Implications for monitoring: Study designs and interpretation of results

Article *in* Canadian Journal of Fisheries and Aquatic Sciences · April 2011
DOI: 10.1139/cjfas-53-11-2629

2 authors:

Roger H Green
The University of Western Ontario

Paul A. Montagna
Texas A&M University - Corpus Christi

Implications for monitoring: study designs and interpretation of results

Roger H. Green and Paul Montagna

Abstract: A Sediment Quality Triad approach was used to integrate concurrently obtained toxicological, sediment contaminant, and benthic ecological data from GOOMEX in a test of coherence of responses. The biological community responded to contaminants by increases in the ratio of "the more tolerant taxon" to "the more sensitive taxon": the nematode to harpacticoid copepod ratio for meiofauna and the polychaete to amphipod ratio for macrofauna. Meiofaunal total abundance declined near platforms whereas macrofaunal total abundance increased — probably because of organic enrichment near platforms. A second approach used small-scale spatial heterogeneity of response as evidence of impact. Cd, Ba, ChemPC1 (first principal component of contaminant information), and the nematode to harpacticoid ratio had significant heterogeneity of variance among distances from platform, with higher variances near platforms. The pattern for Cd and ChemPC1 suggests that the impact may extend out to 200 m. We argue that differing objectives which are often demanded of the same study are often in conflict (e.g., generalization about environmental impact of similar platforms has different design requirements than description of spatial pattern of impact around each platform). Therefore, requiring that a study accomplish both objectives leads to compromises and a suboptimal design for both.

Résumé : Les auteurs ont eu recours, pour l'évaluation de la qualité de sédiments, à une approche triadique leur permettant d'intégrer, dans un test de cohérence des réponses obtenues, des données toxicologiques, d'autres sur des contaminants dans des sédiments et d'autres sur l'écologie en milieu benthique, obtenues concuremment dans le cadre du projet GOOMEX. La réaction de la communauté biologique aux contaminants a conduit à l'augmentation du ratio du « taxon le plus tolérant » au « taxon le plus sensible », soit celui entre les nématodes et les copépodes harpacticoïdes dans le cas de la méiofaune, et celui entre les polychètes et les amphipodes dans celui de la macrofaune. L'abondance totale de la méiofaune a diminué tandis que celle de la macrofaune a augmenté à proximité des plates-formes, probablement à cause d'un enrichissement organique près de celles-ci. Les auteurs ont appliqué une deuxième approche, soit celle axée sur l'hétérogénéité des réponses dans l'espace à petite échelle comme indice des impacts. Dans le cas du Cd, du Ba, du Chem PC1. (premier composant principal de l'information sur les contaminants) et du ratio nématodes au harpacticoïdes, ils ont observé une hétérogénéité d'importance significative au niveau de la variance entre les distances jusqu'aux plates-formes, les variances les plus marquées étant observées à faible distance de celles-ci. Dans le cas du Cd et du Chem PC1, les résultats paraissent indiquer que les effets pourraient se faire ressentir jusqu'à 200 m. Les auteurs soutiennent que les objectifs divergents qu'on se donne dans le cadre d'une même étude sont souvent incompatibles (p. ex., les critères de conception servant à une généralisation sur l'impact environnemental de plates-formes similaires diffèrent de ceux servant à la description de la distribution des impacts dans l'espace autour de chaque plate-forme). Par conséquent, de s'attendre à ce qu'une étude produise les deux types de résultats mène à l'adoption de compromis et à une conception d'expérience qui n'est optimale ni pour l'un ni pour l'autre des objectifs. [Traduit par la Rédaction]

Introduction

The design of an impact/monitoring study and statistical approaches to interpretation of results from that study are inseparable. There should be a logical flow of purpose $\rightarrow$ question $\rightarrow$ hypotheses $\rightarrow$ model $\rightarrow$ study design $\rightarrow$ statistical analysis $\rightarrow$ tests of hypotheses $\rightarrow$ interpretation and presentation of results (Green 1984). On the other hand, some statistical approaches for interpreting results will be used in almost any environmental study regardless of the study design, e.g., regression, correlation, tables of means and standard errors. In this paper the use of two innovative and recent statistical approaches to interpretation and generalization of results from studies of this kind are described (Kennicutt et al. 1996). The first of these is an adaptation of Chapman's Sediment Quality Triad (SQT) concept (see references below), and the second is the use of small-scale spatial heterogeneity of response in several variables as evidence of impact. Finally, the link between study objectives and study design will be discussed with emphasis on the inherent conflict between differing objectives in relation to optimal study design.

Received October 5, 1995. Accepted September 16, 1996. J13413

R.H. Green Department of Zoology, University of Western Ontario, London, ON N6A 5B7, Canada.
P. Montagna. Marine Science Institute, University of Texas, 750 Channelview Dr., Port Aransas, TX 78373, U.S.A.

