Environment Canada Environnement Canada

Report on Biological Toxicity Tests Using Pollution Gradient Studies

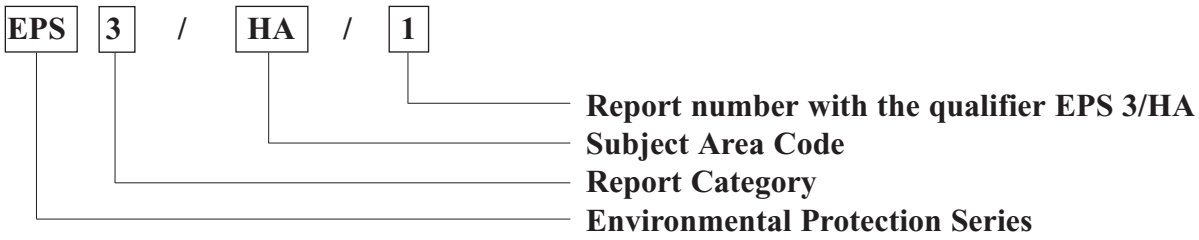
Sydney Harbour

EPS 3/AT/2 - July 2000
Marine Environment Division
Environment Canada

Canada

ENVIRONMENTAL PROTECTION SERIES

Sample Number:



Categories

1	Regulations/Guidelines/Codes of Practice
2	Problem Assessment and Control Options
3	Research and Technology
4	Literature Reviews
5	Surveys
6	Social, Economic and Environmental Impact Assessments
7	Surveillance
8	Policy Proposals and Statements
9	Manuals

Subjects

AG	Agriculture
AN	Anaerobic Technology
AP	Airborne Pollutants
AT	Aquatic Toxicity
CC	Commercial Chemicals
CE	Consumers and the Environment
CI	Chemical Industries
FA	Federal Activities
FP	Food Processing
HA	Hazardous Wastes
IC	Inorganic Chemicals
MA	Marine Pollutants
MM	Mining and Ore Processing
NR	Northern and Rural Regions
PF	Paper and Fibres
PG	Power Generation
PN	Petroleum and Natural Gas
RA	Refrigeration and Air Conditioning
RM	Reference Methods
SF	Surface Finishing
SP	Oil and Chemical Spills
SRM	Standard Reference Methods
TS	Transportation Systems
TX	Textiles
UP	Urban Pollution
WP	Wood Protection/Preservation

New subjects and codes are introduced as they become necessary. A list of EPS reports may be obtained from Environmental Protection Publications, Environmental Protection Service, Environment Canada, Ottawa, Ontario, K1A 0H3, Canada

Report on Biological Toxicity Tests Using Pollution Gradient Studies – Sydney Harbour

by

B.A. Zajdlik
Zajdlik & Associates
R.R.3 Rockwood, Ontario
N0B 2K0

K.G. Doe
Toxicology Laboratory
Environment Canada
P.O. Box 23005
Moncton, New Brunswick
C1A 6S8

L.M. Porebski
Marine Environment Division
Environment Canada
351 St. Joseph Blvd.
Hull, Quebec
K1A 0H3

Report EPS 3/AT/2
July 2000

National Library of Canada Cataloguing in Publication Data

Zajdlik, B.A. (Barry Alan), 1963-

Report on biological toxicity tests using pollution
gradient studies – Sydney Harbour

(Report EPS 3/AT/2)
Issued also in French under title : Rapport sur les essais
biologiques de toxicité au moyen d'études sur le gradient de pollution –
Port de Sydney.
ISBN 0-660-18492-3
Cat. No. En49-24/1-50E

1. Environmental toxicology – Nova Scotia – Sydney Harbour.
2. Aquatic biology – Nova Scotia – Sydney Harbour – Methodology.
3. Toxicity testing – Nova Scotia -- Sydney Harbour – Methodology.
4. Waste disposal in the ocean – Environmental aspects – Nova Scotia
-- Sydney Harbour.

I. Doe, K.G. (Kenneth G.)
II. Porebski, L.M. (Linda M.)
III. Canada. Environment Canada.
IV. Canada. Marine Environment Division.
V. Series : Report (Canada. Environment Canada) ; EPS 3/AT/2.

QH90.8.T68Z32 2001 363.739'463'463'097169 C2001-980121-1

Cover photo; Sydney Harbour. Source: P. Topping

Readers' Comments

Comments regarding the content of this report should be addressed to:

Linda M. Porebski
Marine Sediment Advisor
Marine Environment Division
Environment Canada
351 St. Joseph Blvd.
Hull, Quebec
K1A 0H3

Cette publication est aussi disponible en français. Pour l'obtenir, s'adresser à:

Publications de la Protection de l'environnement
Environnement Canada
Ottawa (Ontario)
K1A 0H3

Review Notice

This report has been reviewed by staff of the Marine Environment Division, Environment Canada, and approved for publication. Mention of trade names or commercial products does not constitute recommendation or endorsement by Environment Canada for use.

Abstract

Biological toxicity tests (responses included survival, fertilization, growth, photoluminescence, and bioaccumulation), sediment and pore water chemistry, and benthic community structure were examined along a known pollution gradient in Sydney Harbour, Nova Scotia, Canada. Major contaminants were PAHs, PCBs, and heavy metals. Relationships between toxicity tests, chemistry, and benthic structure at the different stations were examined. The primary purpose of the study was to assess Environment Canada's Disposal at Sea Program's interim interpretation criteria for each of the toxicity and bioaccumulation assays. This was done by comparing the pass/fail decisions made using current interpretation criteria, for the assays in the battery with chemical guidelines and with the benthic community structure at the stations. A secondary purpose of the study was to identify additional research needs, or refinements to better use the toxicity test battery within the program. Most of the toxicity tests distinguished between the more contaminated sites and the reference sites, some with very good correlations to the major contaminants or the benthic community, or both, although as a whole the three data sets were not significantly correlated. Non-contaminant factors (ammonia, moisture, grain size, etc.) were often correlated with test responses, although to a lesser degree than with contaminants, suggesting a continued need to measure and assess the relative contribution of these factors to the test results.

The choice of reference sites appeared to be critical to whether a station would have "passed or failed" according to program criteria, suggesting that further work on reference site criteria is needed. Amphipod survival, bivalve bioaccumulation, and luminescent bacterial tests, in general, performed well and the interim biological toxicity test interpretation criteria reflected a probable effect level for this data set. The probability of observing toxicity, estimated using mean probable effect level quotients, concurred with the proportion of biological tests actually failing. Echinoid fertilization tests and polychaete growth tests did not appear well correlated with the chemical and benthic results. These tests will require additional research. Other findings of the study were that porewater chemistry may be a valuable addition to the suite of tools used to measure sediment toxicity, and that total PAHs provide a good surrogate for individual PAH measurements.

Résumé

Des essais biologiques de toxicité (portant sur la survie, la fertilisation, la croissance, la photoluminescence et la bioaccumulation), la composition chimique des sédiments et de l'eau interstitielle et la structure de la communauté benthique ont été examinés le long d'un gradient de pollution connu dans le port de Sydney, Nouvelle-Écosse, Canada. Les principaux contaminants étaient les HAP, les BPC et les métaux lourds. Les relations entre les essais de toxicité, la composition chimique et la structure benthique dans les différentes stations ont été examinées. L'étude avait pour but principal d'évaluer les critères provisoires d'interprétation du Programme d'immersion en mer d'Environnement Canada pour chacun des essais de toxicité et de bioaccumulation. Pour ce faire, on a comparé les décisions réussite-échec prises à l'aide des critères actuels d'interprétation, pour la batterie d'essais, les lignes directrices concernant les substances chimiques et la structure de la communauté benthique dans les stations. Son but secondaire consistait à définir les autres besoins en recherche ou les améliorations à apporter pour mieux utiliser la batterie d'essais de toxicité du programme. La plupart des essais de toxicité ont permis de différencier les sites plus contaminés et les sites de référence; pour certains, il existait de très bonnes corrélations avec les principaux contaminants ou les communautés benthiques, ou les deux à la fois, mais en général, les trois ensembles de données n'étaient pas corrélés de façon significative. Les facteurs non contaminants (l'ammoniac, l'humidité, la classe granulométrique, etc.) étaient souvent corrélés avec les réactions aux essais, quoique moins qu'avec les contaminants, ce qui a semblé indiquer la nécessité continue de mesurer et d'évaluer la contribution relative de ces facteurs aux résultats des essais.

Le choix des sites de référence a semblé très important pour déterminer si une station « satisfaisait ou non » aux critères du programme, ce qui a indiqué qu'il fallait peaufiner les critères des sites de référence. En général, les essais portant sur la survie des amphipodes, la bioaccumulation dans les bivalves et les bactéries luminescentes ont donné de bons résultats, et les critères provisoires d'interprétation des essais biologiques de toxicité ont porté à croire qu'il existait une concentration produisant un effet probable pour cet ensemble de données. La probabilité de l'observation d'effets toxiques, calculée à l'aide des quotients moyens des concentrations produisant un effet probable, concordait avec la proportion des essais biologiques se soldant par un échec. Les essais de fertilisation des échinides et de croissance des polychètes n'ont pas semblé bien corrélés avec les résultats pour les analyses chimiques et le benthos. Ces essais nécessiteront une recherche plus poussée. L'étude a aussi démontré que la composition chimique de l'eau interstitielle pouvait être un autre outil valable pour mesurer la toxicité des sédiments et que la mesure des HAP totaux peut remplacer avantageusement celle de chacun de ces composés.

Table of Contents

Section 3

Section 4

List of Tables

List of Figures

Acknowledgements

This document was written by Barry Zajdlik of Zajdlik and Associates with assistance from Ken Doe and Linda Porebski of Environment Canada. The project was organized and implemented by Linda Porebski with assistance in the field by Paul Topping and Adrian MacDonald of Environment Canada, Steve Devitt and Eric Garnier from Arenicola Marine, and by John Simic and Floyd Patttengale captain and owner of the "M. V. Katherine G".

Toxicity testing was conducted by Environment Canada's Ken Doe and Paula Jackman in the Atlantic Region and by Graham van Aggelen and Michelle Fennell in the Pacific and Yukon Region. Benthic community analysis was provided by Pat Pocklington of Arenicola Marine.

Advice was given on the collection of ammonia, sulphide, and redox potential by Barry Hargrave, Department of Fisheries and Oceans.

Analytical Chemistry was overseen by James Doull and Stewart Yee of Environmment Canada. AVS analysis was overseen by Annette Lawson of Dundas Environmental Consulting Inc.

Data analysis and interpretaton was done by Barry Zajdlik.

Funding was provided by Jim Osborne of Environment Canada, through the Ocean Disposal Research Fund.

Specific sections of the document have also been taken from contributors unpublished reports and are greatfully acknowledged as follows:

Steve Devitt. Text in Section 2.2 is extracted from: Devitt, S. 1997. Field acitivity report for collection of marine sediments from Sydney Harbour. Prepared for Ms. L. Porebski, Environment Canada, Marine Environment Division, Ottawa, Ontario.

Paula Jackman and Ken Doe, Environment Canada, Atlantic Region. Text in Sections 2.3.2.1 through 2.3.2.4 is extracted from: Jackman, P. and K. Doe. 1997. Toxicity of sediments from Sydney Habour, NS–Results of the 1997 Pollution Gradient Study. (Unpublished Report), Environment Canada, Atlantic Region, Moncton, NB.

Michelle Fennell and Graham van Aggelen, Environment Canada, Pacific and Yukon Region. Biological toxicity testing. Text in Sections 2.3.3.1 through 2.3.3.4 is extracted from: Fennell, M. and G. van Aggelen. 1997. Biological assessment of sediments collected from Sydney Harbour, Nova Scotia: Site of the 1997 Pollution Gradient Study conducted by Environment Canada's National Disposal at Sea Program. (Unpublished Report).

Pat Pocklington, Arenicola Marine. Text in Section 2.4.2 is extracted from: Arenicola Marine. 1997. Sample collection and analysis for benthic community structure: Sydney Harbour, NS, July 1997.

Many thanks to all the contributors. A special thanks should also go to our reviewers who provided us with many useful comments on both the document and approaches.

Scott Carr, U.S. Geological Survey, Biological Resources Division, Corpus Christi, Texas

Chris Ingersoll, U.S. Geological Survey, Columbia Environmental Research Centre, Columbia, Missouri

Deanna Lee, Environment Canada, Disposal at Sea, North Vancouver, British Columbia

Ed Long, National Oceanic and Atmospheric Administration, Coastal Monitoring and Bioeffects Assessment Division, Seattle, Washington

Abbreviations

AVS	acid volatile sulphides
Eh	redox potential
EC	Environment Canada
EPA	Environmental Protection Agency (United States)
IC25, IC50	the concentration estimated to cause a 25 or 50% inhibition in response, respectively
ISQG	interim sediment quality guideline
PAH	polycyclic aromatic hydrocarbon
PEL	probable effect level
PCB	polychlorinated biphenyl
pH	potenz hydrogen
ppt	parts per thousand
SEM	simultaneously extracted metals
SQG	sediment quality guideline
TEL	threshold level effect
TOC	total organic carbon

Introduction

This is the second of two pollution gradient studies by Environment Canada's Disposal at Sea Program, examining the field performance of chemical and biological tools proposed for the assessment of marine sediments destined for disposal at sea. Toxicity and bioaccumulation tests, sediment and pore water chemistry, and benthic community structure were examined along a known PAH pollution gradient in Sydney Harbour, Nova Scotia, Canada. Relationships among these elements at the different stations were examined. The primary purpose of the study was to assess the program's interim interpretation criteria for the battery of toxicity and bioaccumulation assays. This was done by comparing the assessment of sediment quality using current interpretation criteria, with that obtained using the other assays in the battery, chemical guidelines, and assessment of the benthic community structure at the stations. This report also looks at the functioning of Interim Canadian Sediment Quality Guidelines (CCME, 1999) as chemical benchmarks and at the relevance of other chemical assessment tools, in relation to responses in the toxicity tests and in the benthic community structure.

In Canada, disposal at sea is regulated under the *Canadian Environmental Protection Act* (CEPA, 1999). Before any disposal at sea permit is granted, the material is evaluated according to an international waste assessment framework (EC, 1995c). One of the steps within this assessment process is the characterization of the waste's physical, chemical, and biological properties. Environment Canada is now finalizing a tiered approach for this waste characterization process (EC, 1995c).

Tier 1 screening levels could use a set of Canadian Sediment Quality Guidelines at the Threshold Effects Level (TEL), developed from a co-occurrence database in which synoptic chemical and biological information is evaluated in terms of probability of effects. The Threshold Effects Levels denote chemical concentrations at or below which no adverse biological effects are expected. Levels above these criteria would trigger Tier 2 investigations of sediment quality, in the form of toxicity and bioaccumulation testing.

The battery selected, includes an acute test with amphipods, three sub-lethal tests and one bioaccumulation test (Table 1) (EC, 1992a, b, c; USEPA, 1993). If the sediments or waste materials pass the toxicity/bioaccumulation tests, open water disposal can be considered. Failure in more than one of the tests (or of the acute test alone) disqualifies the material for open water disposal. Interpretation (pass/fail) criteria for the tests have been proposed for use (Table 9) but require field validation (Stebbing, Dethlefsen, and Carr, 1992). The use of sediment quality values at the Probable Effects Level (PEL) (level above which effects are likely to occur) to reject sediments is possible, but is not being considered at this time. Probable Effects Levels are used in this study, however, as chemical benchmarks to assist with the selection of the sampling stations.

The study design followed a sediment triad approach (Chapman, 1992; Stebbing, Dethlefsen and Carr, 1992). The inferences made using each of the triad components may be complementary, contradictory, or uninformative; thus a weight of evidence approach was used to assess the quality of the test sediments in relative terms. It was hoped that this approach could be used to support the toxicity test interpretation criteria being promulgated by Environment Canada.

Table 1 Summary of Biological Tests Being Considered for Regulatory Use

Test/Species	Organism Type	Test Type	Response
Amphiporeia virginiana	amphipod	sediment	Percent survival
Eohaustorius washingtonianus	amphipod	sediment	Percent survival
Eohaustorius estuarius	amphipod	sediment	Percent survival
Rhepoxynius abronius	amphipod	sediment	Percent survival
Microtox® (solid-phase, moisture corrected)	bacteria	sediment	Change in luminescence
Macoma nasuta	bivalve	sediment	Percent survival, bioaccumulation
Dendraster excentricus	echinoderm	porewater	Percent fertilization
Lytechinus pictus	echinoderm	porewater	Percent fertilization
Boccardia proboscidea	polychaete	sediment	Percent survival, growth rate
Polydora cornuta	polychaete	sediment	Percent survival, growth rate

Factors such as total organic carbon, particle size, depth, ammonia, and sulphide may influence organism responses in lab tests and the *in situ* benthic community structure. Thus efforts were made to select test stations having similar geophysical and chemical properties, to mitigate the effects of these known confounding factors. The site also needed to include a wide chemical gradient so that stations below the TELs (a reference station), stations between the TELs and PELs (intermediate effects) and above the PELs (effects likely) could be evaluated. Potential Canadian sites in British Columbia, Quebec, and the Atlantic provinces were examined.

Two sites were selected for gradient studies, one with a predominantly metals gradient at Belledune Harbour, New Brunswick and one with a predominantly organic contaminant gradient at Sydney Harbour, Nova Scotia. This paper focuses only on the Sydney data. The information on the first study in Belledune Harbour, New Brunswick is available as a technical report (Porebski *et al.*, 1998) or shorter paper (Porebski *et al.*, 1999).

The major contaminants in Sydney Harbour stem largely from historical coke oven effluent discharge into Muggah Creek at the mouth of the South Arm of the harbour (Matheson *et al.*, 1983). A 1994 site selection study indicated that Sydney Harbour (Figure 1) would provide a suitable PAH gradient based on historical information. Gradients for PCBs, cadmium, zinc, nickel, and copper were also identified. Evidence of PAH bioaccumulation in lobster was found in 1980 and 1981 studies (BEAK Consultants Ltd., 1996).

