

Analysis of transect data for microhabitat correlations and faunal patchiness

Richard J. Malatesta¹, Peter J. Auster², Bradley P. Carlin³

¹Sea Education Association, PO Box 6, Woods Hole, Massachusetts 02543, USA

²NOAA's National Undersea Research Center, The University of Connecticut at Avery Point, Groton, Connecticut 06340, USA

³Division of Biostatistics, School of Public Health, Box 197, Mayo Memorial Building, University of Minnesota, Minneapolis, Minnesota 55455, USA

ABSTRACT: Distributional studies of benthic megafauna using trawls, direct count, still photographic, and video techniques are all spatially dependent, and traditional analyses which assume independence between observations (e.g. correlation, regression) are not appropriate. We describe a statistical procedure which determines if a species distribution is correlated with microhabitat type and, if so, subsequent patch size analysis is partitioned by microhabitat. Both microhabitat correlation and patch size procedures utilize a randomization protocol which allows for tests of statistical significance, while other spatial statistics have only descriptive value. An example using video transect data taken at a 55 m deep Northwest Atlantic outer continental shelf site (40° 50' N, 70° 55' W) demonstrates the types of ecological questions this analysis can generate.

INTRODUCTION

Factors which control the distribution of benthic megafauna have received a good deal of attention in the recent past (e.g. Scott 1982, Holme & Wilson 1985, Schneider et al. 1987, Kaufmann et al. 1989, Mahon & Smith 1989, Auster et al. 1991). All of these studies, using a variety of sampling methodologies (e.g. trawling, photographic surveys), indicate that sediment type and microhabitat characteristics (e.g. sand wave crests and troughs, biogenic depressions, shell) play a major role in faunal distributions.

Trawl sampling has commonly been used to assess large-scale faunal distributions (e.g. Scott 1982, Colvocoresses & Musick 1984, Overholtz & Tyler 1985, Horne & Campana 1989, Mahon & Smith 1989). Principal component and cluster analysis techniques have been applied to trawl data to determine similarity between samples and between sets of samples in order to define species assemblages over large geographic scales (e.g. Colvocoresses & Musick 1984, Overholtz & Tyler 1985, Horne & Campana 1989, Mahon & Smith 1989). Unfortunately these analytical techniques do not permit significance testing (e.g. Ludwig & Reynolds 1988). Furthermore, the minimum spatial resolution is on the

scale of the tow length, often hundreds or thousands of meters. Faunal studies at this scale often characterize substrates by a minimal number of grab samples (e.g. Horne & Campana 1989) or through comparison of trawl locations with large-scale surficial geology charts (e.g. Scott 1982, Mahon & Smith 1989). These procedures can not discern the effect of small-scale habitat features on faunal distribution.

Still-photographic techniques have been very productive in gaining an understanding of small-scale processes and interactions with specific environmental variables (e.g. Grassle et al. 1975, Smith & Hamilton 1983, Cooper et al. 1987, Schneider et al. 1987, Kaufmann et al. 1989, Langton & Uzmann 1989). However, the use of intermittent still photographs from a towed camera sled or submersible also has its limitations. Noncontiguous photographic quadrats confound statistical analysis, and variability in camera height from the bottom effects photographic resolution. These problems can be eliminated via direct counts and bottom type characterizations using SCUBA along the length of an entire transect (e.g. Luckhurst & Luckhurst 1978, Bell & Galzin 1984, Roberts & Ormond 1987, Yoshioka & Yoshioka 1989). While direct counts and measurements produce high quality data, the

techniques are generally limited by depth constraints and available bottom time.

The use of video technology is one way to increase data acquisition over limited time periods. Video transects eliminate the problem of image separation inherent to still-photographic sampling while also allowing immediate assessment of sample quality. Recent advances in video camera and recording technology allow recording of a continuous stream of images with resolution and color saturation sufficient to identify fauna and determine fine-scale substrate characteristics (e.g. Auster et al. 1989, 1991).