SQT approach to integration and interpretation of results

The SQT concept involves the use of concurrently obtained (synoptic) toxicological, sediment chemistry, and benthic infauna data to determine environmental degredation (e.g., Long and Chapman 1985; Chapman et al. 1987; Chapman et al. 1989). In their application of this concept to evaluate effects associated with an oil platform, Chapman et al. (1991) argued that the SQT "incorporates three essential components: measures to determine the presence and degree of anthropogenic contamination (i.e., bulk sediment chemistry); measures that demonstrate that substances are present that can interfere with the normal functioning of at least some biological organisms tested in the laboratory (i.e., sediment toxicity tests); and, assessments of in situ alteration of resident biological communities (i.e., measures of benthic infaunal community structure).'' They noted that one of the weaknesses of the SQT is "lack of statistical criteria for the combined Triad components.'' Green et al. (1993) also pointed out that appropriate statistical methods had not been used in previous applications of the SQT concept to actual environmental impact situations and that appropriate methods would be those that could relate conceptually different sets of variables in multivariate data (Green 1993*a*). Such methods were then applied to data from a study (Chapman et al. 1989) designed to identify the degree of pollution-induced degradation at each station for 13 sampling stations in Vancouver Harbour, British Columbia, Canada.

Chapman et al. (1991) applied the SQT approach to data obtained from sampling along three radii centered on Platform "C" of the Matagorda 622 field, 22 km offshore in water depths ranging from 24 to 26 m. Benthic infaunal community data obtained from cores provided the biological component. Limited sediment chemistry data provided the contamination component, and the toxicological component consisted of an amphipod bioassay, an assessment of abnormalities in oyster larvae, and Microtox tests. As in earlier applications of the SQT (e.g., Chapman et al. 1989), the statistical analyses were mostly descriptive and applied to each component rather than as criteria for the combined components and were based on a "ratio-to-reference" criterion (the ratio of values at possibly impacted stations to values at an a priori defined reference station). Two other problems are that the experiment-wide error rate is not controlled and that the reasoning tends to be circular (the reference site is chosen and then used for tests of impact).

Here the methods described in Green et al. (1993) are applied to the GOOMEX data described in the preceding articles, in an application of the SQT concept. This is an attempt to integrate the diverse results from various work elements to describe and test for concordance. Most large environmental impact/monitoring studies with numerous principal investigators involved have failed to adequately interrelate the results and interpretations from the different work elements, such as description of the biological community, description of the physical/chemical environment, and laboratory toxicological results. Therefore, interesting and suggestive results often remain in each work element report and are not properly related to other results. This final integration step greatly increases the value of the overall study results.

Of course other paradigms have been described for relating biological response variables to environmental predictor variables (Ter Braak 1986; Clarke 1993; Clarke and Ainsworth 1993). Here we are limiting ourselves to accomplishing this within the SQT paradigm.

For the SQT components, we used the following: *biological*: abundances of two major meiofaunal taxa and two major macrofaunal taxa; *chemical*: the variables total polycyclic aromatic hydrocarbons, unresolved complex mixture, total alkanes, percent sand, percent silt, percent clay, total organic carbon, total inorganic carbon, redox potential, Fe, Cd, Al, and Ba; *toxicological*: a test of porewater toxicity using percent successful development of sea urchin embryos as a criterion.

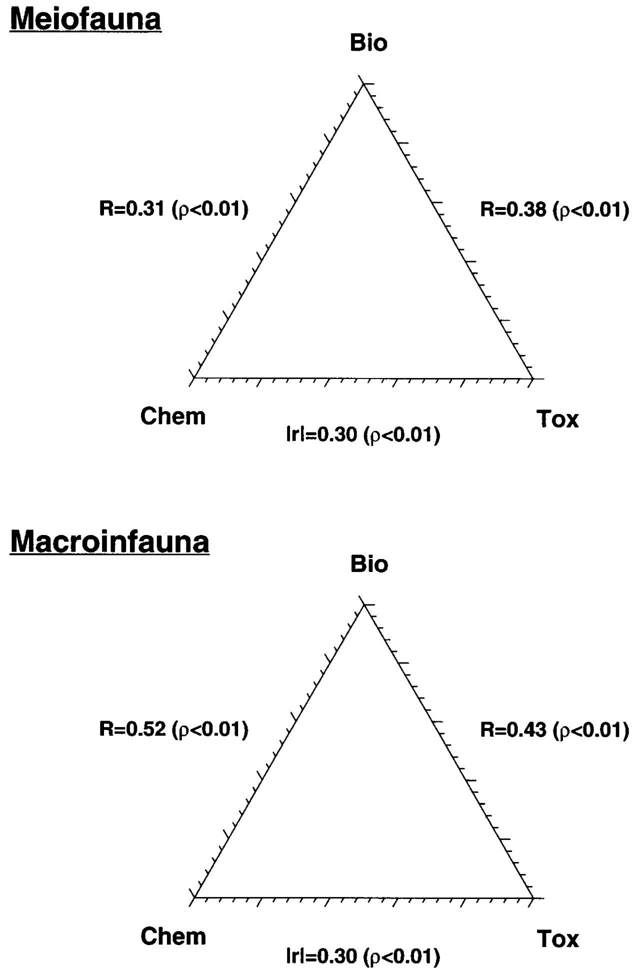
The sea urchin embryo development assay was done for samples from all stations for Cruises 1 and 2 (Carr et al. 1996). Therefore the data analyzed by the SQT approach were from samples taken on the first two cruises at all stations (n = 2 cruises × platforms × 5 radii × 5 distances = 150). The analysis approach was essentially that of section 3.2, pp. 446–450, of Green et al. (1993). There is only one toxicological variable, and the chemical contaminant information has already been reduced to one variable: the ChemPC1 score (see Kennicutt et al. 1996). Two analyses were carried out for the biological component. One was based on meiofaunal community response and the other on macrofaunal community response (Montagna and Harper 1996). For the meiofauna analysis, total harpacticoid abundance and total nematode abundance were used as the variables, reflecting the long-standing use of these taxa and their relative abundances as indicators of environmental impact (see Montagna and Harper 1996). The nematode to copepod (N/C) ratio was originally proposed to detect gradients of organic enrichment (Raffaelli and Mason 1981). The N/C ratio has been criticized as being too simplistic (Coull et al. 1981), but it was a predictor of oil pollution in South Africa (Hennig et al. 1983) and exposure to a gradient of natural petroleum seepage in California (Montagna et al. 1987, 1989). Harpacticoids are usually more sensitive to impact than are nematodes, so at impacted sites the N/C ratio usually increases. The abundances were log-transformed, so a difference between nematode and harpacticoid log abundance represents this ratio.