In November 1996, a preliminary study was conducted in the North and South Arms of Sydney Harbour to help select test stations along the gradient. Chemical analysis, particle-size distribution, TOC, ammonia, solid phase Microtox® tests, and Toxichromopad® tests were done on 12 potential test stations and four reference stations. The results showed a clear gradient for PAH (from 196 to <1 μg/g) decreasing with distance from the mouth of Muggah Creek in the South Arm from Stations 1 to 12. The study also showed Microtox® luminescence increasing along the gradient from the most (1) to the least (12) contaminated

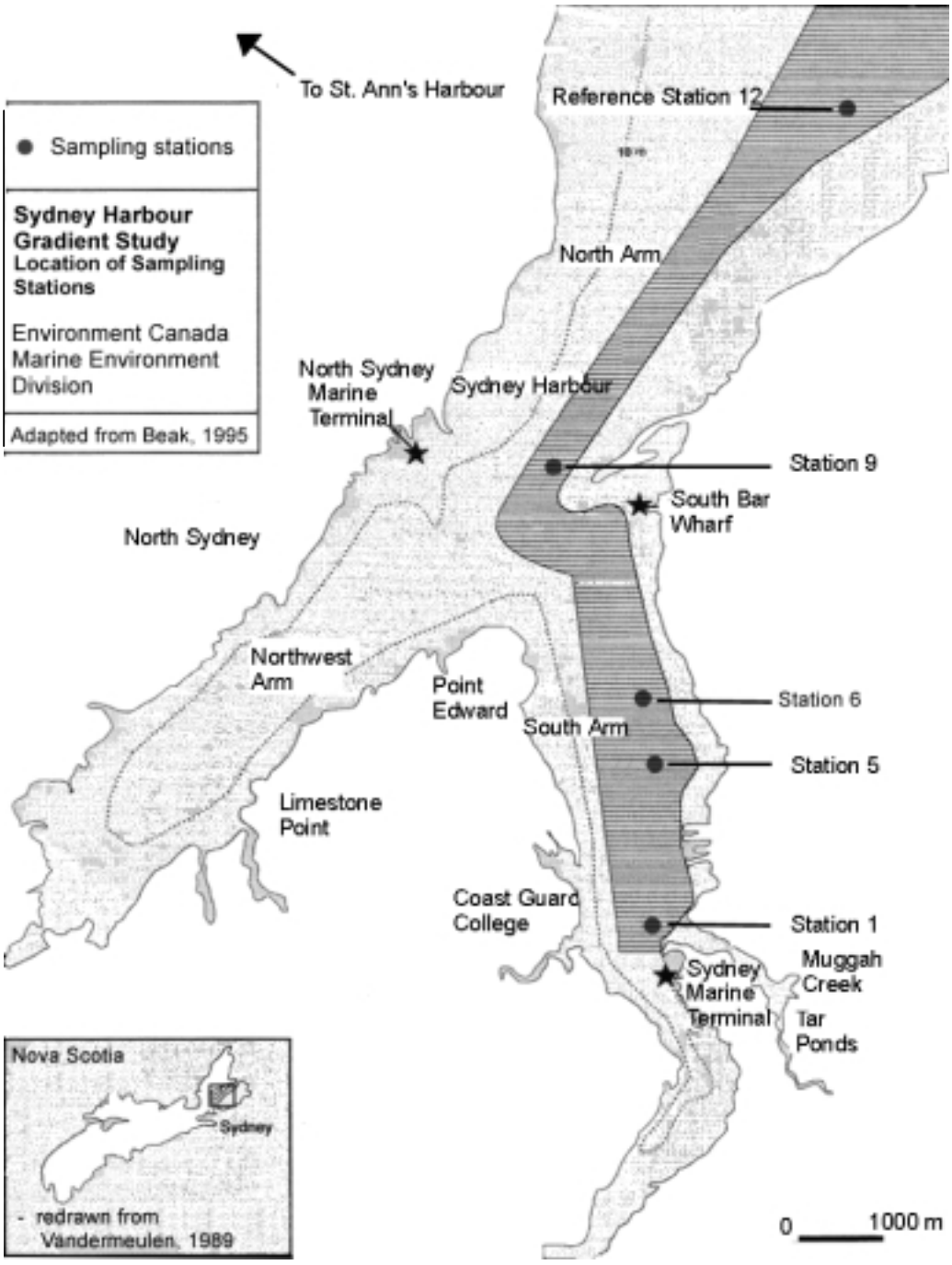


Figure 1 Sample Locations within Sydney Harbour

Stations. The physical parameters were not homogenous along the gradient. Sediments become coarser grained and lower in TOC along the gradient; however, no better site was available. Depths ranged from 10 to 20 metres but did not follow the gradient.

1.1 Purpose/Objectives

This study investigates:

- the suitability of the interim interpretation criteria for biological toxicity tests;
- the effect of confounding factors (TOC, percent moisture, grain size, redox potential, sediment and porewater ammonia, porewater pH, and sulphide) on biological toxicity test interpretation;
- the suitability of recommended species for regulatory use;
- the relationship between sediment chemistry and biological toxicity test results; and,
- the relationship between the *in-situ* benthic macroinvertebrate community and biological toxicity test results.

A series of hypotheses reflecting the study objectives were generated. The hypotheses address issues arising when interpreting the pass/fail status of sediments using three assessment tools: sediment chemistry, biological toxicity tests, and benthic macroinvertebrate community structure. The hypotheses are:

- **H1: Are potential confounding factors homogeneous across the stations?** This hypothesis is tested to verify that the choice of stations achieved the study design goal of minimizing the effect of known confounding factors such as TOC, particle size, ammonia, and Eh.
- **H2a: Do confounding factors affect the biological toxicity test response?** If the confounding factors vary significantly across stations, then differences in responses may be due to the PAH gradient and/or confounding factors.
- **H2b: Are the dose responses predictable?** Hypotheses 2a and 2b are jointly explored using regression techniques. Those variables most correlated with the biological toxicity test responses are used to develop descriptive models.
- **H3: Does a biological toxicity test perform consistently at a given site?** Significant differences in variability between stations may indicate an inconsistent test, or micro-scale differences in sediment physical/chemical quality.
- **H4a: Do all biological toxicity tests characterize the sample in the same way?** A concordance between negative biological test responses provides a powerful weight of evidence regarding a potential impact. However, a lack of concordance may indicate that constituents of the battery are providing complementary rather than redundant information which is the *raison d'être* for a battery.
- **H4b: Do biological toxicity tests rank the stations in the same way?** As the pass/fail criteria dichotomizes the results of toxicity tests, a certain degree of information regarding relative sensitivity is lost. This information may be recovered by analyzing the ranking of the stations. Hypotheses 4a and b address the hypothesis: **Does the suite of biological toxicity tests provide a consistent interpretation of the status of the sediment?**
- **H5: Do the biological toxicity tests indicate an effect when the TELs or PELs are exceeded?** A lack of agreement between the characterization of sediment using biological toxicity tests and SQGs or ISQGs may

necessitate an adjustment of biological toxicity test interpretation criteria.

- **H6: Do the biological toxicity tests indicate an effect when the *in-situ* benthic macroinvertebrate community does?** A lack of agreement between the characterization of sediment using biological toxicity tests and the *in-situ* benthic macroinvertebrate community may necessitate an adjustment of biological toxicity test interpretation criteria.

- **H7: Is there concurrence in the assessment of effect/potential effect when using benthic macroinvertebrate community structure, interim biological toxicity test interpretation criteria, and PELs or TELs?** The degree of agreement in the classification of sediment using the three characterization tools is compared to a comparison of the three data sets using the raw data. (see H9)

- **H8: How strongly are the three data sets correlated?** The study design is a gradient design using a sediment quality triad approach. The constituents of the triad are biological toxicity tests, sediment physical/chemistry, and *in-situ* benthic macroinvertebrate community structure. The degree of correlation between these data sets is explored and compared to the pass/fail characterization of the sediments.

- **H9: Do the interim biological toxicity test interpretation criteria need to be adjusted to match assessments of sediment quality using benthic community structure or PELs or TELs?** Should the previous hypothesis tests indicate that the interim biological toxicity test interpretation criteria do not characterize sediments in the same way as *in-situ* benthic community structure and PELs or TELs; an adjustment to the interim interpretation criteria may be explored.

1.2 Nomenclature and Conventions

At times, the phrase “along the gradient” is used. This phrase refers to the following ordering of stations: Control, 1, 5, 6, 9, 12, and St. Ann’s Harbour. The “top of the gradient” refers to the inner most stations in Sydney Harbour.

Laboratory replicates are synonymous with subsamples and field replicates are synonymous with true replicates or simply replicates. Laboratory replicates are not used in any analyses unless explicitly stated. Wherever possible, field replicates are used. This is not possible when varying levels of replication occur in the same data set. This is the case with combined physical/chemical data sets. Porewater variables were usually only measured once, while and sediment variables were usually measured more often. The use of field replicates is also precluded when sediment variables cannot be identically matched with the samples used to conduct biological toxicity tests. The test’s degrees of freedom are included with all tests in order to clarify the level of replication used in the analysis.

Sydney Harbour sediments in refrigerated storage. Source: K. Doe

Materials and Methods

2.1 Site Selection

This study and the Belledune Harbour study (Porebski *et al.*, 1998; 1999) were conducted to evaluate tools used in assessing the suitability of sediments for ocean disposal. The two areas (Belledune and Sydney Harbour) were chosen for their known metals and organic compound gradients, respectively (BEAK Consultants Ltd., 1996).

The criteria for site selection within the harbours included homogeneity of non-contaminant factors (i.e., depth, temperature, sediment grain sizes, salinity, TOC, ammonia) and suitability of sediments for testing with the biological tests being evaluated. The reference stations were chosen with the same criteria in mind, and also for the absence of the contaminants being investigated. In the present study, samples were collected from five sites within Sydney Harbour and from a single site within St. Ann's Harbour (see Figure 1).

2.2 Sample Collection

The text in Section 4.2 was extracted from the field report by Devitt (1997). The field trip to collect sediment samples and benthic macroinvertebrates was conducted between July 10th and 12th, 1997.

2.2.1 Site Positioning. Sites were located using a using a Garman 75™ handheld global positioning system (GPS) and marked with a small, anchored buoy.

After anchoring, site coordinates were again recorded with the Garman 75™ GPS as well as a Trimble™ handheld GPS for verification. The site coordinates are tabulated in Appendix A, Table A-1 and depicted in Figure 1.

2.2.2 Sampling Methods. Water quality variables including depth, temperature, conductivity, salinity (calculated), pH, dissolved oxygen, and redox potential were measured at 3-min intervals to a depth of 12 m and at 0.5 m above the sediments (see Appendix A; Table A-2). A Hydrolab Water Quality Monitoring System was calibrated according to the manufacturers recommendations using a certified calibration standard for conductivity, certified grade pH buffers for pH, air and a manufacturer-supplied calibration table for dissolved oxygen, and a thermometer calibrated against a National Standards Board thermometer for temperature.

A 0.25 m^2 Van-Veen type grab sampler was used to collect sediments. After being winched to the surface, the overlying water was poured off and sediment placed in a plastic fish tote. Visual observations of consistency, benthic macroinvertebrate organisms, and odour were recorded. Temperature, redox potential, and pH were measured in the upper 5 cm of sediment. An alcohol-filled field thermometer was used to measure the temperature after stabilization. Redox potential and pH were measured using a Barnant 20™ digital pH/ORP/mV meter, calibrated according to the manufacturer's recommendations, using certified grade pH buffers.

Van-Veen grabs used to collect sediment for toxicity and chemical testing. Source: P. Topping

Sampling in Sydney Harbour. Source: P. Topping

To ensure the greatest possible homogeneity, grabs for benthos, toxicity tests, and physicochemistry were taken alternately. The benthic macroinvertebrate community was sampled using a 0.1 m^2 Van-Veen grab sampler. Five grabs were taken at each of the sites. The sediment and overlying water was placed in a fish tote and then sieved through a 0.5-mm nylon screen. The screened material and benthic macroinvertebrates were preserved in buffered formaldehyde.

2.2.3 Sample Homogenization, Storage, and Transport. Collected toxicity samples were manually pre-mixed using a stainless steel spoon. Samples were aggregated until approximately 60 L of sediment were collected for each of three replicates at each site. The replicate sample was homogenized using a ¾-inch, two-speed drill with a stainless steel paddle. The mixed samples were transferred to containers, pre-labelled with blind sample numbers and were stored in either coolers with gel packs or a large Xactit box with ice for the balance of the working day. Blind sample numbers were as in Appendix A; Table A-3: Blind Sample Numbers. Samples of control sediment (sediment where test organisms were collected or reared) were also taken for analysis by each lab.

After each working day, the samples were stored in a walk-in refrigerator at a temperature of 2–3^o C, at Highland Fisheries Ltd., in Glace Bay, Nova Scotia.

Samples destined for chemical and toxicity analyses were shipped from Sydney to Moncton (Mr. K. Doe, Environment Canada Laboratory, Moncton) and to North Vancouver (Mr. S. Yee, Environment Canada Laboratory, North Vancouver). Samples were shipped using refrigerated transport. Samples sent to Moncton were received within three days of shipping but those shipped to North Vancouver were delayed and arrived in a frozen condition 13 days after being shipped. In order to enable testing of fresh sediments in both east and west coast laboratories, the Moncton laboratory sub-sampled each sample they received and shipped those sub-samples by air cargo to the laboratory in North Vancouver to replace the frozen sediments. All tests and analyses were performed using sediments that had never been frozen.

Samples collected for AVS analyses were shipped by air cargo to Burlington (Ms. Annette Lawson, Dundas Environmental Consulting, Inc., Burlington, Ontario).

2.3 Biological Toxicity Test Methods

2.3.1 Data Manipulation. Biological toxicity test data were manipulated as follows:

- non-detected tissue PCB values are replaced with the detection limit of 0.48 ng/g PCBs);
- tissue PAH values below the detection limit are replaced by the sample-specific detection limit;
- the tissue metal means for Stations 1, 6 and 12 are estimated using two laboratory replicates while tissue metal means for Stations 5, 9, and St. Ann's Harbour are estimated using three laboratory replicates;
- a note on the original data spreadsheets indicates there may be an error in Station 6 tissue metal levels, but does not state what it might be; the data is used as is; and,

- a tissue Pb level less than the detection limit of 4 µg/g, was replaced by the detection limit for Station 12 reference site laboratory replicates.

2.3.2 East Coast Analyses. This text, extracted from Jackman and Doe (1997), describes the biological toxicity test methods used at the Environment Canada Toxicology Laboratory (EQL, Moncton, New Brunswick) while conducting biological toxicity tests on *Amphiporeia virginiana*, *Rhepoxynius abronius, Polydora cornuta*, *Boccardia proboscidea*, *Macoma nasuta,* and *Lytechinus pictus*.

Samples were received at the laboratory on July 15, 1997. Each sample was subsampled and shipped by air cargo to replace samples at the West Coast Lab, which had arrived frozen. The remaining samples were placed in a cooler at 4° C until used in the biological toxicity tests.

Sample buckets were removed from the cooler before test initiation and homogenized thoroughly. Adequate subsample was removed for testing purposes and the sample bucket was returned to the cooler until required for another test.

2.3.2.1 Amphipod Toxicity Tests. *A. virginiana* were collected by Environment Canada staff at Martinique Beach, NS on August 5, 1997 and returned to the EQL, Moncton the same day. Animals were at 7–12° C upon arrival and were acclimated to and maintained at 10 ± 2° C until used for testing on August 8, 1997. Samples numbered 2, 6, 11, 15, 23, 24, 33, 34, 38, 40, 48, 71, 73, 75, 80, 84, 87, 100, as well as the control sample number 19 (tested in triplicate) were analyzed.

R. abronius were purchased from Environment Resolution Services. Animals were collected on July 10, 1997 at Whidbey Island, Washington, USA. Animals were received at EQL, Moncton on July 11, 1997 at 16.5° C and maintained at 15 ± 2° C until used for testing on July 18, 1997. Samples numbered: 2, 6, 11, 15, 23, 24, 33, 34, 38, 40, 48, 71, 73, 75, 80, 84, 87, 100, as well as the control sample number 1(tested in triplicate) were analyzed.

On the day before test initiation, each 4-L bucket of test sediment was homogenized and 175-mL portions were added to each of five, 1-L glass mason jars. The jars were then filled with 800 mL of clean seawater (salinity 28 ± 2 ‰), covered, then aerated overnight with oil-free compressed air at a rate of approximately 150 mL/min. Tests were conducted according to Environment Canada (1992a).

The following day, amphipods were sieved from their holding sediment. Twenty animals were added to each of the test vessels. Animals were counted twice for quality assurance. Testing was performed with a 24-h light photoperiod with lighting provided by overhead fluorescent fixtures. *A. virginiana* testing was performed at 10 ± 1° C, while *R. abronius* testing was performed at 15 ± 1° C. Tests were checked daily for observations, aeration, and temperature. Three times a week, a replicate of each sample was monitored for temperature, pH, salinity, and dissolved oxygen. After 10 days, the contents of each jar were sieved through a 0.5-mm sieve. Any immobile animals were observed under a microscope to determine survival, lack of all movement when observed under a dissecting microscope for 5–10 seconds was counted as dead. Any animals missing were assumed to be dead. The mean and standard deviation of each treatment was calculated.

A reference toxicant test was conducted with cadmium chloride (CdCl) on each batch of amphipods using water only exposures for 96-h. Using the survival data at each test concentration, the 96-h LC50 (concentration calculated to cause 50% survival after 96-h exposure) was calculated using the methods of Stephan (1977).

2.3.2.2 Polychaete Survival/Growth Tests. Juvenile *P. cornuta* used in this assay were from laboratory cultures originally collected from Conrad's Beach, NS and maintained for several years in the laboratory at 23 ± 2° C. Juveniles

were 28–30 days old at the start of the test on July 30, 1997. Samples numbered: 6, 36, 57, 66, 68, 92 as well as the control sample number 32 were analyzed.

Juvenile *B. proboscidea* of 30–32 days old were used for the test performed on August 15, 1997. Juveniles were obtained from laboratory cultures originally supplied by Environment Canada, Vancouver, BC and maintained in the laboratory for several years. Samples numbered 35, 66, 68, 92, as well as the control sample number 4 were analyzed.

On the day before test initiation, each 4-L bucket of test sediment was homogenized and 175-mL portions were added to each of five, 1-L glass mason jars. The jars were then filled with 800-mL of clean seawater (salinity 28 ± 2 ‰), covered, then aerated overnight with oil-free compressed air at a rate of approximately 150 mL/min. Tests were conducted according to the draft protocol (EC, 1995a).

The following day, polychaetes were removed from their holding sediment. For *P. cornuta*, five animals were added to each of the test vessels. For *B. proboscidea*, only four animals were used due to an inadequate supply of juveniles. Several juveniles were taken at the start of the tests and washed and dried at 60°C to determine the initial weight of animals. Photoperiod for the testing was 16 hours of light and 8 hours of dark; and salinity was approximately 30 ± 2 ‰. Temperature was maintained at 23 ± 1° C throughout the testing. Animals were fed three times a week with a 1:1 mixture of finely ground *Enteromorpha* (green marine macro-alga) and Tetramin® (commercial fish flakes) at a rate of 5 mg per worm. Tests were monitored daily for temperature, aeration, and observations. Three times a week, a replicate of each sample was checked for pH, dissolved oxygen, temperature, and salinity. Approximately 80% of the water overlying the test sediment was renewed on day seven. The tests were terminated at 14 days and the contents of each jar were sieved through a 0.5-mm sieve. Any polychaetes not found at test termination were considered dead. Immobile animals were checked under a dissecting scope to confirm death. All surviving polychaetes were washed, dried, and weighed. The mean percentage survival of polychaetes in all the replicates was calculated. Mean weights for the five replicates of each treatment were compared to the mean weights of the control worms using the Sigma Stat Statistical Program (Version 1, Windows, 1994) from Jandel Scientific Software.

A reference toxicant test was conducted with CdCl on the *P. cornuta* using water only exposures for 96-h. Using the survival data at each test concentration, the 96-h LC50 was calculated using the methods of Stephan (1977). Due to an inadequate supply of *B. proboscidea*, no reference toxicant test was performed.

2.3.2.3 Echinoid Fertilization Inhibition Assay. White sea urchins, *L. pictus*, tested during the study were from the EP Laboratory stock (received from Marinus Inc. of Long Beach, California, USA, in 1994 and 1996). Testing was performed on July 23, 1997. Samples numbered 2, 6, 11, 15, 23, 24, 33, 34, 38, 40, 48, 71, 73, 75, 80, 84, 87, and 100 were analyzed.

Two, 250-mL portions of each sediment were centrifuged at 3000 rpm for 15 min. The supernatant liquids from the replicates were combined and centrifuged for an additional 15 min at 3000 rpm. Porewater was measured for temperature, pH, salinity, and dissolved oxygen. Four replicates of a dilution series for each porewater were prepared.

The test was conducted according to Environment Canada (1992b). Sea urchins were injected with 1 mL of 0.5 *M* potassium chloride (KCl) solution to induce spawning. Eggs produced from all females were pooled, and the concentration was adjusted to 2000 eggs/mL. Sperm were pooled from all males using the "dry" spawning technique, then stored in a vial on ice. A fixed "sperm-to-egg ratio" of 20 000:1 was used to produce approximately 90% fertilization in the controls. Sperm were activated immediately before test initiation.

Test volume was 10 mL and test temperature was 20 ± 1^{o} C. Sperm were exposed to the test solutions for 10 min, followed by an additional 10 min-exposure of the sperm and eggs. The test was then terminated using 2 mL of 10% formalin per replicate. One hundred eggs were examined from each replicate to determine the percentage fertilized. Percent fertilization data in each concentration of porewater were used to calculate the IC50 and IC25 for each sediment porewater tested. The linear interpolation method as implemented in the ICPIN program of Norberg-King (1993) was used to estimate the endpoints.