Transect data are not easy to deal with statistically. Individual observations (e.g. quadrats) taken along the transect line have an implicit dependence which comes from the distinct spatial relationship between neighboring data points. This invalidates the use of most statistical analyses (e.g. correlations, regressions) which contain the assumption of independence between observations.

A variety of statistical approaches to transect data, pioneered in plant communities (reviewed by Greig-Smith 1983), have been applied to marine benthic studies as well (e.g. Grassle et al. 1975, Jumars 1975, 1976, Langton & Uzmann 1989, Yoshioka & Yoshioka 1989, Langton & Robinson 1990, Langton et al. 1990). The purpose of these techniques is to classify the distribution of a species as overdispersed (even), underdispersed (clumped), or randomly distributed. If clumping can be established, the spatial scale at which clumps occur can be determined by techniques such as variance-to-mean ratios or blocked quadrat variances (Ludwig & Reynolds 1988). These techniques have a variety of descriptive uses, although they do not have any statistical application. A scale of clumping can be established by using one of the above mentioned techniques, for example, but there is no way to test if 2 species clump on significantly different scales. Herein, we describe a statistical procedure to quantify microhabitat-faunal relationships along contiguous linear transects. The procedure provides statistically testable values for correlations of fauna and microhabitat type as well as faunal patch size.

MATERIALS AND METHODS

The analysis covers 2 discrete questions: (1) for a given species, are the number of animals observed correlated with microhabitat features, and (2) given this correlation (or lack thereof) what are the linear scales of aggregation, if any?

The first question must be answered before the second can be addressed, since clumped faunal distributions can be mistaken for correlations between a

species and clumped microhabitat distributions. It is important to separate microhabitat correlations and actual clumping at the start of the analysis. Attempting to perform multiple correlations on data grouped for aggregation analysis introduces a phenomena known as the 'ecological fallacy', a problem first pointed out by Robinson (1950; also see Bendel & Carlin 1990). This paradox is essentially that sample correlation relationships computed at the ecological (aggregate) level are often substantially higher than those same correlations computed at the individual level. Intuitively, what is happening is that by aggregating we are 'averaging out' the variability in our data, replacing the original data pairs by fewer pairs having a more nearly linear relationship. To illustrate the extreme case, suppose the number of quadrats grouped (L) is equal to one-half the number of samples, i.e. $L = n/2$. Then after aggregating we would have only 2 groups, hence 2 data points, and a straight line can always be fit to 2 points, resulting in $r(L) = \pm 1$. The proper way to handle the problem is to answer the correlation question at $L = 1$, and then address the scale of aggregation question in the absence of the microhabitat variable. The approach followed here is to first decide whether a species occurrence is correlated with a microhabitat, and if so, to perform all subsequent analysis by microhabitat type, thus separating these 2 sources of dependence.

Transect line correlation analysis. To account for the lack of independence between successive quadrats, we use the randomization protocol (*sensu* Barnard 1963) outlined below.

For a single species, suppose there are n images producing data pairs (X_i, Y_i) , $i = 1, \dots, n$, where X_i = no. of individuals of Species 1 observed in video frame image i , and Y_i = microhabitat type (i.e. $Y_i = 1$ corresponds to the microhabitat type of interest, $Y_i = 2$ to all other microhabitat types). To test for correlation between X and Y :

$$H_0: \rho_{XY} = 0 \quad \text{vs} \quad H_1: \rho_{XY} \neq 0 \quad (1)$$

the randomization test is applied. That is, the Pearson correlation coefficient r from the original data set is computed, then the frame counts X_i (but not the microhabitats Y_i) are randomly permuted and the sample correlation, r_j is recomputed. The randomization step is repeated $j = 1, \dots, B$ times, and the level α test is to reject H_0 in favor of H_1 if the randomized values r_j exceed the original r in absolute value in at least $(1 - \alpha)B$ of the cases.

