlander, A., Sladek Nowlis, J., Sanchez, J.A., Appeldoorn, R., Usseglio, P., McCormick, C., Bejarano S. & Mitchell-Chui, A. (2003) Designing effective marine protected areas in Seaflower Biosphere Reserve, Colombia, based on biological and sociological information. *Conservation Biology*, 17(6): 1769-1784.

Garza-Perez, J.R., Lehmann, A. & Arias-Gonzalez, J.E. (2004) Spatial prediction of coral reef habitats: Integrating ecology with spatial modelling and remote sensing. *Marine Ecology Progress Series*, 269: 141–152.

Gleason, H. (1926) The individualistic concept of the plant association. *Bulletin of the Torrey Botanic Club*, 53: 7–26.

Gleason, H. (1939) The individualistic concept of the plant association. *American Midlands Naturalist*, 21: 92–110.

Goovaerts, P. (1997) *Geostatistics for natural resources evaluation.* Oxford University Press Inc., New York, 483 pp.

Greenwood, J.J.D. (1996) Basic techniques. In: W.J. Sutherland (Editor), *Ecological census techniques: A handbook*. Cambridge University Press, Cambridge, pp. 11–110.

Hewitt, J.E., Thrush, S.F., Cummings, V.J. & Turner, S.J. (1998) The effect of changing sampling scales on our ability to detect effects of large-scale processes on communities. *Journal of Experimental Marine Biology and Ecology,* 227: 251–264.

Hirzel, A. & Guisan, A. (2002) Which is the optimal sampling strategy for habitat suitability modelling? *Ecological Modelling*, 157: 331–341.

Hurlbert, S. (1984) Pseudoreplication and the design of ecological field experiments. *Ecological Monographs,* 54: 187–211.

James, P. (1967) On the origin and persistence of error in geography. *Annals of the Association of American Geographers*, 57(1): 1–24.

Jassby, A.D., Cole, B.E. & Cloern, J.E. (1997) The design of sampling transects for characterising water quality in estuaries. *Estuarine, Coastal and Shelf Science*, 45: 285–302.

Kendrick, G.A., Aylward, M., Hegge, B.J., Cambridge, M.L., Hillman, K.A., Wyllie, A. & Lord, D.A. (2002) Changes in seagrass coverage in Cockburn Sound, Western Australia between 1967 and 1999. *Aquatic Botany*, 73: 75–87.

Kendrick, G.A., Eckersley, J. & Walker, D.I. (1999). Landscape-scale changes in seagrass distribution over time: A case study from Success Bank, Western Australia. *Aquatic Botany,* 65: 293–309.

Kendrick, G.A., Hegge, B.J., Wyllie, A., Davidson, A. & Lord, D.A. (2000) Changes in seagrass cover on Success and Parmelia Banks, Western Australia between 1965 and 1995. *Estuarine, Coastal and Shelf Science*, 50: 341–353.

Kenny, A.J., Cato I., Desprez M., Fader G., Schuttenhelm R.T.E. & Side J. (2003) An overview of seabed-mapping technologies in the context of marine habitat classification. *Journal of Marine Science*, 60: 411–418.

Kitsiou, D., Tsirtsis, G. & Karydis, M. (2001) Developing an optimal sampling design: A case study in a coastal marine ecosystem. *Environmental Monitoring and Assessment*, 71: 1–12.

Legendre, P. (1993) Spatial autocorrelation: Trouble or new paradigm? *Ecology*, 74(6): 1659–1673.

Legendre, P., Troussellier, M., Jarry, V. & Fortin, M.-J. (1989) Design for simultaneous sampling of ecological variables: From concepts to numerical solutions. *Oikos*, 55(1): 30–42.

McBratney, A.B. & Webster, R. (1983) How big a sample is needed for regional estimation of soil properties? *Soil Science*, 135: 177–183.

McIntosh, R. (1967) The continuum concept of vegetation. *The Botanical Review*, 33: 130–187.

McRea Jr., J.E., Greene, H.G., O'Connell, V.M. & Wakefield, W.W. (1999) Mapping marine habitats with high resolution sidescan sonar. *Oceanologica Acta*, 22(6): 679–686.

Oline, D.K. & Grant, M.C. (2002) Scaling patterns of biomass and soil properties: an empirical analysis. *Landscape Ecology*, 17(1): 13–26.

Plotkin, J.B. & Muller-Landau, H.C. (2002) Sampling the species composition of a landscape. *Ecology,* 83(12): 3344–3356.