Analagously, for the macrofauna analysis we used total amphipod abundance and total polychaete abundance as variables. Amphipods decreased and polychaetes increased close to platforms, so the polychaete to amphipod (P/A) ratio increased as well (Montagna and Harper 1996). Deposit-feeding polychaetes and nematodes increase with proximity to platforms (perhaps caused by trophic enrichment), while amphipods and harpacticoids decrease near platforms (perhaps caused by toxicity). Again, abundances were log-transformed.

The SQT analysis was also performed with the biological component consisting of all four of these taxa. These results were statistically significant and consistent with the results from the analyses of the meiofauna and macrofauna separately. However, the results of the separate analyses were easier to interpret and therefore we will not present or discuss the "combined analysis" results. These SQT analyses were also carried out platform by platform and those results and their implications are summarized .

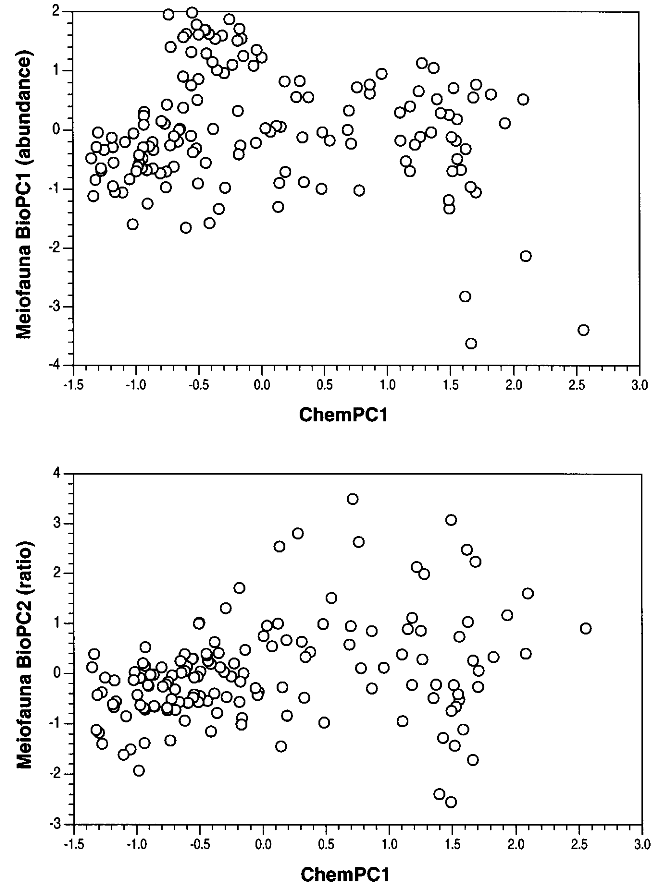
The statistical analysis was the same for all SQT analyses. First the correlations among components were calculated. We used Pearson correlations, but rank correlations could be used and we also calculated them for comparison in cases of apparent nonlinearity. The correlations between the biological and chemical components and between the biological and toxicological components were determined by the multiple correla-

Fig. 1. Significance and strength of relationships among SQT components.



tion ($\sqrt{R^2}$) of the nonbiological component with abundances of the two taxa which defined the biological component. The correlation between the chemical and toxicological components is the simple bivariate correlation coefficient *r*. Thus the significance and strength of each of the three between-component relationships are evaluated, and can be placed on the sides of the triangle representing the SQT (as in Fig. 1). Scatter plots showing these between-component relationships are presented. Second, the entire among-component correlation matrix is tested against the null hypothesis H_0: "no relationships among the three SQT components", using Bartlett's sphericity test. Third, a principal component analysis (PCA) is done on the correlation matrix to evaluate the structural relationship(s) among the three SQT components. This was a two-step procedure. First, a PCA on the two log-transformed taxa abundance variables was done to derive "size" and "shape" PCs. That is, in a PCA on such data, one of the PCs (the "size" PC) is invariably related similarly (same strength of relationship and same sign) to the original variables and represents a tendency for the variables to go up and down together. Here it would represent increases or decreases of abundance of both taxa, and presumably of the entire meiofaunal or macrofaunal community. The other PC (the "shape" PC) will have opposite-signed relationships (one positive and

Fig. 2. Relationship between BioPC1 and BioPC2 of the meiofauna abundance data and the contaminant variable ChemPC1.



one negative) to the original variables, which in these log-transformed data will represent changes in the ratio of the abundances of the two taxa. The PC scores for these two PCs (hereafter referred to as BioPCs) should have these interpretations and can used in the final PCA along with the chemical and toxicological measures. See Pimentel (1979) or Harris (1985) for an explanation of PCA and examples of interpretation of PCA results.