If H_0 is not rejected, the conclusion is that this species is not associated with a microhabitat type, and subsequent analysis may be performed without regard to habitat. If H_0 is rejected, however, then subsequent

analysis must be performed by microhabitat to isolate the question of scale of aggregation from the confounding microhabitat variable Y .

Patch size analysis. Taking a transect with faunal counts and microhabitat types (or, in the case of significant correlation in Eq. 1 above, just the data for a particular microhabitat) a scale of aggregation for the species must be selected to test for significance. The 'patchiness' of a species is measured here (as does Schneider et al. 1987) by comparing the ratio of the sample's variance (S) to mean (\bar{X}) obtained after grouping consecutive frames L at a time. This statistic, otherwise known as the coefficient of dispersion (Blackman 1942), is denoted by

$$I(L) = \frac{S_L^2}{\bar{X}_L} \quad (2)$$

where the no. of adjacent frames grouped, L , serves as a surrogate for the size of the patch. The statistic $I(L)$ has historically been used to measure patchiness since in the case of random Poisson distributions, $I(L) = 1$ for all L . This approach is somewhat unsatisfactory for any $L > 1$ due to the 'starting point dependent' nature of $I(L)$. To illustrate, suppose the frame counts X_i are

{0,0,0,9,9,9,0,1,0,8,9,9}.

In this simple example $L = 3$ is clearly the best choice of a scale of aggregation. By grouping this data 3 points at a time starting with the first observation, we obtain $I(3) = 5.57$. But if the data are altered only slightly to

{0,0,9,9,9,0,1,0,8,9,9,0},

and are again grouped starting with the first observation, $I(3) = 0.67$ is obtained. The solution to this artifact is to recompute $I(L)$ for each of the L possible starting places and take the largest value for the statistic.

Given a hypothesized value l for the true scale of aggregation L , the following may be tested:

H_0 : significant aggregation at scale l
vs H_1 : no significant aggregation at scale l

using a randomization procedure similar to the one following Eq. (1) above. That is, permute the X_i counts, group them l at a time, and obtain the resulting $I_j(l)$. Repeat the process $j = 1, \dots, B$ times, and reject H_0 in favor of H_1 if at least $(1 - \alpha)B$ of the randomized $I_j(l)$ values exceed the original $I(l)$.

Data collection. Video transects were performed with the 2 person submersible DSRV Delta (Delta

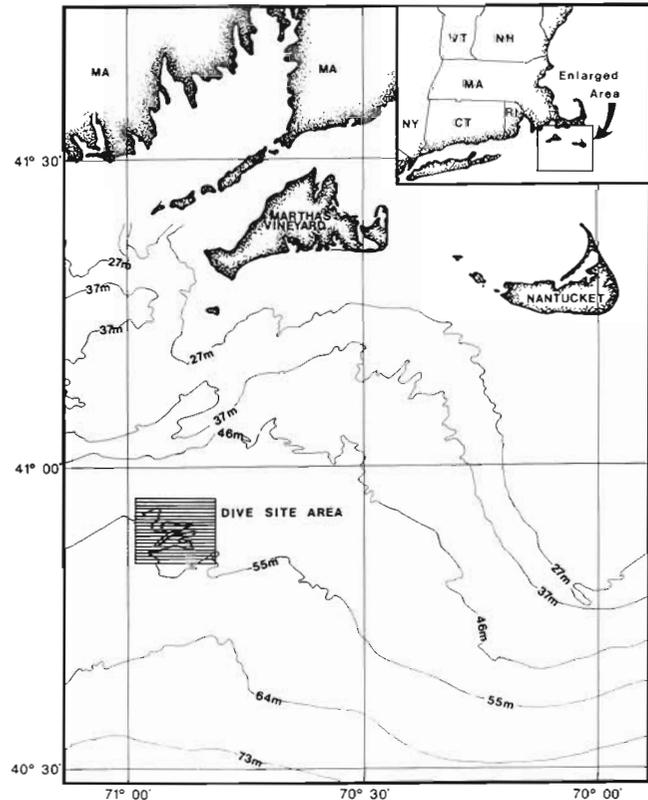


Fig. 1. Location of the submersible transects

Oceanographics, Torrance, California, USA) on October 8, 1989. The study site is an area known locally as 'The Fingers' (approximately 40° 50' N, 70° 55' W; Fig. 1). This site has been characterized in a previous study (Auster et al. 1991). Transects were performed by skidding the submersible over the bottom on a steady compass heading at a speed of ca 25 cm s⁻¹.