Plummer, A., Morris, L., Blake, S. & Ball, D. (2003) *Marine natural values study, Victorian Marine National Parks and Sanctuaries*. Parks Victoria Technical Series No. 1, Parks Victoria, Melbourne.

Rueda, M. (2001) Spatial distribution of fish species in a tropical estuarine lagoon: A geostatistical appraisal. *Marine Ecology Progress Series,* 222: 217–226.

Snedecor, G.W. & Cochran, W.G. (1989) *Statistical methods*. Iowa State University Press, Ames.

Solan, M., Germano, J.D., Rhoads, D.C, Smith, C., Michaude, E., Parry, D., Wenzhofer, F., Kennedy, R., Henriques, C., Battlea, E., Carey, D., Iocco, L., Valente, R., Watson, J. & Rosenberg, R. (2003) Towards a greater understanding of pattern, scale, and process in marine benthic systems: a picture is worth a thousand worms. *Journal of Experimental Marine Biology and Ecology,* 285–286: 313–338.

Stevens, T. & Connolly, R.M. (2004) Testing the utility of abiotic surrogates for marine habitat mapping at scales relevant to management. *Marine and Freshwater Research,* 56:111–123.

Urbanski, J.A. & Szymelfenig, M. (2003) GIS-based mapping of benthic habitats. *Estuarine, Coastal and Shelf Science*, 56: 99–109.

Valavanis, V.D., Georgakarakos, S., Kapantagakis, A., Palialexis, A. & Katara, I. (2004) A GIS environmental modelling approach to essential fish habitat designation. *Ecological Modelling*, 178: 417–427.

Webster, R. & Oliver, M.A. (1992) Sample adequately to estimate variograms of soil properties. *Journal of Soil Science*, 43: 177–192.

Wiens, J.A. (1989) Spatial scaling in ecology. *Functional Ecology,* 3: 385–397.

Zajac, R.N. Lewis, R.S., Poppe, L.J., Twichell, D.C., Vozarik, J., DiGiacomo, M.L. & Cohen, G. (2000) Relationships among sea-floor structure and benthic communities in Long Island Sound at regional and benthoscape scales. *Journal of Coastal Research*, 16(3): 627–640.

Appendix E. Integration across scale: Applying sound spatial sampling principles to mapping for Parks Victoria at Point Addis

Purpose

Design underwater tow video sampling in the deep water habitat mapping effort underway at Point Addis. The major consideration for planning is to collect the best possible dataset for benthic habitat classification while accommodating field logistical issues (number, orientation and length of tracks). All available georeferenced datasets were used to optimise the design in terms of:

(1) characterising the full extent of the study area

(2) capturing the range of habitats present

(3) minimising the effort required for data collection, in terms of ease of vessel steering in relation to prevailing winds and wave direction.

Available datasets used for design

- Existing habitat map (Mnp_habitats_depths.shp; PV archive)
- Park extent (PV archive)
- Bathymetry (2005, Fugro)
- Shallow water video location and interpretation (2005, MAFRI)
- Deep water preliminary video location and interpretation (2005, Fugro, University of Western Australia)

Desired datasets not currently available

- Mosaicked sidescan imagery
- Classified sidescan and bathymetry
- Shallow water habitat classification
- Shallow water video footage
- 2005 aerial photography

Design

Video observations collected for the shallow-water habitat mapping effort (MAFRI, 2005) were included in the planning process, but covered few areas of overlap with the 2005 bathymetry information and areas considered 'deep water' (Figure E1; Table E1A). A small amount of video footage was available for deeper water areas at Point Addis (collected 11 March, 2005) (Figure E1; Table E1B). This included eight short segments, the selection criteria for which have not been described by the field crew. As the existing video information covered a small percentage of the total deep water study area, and it is unknown how representative it is of the park in general, a statistical analysis of the interpreted data was not considered appropriate for designing the larger-scope video sampling plan. Rather, a logical approach to sampling design based on the estimated amount of footage to be collected, known orientation of ecological gradients and geomorphic characteristics and field logistics were used to design the video track layout.