A reference toxicant test was performed simultaneously with the porewater toxicity tests using copper sulphate ($CuSO_4$). IC50s were calculated using the linear interpolation method or ICPIN (Norberg-King, 1993).

2.3.2.4 Bioaccumulation Test Using Baltic Clam. *M. nasuta* were purchased from A.K. Siewers of Santa Cruz, California, USA. The animals were received at EQL, Moncton, NB on July 17, 1997 at 22.5° C. The animals were placed in trays containing aerated seawater and collection site sediment, and were acclimated to 15° C and held until used for the testing on July 30, 1997. Samples numbered 9, 31, 35, 46, 72, 96, as well as the control sample number 39 were analyzed.

Tests were conducted according to USEPA (1993). On the day before test initiation, each 4-L bucket of test sediment was homogenized and 500-g portions were added to fifteen, 1-L glass beakers. Three beakers comprise one replicate for each test sediment and five replicates of each test sediment were performed. The jars were then filled with 500-mL of clean seawater (salinity 28 ± 2 ‰), then aerated overnight with oil-free compressed air at a rate of approximately 150 mL /min.

The following day, three clams were transferred to each beaker. Three sub-samples of nine test organisms were taken at the beginning of the test. The length, weight, and wet tissue weight were recorded, and the tissue was frozen for chemical analysis. Any animals dead or not buried in the first 24 hours were replaced. Daily recording of observations, temperature, and aeration occurred. Three times a week representative test chambers were analyzed for pH, salinity, temperature, and dissolved oxygen. Three times a week, the overlying water was 80% renewed with clean seawater. Lighting was provided by overhead fluorescent fixtures with 16 hours of light and 8 hours of dark, daily. The temperature was maintained at $15 \pm 1°$ C throughout the test.

After 28 days, the clams were removed from the test sediment, rinsed with clean seawater, and placed in clean collection site sediment for 24 hours. The three beakers comprising the one replicate were combined at this point. This depuration period allows for removal of gut content that could interfere with chemical analysis of tissue. The tissue samples were collected and wet weights recorded before submission for chemical analysis. The percentage survival was also computed at the end of the 28-day exposure.

2.3.3 West Coast Analyses. This text, extracted from Fennell and van Aggelen (1997), describes the biological toxicity test methods used at the Pacific Environmental Science Centre while conducting biological toxicity tests on *Eohaustorius washingtonianus*, *Eohaustorius estuarius*, *Boccardia proboscidea*, *Polydora cornuta*, and *Vibrio fischeri*. It should be noted that sediments arrived at the lab frozen and were not used for this study. Replacement sediments for the biological toxicity tests obtained from the East Coast lab were used instead.

The following text outlines the results of four sediment biological toxicity tests performed in July and August, 1997: 10-day survival tests using two amphipod species; 14-day growth and survival tests using two polychaete species; the Microtox® solid-phase metabolic-inhibition test; and an echinoid fertilization-inhibition test.

2.3.3.1 Amphipod Toxicity Tests. Amphipod sediment testing was performed using two species of infaunal amphipods, *E. washingtonianus* and

E. estuarius. The sediment samples were numbered 2, 6, 11, 15, 23, 24, 33, 34, 38, 40, 48, 71, 73, 75, 80, 84, 87, and 100, as well as the control sample numbers 29 and 13, respectively.

Amphipod test set up in EC Toxicology Laboratory in Moncton, NB. Source: K. Doe

E. washingtonianus were field collected at Esquimalt Lagoon, Vancouver Island by Biologica Environmental Services. *E. estuarius* were field collected at Beaver Creek, Oregon by Northwestern Aquatic Sciences. Amphipods were held in control sediment (i.e., collection site sediment) under continuous light and aeration and were acclimated at 15 ± 1° C over two or three days before test initiation.

Static 10-day acute survival tests were performed according to the procedures outlined in Environment Canada (1992a). The control sediment used in these tests was homogenized and wet sieved through a 0.5-mm stainless steel sieve to remove native organisms. Each test sediment sample was homogenized by hand. Five acid-washed 1-L jars were prepared for each control and test sediment. Approximately 175 to 200 g of sediment (to a height of 2 cm) was added to each jar. Each container was then carefully filled with a fresh laboratory supply of sand-filtered seawater from Burrard Inlet, being careful not to disturb the sediment layer. The test containers were aerated and allowed to settle overnight. Twenty randomly selected amphipods were added to each of five replicate jars per sediment. Water quality (temperature, pH, salinity and dissolved oxygen) was monitored periodically throughout the test in replicate A. The biological toxicity tests were conducted in an environmental chamber at 15±1° C under continuous light. At the conclusion of the biological toxicity tests, the total number of emergent amphipods on the sediment surface (or swimming in the water column) of each test container was recorded. The sediments were then wet-sieved through a 0.5-mm stainless steel screen, and total surviving, dead and missing amphipods were recorded. Means and standard deviations were calculated for percent survival and percent emergent[1].

In addition, 96-h LC50 positive control tests were run concurrently, using various concentrations of the reference toxicant CdCl in seawater, to assess the acceptability of test conditions and amphipod sensitivity in reference to historical performance under the same conditions (including such conditions as darkness and absence of substrate). The LC50 values (and associated 95% confidence limits) for the positive reference toxicant tests were determined using the Environment Canada computer program following Stephan (1977).

2.3.3.2 Polychaete Survival/Growth Tests.
Fourteen-day polychaete sediment biological toxicity tests and concurrent 96-h positive control reference toxicant tests with cadmium were performed using *B. proboscidea* and *Polydora cornuta*, two species of spionid polychaetes cultured in-house. When juveniles were almost three to three-and-one-half weeks old, they were considered ready for use in toxicity tests.

Control sediment consisted of sieved (500 µm) and rinsed (clean natural seawater) sediment from the polychaetes' natural environment. *B. proboscidea* sediment was collected from Witty's Beach, Vancouver Island, BC, and *P. cornuta* control sediment was collected from Conrad's Beach, NS. Samples numbered 6, 66, and 68 as well as the control sample numbers 32 and 4 were analyzed. Tests were conducted

[1] Not analyzed.

according to Environment Canada's draft protocol (1995a).

A controlled environment room was set to uniformly maintain 23 ± 1° C and a photoperiod of 16 hours light to eight hours dark. The test vessels were prepared the day before the polychaete introduction (Day-1). Five acid-washed 1-L glass mason jars were each filled with 175–200 mL (to a height of 2 cm) of a test sediment, to which 750–800 mL of clean control/dilution water was added (fresh laboratory supply of sand-filtered natural seawater from Burrard Inlet). The jars were aerated with filtered, oil-free compressed air overnight, and for the duration of the test, at a steady rate of approximately 150 mL/min/L through plastic aquarium airline tubing and pre-cut, disposable 1-mL polystyrene pipettes.

On Day 0 (test initiation day) rearing vessels containing polychaetes of appropriate testing age were sieved. Five polychaetes were added per test replicate chamber. Pre-weighed aluminum pans containing a known number of test age juveniles were dried overnight (60° C) for initial weight determination.

Water quality parameters (pH, dissolved oxygen, temperature, and salinity) were measured periodically throughout the 14-d test period. Approximately 80% of the overlying water was replaced on Day 7. Test organisms were fed every Monday, Wednesday, and Friday 500 µL per test vessel (= 5 mg per worm). The food consisted of a 50:50 by weight blend of ground *Enteromorpha* spp.: Tetramarin® (a green alga and commercial fish flake, respectively), each ground to a fine powder and mixed up into a seawater slurry (2.5 g *E*:T/50 mL seawater).

On Day 14, test vessels were sieved (500 µm) and the numbers of surviving, dead, and missing polychaetes were recorded. Surviving polychaetes were rinsed in de-ionized water before placement in pre-weighed aluminum pans for final dry weight determinations.

The means and standard deviations for percent survival and growth achievement of polychaetes exposed to contaminated test sediment samples and negative control sediment were calculated. The negative control (sediment from original polychaete collection site) provides not only a basis for interpreting data obtained from any reference and test sediments, but also provides evidence of the relative quality of the test organisms and suitability of test conditions and procedures. An acceptable (≥ 90%) control survival level must be achieved for a test to be considered valid.

The LC50 values (and associated 95% confidence limits) for the positive control reference toxicant tests were determined using the Environment Canada computer program based on Stephan (1977) and were compared with values derived from previous reference toxicant testing.

2.3.3.3 Marine Photoluminescent Bacterium Acute Toxicity Test. A marine bioluminescent bacterium, *Vibrio fischeri,* was used to assess the toxicity of the test sediments using the Microtox® test system. Vials of freeze-dried *V. fischeri* stored at –20 ± 2° C were reconstituted in 1.0 mL of distilled water and incubated at 5.5 ± 1° C for no less than 20 min before use in solid-phase tests. Test results were based on measured light output in the presence of various levels of test substance in aqueous solutions, which were compared with light output of a control blank (i.e., bacterial cell suspension in diluent only). Light output is a product of the electron transport system and relates directly to the metabolic state of the bacteria (Schiewe *et al*., 1985). The degree of light loss (degree of metabolic inhibition in the bacteria) indicated then the degree of toxicity of the sample.

The sediment remaining in one polystyrene 50-mL tube following centrifugation was homogenized before solid-phase testing carried out according to methods outlined by Microbics Corporation (1992). Bacteria were incubated for 20 min at ambient room temperature in a series of aqueous solutions of various concentrations made up of the sediment sample and a 3.5%

solution of Reagent Grade NaCl crystals dissolved in de-ionized water. Following this incubation period of direct bacterium-particle interaction, the solutions were filtered and 500 µL of each filtrate was transferred to a corresponding glass cuvette within the incubation unit. After a further five-minute incubation period at 15.0 ± 0.5° C, light emission from each concentration was measured.

A Microtox® model 500 Toxicity Analyzer (Beckman Instruments, Carlsbad, CA) controlled by the appropriate Microtox® software (version 7.03) was used for all procedures. A dose-response curve was determined by Microbics software (version 7.03 for solid-phase), on which the IC50 was located. A 95% confidence range was also reported. The IC50 is the inhibiting concentration of a sample causing a 50% decrease in the bacterial light output under defined conditions of exposure time and test temperature. IC50s derived from solid-phase testing were corrected for moisture content by standard laboratory procedures based on Microbics Corporation (1992) using the remaining sediment in the tube on the day of testing and oven-drying (overnight at 100 ± 5° C) three replicates of 5.0 ± 0.2 g per sediment sample.

Microtox® Model 500 Toxicity Analyzer. Source: K. Doe

2.3.3.4 Echinoid Fertilization Inhibition Assay. Fertilization inhibition tests were performed using the gametes of the echinoderm *Dendraster excentricus* (eccentric sand dollar). Sand dollars spawned for collection of gametes were field-collected in May 1997 at low tide from Crescent Beach, White Rock and were held at the laboratory in an outside tank with a 7–8 cm bed of Crescent Beach sand and a source of flowing seawater. Testing procedures were those outlined in Environment Canada (1992b) and British Columbia Ministry of Environment, Lands and Parks (1994). Samples numbered 2, 6, 11, 15, 23, 24, 33, 34, 38, 40, 48, 71, 73, 75, 80, 84, 87, and 100 were analyzed.

Five full 50-mL polystyrene tubes per test sediment were centrifuged for 30 min at 4000 rpm and 4° C to extract the pore water from the sediment. The interstitial water was collected into beakers and water quality parameters were measured.

Control/dilution water was a laboratory supply of sand-filtered seawater from Burrard Inlet which was subsequently filtered through a 0.8-µm filter, adjusted with natural brine salts to match the salinity of the most saline pore water samples within 2 ‰ to a minimum of 28 ‰, aerated gently, and held at 15°C.

A positive reference toxicant test using a range of $CuSO_4$ concentrations was run concurrently to measure species sensitivity and acceptability of test conditions.

Following wet spawning of the sand dollars, collected sperm and eggs were kept separate to avoid gamete contamination. For each gender, gametes from at least three individuals were pooled and after density determinations, dilutions were made to achieve a final sperm-to-egg ratio of 2000:1 in a 2.0-mL test volume.

Initially, sperm were exposed for 10 min to three replicates of full strength (100%) pore water obtained from each sediment sample. Following the 10-min sperm-only exposure, eggs were added for an additional 10-min exposure period. Immediately thereafter, the samples were

preserved with 10% buffered formalin to fix the eggs.

Fertilization rates in each pore water sample were determined by calculating the average for all replicates of the number of eggs with fertilization membranes counted out of the first 100 eggs encountered under a microscope for each replicate.

The results from the positive reference toxicant test were adjusted using Abbott's formula (Finney, 1971) to correct all values for mean percent unfertilized eggs at test end, in keeping with the variable and gamete-dependent differences from test to test with respect to fertilization success rate and the associated percentage of unfertilized control eggs. Thereafter, Environment Canada's statistical package for calculating LC50s, based on Stephan (1977) was used to compute the IC50 (and its associated 95% confidence limits).

2.4 Benthic Macroinvertebrate Survey

This text was extracted from the report by Arenicola Marine (1997).

2.4.1 Data Manipulation. Data was manipulated as follows:

- Taxa with no entries were deleted. The only taxon included in the taxa list that did not have any entries was *Tharyx marioni*.

2.4.2 Data Presentation. The following indices of benthic community structure were calculated and are presented graphically in Section 3.2, total number of organisms, number of species, Simpson's Diversity Index, Pielou's J (evenness), McIntosh's Index (evenness), Margalef's Index (diversity), and Shannon's H (diversity).

The total number of organisms and the number of taxa per square metre are summarized for each station. It is generally thought that a community suffering a deleterious impact will be characterized by a relatively small number of organisms belonging to a few taxa.

Simpson's Index (1949) is a diversity index measuring the probability that two organisms chosen at random from a population will belong to the same taxa. A diverse community will produce a low value of Simpson's Index. This index does not account for the non-uniform distribution typical of benthic macroinvertebrates. The index is also a function of sample size; thus comparison of samples from different locations or samples collected using different methods are confounded by sample size.

Shannon's H´ as described in Shannon and Weaver (1949) is a diversity index founded on information theory. The validity of this measure of diversity is largely a consequence of different interpretations of the term "diversity." It has been shown that diversity may increase even when the species numbers decrease, if evenness increases (Hurlbert, 1971). Wilhm (1970) suggests that diversity indices > 3 are characteristic of diverse sites, while values < 1 suggest gross pollution.

Evenness is a measure of how evenly species are distributed across taxa. Pielou's J is a measure of evenness and is often given by H´/ H´max. It has a maximum value of 1. Diversity and evenness measures are usually highly correlated.

McIntosh's Index is a measure of equitability or evenness similar in form to that of Pielou's J with an index being divided by the maximum value attainable by that index. The index is based on the Euclidean measure of distance; consequently, it is also known as McIntosh's Ecological Distance (McIntosh, 1967).

Margalef's Index (1958) is another diversity index and assumes a linear relationship between species abundance and number of species.

2.5 Sediment Physical/Chemical Analyses

2.5.1 Data Manipulation.

- All sediment samples were subsampled to produce two or three pseudoreplicates. One of the two sediment subsamples for PAHs for

the St. Ann's Harbour reference station, was further split into two sub-subsamples. The two sub-subsamples were averaged and then averaged with the single subsample. When present, subsample values were averaged and averages were used in subsequent calculations.

- Metal levels below the detection limit were replaced by the detection limit.

2.5.2 Oxidation Reduction (Redox) Potential, Ammonia, and Sulphide. Sediment samples were thoroughly homogenized and subsampled for analysis of sulphide, redox potential (Eh), and ammonia by specific ion electrode according to the manufacturer's instructions and advice by Dr. B. Hargrave (Dept. of Fisheries and Oceans, Scotia–Fundy Region). These analyses were conducted in triplicate. Results for sediments are expressed as μg S/g dry weight of sediment for sulphide, μg NH_3-N/g dry weight of sediment for ammonia and millivolts corrected for the normal hydrogen electrode for redox potential. Other subsamples of these sediments were centrifuged at 3000 rpm for 15 min, the porewater was decanted off and analyzed for ammonia and pH. Porewater ammonia is expressed as mg NH_3-N/L, the pH is expressed in pH units. Testing was conducted on July 30 to August 1, 1997. Samples numbered 2, 6, 11, 15, 23, 24, 33, 34, 38, 40, 48, 71, 73, 75, 80, 84, 87, and 100 were analyzed.

2.5.3 Sediment Metal Concentrations. Total sediment metal concentrations were measured using inductively coupled plasma atomic emission, graphite furnace atomic absorption, and cold vapour atomic fluorescence for mercury (Hg) (PESC, 1999a).

2.5.4 Porewater Metal Concentrations. Total porewater metal concentrations (with the exception of porewater Hg) were measured using inductively coupled argon plasma atomic emission spectrometry, and graphite furnace atomic absorption spectrometry (PESC, 1999b). Porewater Hg was measured using cold vapour atomic fluorescence spectrometry following acid digestion (PESC, 1999c).

2.5.5 Sediment Organic Compound Concentrations. Sediment organic compounds (PAHs and PCBs) were measured by the Atlantic Region Monitoring and Evaluation Branch (EC, 1997) method. This involves extraction into 1:1 mixture of hexane and ethane while sonicating. The extract is mixed with acidified water, and back-extracted using hexane. The second stage extract is dried using anhydrous sodium sulphate, and cleaned with a silica-gel mini-column. Further cleanup using toluene follows. The extract is made up to volume and is analyzed using gas chromatography with mass spectrometric detection for PAHs, and gas chromatography with electron capture detection for PCBs.

2.5.6 Porewater Organic Compound Concentrations. The analysis of PCBs and PAHs was conducted following methods outlined in the Atlantic Region Environmental Quality Laboratories (EC, 1992). Porewater was extracted by centrifugation (Jackman and Doe, 1997). Organic compounds are extracted into hexane, dried by passing through anhydrous sodium sulphate, then cleaned if necessary on a silica-gel mini-column. The extract is made up to volume and analyzed using gas chromatography with mass spectrometric detection for PAHs, and gas chromatography with electron capture detection for organochlorine compounds, PCBs, and chlorinated benzenes.

2.5.7 AVS and SEM Metals. The extraction procedure for AVS and SEM metals follows Allen *et al.*, (1993) and was performed by Dundas Environmental Services, Burlington, Ontario. Unfortunately, SEM Ni was not measured by the contractor and is therefore not included in the estimation of total SEM.

2.6 Tests of Hypotheses

The methods used to test each hypothesis are presented in Section 4 "Tests of Hypotheses."

Survey Observations and Biological Toxicity Test Results

3.1 Biological Toxicity Tests

The results of biological toxicity tests are best presented using Box and Whisker plots. Due to the lack of this type of plot in Microsoft Excel®, a similar plot is presented without the inclusion of boxes. These graphics show the maximum, minimum, median and 25th and 75th percentiles.

The following plots use the subsamples as raw data. Thus the sample size for a site with three field replicates and five lab replicates or subsamples, is 15. This procedure is used to present the raw data and is not used during data analyses in Section 4.

3.1.1 Acute Tests for Sediment Toxicity Using Marine Amphipods. Results from these tests are shown in (Figures 2, 3, 4, and 5).