Each video transect was treated as a series of non-overlapping, adjacent, trapezoidal quadrats 1.54 m² (each quadrat being 1.12 m along the axis of the transect). Each adjacent video frame image was enumerated for all megabenthic species (i.e. all visible fauna) and given a categorical score for microhabitat type. Four microhabitat types were common in the area of study: flat sand with amphipod tubes, sand wave crest, shell (dense patches and individual valves of the ocean quahog *Arctica islandica*), and biogenic depression (in flat sand or around shell). Three transects, spanning the range of microhabitat types, are analyzed here (Fig. 2). Transect 1 is composed of sand and sand waves; Transect 2, sand and shell hash; and Transect 3, sand, sand wave, shell as well as biogenic depression microhabitats. Significant associations of megafaunal organisms to these microhabitat types have been previously demonstrated (Auster et al. 1991).

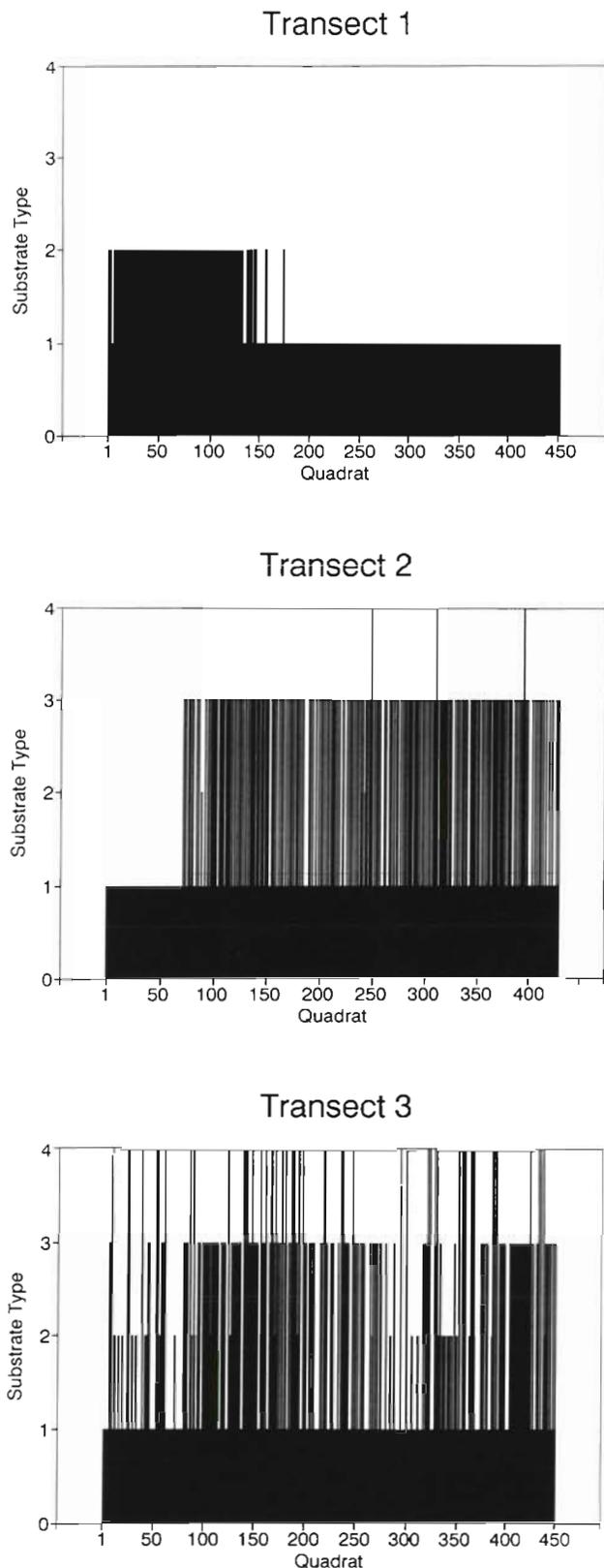


Fig. 2. Histograms of microhabitat types which occurred along the 3 transects. Four types were discerned: 1. flat sand; 2. shell; 3. sand wave; and 4. biogenic depression

RESULTS

While all epibenthic megafauna were quantified during video analysis, only 2 species groups occurred at sufficiently high densities to analyze for microhabitat correlation and spatial pattern. These were the asteroid starfish *Asterias vulgaris* and the caridean shrimps *Crangon septemspinosa* and *Dichelopandalus leptocerus*. Resolution was not sufficient to differentiate the 2 shrimp species so they were grouped as a single taxon.

Faunal-microhabitat correlations were performed on both taxa for all 3 transects (Table 1). Significant positive correlations were found for starfish with sand in Transects 1 and 3, and with sand waves in Transect 2. Shrimp showed a positive correlation for sand and a corresponding negative correlation for shell in Transect 1 but showed no other microhabitat affinities in the other transects.

Patch size analysis was performed in 2 steps. First the transect data were analyzed using the 2-term local quadrat variance (TTLQV) method (Hill 1973). Block sizes no longer than 10 % of the total transect length were used, due to this test's lack of reliability beyond this range (Ludwig & Reynolds 1988). While the TTLQV procedure showed areas of possible clumping, it did not provide the ability to test for statistical significance. Use of this analysis did provide an excellent *a priori* block size to secondarily apply our randomization technique.

