CRC for Coastal Zone Estuary and Waterway Management
Project CB3: Benthic Biology and Habitat Mapping
Task 2.1 Milestone Report, May 2004

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

Researchers: Karen W. Holmes[1,2], Kimberly Van Niel[2], and Katrina Baxter[1,2]
Sub-project Leader: Gary Kendrick[1]
[1]School of Plant Biology
[2]School of Earth and Geographical Sciences
University of Western Australia, 35 Stirling Highway, Crawley, WA 6009

Table of 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 characterize 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. Finally, we offer general recommendations to optimise sampling protocols for research projects that involve mixed data types in the coastal zone.

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 and Harborne, 1999), although there also exists a large number of more anecdotal descriptions of the application of acoustic (Brown et al., 2002; Cochrane and Lafferty, 2002; Kenny et al., 2003; Kvitek et al., 1999; McRea Jr. et al., 1999; Penrose and 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 near shore habitat mapping over 100's to 1000's of square kilometers. 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 (eg., IKONOS, with 4m pixels) is now available, but at a greater cost. Manned systems (eg. SKYLAB and space shuttles) can also provide useful data, but lack temporal repeatability. Highly specialized 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 specialized 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

Side scan 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 200m 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 20cm resolution at a 50m range drops to 40cm at the 100m 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 / 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: Cost estimates are approximate. Expense varies greatly depending on start up costs, minimum purchase requirements, travel costs, etc.

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 and Edwards, 2002) Best for large areas (> 500km^2)
SPOT	$ A$0.10/ km^2	100 +	60 x 60 km	20 m	Georef. ~10 m	800 - 2000 m^2	(Mumby and Edwards, 2002; Mumby et al., 1998) Better than Landsat for < 3600 km^2
IKONOS	$$ A$45 - 69 / km^2	100+	1 km^2, min 100 km^2	4 m, 1 m panchromatic	Georef. ~ 3m	32 - 80 m^2	(Mumby and Edwards, 2002) Good for identifying boundaries, not better than Landsat for habitats.
Airborne Sensors							
Aerial photography	$$	10+	W = depends on survey	0.05 - 1m	GPS, Ground control	0.2 – 1 m^2	
CASI	$$$	10+	W = 512 m	1 – 10m, typically 1 x 2 m	GPS, 2 m	4 - 10 m^2	(Kvitek et al., 1999; Mumby and Edwards, 2002) Best for areas < 100km^2 Water < 30m deep
LIDAR	$$$$ US$3100-3900 /$km^2$	10+	W = 50m 8 -32 km^2 /hr	4 m	3m horiz (improved with DGPS) 0.15m vertical	32 - 80 m^2	(Kvitek et al., 1999) < 20 m water depth, to 60m possible

LLS (TOWED, Not airborne)	$$$$$?	10	W = 4 – 65 (depends on water clarity)	0.1 - 3 cm	DGPS error	< 1cm^2	(Kvitek et al., 1999; Solan et al., 2003) >3m water, < 1500 m Depends on field conditions
Acoustic Methods							
Side Scan Sonar	$$	5	W = up to 300 m	0.5 m	GPS, layback	2 - 5 m^2	(Kvitek et al., 1999) Distortion across track >2m water depth, gap under fish Easily distorted by waves, etc. Depends on frequency used and field conditions
Multibeam Sonar	$$$$$	3.5	W = 2 x water depth	0.10 m	DGPS error	0.08 – 0.20 m^2	(Penrose and Siwabessy, 2001) (Kvitek et al., 1999) > 3m water depth Depends on equipment and field conditions
Field remote sensing*							
Videography	$$	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 – 25m^2	Small	GPS	Depends on FOV, small	(McDonald, 1998) < 30 m water depth vid 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 not only interested 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 and 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 and Cox, 1957; Snedecor and 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 fertilizer 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 and 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 and 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 neighboring 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 and 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 and 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 and 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 and 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 and 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 and 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 realization of a random field (Borgman and Quimby, 1988; Brus and 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 Regionalized 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 (regionalized variable); and (3) spatially uncorrelated random noise, or residual error (Burrough and 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 and 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, permit spatial analysis, and 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 and 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 (eg. seed propagation method).

Random sampling designs

Simple random sampling: Locations for sample collection are selected randomly, using a random number generator (see Van Niel and 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 and 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 under sampled, 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 maximize the variation between units, and minimize the variation within each unit. The one or more strata selected are expected to be major drivers of the system under study, i.e. 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 and 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 strata are sampled. For example, the field area consists of 33 sand patches and 24 seagrass patches, but only 5 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 and 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 randomizing 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 and Webster, 1983). They can also be used to assess variability of map units, or boundary spacing (Burgess and Webster, 1984a). Nested transects have proven useful for exploring differences in process and pattern over multiple scales (Oline and Grant, 2002).