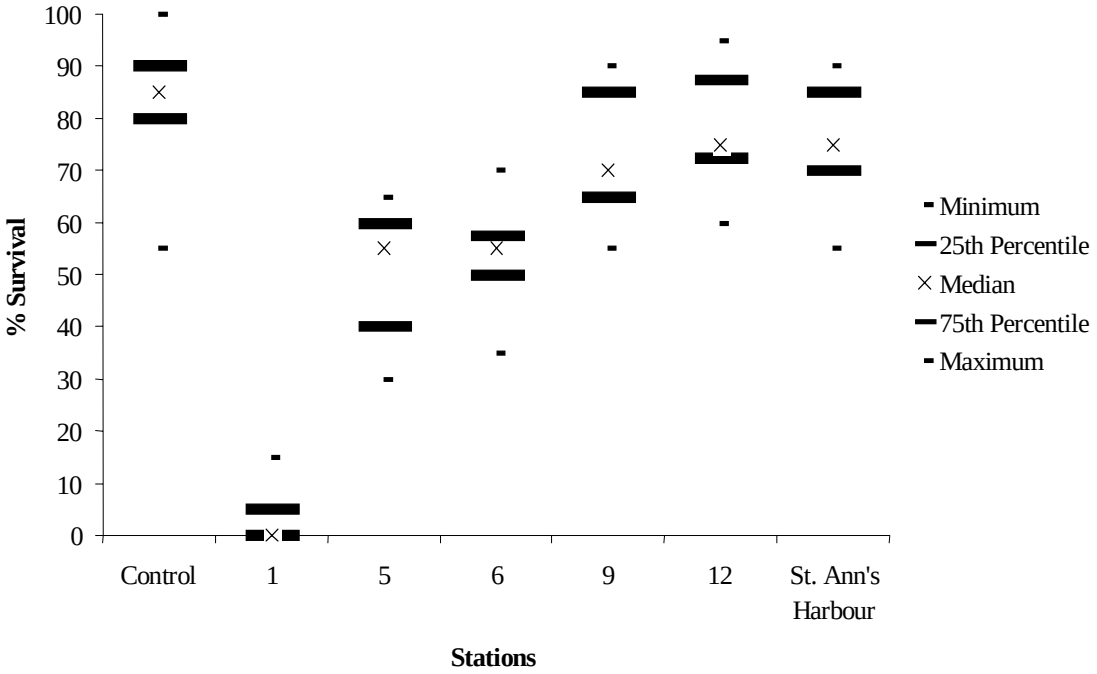


Figure 2 Percent Survival for *Amphiporeia virginiana.* Survival is very low at Station 1, but gradually improves along the gradient. Note that survival in the reference site is still not as high as the control sediment.

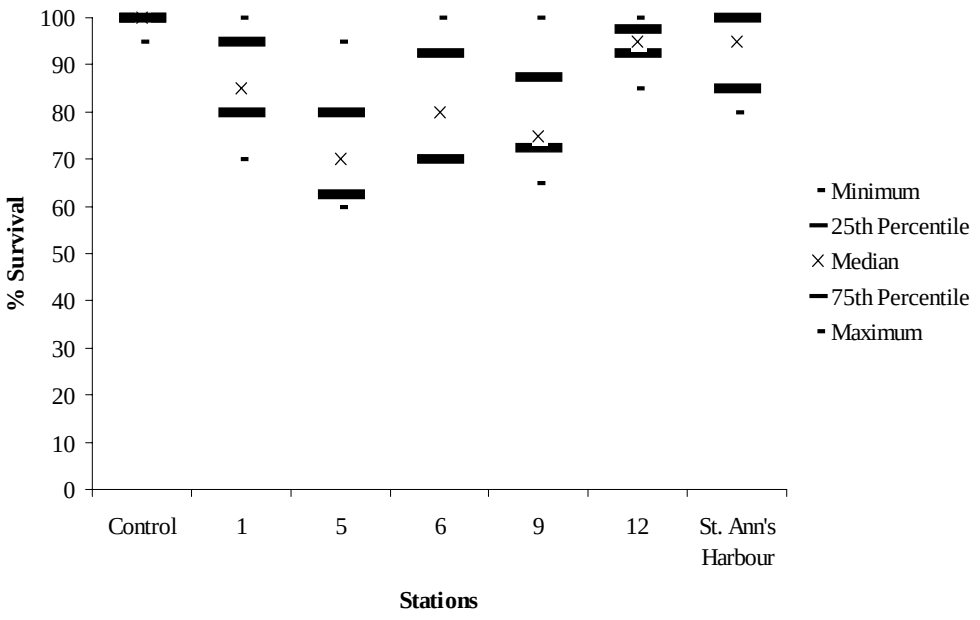


Figure 3 Percent Survival for *Rhepoxynius abronius*. A minimum survivorship is seen at Station 5 with a gradual improvement in survivorship to the reference station. This response is not as strong as that observed for the other amphipods and is obscured by variability.

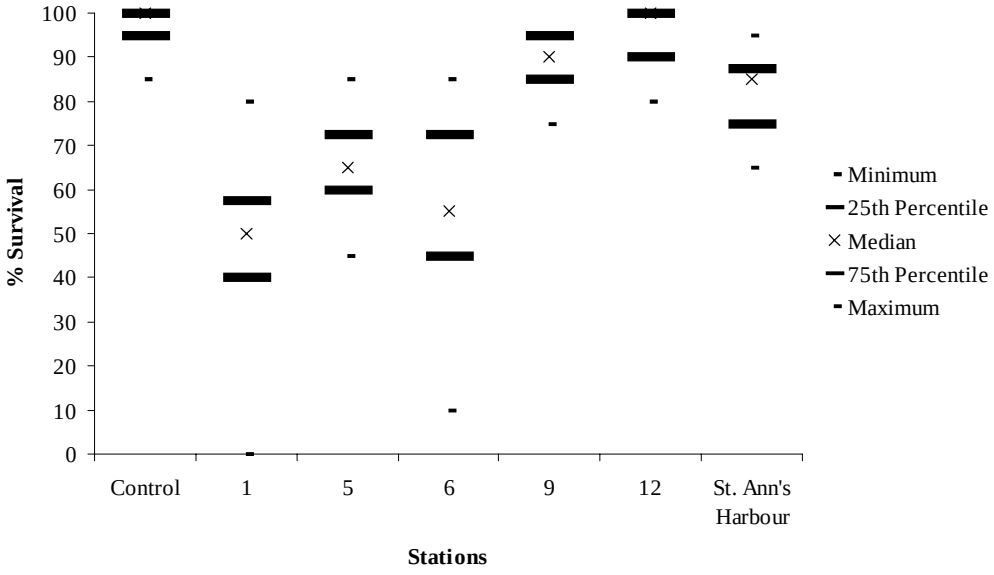


Figure 4 Percent Survival for *Eohaustorius washingtonianus.* The survival of *Eohaustorius washingtonianus* is similar to that of the other amphipods.

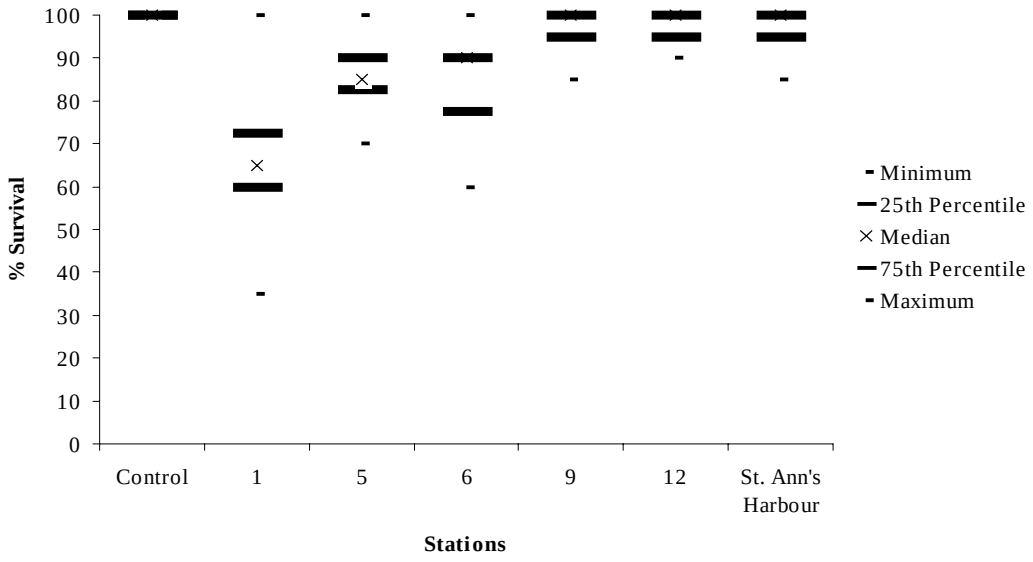


Figure 5 Percent Survival for *Eohaustorius estuarius*. Again, there is a decreased survivorship at Station 1 with a gradual increase in survivorship to the reference station.

3.1.2 Sublethal Toxicity Test for Metabolic Inhibition Using a Marine Bacterium.

The Microtox® solid-phase toxicity test was performed on both a wet-weight and moisture-corrected basis. Moisture-corrected IC50s are presented in Figure 6.

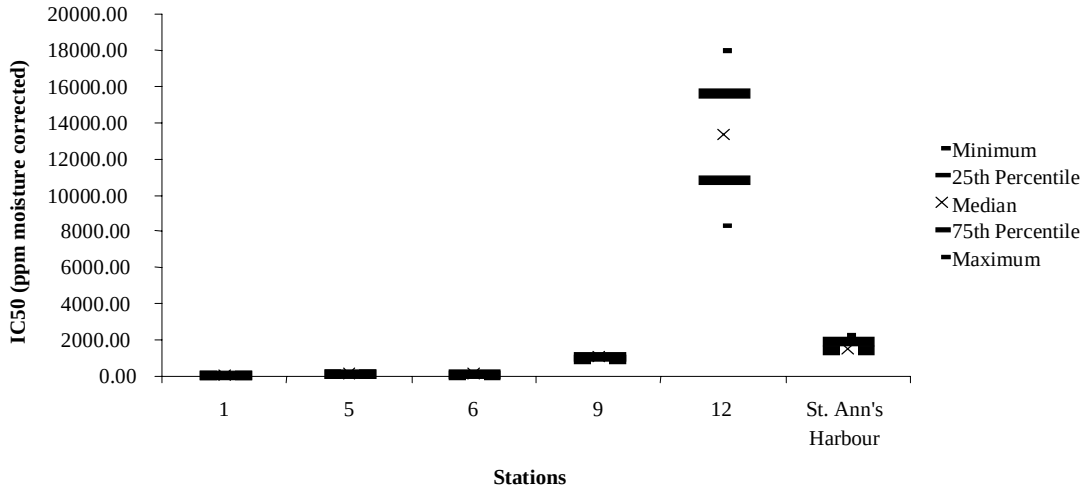


Figure 6 Photoluminescent Bacteria Light Inhibition. The Microtox® assay shows an increase in IC50 along the gradient with greatly elevated IC50 at the Station 12 reference site but a decrease at the St. Ann's Harbour reference site. The mean IC50 for stations, 1, 5, 6, 9, 12 and St. Ann's Harbour are 97, 122.67, 144.67, 1009.33, 13200, and 1733.333 ppm, respectively.

3.1.3 Sublethal Test for Sediment Toxicity Using Marine Polychaetes. The results of toxicity tests conducted by the East coast laboratory using *Polydora cornuta* and *Boccardia proboscidea* are presented in Figures 7, 8, 9, and 10. Polychaete testing was also conducted at some stations along the gradient by the Pacific Environmental Science Centre where sufficient sample was available. These data have not been incorporated into this analysis but are presented in Appendix A; West Coast Polychaete Analyses.

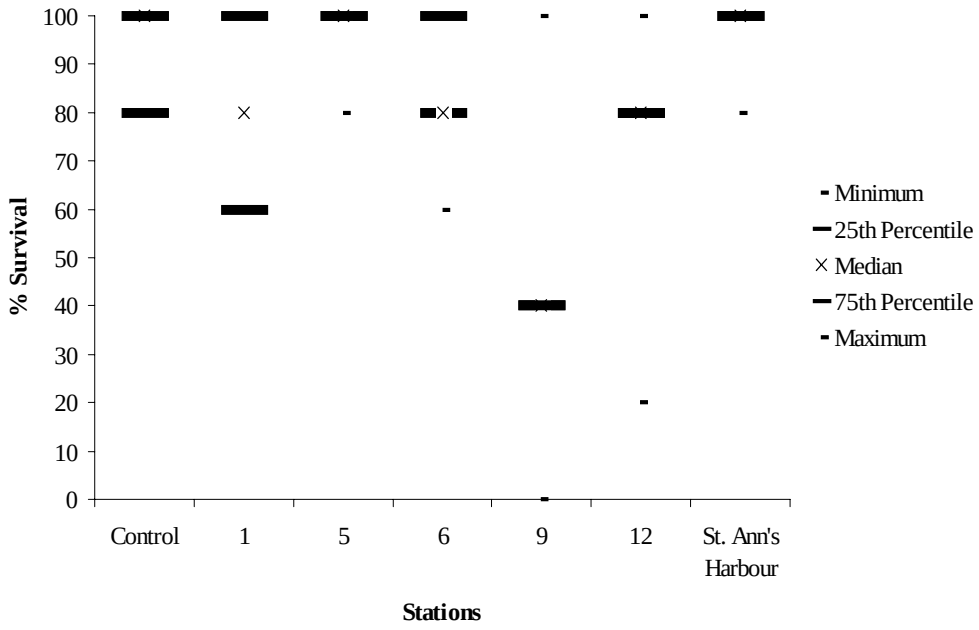


Figure 7 Percent Survival for *Polydora cornuta*. The survivorship for *Polydora cornuta* is reduced at Station 9 and is extremely variable at Stations 9 and 12. Note that approximately half (17/35) of the tests showed complete survival.

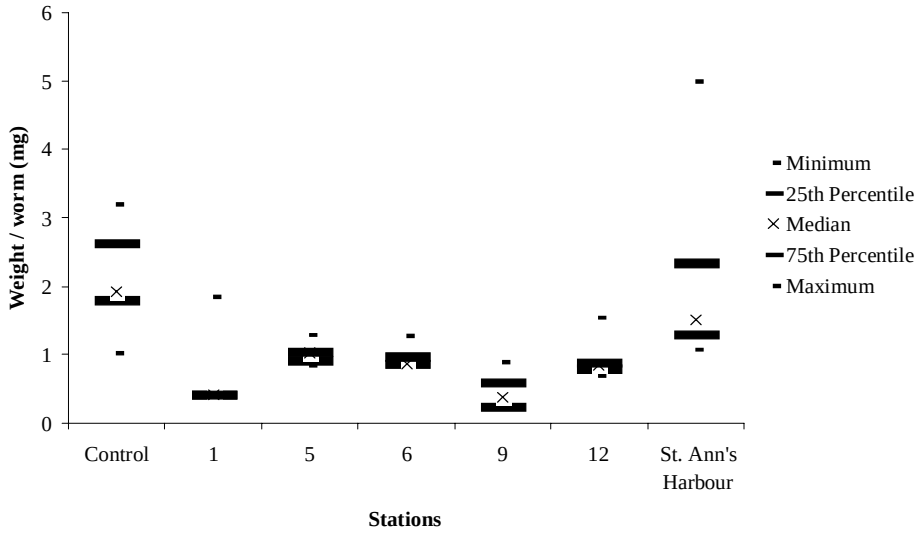


Figure 8 Growth for *Polydora cornuta.* Growth is depressed at intermediate stations relative to the control station and St. Ann's Harbour.

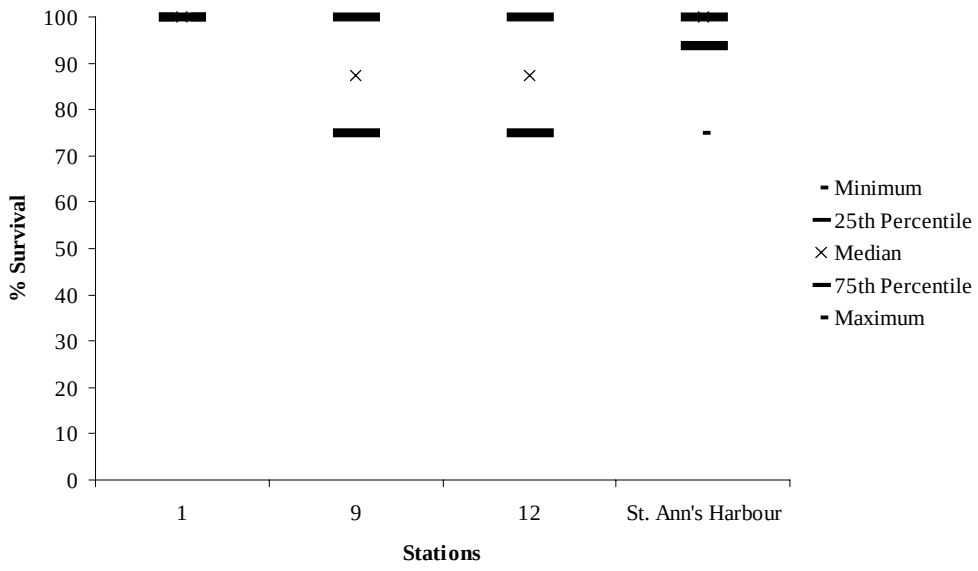


Figure 9 Percent Survival for *Boccardia proboscidea*. As for *Polydora cornuta*, most organisms have high survival in the sediments.

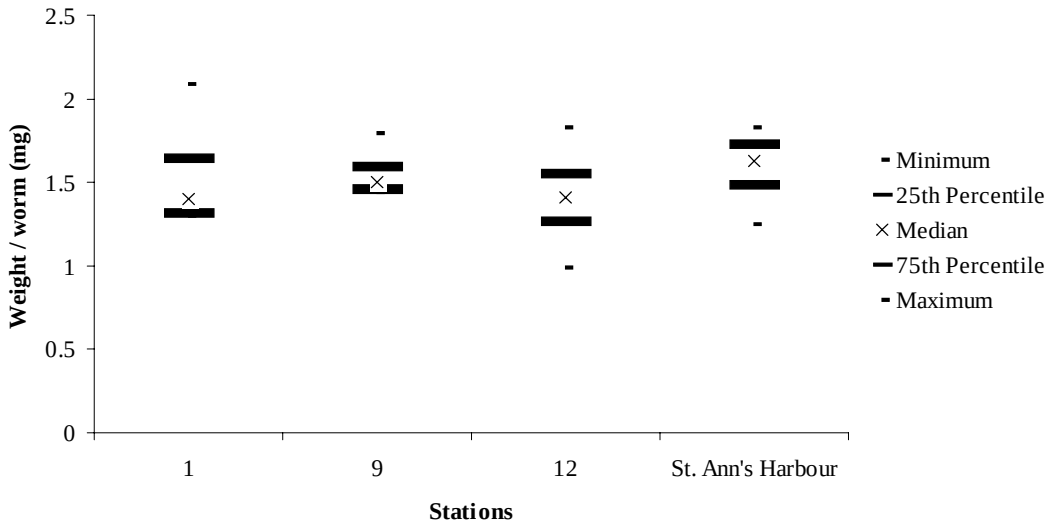


Figure 10 Growth for *Boccardia proboscidea.* There seems to be no trend in *Boccardia proboscidea* growth.

3.1.4 Sublethal Toxicity Test for Echinoid Fertilization Inhibition. Fertilization inhibition is presented for *Dendraster excentricus* (Figure 11) and *Lytechinus pictus* (Figure 12).

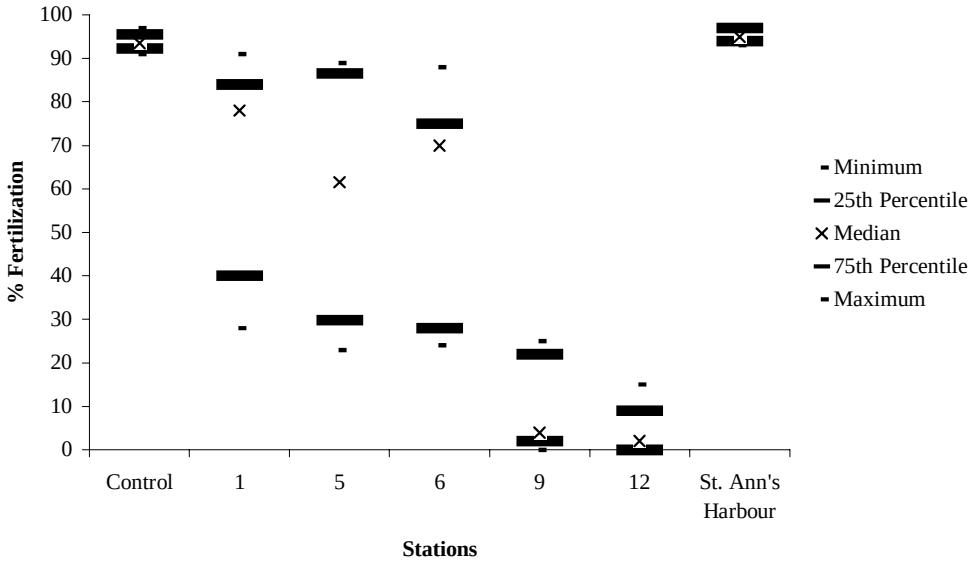


Figure 11 Percent Fertilization for *Dendraster excentricus.* The percent fertilization decreases along the gradient and is a maximum at the reference station. This is the opposite of what is expected along an organic contaminant gradient. Note that the St. Ann's harbour station performs similarly to the control sediment whereas the Station 12 reference sediment does not. Note the large degree of variability.