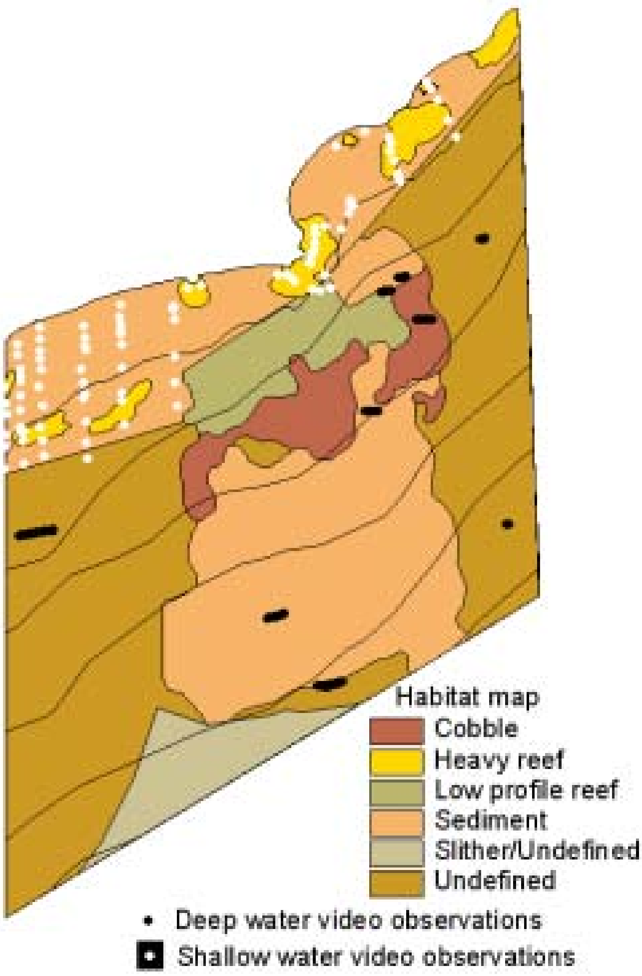

Figure E1. Existing habitat map for Point Addis National Marine Park showing the locations of underwater tow video observations

Table E1. Summary of the broadscale habitat classifications from video observations in (A) shallow and (B) deep water

A. Shallow-water video classification

Substratum type	Substratum category		No. of observations
Rock/reef	High profile reef		20
	Low profile reef		55
		Subtotal	75
Sediment	Unvegetated sediment		24
	Vegetated sediment		3
		Subtotal	27
		TOTAL	**102**

B. Deep-water video classification

Substratum type	Substratum category		No. of observations
Rock/reef	High profile reef		39
	Low profile reef		57
		Subtotal	96
Rock/reef – Sediment	High profile reef		4
	Low profile reef		3
	Low profile reef/Unvegetated sediment		5
	Low profile reef/Vegetated sediment		2
		Subtotal	14
Sediment	NA		1
	Unvegetated sediment		105
	Vegetated sediment		38
		Subtotal	143
		TOTAL	**254**

The basic principles behind the sample layout are to get fairly regular, unbiased coverage of the entire park, to gather seafloor information over a variety of spatial scales in multiple directions, particularly those with the most frequent (spatial) changes in habitat, and to collect footage in areas of interest identified from the preliminary bathymetric maps. Perhaps more important than sample distribution is sample georeferencing. This will allow direct comparison of the video observations with the other available datasets as well as predictive mapping of marine habitats.

(1) Characterising the full extent of the study area

A large amount of field time was allocated for the video collection because of the potentially poor weather in May. Assuming 5 hours of video collection for 5 days, with the boat travelling on average 1.5 knots, 70 km of video footage can realistically be collected. Point Addis covers a 5 km x 8 km area, so to get regular coverage over the full extent of the park, transects would be placed approximately 1 km apart in a regular grid. Because more rapid change in ecological and geomorphic features is expected perpendicular to the coastline, a larger percentage of the video tracks were planned in that orientation (65% of estimated total video footage), and a smaller amount oriented orthogonally (35%). These numbers have been used as rough estimates to guide video track placement (Figure E2, Table E2). The lengths of tracks plotted are not overly precise, as the actual data logged in the field will likely deviate from the plan somewhat.

(2) Capturing the range of habitats present

The large change in bathymetry perpendicular to the coast is considered the major environmental factor likely to influence seafloor habitat characteristics, in terms of physical oceanographic and biological processes controlling bedforms and distribution of benthic flora and fauna. The orientation of the video tracks takes this into account by including a larger proportion of video in the direction of maximum change (see Point 1 above).

The preliminary bathymetry dataset, roughly interpreted by overlaying the 2005 deep water classified video footage, shows a variety of textural features likely related to patterns of bedrock weathering, including paleo-shoreline features, and bedforms in unconsolidated sediments. Ideally, these geomorphic features would be classified from the hydroacoustics, and video coverage proportionately allocated to provide information about all geomorphic and depth classes. The bathymetry and sidescan mosaic will eventually be classified to more clearly visualise the different physiographic regions, but the methods have not yet been sufficiently refined to produce a proto-habitat map before the video sampling is scheduled to begin. However, a shaded relief map of the bathymetry was consulted to ensure all textural features of interest would be recorded by video.