The results for the meiofauna-based analysis and for the macrofauna-based analysis are similar. The strength and significance of between-component SQT relationships in the two analyses are shown in Fig. 1. All three relationships in both analyses are highly significant ($p < 0.01$). Scatter plots showing the between-component relationships are shown in Fig. 2–4. Some (e.g., Fig. 4) are obviously nonlinear, or "threshold", relationships. For those, we also calculated rank correlation coefficients, which showed similar relationships (both strength and direction). For both analyses, Bartlett's sphericity test is highly significant, with χ^2 (6 df) = 43.9 ($p < 0.01$) for the meiofauna-based analysis and χ^2 (6 df) = 78.6 ($p < 0.01$) for the macrofauna-based analysis. For those concerned about the assumptions of Bartlett's sphericity test, there is an analogous randomization test called ANOSIM (Clarke and Green 1988).

The PCA results are shown in Table 1. In the meiofauna-based analysis the first BioPC was a "size" PC" and the sec-

Fig. 3. Relationship between BioPC1 and BioPC2 of the macrofauna abundance data and contaminant variable ChemPC1.

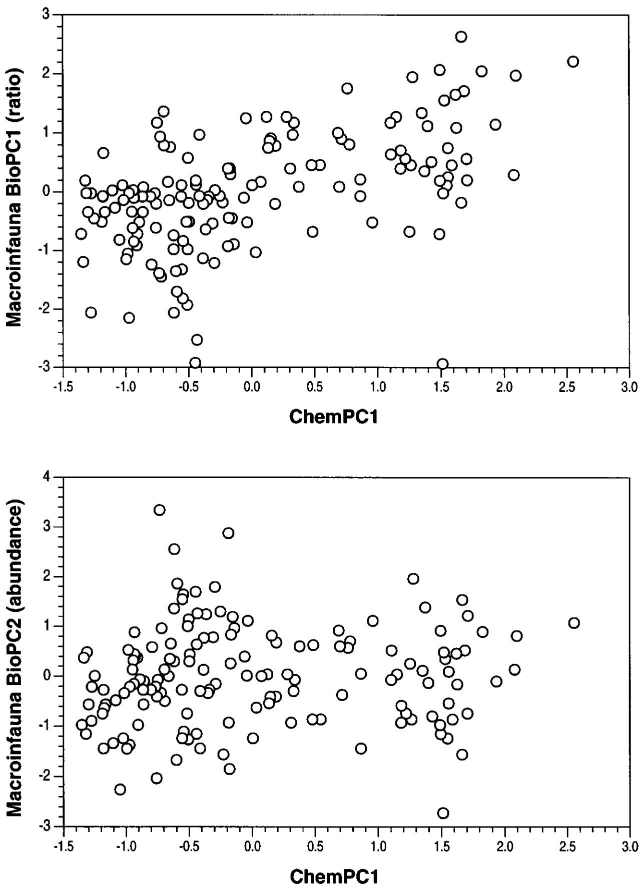


Fig. 4. Relationship between percent successful development (MDEVEL) in sea urchin embryo pore water assays and contaminant variable ChemPC1.

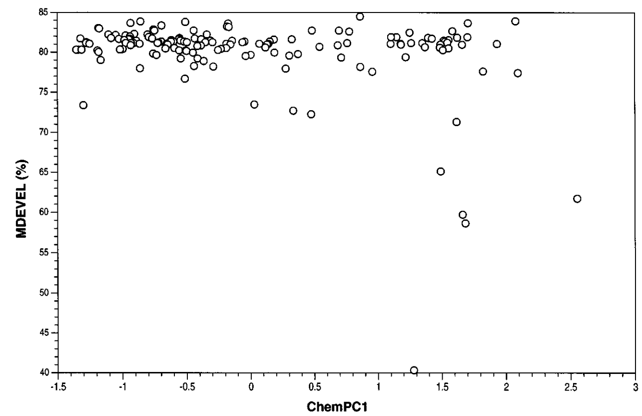


Table 1. Summary of results of PCA on correlations among SQT components.

Meiofauna	PC1 (41.5%)
Chem component	0.703
Tox component	−0.758
Bio component (total abundance)	−0.267
Bio component (NEMA/HARP)	0.721
Macrofauna	PC1 (45.6%)
Chem component	0.782
Tox component	−0.706
Bio component (total abundance)	0.255
Bio component (POLY/AMPH)	0.805

ond a "shape" PC, indicating that fluctuations in abundance of both taxa together dominate the abundance variation, with relative abundance changes (changes in the N/C ratio) being smaller in magnitude. However, in the macrofauna-based analysis the first BioPC was the "shape" PC and the second the "size" PC, indicating that the relative abundance changes (the P/A ratio) dominate the fluctuations in overall abundance of both taxa. For both of the SQT summary analyses, there is a dominant PC1 (41.5% of the structure in the meiofauna-based analysis and 45.6% of the structure in the macrofauna-based analysis) that can be interpreted as follows: As chemical contamination increases (positive relationship with ChemPC1) the percent successful development of sea urchin embryos decreases (negative relationship with that toxicological measure), and the biological community responds by changing the relative abundances of the taxa in the direction of increasing the ratio of the abundance of the more tolerant taxon to the abundance of the more sensitive taxon.