Fig. 3 illustrates the results of the TTLQV analysis of faunal blocking by transect. Transect 1 had a continuous segment of sand and another segment of shell with very little mixing of the 2 microhabitats (Fig. 2). As such, the transect was divided into 2 segments and analyzed separately for patch size. Transects 2 and 3 had a mix of microhabitat types over very short linear distances (Fig. 2). Ultimately, this was not a problem as all microhabitat types, including shell 'patches', showed no blocking at the scales we examined, although a steady increase in variance indicates that they may block at some larger scale.

Shrimp consistently showed significant clumping at smaller scales, between block size 2 (2.24 m linear distance along the transect) and 9 (10.08 m), over all transects. Starfish showed no tendency to clump in Transects 1 and 3 but showed significant clumping at a larger scale, block size 21 (23.52 m), in Transect 2.

DISCUSSION

The analysis of spatial patterns in transect data has a long history in the botanical literature (Ashby 1935, Greig-Smith 1952, 1983, Dale & MacIsaac 1989) and is

Table 1. Correlation matrix for species-microhabitat comparisons by transect. (+) and (-): significant positive or negative faunal-substrate associations (p values appear in parenthesis); NS: no significant correlation (B = 500 for the randomization correlation analysis)

Transect no.	Taxa	Flat sand	Shell	Sand wave crests	Biogenic depressions
1	Starfish	+ (<0.001)	- (<0.001)		
	Shrimp	+ (<0.001)	- (<0.001)		
2	Starfish	- (0.002)		+ (0.004)	
	Shrimp	NS		NS	
3	Starfish	+ (0.012)	NS	NS	- (0.040)
	Shrimp	NS	NS	NS	NS

increasingly being performed as one aspect of many benthic studies (e.g. Schneider 1987, Schneider et al. 1987, Yoshioka & Yoshioka 1989, Langton & Robinson 1990, Langton et al. 1990). Transect procedures are especially attractive for deep water work since continuous video image data are relatively easy to acquire during towed sled, remote-operated vehicle and manned submersible operations. Contiguous quadrat data are necessarily non-independent, however, and parametric analyses are inappropriate. Instead, the nonparametric 'randomization-type' techniques introduced here are more suitable for this data. These