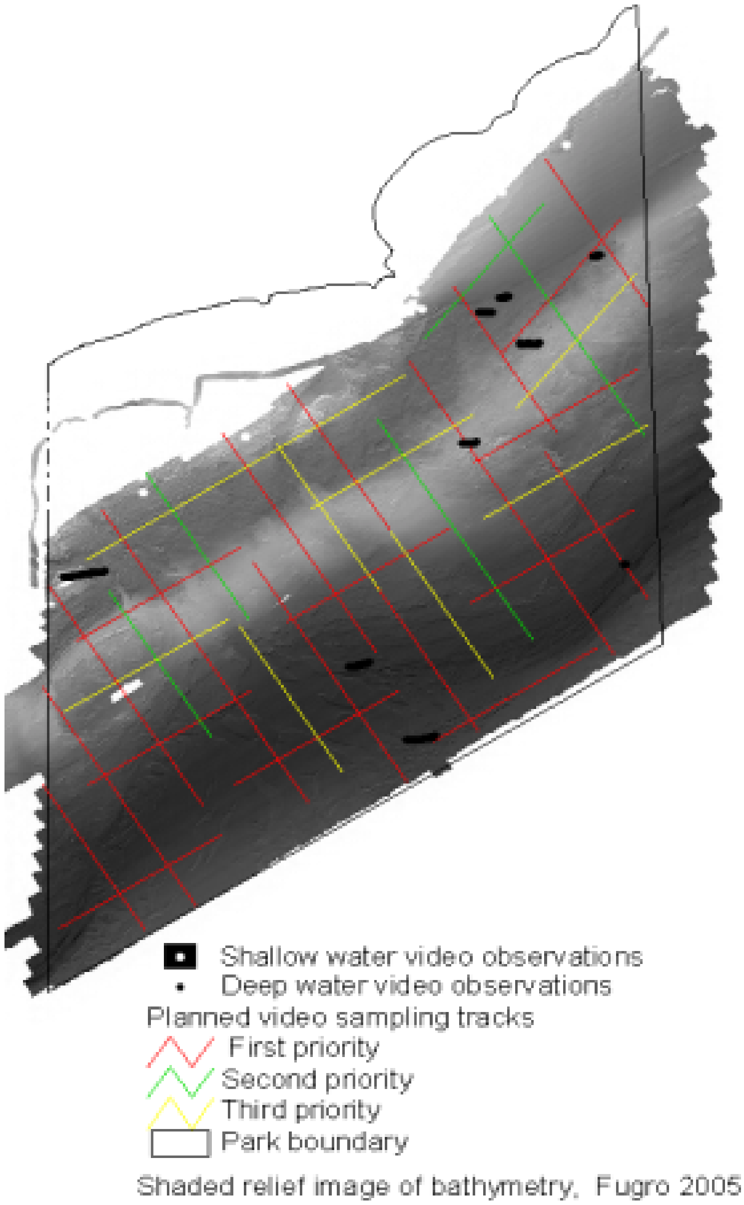

Figure E2. Sampling plan for underwater tow video coverage at Point Addis National Marine Park, May 2005

Table E2. Amount (metres) of planned video collection summarised by priority and orientation

Priority	Orientation to coastline	Distance (m)	Proportion (%)	Total distance (m)
1	Perpendicular	32 530	38	
	Parallel	22 799	27	55 329
2	Perpendicular	10 170	12	
	Parallel	2 051	2	12 221
3	Perpendicular	6 102	7	
	Parallel	12 374	14	18 476
	TOTAL		**100**	**86 026**

(3) Minimising the effort required for data collection

All video tracks were designed to run parallel and perpendicular to the shoreline line, as was done for the previous hydroacoustic survey (parallel to coast) and the shallow-water video collection (perpendicular to the coast, under the assumption that wave conditions travelling in these directions will be suitable for data collection. The planned tracks have been prioritised so those areas considered critical can be targetted first, in case of weather or equipment failure which could limit the duration of data collection. Tracks estimated to require >20 hours in the field have been categorised as high priority, 5–20 hours as medium priority, and <5 hours as lowest priority. Straight lines and parallel tracks have been plotted as the planned track layout, although the actual field data collected is not expected to be as neatly aligned. This is a natural consequence of having to adapt data collection to the weather and wave conditions.