In other words, the same relationships among the SQT components are seen in the meiofauna-based and macrofauna-based SQT analyses, which strengthens the interpretations and conclusions. Variations in meiofaunal abundances are dominated by changes in total abundance (e.g., nematodes and harpacticoids going up and down together) whereas variation in macrofaunal abundances is dominated by relative abundance changes (changes in the P/A ratio). Although neither community shows a strong relationship between ChemPC1 and the total abundance of both taxa (it is mostly between ChemPC1 and the relative abundances of the taxa), such relationship as there is is a negative one for the meiofauna and a positive one for the macrofauna. Organic enrichment of the macrofaunal community near the platform may cause increased total macrofaunal abundance there (Montagna and Harper 1996; Peterson et al. 1996). If such enrichment does not affect the meiofaunal community, then there would be a contaminant-driven decrease in total abundance near platforms. This is in fact what we see. In summary, both communities show impact effects in their ratios of "tolerant to sensitive taxa abundances". The meiofaunal community also shows some negative response in total abundance, but in the macrofaunal community, any such response is overwhelmed by enhancement of total abundance (mostly of polychaetes), probably by nutrient enrichment near platforms.

The platform by platform analyses indicate that most of the pattern we see in these results is driven by what is going on at HI-A389. The patterns at the other two platforms are not obviously different from HI-A389, but they are certainly weaker. At HI-A389, Bartlett's sphericity test is highly significant and all three SQT relationships are highly significant, for both meiofauna and macrofauna. At MU-A85, there is no Tox-Chem relationship and no Bio-Tox relationship for either meiofauna or macrofauna. There is a significant ($p < 0.01$) Bio-Chem relationship for macrofauna and a marginally non-significant one ($p = 0.11$) for meiofauna. However, the overall test of whether there are any significant SQT relationships, Bartlett's sphericity test, is not significant for either meiofauna or macrofauna at MU-A85. At MAI-686, Bartlett's sphericity test is highly significant for both meiofauna and macrofauna, but only the Bio-Chem relationship is significant ($p < 0.01$, for both meiofauna and macrofauna). The Bio-Tox and the Tox-Chem relationships are nonsignificant for both meiofauna and macrofauna at MAI-686.

Finally, it should be remembered that the ChemPC1 score we are using as our measure of the chemical contaminant SQT component is a measure of the confounded chemical and substrate composition gradients related to distance from platform. All interpretations of these results must be made with this reality in mind. However, the same can be said about any environmental impact study based solely on observational data. For consideration of the sensitivity of multivariate methods in detection of marine environmental impacts, see Gray et al. (1990) and Olsgard and Gray (1995).

Heterogeneity of variance as an impact response

It is usual to formulate response variables for impact studies in terms of mean values of contaminant concentrations, taxa abundances, etc. However, data collected and analyzed in pollution impact studies suggest that variances as well as means typically respond to impacts. The response is usually "patchy", since the small-scale spatial distribution of the impacting contaminants is rarely uniform. This is visually obvious in an intertidal area after an oil spill. An increased variance among small-scale spatial replicate samples (cores, quadrats) is the result. This is most often observed when the data are examined to see whether they satisfy assumptions for analysis of variance (ANOVA), since one of the ANOVA assumptions is "homogeneity of error variance". Concern is usually focused on whether this ANOVA assumption is satisfied, for testing whether means differ, without realizing that such variance heterogeneity can itself be evidence of impact. Green (1993*b*) discussed this question, and Underwood (1992, 1993) described an ANOVA design that incorporates a test of impact-caused increased variation. Warwick and Clarke (1993) described increased variation in response to stress in marine communities, but Chapman et al. (1995) were unable to demonstrate increased variability in polluted versus control assemblages.

In deciding how to test for increased spatial variation attributable to impact, two things must be kept in mind. One is that the appropriate spatial scale of sampling must be used, and the other is that one must use the appropriate scale for measuring the response. Regarding the first, the scale of sampling must be the residual spatial error variation in some sense and not a scale of variation which is included in the ANOVA model. For example, it would not be appropriate to use the variance among boxcores within a particular distance-from-platform (within Platform[1] and Cruise). Because there is only one boxcore taken at each Distance and on each Radius, all of that variance would be among-Radius variance. An increased near-the-platform variance of that kind would only be evidence of directionality of impact, which we are already testing by testing the Distance·Radius interaction (Kennicutt et al. 1996). In the GOOMEX study design the small-spatial-scale variance is the "between-subcores within-boxcores" variance. (If replicate boxcores had been taken, then it would have been the "among replicate boxcores at each Distance and Radius" variance.)

The appropriate scale for measuring the response is usually the transformed scale. If percent variation is appropriate for evaluation of response in mean values, and we therefore usually log-transform the response variable data, then it will usually also be appropriate for evaluation of variance magnitude. If the variance of untransformed contaminant concentration is higher near a platform (where contaminant mean concentration is higher), it could be argued that it is only caused by the relationship between the error variance and the mean, which always exists in (untransformed) data reflecting percent variation. The raw (replicate subcore) observations in GOOMEX response variable data are typically log-transformed, e.g., for contaminant concentrations and taxa abundances (Kennicutt et al. 1996). The PC1 scores that we are using to represent the contaminant gradient are calculated from a linear additive function of appropriately transformed variables, so ChemPC1 as a synthetic variable is also being measured on the appropriate scale.