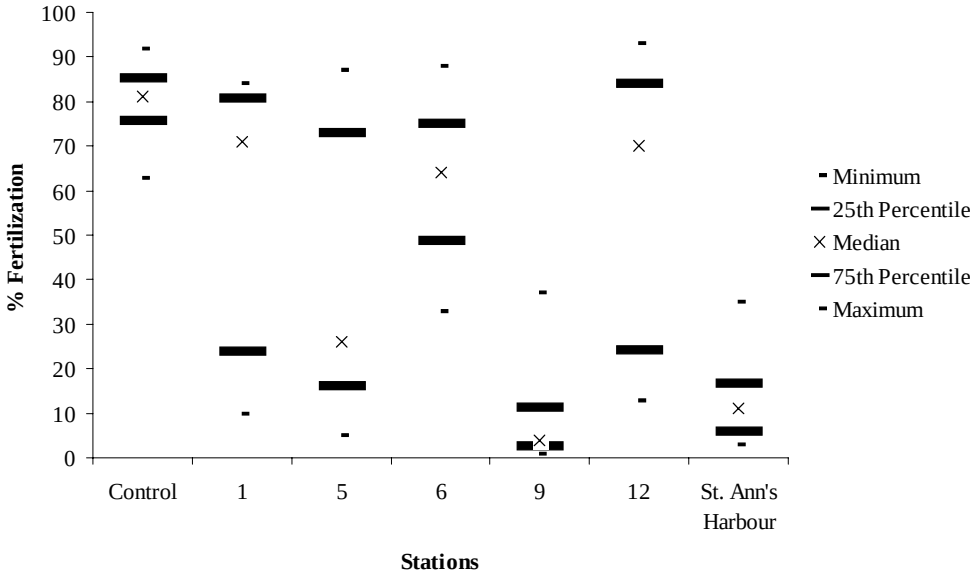


Figure 12 Percent Fertilization for *Lytechinus pictus*. The results are generally variable with lowest fertilization rates seen at Station 9. Like the results with *Dendraster excentricus* the reduced fertilization rate at Station 9 relative to Stations 1, 5, and 6 is unexpected. Note that the St. Ann's harbour station does not perform as well as the station 12 reference station for *Lytechinus pictus*. Again, note the large degree of variability.

3.1.5 Summary Plots and Statistics for Biological Toxicity Tests. The biological toxicity test mean responses are plotted jointly in Figure 13 to show how organisms react along the gradient. Each of the biological test responses is scaled by the maximum value of the response for that biological toxicity test. Thus a value of 1 corresponds to the maximum for each test. This allows the reader to view the relative change from station to station.

The mean biological test responses are also presented in Table 2.

Sediment pass/fail status is discussed in Section 4.4.

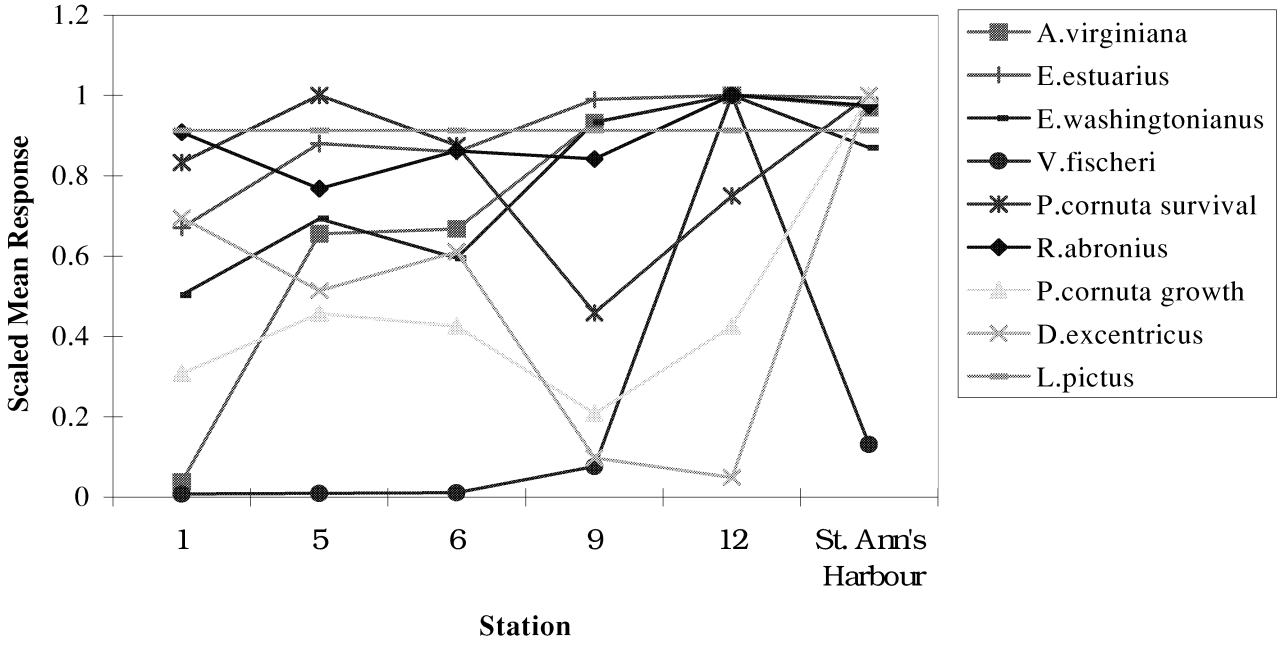


Figure 13 Scaled Biological Test Responses. The amphipods show similar relative responses with some discrepancies at Station 1. The *V. fischeri* IC50 values appear to correspond with the proportion of surviving *A. virginiana*, *E. estuarius,* and *E. washingtonianus*. Note the extreme variability in the fertilization rates for *Lytechinus pictus* results in a relatively stable mean response across all stations. *D. excentricus* percent fertilization and *P. cornuta* growth and survival show relative trends that do not coincide with the other biological test responses.

Table 2 Summary of Mean Biological Test Responses

Response	Test	Station						
		Control	1	5	6	9	12	St. Ann's Harbour
Percent Survival	*A. virginiana*	83.33	3.00	52.00	53.00	74.00	79.33	77.00
Percent Survival	*R. abronius*	99.33	86.00	72.67	81.67	79.67	94.67	92.33
Percent Survival	*E. estuarius*	100.00	65.33	85.67	83.67	96.33	97.33	96.67
Percent Survival	*E. washingtonianus*	97.00	47.67	65.67	56.33	88.33	94.67	82.33
IC50 (ppm moisture corrected)	*V. fischeri*		97	122.67	144.67	1009.33	13200	1733.333
Percent Survival	*P. cornuta*	92.00	80.00	96.00	84.00	44.00	72.00	96.00
Growth (mg)	*P. cornuta*	2.11	0.69	1.03	0.95	0.47	0.95	2.24
Percent Survival	*B. proboscidea*		100.00			87.50	87.50	93.75
Growth (mg)	*B. proboscidea*		1.55			1.56	1.41	1.58
Percent Fertilization	*D. excentricus*	93.83	66.11	48.89	58.11	9.33	4.67	95.11
Percent Fertilization	*L. pictus*	80.67	56.50	39.00	61.92	9.50	58.83	12.50

3.1.6 Bioaccumulation Test Using Bivalves. The survival results for the *Macoma nasuta* bioaccumulation study are presented in Figures 14, 15, 16, and 17. Tissue levels for contaminants of concern are presented below. Relationships between tissue contaminants and sediment and porewater contaminants are examined in Section 4.2.4.

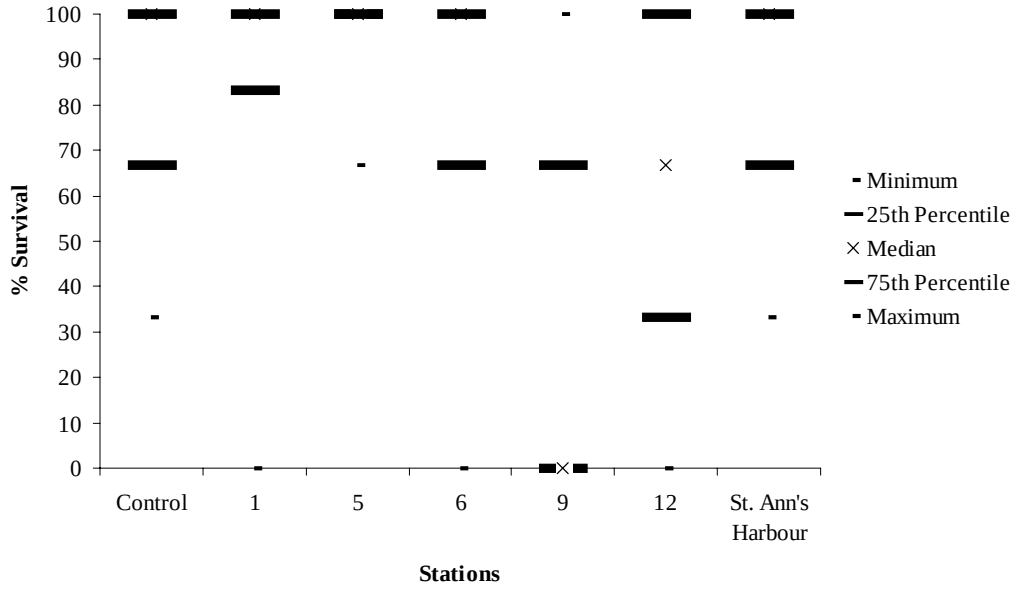


Figure 14 Percent Survival for *Macoma nasuta*. Survival is quite variable, but 100% survival occurs in at least one test from every site. Also, decreased survival is observed at Station 9, which is unexpected, given the organic contaminant gradient.

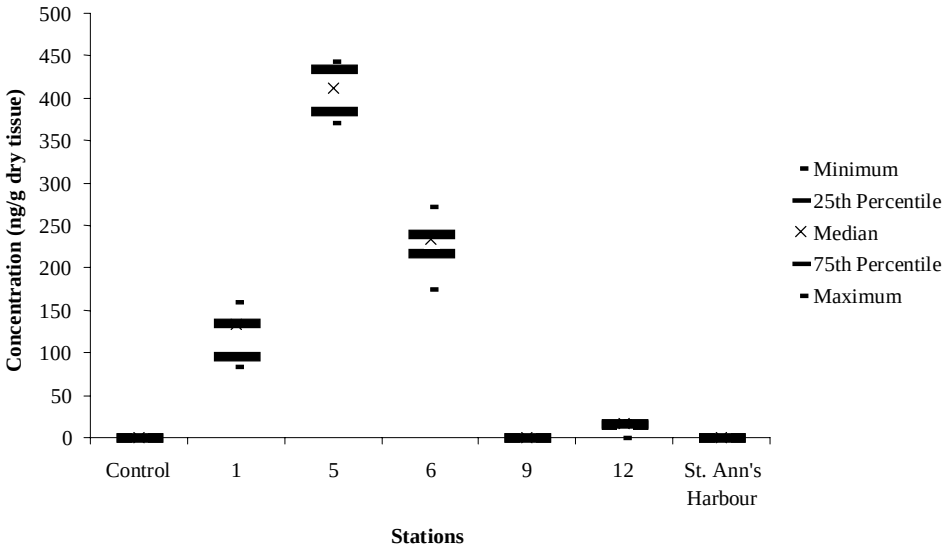


Figure 15 PCB Tissue Concentrations in *Macoma nasuta.* PCB tissue levels reach a maximum at Station 5 and decline to control levels at the reference station. Note that as survival at Station 9 was minimal, bioaccumulation at that station was not expected.

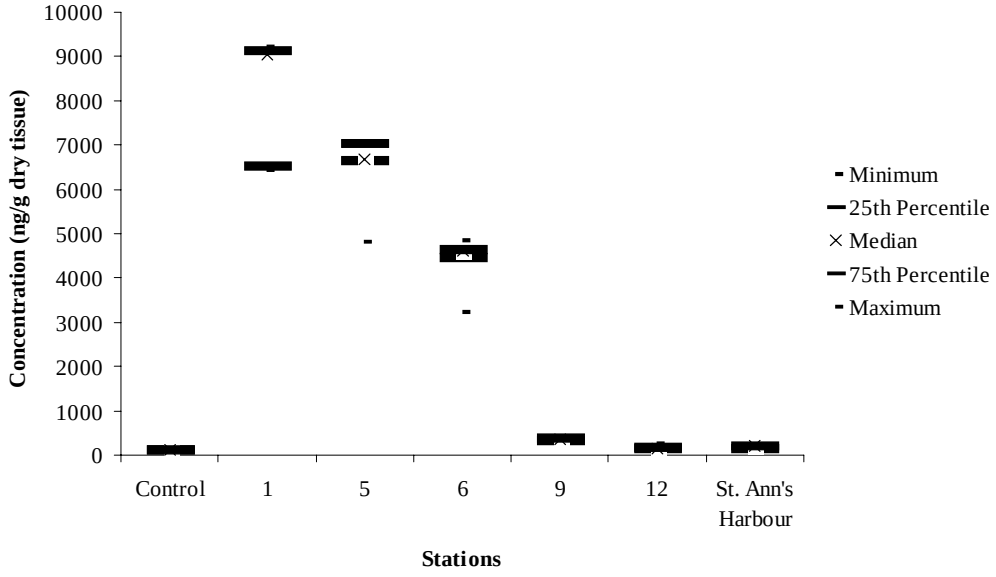


Figure 16 PAH Tissue Concentrations in *Macoma nasuta.* PAH tissue concentrations decrease monotonically from Station 1 closely following the trend in sediment PAHs.

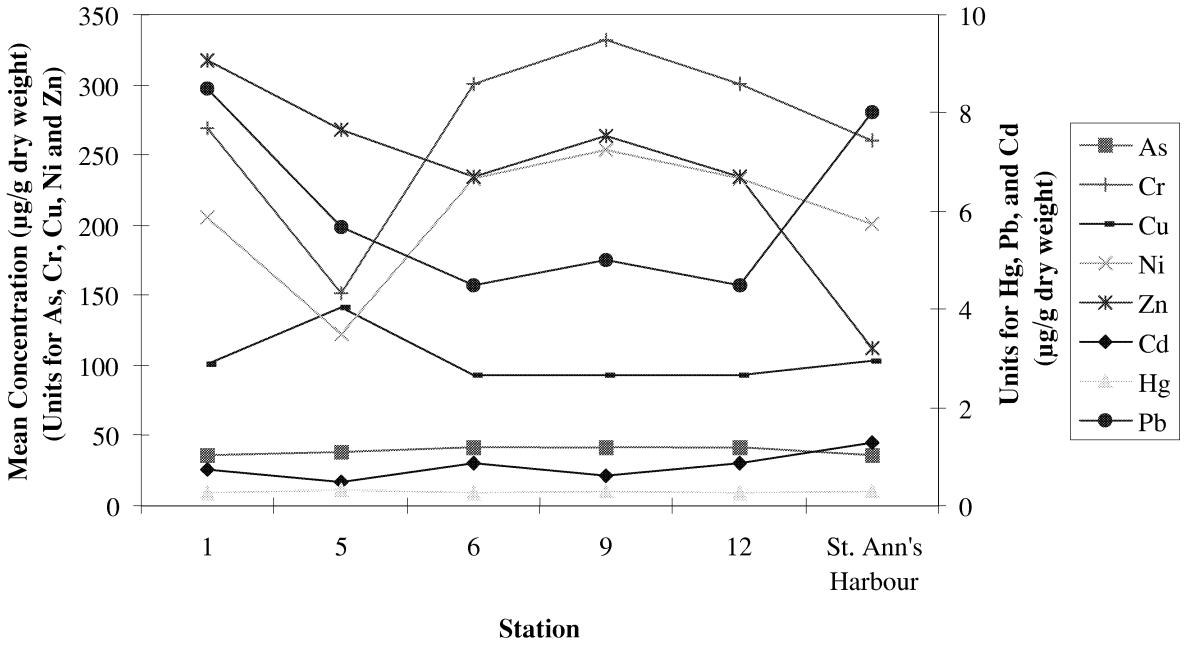


Figure 17 **Tissue Metal Levels in *Macoma nasuta*.** Little evidence of trends for tissue levels of As, Cd, and Hg is seen. Cr and Ni tissue levels exhibit some degree of structure; however, a superimposed line of best fit would indicate no trend. Pb values show an inverse quadratic trend from Muggah Creek (Station 1) to St. Ann's Harbour. Zinc levels drop from Muggah Creek, while Cu values show a spike at Station 5. Ni, Cr and Zn show similar patterns. The tabulated data is presented in Table 3.

Table 3 **Summary of *Macoma nasuta* Mean Tissue Metal Levels (µg/g dry weight)**

	Station					
Parameter	1	5	6	9	12	St. Ann's Harbour
As	35.5000	37.6667	41.5000	41.0000	41.5000	35.6667
Cd	0.7500	0.4667	0.8500	0.6000	0.8500	1.2667
Cr	268.8000	151.9000	301.0000	332.2000	301.0000	260.4333
Cu	101.1500	140.9167	92.9700	93.3600	92.9700	102.8000
Hg	0.2685	0.3067	0.2480	0.2757	0.2480	0.3013
Ni	205.0000	122.8333	233.0000	253.3333	233.0000	201.0000
Pb	8.5000	5.6667	4.5000	5.0000	4.5000	8.0000
Zn	317.0500	268.4000	234.6000	263.1000	234.6000	112.1000
Moisture %	82.5000	81.5667	41.6150	81.3000	82.7000	78.9700

3.2 Benthic Macroinvertebrate Survey

Figures 18, 19, and 20 represent the benthic macroinvertebrate community richness, abundance, and structure indices.

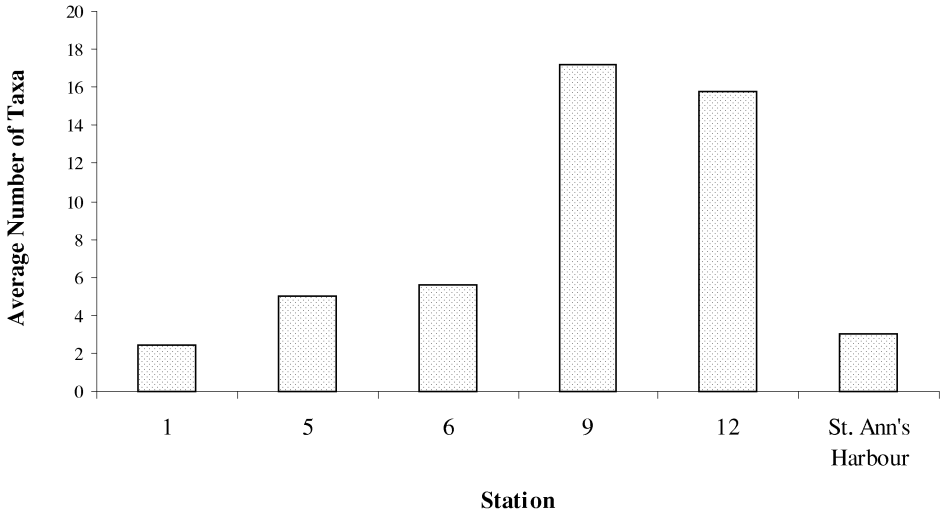


Figure 18 Benthic Macroinvertebrate Community Richness. The average number of taxa increases with increasing distance from Muggah Creek (Station 1). The St. Ann's Harbour reference station shows a relatively low number of taxa while Stations 9 and the Station 12 reference site show relatively high numbers of taxa.