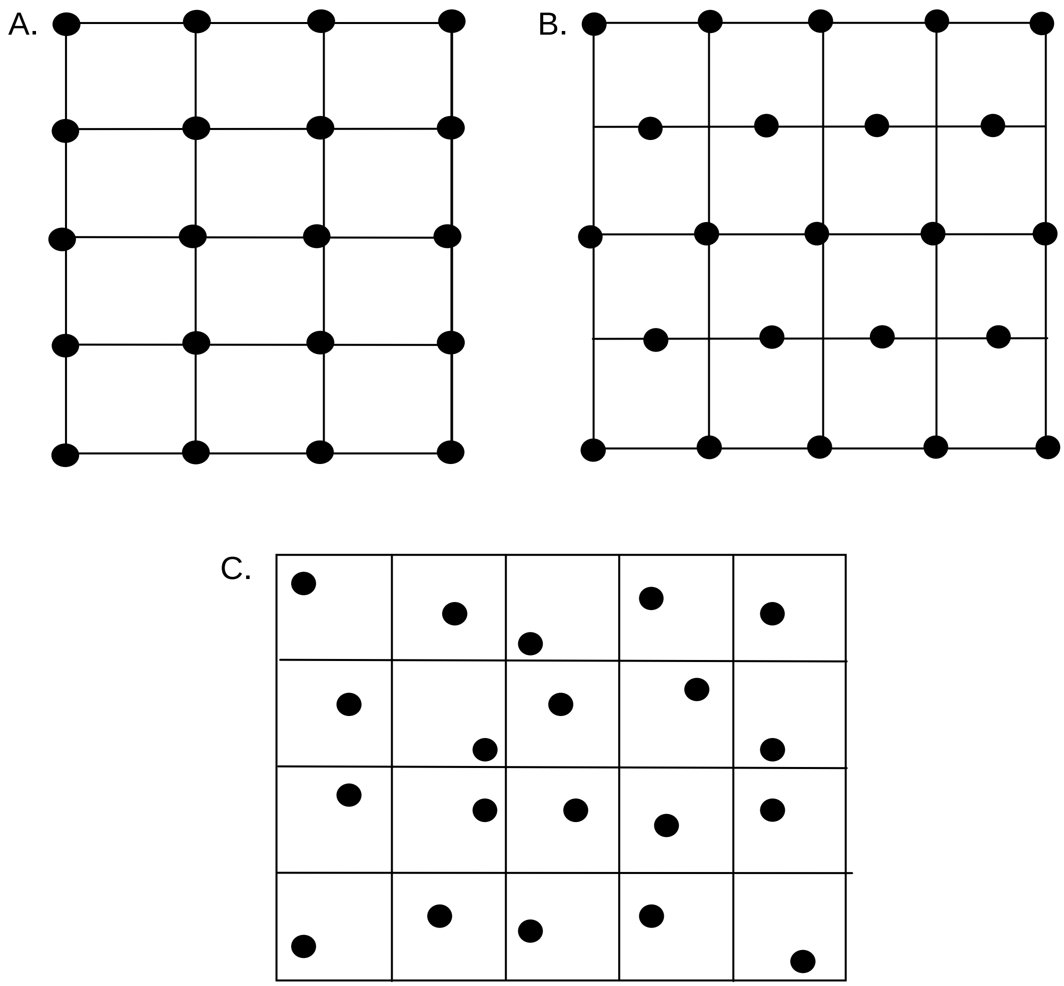

Figure 2. Grid types: A) square grid; B) equilateral triangle grid; and C) unaligned grid.

Number of samples

The number of samples necessary to characterize 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 and Cochran, 1989; p82 in Greenwood, 1996; Burrough, 1995. If sampling cost can be estimated, it can be also be taken into consideration to optimize 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 and Oliver, 1992; Webster and Oliver, 2001), which can be extremely useful for optimizing 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 optimizing sampling efficiency in terms of global estimation or testing for minimum sample support (Heidelbaugh and Nelson, 1996). Sample support can be adjusted through aggregation and disaggregation procedures.

Pre-sampling, or two-phase sampling

No sampling design can be optimized without prior information concerning the variability of the variable of interest, and particularly for spatial sampling two phases of sampling are recommended (Borgman and Quimby, 1988; Legendre et al., 1989). Stratified random nested sampling is probably the best approach for gathering sufficient data to determine spatial variability (Webster and 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 optimize 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 modeled 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):

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

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*.