techniques are flexible, and many statistical procedures (e.g. regressions, ANOVA's) can be adapted to this technique (Schneider et al. 1987, Bendel & Carlin 1990). Applying these procedures provides a powerful statistical tool since patterns can now be examined for statistical significance. An example of this is seen in the second transect. In this sand and shell environment, 2 peaks in the TTLQV are shown (Fig. 3). Starfish clump significantly ($p < 0.001$) at a block size of 21 (23.52 m) while a similar appearing block size of 9 (10.01 m) for shrimp is not significant ($p = 0.516$). Fine-scale gradations in pattern, such as the 2 peaks

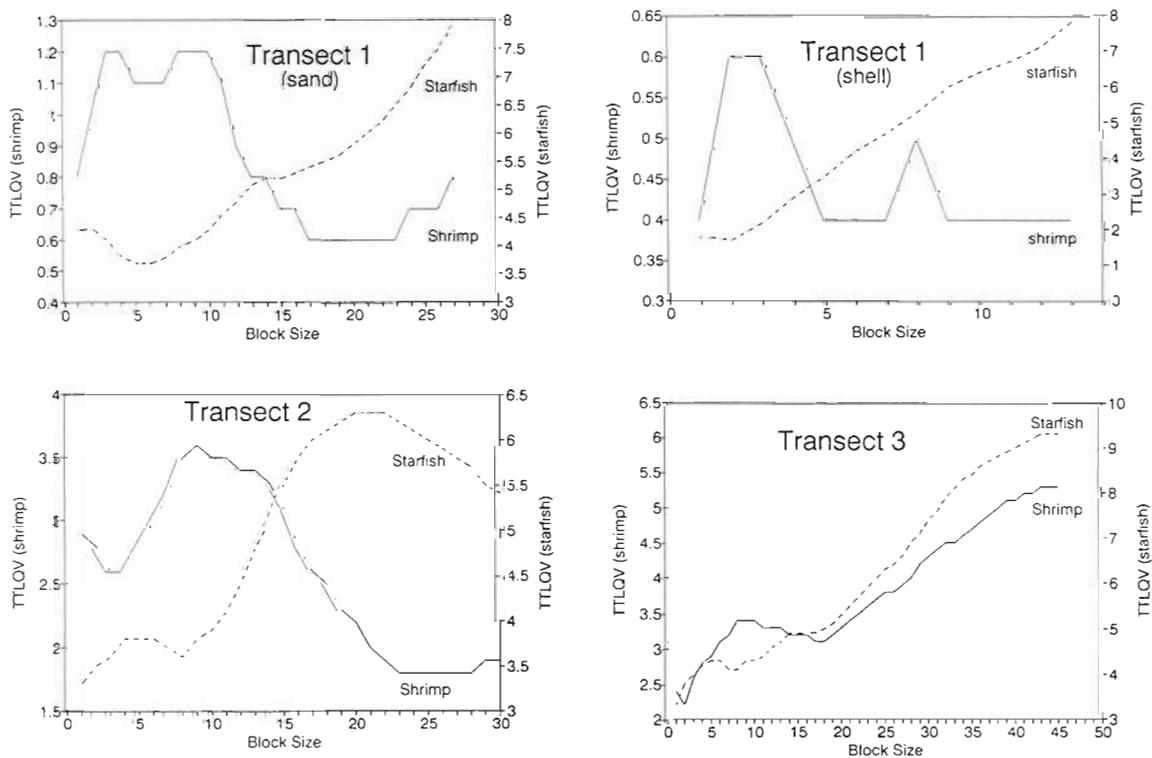


Fig. 3. Two-term local quadrat variance (TTLQV) analysis of Transects 1 to 3. Transect 1 is split into sand and shell microhabitats. Transects 2 & 3 were not split (see text). Peaks in variance estimates represent points which are analyzed for significant blocking

shown by the shrimp on sand areas of Transect 1 can also be examined. Here block size 4 (4.48 m) is significant ($p = 0.014$) while a second peak at block size 9 (10.08 m) is only marginal ($p = 0.088$).