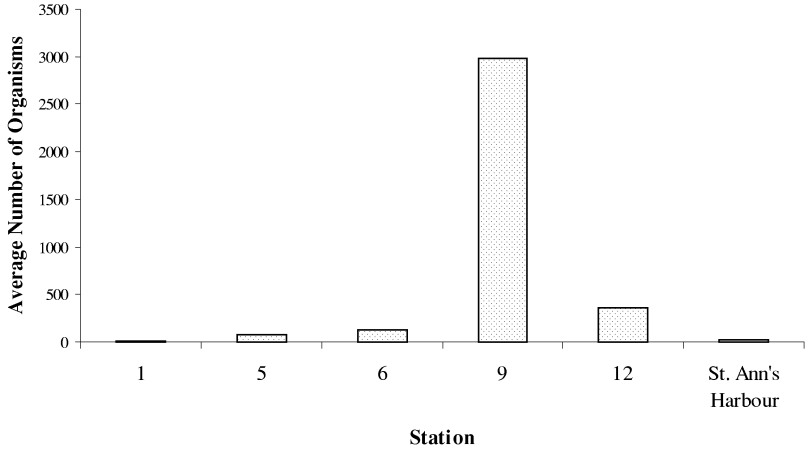


Figure 19 Benthic Macroinvertebrate Community Abundance. Station 9 shows a relatively large number of organisms due to the abundance of *Polydora quadrilobata*.

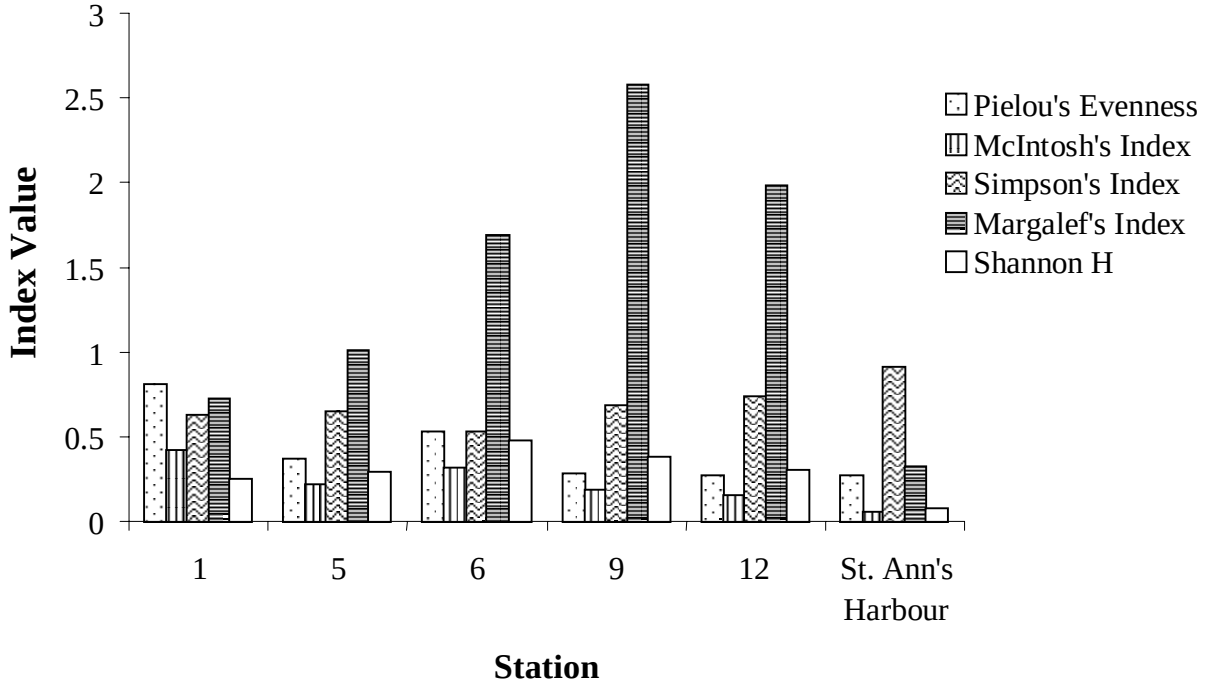


Figure 20 Benthic Macroinvertebrate Community Structure Indices. The distribution of organisms across taxa is most even at Station 1 and 6 with the other stations showing lower values for Pielou's evenness. McIntosh's index, which is another measure of evenness, provides the same interpretation as Pielou's index. Note that the St. Ann's reference station shows the lowest evenness using McIntosh's index.

Simpson's diversity index shows that the St. Ann's reference station is the least diverse station. Margalef's and Shannon's diversity indices also support this interpretation and show that Station 9 (Margalef's) and Station 6 (Shannon's) are the most diverse stations.

3.3 Summary of Available Biological Responses

A listing of the available responses is presented in Table 4. The column entitled "Obvious Response" comments on the presence of a visually obvious response. There may be no response, a response but no trend, or a trend in response. The term "along gradient" is directional and refers to stations in the following order 1, 5, 6, 9, 12, and St. Ann's Harbour. Thus an increase in response along the gradient is interpreted as an increase in the response from Stations 1 through to St. Ann's Harbour.

Table 4 Summary of Available Biological Responses

	Test/Species	**Response**	**Obvious Response?**
Acute Responses	*A. virginiana*	Percent survival	Increasing trend.
	E. estuarius	Percent survival	Increasing trend with asymptote at Station 9.
	E. washingtonianus	Percent survival	Increasing trend, with reference response lower than control response.
	R. abronius	Percent survival	A quadratic response curve with a minimum at Station 5.
	B. proboscidea	Percent survival	Little response overall, decreased survivorship at Station 12 reference site.
	P. cornuta	Percent survival	The response is variable with a minimum occurring at Station 9.
	M. nasuta	Percent survival	Extremely variable with a minimum at Station 9.
Sublethal Responses	*B. proboscidea*	Growth rate	No trend seen.
	P. cornuta	Growth rate	Growth is elevated at the St. Ann's Harbour station and depressed at Station 1.
	D. excentricus	Percent fertilization	Decreasing trend, with large variability among all responses.
	L. pictus	Percent fertilization	Variable with minimum percent fertilization at Stations 9 and St. Ann's Harbour.
	Microtox® (solid-phase, moisture corrected)	Change in luminescence	Peak IC50 found at Station 12 reference sites with IC50s higher at reference stations than innermost stations.
Bio-accumulation	*M. nasuta*	Bioaccumulation	A well-defined peak in PCB tissue concentrations at Station 5 declining monotonically to St. Ann's Harbour. PAH tissue concentrations decline monotonically to St. Ann's Harbour. As, Cd, and Hg tissue concentrations are invariant with respect to stations while Cr, Cu, Ni, Pb and Zn levels change with location.
In Situ Response*	Benthos	Abundance	Peak abundance at Station 9.
		Diversity	Intermediate stations most diverse, reference station, least diverse.
		Evenness	Generally decreasing.
		Richness	Peak richness at Station 9.

* The compositing procedure precludes enumeration of benthic samples. Therefore, in the strict sense of the term, the benthic samples are not synoptically collected.

Most of the toxicity tests exhibit responses; often trends in the responses are seen along the gradient. Station 9 seems to be a pivotal station for echinoids, *P. cornuta* and *M. nasuta* with minimum responses occurring. This is contrasted with maximal richness and abundance of benthic macroinvertebrates at Station 9. The greatest decrease in amphipod survival and bacterial photoluminescence was observed in organisms exposed to sediments collected from Station 1.

3.4 Validity of Test Sediments and Biological Toxicity Tests

3.4.1 Suitability of Test Methods for Marine and Estuarine Amphipods. Various criteria are being developed to ensure the validity of interpretations using Environment Canada's reference methods (EC, 1998a) for marine and estuarine amphipods. The limits for physicochemical characteristics as of July 1998 are shown in Table 5. Table 6 summarizes the grain size distribution at the stations sampled during the Sydney harbour study.

Table 5 Species-specific Application Limits for Reference Method (EC, 1998a)

Test Species	Acceptable Physicochemical Characteristics of Test Sediment			
	Porewater salinity (‰)	Sediment Grain Size		
		Percent very coarse-grained (> 1mm)	Percent Fines (< 0.063 mm)	Percent Clay (<0.004 mm)
R. abronius	Must be 25 to 35	0 to 100 is acceptable	Must be < 90	Must be < 40
E. washingtonianus	Must be 15 to 35	Must be < 25	Must be < 80	Must be < 20
E. estuarius	Must be 2 to 35	Must be < 90	0 to 100 is acceptable	Must be < 70
A. virginiana	Must be 15 to 35	0 to 100 is acceptable	Must be < 90	Must be < 35

Table 6 Summary of Grain Size Application Limits

Station	Sediment Grain Size		
	Coarse*	Fines	Clay
1	2.8	72.3	14.2
5	2.3	71.5	12.4
6	7.4	62.7	13.2
9	3.2	80.7	19.1
12	1.0	59.4	5.2
St. Ann's Harbour	4.4	68.5	10.5

* "Coarse" sediments consist of gravel + percent passing through a size 30 mesh. This is slightly at odds with the definition presented in Table 5, but was used as the Sydney harbour study was conducted before sediment grain size application limits were established.

The only failure of the stations with respect to application limits occurs at Station 9 where the percent fines criteria for *E. washingtonianus* is very slightly exceeded. As the exceedance is so slight as to be negligible and the percent survival for *E. washingtonianus* is high at this station, the complete *E. washingtonianus* data set is used in subsequent analyses (K. Doe, pers. comm., Environment Canada, Moncton, NB, 1999).

3.4.2 Suitability of Test Sediments for Comparison with Reference Site. A sediment should only be compared to a reference site if the mean 10-day survival in reference site sediments is > 80% for *R. abronius* and *E. estuarius*, > 75% for *E. washingtonianus*, and > 70% for *A. virginiana*. This proviso ensures that a protective benchmark for comparison of biological toxicity test results is used when making ocean disposal decisions.

Replicate # 1 for *A. virginiana* has a mean 10-day survival of 69%. Therefore Replicate # 1 would not pass the criterion for comparison of reference sites to exposure sites. However K. Doe (pers. comm., Environment Canada, Moncton, NB, 1998) and L. Porebski (pers. comm., Environment Canada, Marine Environment Division, Ottawa, ON, 1998) state that the mean of interest when a client submits an application for disposal of dredged materials would be the mean of all samples and subsamples for a given station. In this case, the overall mean survival rate for the St. Ann's Harbour reference station is 77% and the comparison of exposure sites to this site would be valid according to current criteria. All site means for the Station 12 reference station meet the criteria for minimum survival.

3.4.3 Species Specific Criteria for Validity of a Test. An amphipod toxicity test must meet certain survival criteria in control and reference sediments to ensure that inferences made using test results are valid (Table 7).

3.5 Station Pass/Fail Status

This section compares the biological toxicity test responses with a control test response in accordance with the pass/fail criteria generated by Environment Canada personnel. The interim interpretation criteria are summarized in Table 8.

The Environment Canada Atlantic Region Toxicology Laboratory interpreted the biological toxicity tests according to the Environment Canada interim interpretation criteria (Environment Canada, 1996). These results are presented in Table 9 alongside analyses conducted by Zajdlik & Associates. Note that only laboratory replication of the polychaete tests was conducted; therefore, comparison among stations is not possible.

Comparisons to the control are made using a simulation test (Edwards and Berry, 1987), similar in principal with Dunnett's test, at the $\alpha = 0.05$ level using a one-sided test. Subsampling error is incorporated into the overall error variance. However, the Environment Canada Atlantic Region Toxicology Laboratory used subsamples or laboratory replicates as true replicates, artificially decreasing the estimated variance.

Table 7 Summary of Required Control Survival Proportions (EC, 1998a)

Biological Toxicity Test Species	Minimum Proportion Survival (%)	
	Control	Reference
A. virginiana	80	70
E. estuarius	90	80
E. washingtonianus	85	75
R. abronius	90	80

All amphipod toxicity tests listed in Table 7 met the criteria for control sediment and reference station survival.

Table 8 Interim Interpretation Criteria

Test	Interim Interpretation Criteria
Amphipod Survival*	A statistically significant decrease in survival of at least 20% in test sediments as compared to reference sediments, or 30%, when compared to control sediments.
Polychaete Growth and Survival	Test under development.
Photoluminescent Bacteria* (solid phase)	An IC50 < less than 1000 mg dry solids/L diluent (ppm).
Echinoid Fertilization*	A statistically significant decrease in fertilization of at least 25% in test sediment pore water as compared to control water.
Bioaccumulation	A statistically significant difference in tissue bioaccumulation from control or reference sediment.

* Refer to Environment Canada (1996) for criteria to ensure test validity.

of the overall error variance. This has the effect of making the test artificially more powerful. In this case, stations that are not truly different from the control may be shown to be statistically different (from the control). This practice, although commonly encountered is not recommended. In the following table, Environment Canada's (Atlantic Region Toxicology Laboratory) pass/fail decisions are provided along with those made by Zajdlik & Associates.

Table 9 shows that the biological responses of organisms exposed to Sydney Harbour sediments were in general accord. Stations 1, 5, and 6 usually failed the Environment Canada pass/fail criteria using the battery of tests, while Station 9 often failed. *L. pictus* and *D. excentricus* seem to be the most sensitive species while *R. abronius* may be the least sensitive species. When St. Ann's Harbour is used as the reference station, *D. excentricus* is the most sensitive species and the *L. pictus* and *R. abronius* tests are the least sensitive. This reversal of relative sensitivity of *Lytechinus pictus* is due to very low percent fertilization in St. Ann's Harbour sediments. When Station 12 is used as the reference station, the tests using *E. washingtonianus*, *A. virginiana*, and *V. fischeri* are the most sensitive and *D. excentricus* and *R. abronius* are the least sensitive tests.

There is excellent concurrence in the pass/fail status of a station when test sediments are compared with either control sediments or reference sediments. The only exceptions occurred when the performance at the St. Ann's Harbour reference station was poor relative to test sediments as with the *L. pictus* test. When Station 12 becomes the reference station very low percent fertilization rates for *D. excentricus* reverses it's status as the most sensitive test, making it and the *R. abronius* tests, the least sensitive.

Also, although not shown in the table, effect or cutoff criteria were stricter than the statistical test of significance. For example, the use of only the statistical test of significance would have resulted in an increase in the number of stations found to be different from either the control or reference sediments. A discussion regarding the combination of expert judgement and statistical objectivity is provided in Section 6.6.

Using current disposal at sea guidance, it should be noted that only one species would have to be used for each test endpoint. Thus passing decisions (using the pass/fail interpretations in Table 9) would have been made as in Table 10, bearing in mind that field replication is not always required (thus each replicate can provide a separate decision).

Decisions assume that the species that passed was selected for the test. Decisions also assume (for simplicity) that the bioaccumulation result from replicate 1 would have been the same for other replicates (in practice this test would be done to confirm). Mitigation assumes that the substance can be considered for disposal at sea with special handling.

Table 9 Summary of Sediment Toxicity Test Failures

Test/Species	Cutoff Value[a] (percent or as stated)	Station/Replicate Failing Pass/Fail Criteria (Jackman and Doe, 1997[b])	Station Failing Pass/Fail Criteria with Respect to:		
			Control Sediment	Station 12 reference	St. Ann's Harbour
A. virginiana [b,c]	53.33/59.33/57.00	1/1,2,3, 5/3, 6/2,3	1, 5, 6	1, 5, 6	1, 5, 6
E. estuarius [b,c]	70.00/77.33/76.67	1/1,2,3	1	1	1
E. washingtonianus [c]	67.00/74.67/62.33	1/1,2,3, 5/1,3, 6/1,2,3	1, 5, 6	1, 5, 6	1, 6
R. abronius [b,c]	69.33/74.67/72.33	5/1	None	None	None
D. excentricus [b]	68.83/-20.33/70.11	1/1, 5/1,3, 6/1,3, 9/1,2,3, 12/1,2,3	1, 5, 6, 9, 12	None	1, 5, 6, 9, 12
L. pictus	55.67/33.83/-12.50	1/1, 5/2, 5/3, 6/2, 9/1, 9/2, 9/3, 12/3, St. Ann's Harbour/1, St. Ann's Harbour/2, St. Ann's Harbour/3	5 , 9, St. Ann's Harbour	5, 9	None
B. proboscidea growth and survival		None[e]	No replication.	No replication.	No replication.
P. cornuta survival		None[e]	No replication.	No replication.	No replication.
P. cornuta growth		1, 9[e]	No replication.	No replication.	No replication.
Photoluminescent Bacteria (Solid Phase)[d]	EC50 <1000 mg/L	1/1,2,3, 5/1,2,3, 6/1,2,3, 9/1	1, 5, 6	1, 5, 6	1, 5, 6
M. nasuta total tissue PAH[f]	statistically significant difference		1, 5, 6 but no replication	1, 5, 6 but no replication	1, 5, 6 but no replication

[a] The cut-off value is the minimum absolute change in response required in Table 8, applied to the response at the station being considered as a reference station. Thus the first entry of 53.33% = 83.33 - 30% survival in control sediments. The second value of 59.33% corresponds to the cutoff value when Station 12 is used as the basis of comparison and the value of 57.00% represents the cutoff value when St. Ann's harbour is used as a reference station.

[b] Jackman, P. and K.G. Doe. 1997. Toxicity of Sediments from Sydney Harbour, NS - Results of the 1997 Pollution Gradient Study. OR Calculations performed by K. Doe, (pers. comm., Environment Canada, Moncton, NB, 1999) using methods described herein.

[c] Pass/fail decision based on a statistically significant decrease in survival of 30% when comparing to control sediment or a 20% when comparing to reference sediment. Format is control sediment cutoff/Station 12 reference site sediment cutoff/St. Ann's Harbour reference sediment cutoff.

[d] Mean moisture corrected IC50s for each station compared to pass/fail criteria of 1000 mg dry solids/L diluent (~ ppm).

[e] Based on a statistical comparison of subsamples or laboratory replicates from exposure and control sediments (Jackman and Doe, 1997).

[f] **This comparison uses <u>subsamples rather than replicates;</u> therefore, the comparisons are artificially powerful.**

Table 10 Theoretical Pass/Fail Decisions by Field Replicate*

Replicate	1	5	6	9	12
1	All amphipods - **F** All echinoids - **F** Microtox® - **F** Bioaccumulation - **F**	*A. virgin* & *E. est* - **P** *L. pictus* - **P** Microtox® - **F** Bioaccumulation - **F**	*A. virgin*, *R. abron.* & *E. est* - **P** *L. pictus* - **P** Microtox® - **F** Bioaccumulation - **F**	All amphipods - **P** All echinoids - **F** Microtox® - **F** Bioaccumulation - **P**	All amphipods - **P** *L. pictus* - **P** Microtox® - **P** Bioaccumulation - **P**
2	*Repoxinius* - **P** All echinoids - **P** Microtox® -**F**	All amphipods - **P** *D. exentric*.- **P** Microtox® - **F**	*R. abron.* & *E. est* - **P** *D.excentric.* - **P** Microtox® - **F**	All amphipods - **P** All echinoids - **F** Microtox® -**P**	All amphipods - **P** *L. pictus* - **P** Microtox® - **P**
3	*Repoxinius* - **P** All echinoids - **P** Microtox® - **F**	*R. abron.* & *E.est*- **P** All echinoids - **F** Microtox® - **F**	*R. abron* & *E. est* - **P** *L. pictus* - **P** Microtox® - **F**	All amphipods - **P** All echinoids - **F** Microtox® -**P**	All amphipods - **P** All echinoids - **F** Microtox® - **P**
DECISION	No disposal at 1 No disposal at 2 No disposal at 3	No disposal at 1 No disposal at 2 No disposal at 3	No disposal at 1 No disposal at 2 No disposal at 3	No disposal at 1 Mitigation at 2 Mitigation at 3	Disposal at 1 Disposal at 2 Mitigation at 3

* Only one sample was collected per station for bioaccmulation tests. Therefore the level of replication is on the subsample level. The statisical comparison uses subsamples rather than replicates; therefore, the comparisons are artificially powerful. The pass/fail decision for bioaccumulation is based on the same data set. The bioaccumulation pass/fail decision is only presented in the first row of Table 10.