The between-subcore within-boxcore variance was estimated for each ring (Distance) and direction (Radius), pooled over Platforms and Cruises, on the appropriate response scale, for Ba, Cd, ChemPC1, nematode abundance (NEMA), harpacticoid abundance (HARP), and N/C. Macrofauna data were not used because it seemed likely that the within-boxcore spatial scale would be too small for proper assessment of macrofauna patchiness. Using these variance estimates as the response variables in among-distances ANOVAs, we found significant ($p < 0.01$) heterogeneity of variance for Ba, Cd, and ChemPC1, but not for NEMA and HARP ($p > 0.05$). For the N/C ratio, heterogeneity of variance was significant at $p < 0.05$. Because of some concern about the distributional properties of estimated variances used as ANOVA response variables, we repeated these ANOVAs as nonparametric ANOVAs (Kruskal–Wallis tests of medians). The significances and nonsignificances were unchanged. The estimated median variances for each distance for these five variables is presented in Table 2. For the four variables with significant heterogeneity of variance the pattern is one of increased variance near the platform. Variance as a function of distance from the platform is presented in Fig. 5 for these variables.

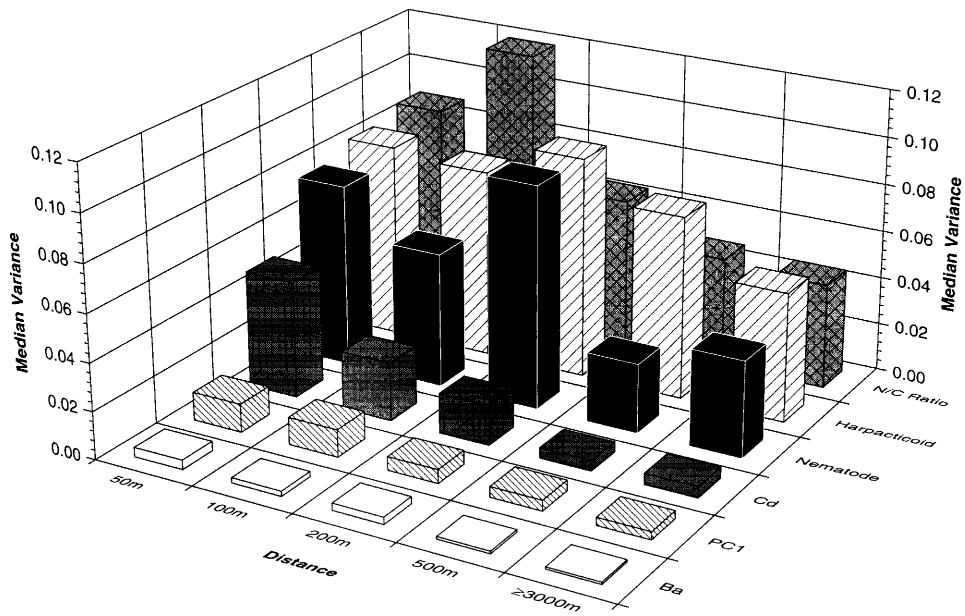
These results provide additional evidence for platform-derived impacts. The pattern for Cd and ChemPC1 suggests that the impact may extend to the third ring (200 m, Fig. 5). However, the impact could be caused by sediment changes rather than contaminants, which could easily increase the variance of both these variables. The absence of heterogeneity of variance for the two meiofaunal taxa taken by themselves contrasts with the heterogeneity of variance for the N/C ratio. This provides additional support for the use of the "ratio of a tolerant to a sensitive taxon" as an indicator of biological impact. Ratios should not be calulated from the raw data for each sample and then used as a response variable in

[1] Throughout this paper, we have capitalized the first letter of Platform, Cruise, Distance, and Radius where these terms are used in the sense of a study design element.

Table 2. Median subcore variances for selected variables.

Distance from platform (m)	Cd	ChemPC1	Ba	NEMA	HARP	N/C
50	0.0491	0.0119	0.0039	0.0770	0.0822	0.0870
100	0.0245	0.0115	0.0021	0.0566	0.0793	0.1184
200	0.0167	0.0057	0.0029	0.0940	0.0931	0.0628
500	0.0043	0.0044	0.0010	0.0293	0.0770	0.0463
≥3000	0.0046	0.0033	0.0009	0.0406	0.0543	0.0442

Fig. 5. Median variance of selected variables between subcores within a boxcore versus distance from platform.



statistical analysis because such derived ratio variables do not behave well as response variables (Green 1979, 1986). The use of log–log linear models (regression, analysis of covariance, or PCA as used here) to imply ratios when coefficients have opposite signs should be the preferred practise.