The ability to ask questions regarding correlation with microhabitat type and scales of patchiness allows ecologically relevant questions to be formed. The most interesting of these which arose during this analysis was the apparent 'flip' which starfish exhibited in their association with flat sand (Table 1). Starfish were positively associated with sand in Transects 1 and 3, while showing a negative association in Transect 2. This 'flip' could be interpreted as unrelated to sand per se, but rather as a negative response to biogenic structures (such as shells and depressions) and a positive response to sand waves. Starfish may avoid biogenic structures such as shell due to difficulties maneuvering in and around them. Additionally, starfish may tend to clump on the down-current portion of sand waves due to differential prey abundance on those structures.

Acknowledgements. This study was funded by NOAA's National Undersea Research Center at the University of Connecticut at Avery Point (NOAA Award No. NA90AA-DUR128). The views expressed herein are those of the authors and do not necessarily reflect the views of NOAA or any of its sub-agencies. The authors acknowledge the excellent work of the crews of both the RV 'J.W. Powell' and DSRV 'Delta'. The tedious job of translating video images into data was performed with much patience by Susan LaRosa. An anonymous reviewer provided much appreciated comments on the manuscript.

LITERATURE CITED

- Ashby, E. (1935). The quantitative analysis of vegetation. *Ann. Bot.* 49: 779-802
- Auster, P. J., Stewart, L. L., Sprunk, H. (1989). Scientific imaging with ROVs: tools and techniques. *Mar. Technol. Soc. J.* 23(3): 16-20
- Auster, P. J., Malatesta, R. J., LaRosa, S. C., Cooper, R. A., Stewart, L. L. (1991). Microhabitat utilization by the megafaunal assemblage at a low relief outer continental shelf site - Middle Atlantic Bight, USA. *J. Northw. Atl. Fish. Sci.* 11: 59-69
- Barnard, G. A. (1963). Comment on 'The Spectral Analysis of Point Processes.' by M. S. Bartlett. *J. R. statist. Soc. B.* 25: 294
- Bell, G. D., Galzin, R. (1984). Influence of live coral cover on coral-reef fish communities. *Mar. Ecol. Prog. Ser.* 15: 265-274
- Bendel, R. B., Carlin, B. P. (1990). Bayes methods in the ecological fallacy context: estimation of individual correlation from aggregate data. *Commun. Statist., Ser. A* 19: 2595-2623
- Blackman, G. E. (1942). A study by statistical methods of the distribution of species in grassland associations. *Ann. Bot.* 49: 749-777
- Colvocoresses, J. A., Musick, J. A. (1984). Species associations and community composition of Middle Atlantic Bight continental shelf demersal fishes. *Fish. Bull. U.S.* 82: 295-313
- Cooper, R. A., Shepard, A. N., Valentine, P., Uzmann, J. R., Hulbert, A. (1987). Pre and post drilling bench marks and monitoring data of ocean floor fauna, habitats, and contaminant loads on Georges Bank and its submarine Canyons. *Symp. Ser. for Undersea Res. NOAA, Office of Undersea Research, Rockville, MD*, 2(2): 17-48
- Dale, M. R. T., MacIsaac, D. A. (1989). New methods for the analysis of spatial pattern in vegetation. *J. Ecology* 77: 78-91
- Grassle, J. F., Sanders, H. L., Hessler, R. R., Rowe, G. T., McLellan, T. (1975). Pattern and zonation: a study of the bathyal megafauna using the research submersible 'Alvin'. *Deep Sea Res.* 22: 457-481