3.6 Sediment Physical/Chemical Analyses

The average results for the measured sediment and porewater physical/chemical analyses are presented in this section. Table 11 summarizes the sediment-related variables. The cell format is mean/standard deviation.

Table 12 summarizes the porewater variables. Where field replicate data is available, the standard deviation is also presented using the format “mean/standard deviation.”

Metals are generally below the detection limit with the exception of Hg at Station 1 and Zn at St. Ann’s Harbour. SEM/AVS ratios never exceed 1 indicating that porewater metals are not in sufficient quantity to exceed the AVS binding capacity of the porewater. Ni was inadvertently omitted from the SEM analysis and thus the total SEM has been underestimated. However as the measured porewater Ni is below the detection limit of 2 mg/L, it is anticipated that the true SEM/AVS ratios will remain below 1. PCBs are only detectable in the first three stations along the gradient while PAHs are detected in only the first four stations.

3.6.1 Summary of Available Physical/Chemical Measurements (Table 13).

Station depths, salinity, temperature, conductivity, and dissolved oxygen concentration of overlying water in increments of 3 m from the surface to the bottom, global positioning system coordinates, and a qualitative description of the sediments collected are also available (see Appendix A; Field Data).

Table 11 Summary of Sediment-related Variables

Variables	Stations					
	1	5	6	9	12	St. Ann's Harbour
Metals (µg/g dry weight)						
Hg	0.708/0.305*	0.486/0.03	0.333/0.01	0.038/0.001	0.024/0.006	0.055/0.001
As	41/3	39.333/3.786	32.667/2.517	10/0	9.667/0.577	15.667/5.132
Cd	1.167/0.153	0.927/0.085	0.467/0.141	0.147/0.081	0.08/0	0.247/0.012
Cr	81.233/7.217	86.567/22.938	61.7/5.789	33.9/3.534	24.1/1.153	41.933/2.916
Cu	101.333/3.512	73.667/4.637	53.7/4.158	22.7/0.361	13/1.732	37.3/0.4
Pb	285.667/26.312	214/7.55	133.333/5.774	32/2.646	21/1.732	37/2.646
Ni	34.667/0.577	37/2	37.667/1.155	24.667/1.155	18/1	27.333/0.577
Ag	2/0	2/0	2/0	2/0	2/0	2/0
Zn	516.267/50.816	865.667/749.436	281.567/19.798	91.1/1.54	56.2/3.812	84.233/0.351
PCBs (ng/g dry weight)						
Total PCBs	2095.023/861.506	1186.23/348.893	642.735/68.498	-69.131/4.049	-63.561/1.116	-30.847/120.013
Dry Wt (g)	4.853/0.423	4.834/0.291	5.197/0.365	7.202/0.426	7.792/0.158	4.841/0.314
PAHs (ng/g dry weight)						
Naphthalene	3982.435/886236.982	1617.295/228888.665	599.752/3750.125	59.352/108.958	21.797/13.866	4.785/1.021
2-Methyl-Naphthalene	1161.378/84039.966	710.909/22090.281	361.33/2907.549	88.167/262.524	39.795/5.591	5.56/0.257
1-Methyl-Naphthalene	735.118/28447.895	485.867/6977.545	260.151/1466.355	67.07/143.448	29.578/1.222	6.363/0.361
2,6-Di-Methyl-Naphthalene	716.182/24587.576	449.565/5786.11	238.984/989.323	56.558/93.875	27.785/3.608	0/0
Acenaphthylene	690.127/22186.793	372.337/2627.082	202.436/2349.982	7.209/2.019	1.13/0.179	0/0
Acenaphthene	419.286/9481.441	259.197/3569.849	152.429/1997.504	27.619/148.56	2.671/0.094	0.868/0.007
2,3,5-Tri-Methyl-Naphthalene	NA	NA	NA	NA	NA	0/NA
Fluorene	1636.157/83833.504	901.975/42009.474	432.14/484.075	50.378/135.725	12.413/0.389	5.144/0.199
Phenanthrene	8839.264/2408925	4929.827/953506.5	2427.151/61099.72	340.478/11139.88	50.542/196.316	13.347/3.517
Anthracene	5498.471/1204490	2787.591/312358	1362.971/32028.61	125.853/1312.1	41.734/7.663	9.373/1.342
1-Methyl-Phenanthrene	932.712/6463.562	567.783/6774.507	264.325/6701.992	58.653/121.688	15.919/3.593	0.979/2.874
Fluoranthene	13651.301/4564063	6054.457/942833.6	2950.918/41360.58	337.512/7905.176	41.23/262.417	31.941/4.269
Pyrene	19589.421/22524310	6888.913/945672.3	3450.766/25696.41	289.418/5033.69	34.936/51.939	24.558/7.272
Benzo*(a)*Anthracene	14229.563/5114055	6848.022/1050705	3147.776/63196.41	283.823/3576.512	24.931/3.463	22.432/43.278
Chrysene	17391.319/8247566.82	8184.197/1500450.97	3798.197/51278.354	312.285/2220.575	30.772/17.565	23.002/32.062
Benzo*(b)*Fluoranthene	27017.223/31935430	9470.852/819391.1	3476.122/27101.88	261.862/1604.273	19.882/2.329	37.435/1.487
Benzo*(k)*Fluoranthene	10760.32/4323773	4222.973/167416.1	1621.6/187614.8	81.708/184.096	4.979/0.32	8.948/9.41

Variables	Stations					
	1	5	6	9	12	St. Ann's Harbour
Benzo*(e)*Pyrene	12006.685/5277013	5073.646/684284.1	2005.741/21176.32	108.579/288.475	9.162/1.318	17.59/2.361
Benzo*(a)*Pyrene	23464.113/24031510	8502.341/1219261	3352.986/47869.2	193.527/1247.788	11.375/2.342	17.88/2.015
Perylene	10804.859/4310743	4158.546/470902.1	1529.271/8079.207	84.455/373.671	8.977/3.872	65.646/40.565
Indeno(123-cd)Pyrene	20076.987/17091900	7092.662/38102.6	2577.197/49100.06	178.978/736.443	9.295/5.634	40.749/14.372
Dibenzo*(a,h)*Anthracene	4223.435/620459.431	1871.473/4829.516	671.267/2868.767	44.725/75.597	0/0	0/0
Benzo*(g,h,i)*Perylene	14390.661/8315889	5476.129/37894.87	2030.201/14800.92	121.207/835.753	7.296/10.961	29.818/6.151
Total.PAHs	212217.02/1725983000	86926.557/101934100	36918.038/3800310	3179.414/336709.3	446.198/3533.43	373.617/834.271
NH3 (mg NH_3–N/g dry wt.)	43.083/196.521	36.828/3.139	38.165/52.862	32.568/21.234	13.573/3.154	45.638/56.943
Sulphide (mg S/g dry wt.)	95.776/1978.824	103.817/1770.73	61.146/525.56	18.009/38.957	27.267/61.652	116.956/3950.251
Redox Potential (mV)	-78.867/694.86	-64.322/3836.587	-41.044/139.03	-21.067/2437.208	78.433/7215.678	48.356/32589.205
Moisture %	69.1/1.1	69.733/1.582	68.067/2.203	37.133/0.503	26.733/1.102	75.167/0.961
TOC (µg/g)	96066.67/3544.479	81200/14988.996	47900/1808.314	18800/1126.943	8060/1338.768	35933.33/1305.118
Gravel %	0.1/0	0.1/0	0.1/0	0.1/0	0.1/0	0.1/0
Sand %	27.7/1.609	28.5/1.493	37.1/3.005	18.433/0.586	39.8/1.375	31.5/0.656
Silt %	58.1/1.664	59.133/3.266	49.533/3.007	61.6/0.854	54.2/1.493	58/2.858
Clay %	14.233/0.493	12.4/2.563	13.233/3.98	19.133/1.258	5.2/0.625	10.533/2.344

* The cell format is mean/standard deviation.

Metals generally decrease along the gradient but rise again at the reference station although metal levels at St. Ann's Harbour are not as high as Station 1. PCBs and PAHs generally decrease monotonically along the gradient.

Table 12 Summary of Porewater-related Variables

Station	1	5	6	9	12	St. Ann's Harbour
Ammonia (NH_3–N mg/L)	8.732/6.93	7.897/0.996	8.779/2.006	27.617/4.72	25.561/6.93	7.401/0.365
pH	7.66/0.001	7.663/0.004	7.697/0	7.723/0.022	7.833/0.036	7.46/0.006
Metals (mg/L, except Hg as µg/L)						
Hg	0.121	0.06	na	< 0.05	0.06	< 0.05
As	< 6	< 6	na	< 6	< 6	< 6
Cd	< 0.6	< 0.6	na	< 0.6	< 0.6	< 0.6
Cr	0.7	< 0.6	na	0.6	< 0.6	1.3
Cu	< 0.6	< 0.6	na	< 0.6	< 0.6	< 0.6
Pb	< 6	< 10	na	< 6	< 6	< 6
Ni	< 2	< 2	na	< 2	< 2	< 2
Ag	1	2	na	1	2	1
Zn	< 0.2	< 0.2	na	< 0.2	< 0.2	2.1
SEM/AVS variables (µmol/g dry)						
SEM-Cu	0.131	0.110	0.079	0.138	0.066	0.100
SEM-Zn	5.11	4.53	2.88	0.603	0.255	0.572
SEM-Pb	0.757	0.022	0.180	0.032	1.65	0.057
SEM-Cd	9.29	8.57	4.51	1.04	0.42	2.13
SEM-Hg	0.020	0.051	0.049	0.012	0.008	0.026
Total SEM(µmol/ dry)*	15.309	13.283	7.697	1.825	2.401	2.882
AVS (µmol/g dry)	36.04	34.69	26.77	13.95	4.40	23.6
SEM/AVS Ratio	0.42	0.38	0.29	0.13	0.55	0.12
PCBs (µg/L)						
PCB	0.76	0.97	0.46	< 0.48	< 0.51	< 0.44
PAHs (µg/L)						
Total PAH	21.11	11.16	6.10	1.50	< 0.33	< 0.71
Naphthalene	0.04	0.03	0.04	0.03	0.03	0.02
2-Methyl-Naphthalene	0.02	0.02	0.02	< 0.01	< 0.01	< 0.01
1-Methyl-Naphthalene	0.02	0.02	0.02	< 0.01	< 0.01	< 0.01
2,6-Di-Methyl-Naphthalene	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01
Acenaphthylene	< 0.03	< 0.03	< 0.05	< 0.05	< 0.04	< 0.04
Acenaphthene	< 0.02	< 0.02	< 0.02	< 0.02	< 0.02	< 0.02
2,3,5-Tri-Methyl-Naphthalene	< 0.02	< 0.02	< 0.02	< 0.02	< 0.02	< 0.02
Fluorene	0.03	0.02	0.02	< 0.01	< 0.01	< 0.01
Phenanthrene	0.23	0.17	0.12	0.04	0.01	0.01
Anthracene	0.33	0.26	0.16	0.05	< 0.02	0.02
1-Methyl-Phenanthrene	< 0.02	< 0.02	< 0.03	< 0.01	< 0.01	< 0.01
Fluoranthene	0.62	0.28	0.19	0.11	0.04	0.05
Pyrene	2.97	1.35	0.77	0.10	0.01	0.03
Benzo*(a)*Anthracene	0.70	0.30	0.22	0.09	0.01	0.02
Chrysene	1.03	0.56	0.30	0.08	< 0.01	0.02
Benzo*(b)*Fluoranthene	3.31	1.82	0.94	0.16	0.01	0.06
Benzo*(k)*Fluoranthene	1.28	0.59	0.31	0.06	< 0.01	0.02
Benzo*(e)*Pyrene	1.73	0.98	0.42	0.09	< 0.01	0.04
Benzo*(a)*Pyrene	2.99	1.56	0.80	0.11	< 0.01	0.03
Perylene	0.74	0.22	0.13	0.03	< 0.01	0.09
Indeno(123-cd)Pyrene	2.53	1.50	0.79	0.16	< 0.01	0.07
Dibenzo*(a,h)*Anthracene	0.57	0.25	0.15	0.04	< 0.01	0.01
Ben*(g,h,i)*Perylene	1.86	1.13	0.57	0.21	< 0.01	0.10

* SEM Ni was not measured by the contractor and is therefore not included in the estimation of total SEM. However the porewater Ni values for total extractable Ni is below the detection limit of 2 mg/L.

Table 13 Summary of Available Physical/Chemical Measurements

Measurement Type	Number of Field Replicates/Station
Sediment Metals	3 (Station 12 reference site, missing rep 1, Cd)
Sediment PCBs	3
Sediment PAHs	3
Sediment Sulphide, Ammonia	3
Sediment Particle Sizes	3
Sediment Total Organic Carbon	3
Porewater Metals	1 (Station 6 data missing)
Porewater PCBs	1
Porewater PAHs	1
Porewater Ammonia	3
Acid-Volatile Sulphides	1
Simultaneously Extractable Metals	1
Redox Potential	3
Porewater pH	3
Porewater Salinity	3

Tests of Hypotheses

During the analyses, it became apparent that the reference station at St. Ann's Harbour was not behaving as a reference station with respect to the *in-situ* benthic community structure. As the interpretation of the relative performance of the sediment evaluation tools depends, in part, on how a tool performs relative to a reference station, the definition of the reference condition is discussed.

Many definitions for a reference station have been established, but possibly the most pragmatic is:

A reference condition is identical as far as possible, to the exposure condition except for the intervention being considered.

In the Sydney Harbour study, the intervention of interest was the contaminant gradient, which was primarily PAHs and PCBs. Thus, St. Ann's Harbour was chosen as a reference station as there were no known PAH or PCB inputs other than atmospheric deposition. Attempts were also made to ensure that non-contaminant factors such as grain size and TOC were similar between the reference and exposure conditions.

Even though St. Ann's Harbour met the preceding definition of a reference condition, data analyses showed that the *in-situ* benthic community there was similar to the community found at the top (most contaminated part) of the gradient. Additionally, Arenicola Marine (1997) concluded that the *in-situ* benthic community at the St. Ann's Harbour station was degraded.

Thus, evaluation of the relative performance of the sediment assessment tools is made using two different reference sites (i.e., St. Ann's Harbour and Station 12 reference site). The dual evaluations begin at hypothesis 4. The following sections describe the methods used to test each hypothesis and the conclusions reached.

4.1 H1: Homogeneity of Confounding Factors

This section addresses the null hypothesis Ho: **Are potential confounding factors homogeneous across the stations?** This is tested to verify that the choice of stations achieved the study design goal of minimizing the effect of known confounding factors such as TOC, particle size, ammonia, and Eh.

The hypothesis is first tested using the non-parametric Kruskal-Wallis test (Tables 14, 15, and 16). Variables that are significantly different among stations (p-value < 0.05) are highlighted.

The analysis shows that all variables in Table 14, except porewater pH, vary significantly among the stations. As there is some concern regarding the choice of reference stations, the hypothesis of homogeneity of confounding variables is retested, by separately removing the St. Ann's Harbour and Station 12 reference sites and re-testing the null hypothesis: Are potential confounding factors homogeneous across the stations?

All variables in Table 15, except porewater pH, vary significantly among the stations. Similarly, in Table 16, all variables, except porewater pH and percent clay, vary significantly among the stations.

The combined analyses show that porewater pH remains constant among stations. When the Station 12 reference site is omitted, the percent clay composition does not vary significantly among stations, implying that the Station 12 reference site has a significantly different clay composition than other stations. Due to their importance as potential confounding factors, the TOC, Eh, grain sizes, pH, ammonia, and sulphide levels are examined in more detail.

The potential confounding variables across the gradient are presented in graphically in Figures 21–26.

Table 14 Summary of Kruskal-Wallis Tests: All Stations

Variable	Kruskal-Wallis χ^2	Degrees of Freedom	P-value
TOC	16.2156	5	0.0063
Percent Moisture	15.0331	5	0.0102
Sand	15.8175	5	0.0074
Silt	13.4174	5	0.0198
Clay	13.6866	5	0.0177
Eh	30.5927	6	< 0.0001
Sediment Ammonia	40.7095	6	< 0.0001
Porewater Ammonia	46.6413	6	< 0.0001
Porewater pH	11.1025	6	0.0853
Sulphide	47.4889	6	< 0.0001

Table 15 Summary of Kruskal-Wallis Tests: St. Ann's Harbour Omitted

Variable	Kruskal-Wallis χ^2	Degrees of Freedom	P-value
TOC	12.9815	4	0.0114
Percent Moisture	11.2952	4	0.0234
Sand	12.4138	4	0.0145
Silt	11.9296	4	0.0179
Clay	11.1449	4	0.025
Eh	32.4888	5	< 0.0001
Sediment Ammonia	34.7565	5	< 0.0001
Porewater Ammonia	41.0016	5	< 0.0001
Porewater pH	5.2082	5	0.391
Sulphide	42.149	5	< 0.0001

Table 16 Summary of Kruskal-Wallis Tests: Station 12 Reference Site Omitted

Variable	Kruskal-Wallis χ^2	Degrees of Freedom	P-value
TOC	12.9815	4	0.0114
Percent Moisture	11.2952	4	0.0234
Sand	12.8813	4	0.0119
Silt	9.5253	4	0.0492
Clay	9.3751	4	0.0524
Eh	21.9077	5	0.0005
Sediment Ammonia	25.0282	5	0.0001
Porewater Ammonia	34.4793	5	< 0.0001
Porewater pH	8.8385	5	0.1157
Sulphide	36.0984	5	< 0.0001

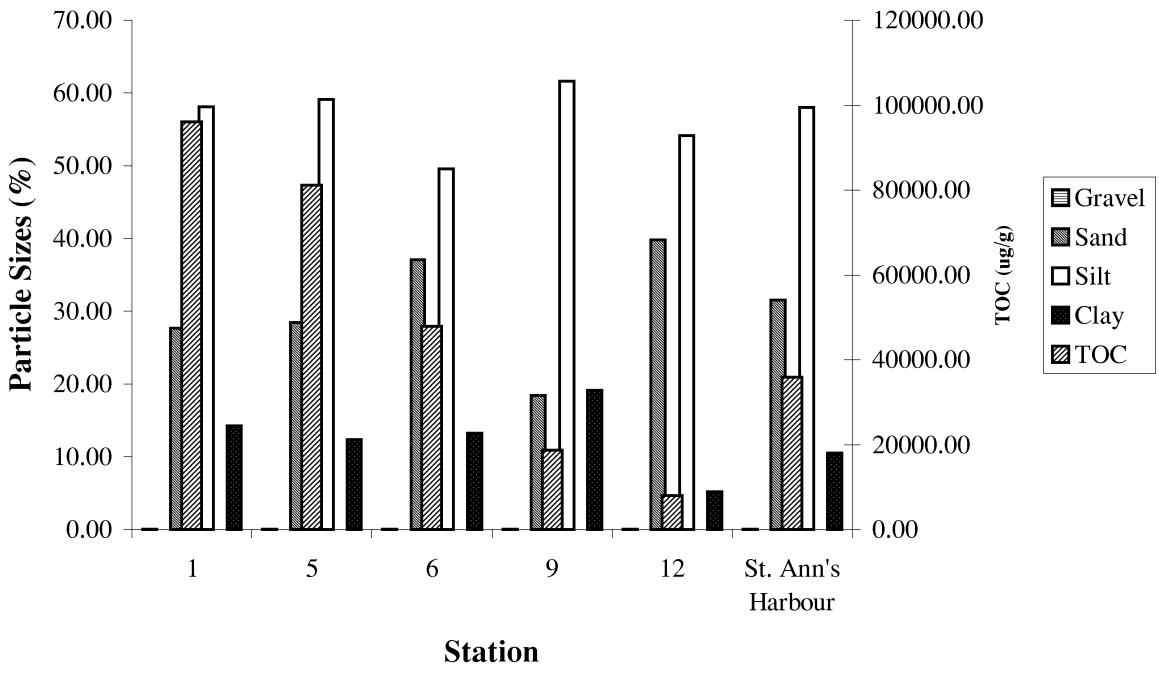


Figure 21 Sediment Physical Characteristics. The TOC follows the PAH gradient, while the sediment physical parameters do not.