Recommended sampling designs for future studies

The GOOMEX Phase I study is based on a complex sampling design described in Kennicutt et al. (1996). It was used throughout this study for hypothesis testing and interpretation. The sampling design contains the following components: three Platforms, four Cruises (times, at half-year intervals), five Distances from each Platform (for some variables only two Distances), and five Radii (directions from Platforms). This sampling design was developed to satisfy the requirement that there be intensive sampling around Platforms to describe direction and distance (spatial scale) of impacts in relation to the point source. That is all very well, but any project has a fixed budget and any increase of "oversampling" around each point-source is inevitably a trade-off against including more point-sources in the study design.

Two quite different objectives are in conflict. Elucidating pattern of impact (direction and distance) around a point-source calls for a different study design than does generalizing about impacts from point-sources. Requiring that a study accomplish both inevitably leads to compromises and to a nonoptimal design for both. The GOOMEX Phase I study design emphasizes the first objective. Impact patterns around each of the three platforms were adequately described. However, with only three platforms, each of which is environmentally unique in some ways, it is difficult to generalize about impacts of "platforms", and any attempt to relate impact differences among platforms to historical production differences among platforms or to natural environment differences among platforms leads to a regression or correlation based on three data points. Even if there were a perfect rank correlation between the extent of impact-related responses and discharges from platforms, with only three platforms that perfect rank correlation (in the direction expected, i.e., one-tailed) would occur by chance one time out of six (0.17 probability). Typically, 0.05 is required for statistical significance. With four platforms, it would be 1 in 24, or 0.04. With five, it would be 1 in 120, or 0.008. Thus the likelihood of confounding with natural environment patterns decreases rapidly. In our three-platform study the rank order of contaminant levels is also a perfect rank correlation with water depth, distance from shore, and all environmental variables related to them.

In other words, the study design describes and tests the spatial pattern of impact around these three platforms (i.e., platforms treated as fixed-effect blocks, or strata), but there is little ability to generalize to impacts caused by "platforms" as a generic class or to describe differences in impacts due to the kind of platform. If the platforms are thought of as a sample of platforms of a kind in the region, then three is too few and these platforms are too different for any meaningful conclusion to be reached about "platforms". If each of these three platforms is considered a representative of a platform type (e.g., contaminant levels, depth and distance from shore), then there is no replication within platform type and impact differences among platform types cannot be tested.

To generalize the results, rather than just describe what happens around each platform, an increase in the number of platforms and a corresponding decrease in effort allocated around each is needed. For example, the design might consist of 12 or more platforms with sampling at near and far distances (e.g., 50 and 3000 m) along one radius in a direction and at distances from each platform that were established from a previous study, or preliminary sampling, to be the average direction of maximum contaminant movement and the appropriate distances from the platform in that direction. This would result in a paired design, in which the "near versus far" difference would be tested based on a sample of $n \geq 12$ platforms. To reduce costs, one sampling method could be used and one response variable measured (again established from preliminary sampling). The procedure for selecting platforms is critical. The region should be defined a priori, and objective criteria for one or more types of platform should be stated. Then, all platforms of that type occurring in that region should be catalogued and the ones to be included in the study randomly selected from them.

Twelve or more platforms are recommended, for two reasons. First, robustness of statistical tests can be assured by having at least 10 error degrees of freedom for the test (Green 1979; Harris 1985). In such a paired design the error is the among-platform variation in the near versus far difference. Therefore the error degrees of freedom will be the number of platforms less one (or if there is more than one type of platform, it will be the number of platform types times one less than the number of platforms of each type). Second, power in tests (ability to detect differences that really exist, e.g., near versus far) increases with sample size (number of platforms). Simulations were carried out with specified mean near versus far differences and standard deviations for those differences. Power initially increases rapidly with number of platforms but then begins to flatten out around n = 10–12 platforms.

If a study must be dual-purpose, requiring a compromise design capable of elucidating pattern of impact around point-sources as well as generalizing about impacts of that kind, then the best approach will be to have a paired study design with $n \geq 12$ platforms for the generality, with one of those platforms dedicated to the intensive spatial pattern sampling which would include the "near versus far in one direction" sampling compatible with the other platforms.

References

Carr, R.S., Chapman, D.C., Presley, B.J., Biedenback, J.M., Robertson, L., Boothe, P., Kilada, R., Wade, T., and Montagna, P. 1996. Sediment porewater toxicity assessment studies in the vicinity of offshore oil and gas production platforms in the Gulf of Mexico. Can. J. Fish. Aquat. Sci. **53**. This issue.

Chapman, M.G., Underwood, A.J., and Skilleter, G.A. 1995. Variability at different spatial scales between a subtidal assemblage exposed to the discharge of sewage and two control assemblages. J. Exp. Mar. Biol. Ecol. **189**: 103–122.

Chapman, P.M., Dexter, R.N. and Long, E.R. 1987. Synoptic measures of sediment contamination, toxicity and infaunal community structure (the Sediment Quality Triad) in San Francisco Bay. Mar. Ecol. Prog. Ser. **37**: 75–96.

Chapman, P.M., McPherson, C.A., and Munkittrick, K.R. 1989. An assessment of the ocean dumping tiered testing of ocean using the Sediment Quality Triad. Prepared for Department of Fisheries and Oceans, Institute of Ocean Sciences, Sidney, B.C. DSS Project No. 06SB.FP941-8-9060.

Chapman, P.M., Power, E.A., Dexter, R.N., and Andersen, H.B. 1991. Evaluation of effects associated with an oil platform, using the Sediment Quality Triad. Environ. Toxicol. Chem. **10**: 407–424.