- Greig-Smith, P. (1952). The use of pattern analysis in ecological investigations. Recent advances in botany 2. University of Toronto Press, Toronto
- Greig-Smith, P. (1983). Quantitative plant ecology. University of California Press, Berkeley
- Hill, M. O. (1973). The intensity of spatial pattern in plant communities. *J. Ecology* 61: 225-236
- Holme, N. A., Wilson, J. B. (1985). Faunas associated with longitudinal furrows and sand ribbons in a tide-swept area in the English Channel. *J. mar. biol. Ass. U.K.* 65: 1051-1072
- Horne, J. K., Campana, S. E. (1989). Environmental factors influencing the distribution of juvenile groundfish in nearshore habitats of southwest Nova Scotia. *Can. J. Fish. Aquat. Sci.* 46: 1277-1286
- Jumars, P. A. (1975). Methods for measurement of community structure in deep-sea macrobenthos. *Mar. Biol.* 30: 245-252
- Jumars, P. A. (1976). Deep-sea species diversity: does it have a characteristic scale? *J. mar. Res.* 34: 217-246
- Kaufmann, R. S., Wakefield, W. W., Genin, A. (1989). Distribution of epibenthic megafauna and lebensspuren on two central North Pacific seamounts. *Deep Sea Res.* 36: 1863-1896
- Langton, R. W., Langton, E. W., Theroux, R. B., Uzmann, J. R. (1990). Distribution, behavior, and abundance of sea pens, *Pennatula aculeata*, in the Gulf of Maine. *Mar. Biol.* 107: 463-469
- Langton, R. W., Robinson, W. E. (1990). Faunal associations on scallop grounds in the western Gulf of Maine. *J. exp. mar. Biol. Ecol.* 144: 157-171
- Langton, R. W., Uzmann, J. R. (1989). A photographic survey of the megafauna of the central Gulf of Maine. *Fish. Bull. U.S.* 87: 945-954
- Luckhurst, B. E., Luckhurst, K. (1978). Analysis of the influence of substrate variables on coral reef fish communities. *Mar. Biol.* 49: 317-323
- Ludwig, J. A., Reynolds, J. F. (1988). Statistical ecology. Wiley-Interscience, New York
- Mahon, R., Smith, R. W. (1989). Demersal fish assemblages on the Scotian Shelf, Northwest Atlantic: spatial distribution and persistence. *Can. J. Fish. Aquat. Sci.* 46: 134-152
- Overholtz, W. J., Tyler, A. V. (1985). Long-term responses of the demersal fish assemblages of Georges Bank. *Fish. Bull. U.S.* 83: 507-520
- Roberts, C. M., Ormond, R. F. G. (1987). Habitat complexity and coral reef fish diversity and abundance on Red Sea fringing reefs. *Mar. Ecol. Prog. Ser.* 41: 1-8
- Robinson, W. S. (1950). Ecological correlations and the behavior of individuals. *Am. sociol. Rev.* 15: 351-357
- Schneider, D. C. (1987). Patch structure of benthic populations on an intertidal sandflat. *Oceanol. Acta* 10: 469-473

- Schneider, D. C., Gagnon, J. M., Gilkinson, K. D. (1987). Patchiness of epibenthic megafauna on the outer Grand Banks of Newfoundland. *Mar. Ecol. Prog. Ser.* 39: 1–13
- Scott, J. S. (1982). Selection of bottom type by groundfishes of the Scotian Shelf. *Can. J. Fish. Aquat. Sci.* 39: 943–947
- Smith, C. R., Hamilton, S. C. (1983). Epibenthic megafauna of a

- bathyal basin off southern California: patterns of abundance, biomass, and dispersion. *Deep Sea Res.* 30: 907–928
- Yoshioka, P. M., Yoshioka, B. B. (1989). A multispecies, multi-scale analysis of spatial pattern and its application to a shallow-water gorgonian community. *Mar. Ecol. Prog. Ser.* 54: 257–264

This article was presented by K. Sherman, Narragansett, Rhode Island, USA

Manuscript first received: April 2, 1992
Revised version accepted: August 25, 1992