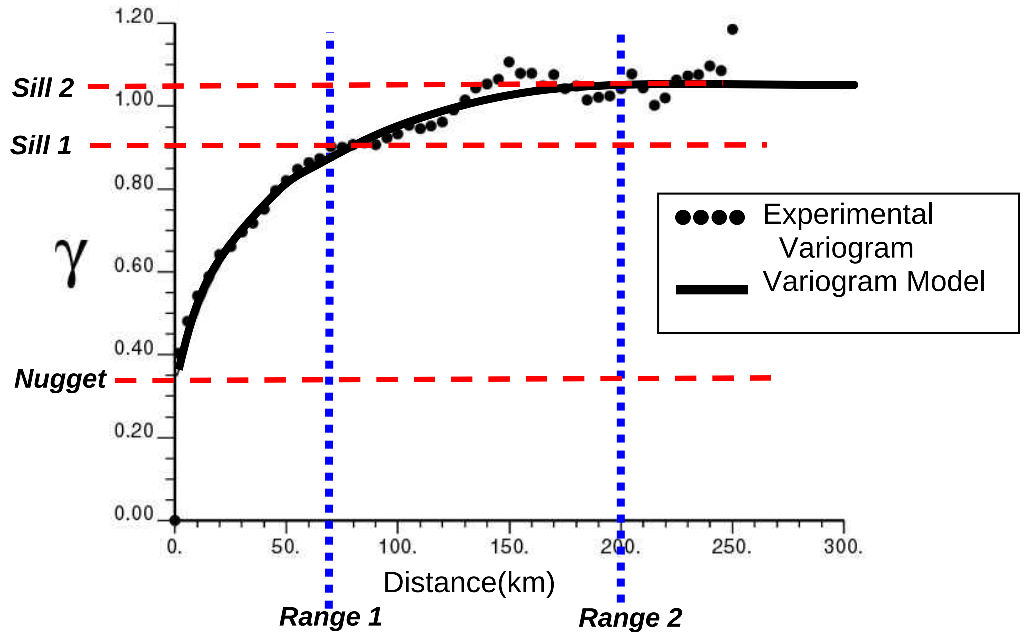

Figure 2: Example of an experimental variogram (black dots) which has been fitted using a spherical model (black line). In this case, two model structures were fitted, resulting in a nested variogram with two sills and two ranges, as shown by the red and blue lines, respectively.

Spatial correlation can be visualized and measured (modeled) 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 ($\gamma = 0.34$) which accounts for more than a third of the overall variance (sill2, $\gamma = 1.02$). 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 analyzing 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, 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, 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 and Webster, 1983), nested transects (Oline and Grant, 2002), and stratified nested sampling (Burrough and McDonnell, 1998; Webster and 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 finalizing 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 grainsize, 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 optimize 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 and 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 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 optimize sampling design are found in the literature, including: characterizing 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 and Robertson, 2003), optimizing 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 and 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 behavior resulting from non-linear 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 and 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 artifacts 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, generalizing 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 following sections. In Appendix A we provide a brief introduction to scaling extent and grain using arithmetic, geostatistical, process modeling, 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.

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 modeling 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 modeling, 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 behavior 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 behavior; (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:

$$l = m(z) \tag{2}$$

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

$$l = m(g(x, y)) = f(x, y) \tag{3}$$

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:

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

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 centered on location x_0 :

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

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*):

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

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 emphasize 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. Neighboring values are used to estimate aggregated area (larger support, or block) mean values in areas where little or no measurements exist (Goovaerts, 1997). While this approach optimizes estimation in areas with few measurements, it also leads to over smoothing of the data, because neighboring 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 over smoothing, and may not represent the true correlation. The preferable method for extending sample support is conditional simulation. Conditional simulation utilizes 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 realizations 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 modeling. 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 characterize non-linear behavior (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 characterize 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 modeling 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 regularizing data support for all levels may be difficult.

IV. Nested hierarchical scaling

The landscape (and seascape) has commonly been conceptualized as a hierarchy which contains different object sizes and patterns at particular scales resulting from different processes acting over defined ranges of scale (Allen and 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 and 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, and detecting and modeling system non-linearities. Wu and 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 modeling at focal levels; and (3) extrapolating information across the domains of scale.