Clarke, K.R. 1993. Non-parametric multivariate analyses of changes in community structure. Aust. J. Ecol. **18**: 117–143.

Clarke, K.R., and Ainsworth, M. 1993. A method of linking multivariate community structure to environmental variables. Mar. Ecol. Prog. Ser. **92**: 205–219.

Clarke, K.R., and Green, R.H. 1988. Statistical design and analysis for a biological effects study. Mar. Ecol. Prog. Ser. **46**: 213–226.

Coull, B.C., Hicks, G.R.F., and Wells, J.B.J. 1981. Nematode/copepod ratios for monitoring pollution: a rebuttal. Mar. Pollut. Bull. **12**: 378–381.

Gray, J.S., Clarke, K.R., Warwick, R.M., and Hobbs, G. 1990. Detection of initial effects of pollution on marine benthos: an example from the Ekofisk and Eldfisk oilfields, North Sea. Mar. Ecol. Prog. Ser. **66**: 285–289.

Green, R.H. 1979. Sampling design and statistical methods for environmental biologists. Wiley, New York.

Green, R.H. 1984. Statistical and nonstatistical considerations for environmental monitoring studies. Environ. Monit. Assess. **4**: 293–301.

Green, R.H. 1986. Some applications of linear models for analysis of contaminants in aquatic biota. *In* Statistical aspects of water quality monitoring. *Edited by* A.H. El-Shaarawi and R.E. Kwiatowski. Elsevier, New York.

Green, R.H. 1993*a*. Relating two sets of variables in environmental studies. *In* Multivariate analysis: future directions. *Edited by* C.R. Rao. Elsevier, Amsterdam, The Netherlands. pp. 151–165.

Green, R.H. 1993*b*. Application of repeated measures designs in environmental impact and monitoring studies. Aust. J. Ecol. **18**: 81–98.

Green, R.H., Boyd, J.M., and Macdonald, J.S. 1993. Relating sets of variables in environmental studies: the Sediment Quality Triad as a paradigm. Environmetrics, **4**: 439–457.

Harris, R.J. 1985. A primer of multivariate statistics. 2nd ed. Academic Press, New York.

Hennig, H.F-K., Eagle, G.A., Fielder, L., Fricke, A.H., Gledhill, W.J., Greenwood, P.J., and Orren, M.J. 1983. Ratio and population density of psammolittoral meiofauna as a perturbation indicator of sandy beaches in South Africa. Environ. Monit. Assess. **3**: 45–60.

Kennicutt, M.C. II, Boothe, P.N., Wade, T.L., Sweet, S.T., Rezak, R., Kelly, F.J., Brooks, J.M., Presley, B.J., and Wiesenburg, D.A. 1996. Geochemical patterns in sediments near offshore production platforms. Can. J. Fish. Aquat. Sci. **53**. This issue.

Long, E.R., and Chapman, P.M. 1985. A sediment quality triad: measures of sediment contamination, toxicity and infaunal community composition in Puget Sound. Mar. Pollut. Bull. **16**: 405–415.

Montagna, P.A., and Harper, P.E. Jr. 1996. Benthic infaunal long-term response to offshore production platforms in the Gulf of Mexico. Can. J. Fish. Aquat. Sci. **53**. This issue.

Montagna, P.A., Bauer, J.E., Hardin, D., and Spies, R.B. 1987. Temporal variability and the relationship between benthic meiofaunal and microbial populations of a natural coastal petroleum seep. J. Mar. Res. **45**: 761–789.

Montagna, P.A., Bauer, J.E., Hardin, D., and Spies, R.B. 1989. Vertical distribution of microbial and meiofaunal populations in sediments of a natural coastal hydrocarbon seep. J. Mar. Res. **47**: 657–680.

Olsgard, F., and Gray, J.S. 1995. A comprehensive analysis of the effects of offshore oil and gas exploration and production on the benthic communities of the Norwegian continental shelf. Mar. Ecol. Prog. Ser. **122**: 277–306.

Peterson, C.H., Kennicutt, M.C. II, Green, R.H., Montagna, P., Harper, D.E. Jr., Powell, E.N., and Roscigno, P.F. 1996. Ecological consequences of environmental perturbations associated with offshore hydrocarbon production: a perspective on long-term exposures in the Gulf of Mexico. Can. J. Fish. Aquat. Sci. **53**. This issue.

Pimentel, R.A. 1979. Morphometrics: the multivariate analysis of biological data. Kendall/Hunt, Dubuque, Iowa.

Raffaelli, D., and Mason, C.F. 1981. Pollution monitoring with meiofauna, using the ratio of nematodes to copepods. Mar. Pollut. Bull. **12**: 158–163.

Ter Braak, C.F.J. 1986. Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis. Ecology, **67**: 1167–1179.

Underwood, A.J. 1992. Beyond BACI: the detection of environmental impacts on populations in the real, but variable world. J. Exp. Mar. Biol. Ecol. **161**: 145–178.

Underwood, A.J. 1993. The mechanics of spatially replicated sampling programmes to to detect environemntal impacts in a variable world. Aust. J. Ecol. **18**: 99–116.

Warwick, R.M., and Clarke, K.R. 1993. Increased variability as a symptom of stress in marine communities. J. Exp. Mar. Biol. Ecol. **172**: 215–226.