The approach accommodates all types of modeling and extrapolation methods, but 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 categorize 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 and Edwards, 2002; Mumby and Harborne, 1999), sidescan and automated classification at intermediate scales (Brown et al., 2002; Cochrane and Lafferty, 2002; Kenny et al., 2003; Pinn and 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 analyzed 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 a 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 customized for a given field area, research objectives, and method of data collection, but there are some general field sampling principles and theoretical guidelines which can help structure sampling design. The purpose of the survey and existing data availability typically determine whether one or more sampling campaigns are necessary and the appropriate sample layout for the planned analyses. Unfortunately sampling costs, in terms of both money (equipment rental, field and lab personnel, sample analysis) and time, are often the main factor that determines the number of samples to be collected.

- 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, 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 minimize sample collection and laboratory analysis.
- If using multiple sampling techniques, overlap the data collection in each stratified zone (if identified) to allow full characterization of the field area, calibration of the different data types, and comparison of variance.
- Allow for time after the first sampling campaign to analyze the data and define second sampling phase based on gaps in data or high variance.
- Use semivariogram analysis to determine sample spacing, or optimize selection of few locations for intensive temporal monitoring (Flatman et al., 1988; Webster and Oliver, 2001).
- Track the measurement and aggregation error in the data, and consider stochastic simulation to assess error propagation through modeling 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. and Costa, M.H., 2002. Characterizing patterns of agricultural land use across Amazonia by merging satellite images and census data. Global Biogeochemical Cycles, 16(3): no. 1045.

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

Zhu, J., Morgan, C.L.S., Norman, J.M., Yue, W. and 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 characterization of high-resolution NDVI patterns for precision agriculture. International Journal of Applied Earth Observation and Geoinformation, 3, 121-132.

References

Allen, T.F.H. and 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 and 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. and 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. and 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. and 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. and 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. and Webster, R., 1984a. Optimal sampling strategies for mapping soil types. I. Distribution of boundary spacings. Journal of Soil Science, 35: 641-654.

Burgess, T.M. and Webster, R., 1984b. Optimal sampling strategies for mapping soil types. II. Risk functions and sampling intervals. Journal 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 and O.F.R. Van Tongeren (Editors), Data Analysis in Community and Landscape Ecology. Cambridge University Press, Cambridge, pp. 213-251.

Burrough, P.A. and McDonnell, R.A., 1998. Principles of Geographical Information Systems. Spatial Information Systems. Oxford University Press, New York, 333 pp.

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

Cochran, W.G. and Cox, G.M., 1957. Experimental Designs. John Wiley & Sons, Inc., New York.

Cochrane, G.R. and 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. and 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. and 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. x -.

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

Heidelbaugh, W.S. and 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. and Roberts, D.A., in review. Tropical land-cover change impacts on multi-scale soil nutrient variability. Biogeochemistry.

Isaaks, E.H. and 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. and Cloern, J.E., 1997. The design of sampling transects for characterizing water quality in estuaries. Estuarine, Coastal and Shelf Science, 45: 285-302.

Kenny, A.J. et al., 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 and R.H. Gardner (Editors), Quantitative Methods in Landscape Ecology. Ecological Studies 82. Springer-Verlag, New York, pp. 479-517.

Kitsiou, D., Tsirtsis, G. and 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. and Tzilkowski, W.M., 2002. Autologistic regression modeling of American Woodcock habitat use with spatially dependent data. In: M.J. Scott et al. (Editors), Predicting Species Occurrences: Issues of Accuracy and Scale. Island Press, Washington D.C., pp. 335-344.

Kvitek, R. et al., 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. and 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. and 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, 155pp pp.

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

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

Mumby, P.J. and 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. and Turner, T., 1997. Estimating basal area coverage of subtidal seagrass beds using underwater videography. Aquatic Botany, 58: 269-287.

Oline, D.K. and 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. and 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. and 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. and 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. and Stanford, M., 2000. Optimizing remotely sensed solutions for monitoring, modeling, and managing coastal environments. Remote Sensing of the Environment, 73: 117-132.

Pinn, E.H. and 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. et al., 2003. The ups and downs of benthic ecology: considerations of scale, heterogeneity and surveillance for benthic-pelagic coupling. Journal of Experimental Marine Biology and Ecology, 285-286: 191-203.

Rastetter, E.B. et al., 1992. Aggregating fine-scale ecological knowledge to model coarser-scale attributes of ecosystems. Ecological Applications, 2(1): 55-70.

Snedecor, G.W. and Cochran, W.G., 1989. Statistical Methods. Iowa State University Press, Ames.

Solan, M. et al., 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. et al., 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. and 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. and 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. and Oliver, M.A., 1992. Sample adequately to estimate variograms of soil properties. J. Soil Sci., 43: 177-192.

Webster, R. and 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. and David, J.L., 2002. A spatially explicit hierarchical approach to modeling complex ecological systems: theory and applications. Ecological Modelling, 153: 7-26.