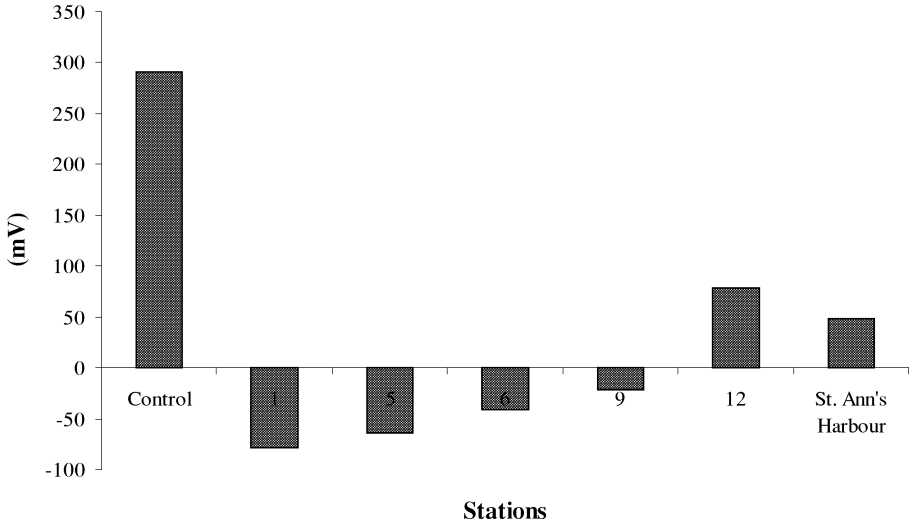


Figure 22 Redox Potential Against Stations. The redox potential chart shows that reducing condition occur at the top of the gradient gradually switching to oxidizing conditions at Station 12 reference site. Eh potential drops in anoxic sediments but may also be reduced by humic acids and some metals, notably ferrous Fe. A rank correlation test between porewater Fe and Eh potential is not significant (Spearman's rho = 0.8, with p-value = 0.134). Thus, oxygen or humic acid levels may drive redox potential. Note that the Eh is calibrated to neutral pH, thus the following plot of pH versus station should be examined in conjunction with the plot of redox potential.

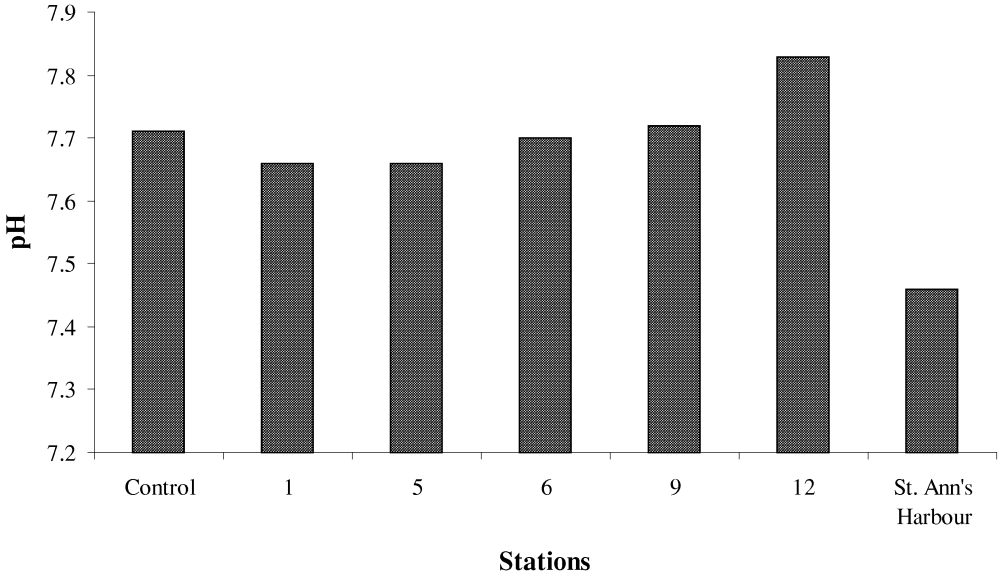


Figure 23 Porewater pH Against Stations. Porewater pH is lowest at the St. Ann's Harbour reference station, but not significantly so.

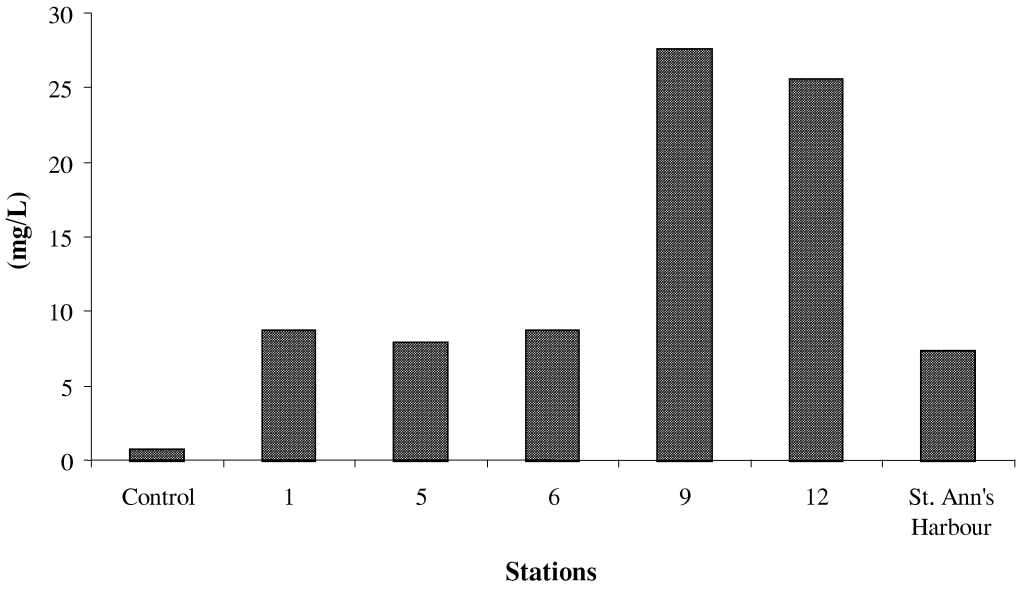


Figure 24 Porewater Ammonia Against Stations. Porewater ammonia is highest at Stations 9 and 12.

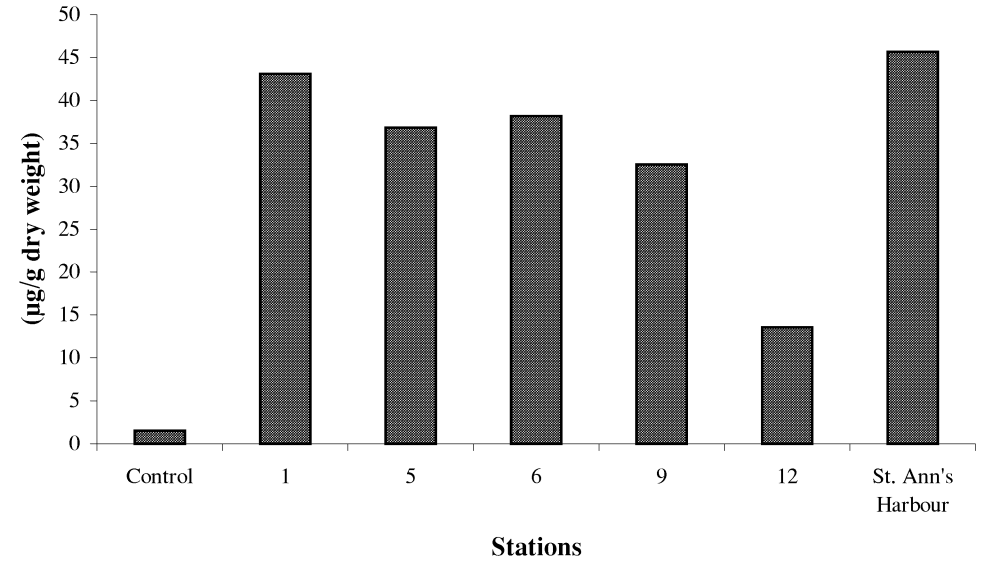


Figure 25 Sediment Ammonia Against Stations. Sediment ammonia decreases along the gradient but increases significantly at the St. Ann's Harbour reference station.

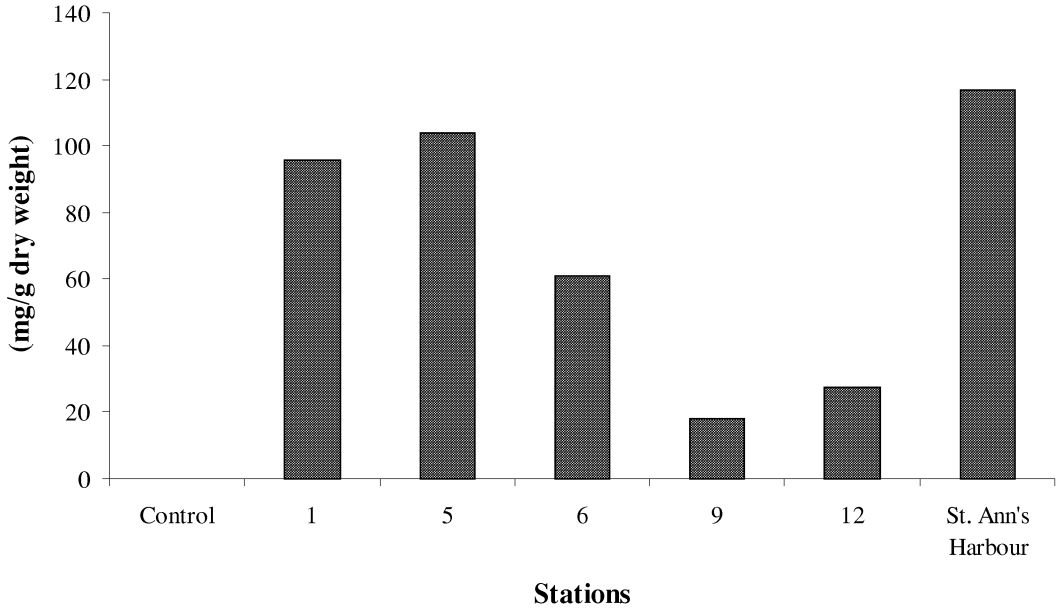


Figure 26 Sediment Sulphide Against Stations. Sulphide, like sediment ammonia, decreases along the gradient but increases significantly at the St. Ann's Harbour reference station.

Generally, the St. Ann's Harbour station exhibits characteristics (TOC, sediment, and porewater NH_3–N, and sulphide) similar to stations at the top of the gradient. However, porewater pH is markedly lower at the St. Ann's Harbour station, although this was not statistically significant.

4.2 H2: Predicting Biological Toxicity Test Responses

This section addresses the following null hypotheses:

- Ho 2a: **Do confounding factors affect the biological toxicity test response?** and,

- Ho 2b: **Are the dose responses predictable?**

The previous section has shown that some of the confounding factors vary significantly across stations. Thus, differences in responses may be due to the PAH gradient and/or confounding factors. The relationship between the independent variables and the nine biological toxicity tests is now examined.

4.2.1 Correlation Between Independent Variables. The correlation between toxicants, sediment physical/chemical characteristics among stations is examined to provide information regarding how these parameters covary. This is extremely important in model building and interpretation. An example follows.

Consider two sediment variables, [Cu] and percent Clay. If a very high correlation exists between these two variables, then a model such as Mortality = a + b * [Cu], is functionally the same as Mortality = a + b * percent Clay, i.e., both models will produce similar estimates of mortality. A stepwise model building algorithm will choose either [Cu] or percent Clay as an independent variable, according to various criteria, but not both, due to their high correlation. Depending on the criteria used in the statistical package, either model could be produced.

The interpretation of the two models, however, varies considerably. In one case, it could be said that the toxicant does not affect mortality or alternatively, the response is explicable by environmental variables. The other model states that the observed mortality is related to sediment [Cu].

Thus if the correlation between the independent variables is known, *a priori*, contending models can be better interpreted. Appendix A; Table A-19 contains the rank correlation (r) between pairs of sediment contaminants whose absolute value is greater than an arbitrary value of 0.90. From this table we see that the PAHs are highly correlated with one another. This degree of correlation implies that there is a large amount of redundant information in the data set. This type of correlation makes the model fitting and evaluation process difficult. We proceed by substituting total PAHs for the individual PAHs. Table 17 contains the rank correlation (r) between pairs of independent variables (with individual PAHs removed) whose absolute value is greater than an arbitrary value of 0.90. Note that other pairs of variables are also correlated but not as strongly.

The metals here are highly correlated with one another. As there is currently no useful generic summary for metals the following sediment variable data set is used in subsequent analyses: As, Cd, Cr, Cu, dry weight (of sediment) Hg, Ni, PCBs, Pb, Total PAHs, and Zn.

The correlation between porewater variables is investigated. As porewater metals, PAHs, and PCBs were measured only once per station while other porewater variables were replicated, averages of the replications are used in the estimation of the correlations. Table A-20 (Appendix A) contains the rank correlations (r) between pairs of independent porewater variables with absolute values greater than an arbitrary value of 0.90.

In this table, as with the sediment variables, the porewater PAHs are highly correlated. Total PAHs are used to summarize the PAH data set and re-estimate the correlations among the porewater variables. Table 18 contains the rank

Table 17 Summary of Rank Correlations among Reduced Sediment Variables with |r|≥0.90

Variable 1	Variable 2	Rank Correlation
As	Cd	0.9074
As	Cu	0.9254
As	Pb	0.9392
Cd	Cu	0.9608
Cd	PCB	0.9577
Cd	Pb	0.9698
Cu	Hg	0.9071
Cu	PCB	0.9203
Cu	Pb	0.9736
Hg	Pb	0.9326
PCB	Pb	0.9345

Table 18 Summary of Rank Correlations among Reduced Porewater Variables with |r|≥0.90

Variable 1	Variable 2	Rank Correlation
Cr	Zn	0.9899
Cr	pH	-0.9837
PCB	Eh	-0.9187
PCB	PAH	0.9813
Zn	pH	-0.9530
Ammonia	Sulphide	-0.9142

correlations (r) between pairs of independent variables (with individual PAHs removed) with absolute values greater than an arbitrary value of 0.90.

There is a marked decrease in the number of correlated variables when PAHs are summarized as total PAHs. The following regression model building procedures will use total porewater PAHs as a surrogate for individual porewater PAHs. Note that a negative sign preceding the estimated correlation indicates an inverse correlation. For example, when ammonia increases, sulphide decreases. The magnitude of the absolute value of the correlation, determines how strong the relationship is. Correlations close to 1 or –1 are strong, those close to zero, are weak.

4.2.2 Automated Model Building for Biological Toxicity Test Responses. The responses on the independent variables are regressed using an automated model building approach due to the large number of potential independent variables (12 porewater and 12 sediment). Note that automated model building

procedures are generally frowned upon by statisticians due to failure of the algorithm to include known predictor variables in the model. In the case of process control studies this is often due to tight control of known predictor variables causing little variability in values; hence the failure to be included in the model. Other criticisms of automated variable selection techniques center around the circumvention of thought involved in the model building process.

Herein, there are 24 potential independent variables and nine biological toxicity tests, resulting in nine models that must be fit. If consideration is restricted to linear models, with terms of the first degree, and no interactions, models with 10 predictors will result in examination of 1 961 256 models for each of the nine biological toxicity tests. Due to the large number of responses and potential predictor variables, automated variable selection procedures are used. The results should be viewed as exploratory, other variable selection schemes may result in different models being produced. The comparison of models across responses is robust. However, comparison of these models to other models in the literature may not be.

An automated method that is theoretically attractive is the stepwise regression procedure. Other methods such as forward and especially backwards regression tend to include far too many variables in the model especially when variables are highly correlated. Table 19 provides coefficients for those variables included in the final model. Models summarizing a significant proportion of the data set are highlighted in the row "Model P-value."

Most biological toxicity test responses are well described as shown by the large to very large goodness of fit indicators. An exception to this occurs with *R. arbonius*. The *R. arbonius* data set is quite variable (see Figure 3) with only a minimal depression in survival at Stations 5 and 6. Although the model P-value is significant (i.e., $< \alpha = 0.05$) this model does not fit the data well and consequently should be ignored. The *L. pictus* data set is also extremely variable. The large pseudo-R^2 value or goodness of fit indicator is an artifact of using many variables to represent the noisy data set. This model should also be ignored. The model for *D. excentricus* should also be ignored as it does not describe a significant proportion of the variability in the data set. The very large goodness of fit indicator is an artifact of using many variables to represent the noisy data set.

The model for *A. virginiana* almost describes a significant proportion of the total variability in the data set with a p-value only slightly larger than the generally accepted cut-off for significance of $\alpha = 0.05$.

More porewater variables than sediment variables are correlated with the observed responses among the 5 (or 6 if *A. virginiana* is included) well described data sets. The confounding factors do not play a large role in the observed responses despite their heterogeneity across stations. Only the *V. fischeri* test is correlated with sediment ammonia and the *E. estuarius* test is correlated with percent silt.

How to Read Table 19:

The columns constitute a model for a given response. The rows containing variable names on the left have coefficients inserted only if that variable was included in the model. The final three rows provide a measure of the goodness of fit, the type of model used to describe the response, and a test of how well the model describes the variability in the response. Thus the model for A. virginiana *is found by following the first response column down. The variables that were useful in predicting* A. virginiana *survival are Pb with a coefficient of 0.309 and total PAHs with a coefficient of -0.223. The model describes (approximately, see footnote below table) 81.6% of the variation around the percent survival. The significance of the model is 0.0718%. It can be concluded that the logistic model, although describing a reasonable amount of variability in the response (81.6%) does not account for a significant proportion of the variability as the p-value of 0.0718 is not less than the usually accepted value of 0.05.*

Table 19 Summary of Regression of Biological Toxicity Test Responses on Sediment and Porewater Variables

Variable	Survival				Percent Fertilization		Weight Gain		IC50
Biological Toxicity Test	*A. virginiana*	*E. estuarius*	*E. washingto nianus*	*R. abronius*	*D. excentricus*	*L. pictus*	*B. proboscidea*	*P. cornuta*	*V. fischeri*
Intercept	-0.582	-1.290	3.151	0.854	397.939	-346.373	1.718	0.610	14323.090
Porewater Variables									
Ag			0.764		-42.416		-0.155		
Ammonia						-1.646			
Cr									8568.085
Hg				17.510	-273.379				
Eh				0.01361					
Pb	0.309								
PCB									
pH						53.834			
Salinity					-9.421				
Sulphide									
Total PAHs	-0.223	-0.131							
Zn								0.767	
Sediment Variables									
Ammonia					-2.467	0.116			-517.984
As			-0.0434		0.316				
Cd					-4.307	1.875			
Cr					-0.167				
Cu						-0.414			
Dry Weight					-1.624				
Hg					4.817	12.471			
Ni			-0.0639		1.303				
PCB					-0.00161	0.00772			
Pb					0.0404	-0.135			
Total PAHs									
Zn					0.00742	-0.0135			
Moisture					0.366	-0.439			
TOC						0.000488			
% Sand						-0.454			
% Silt		0.0787							
% Clay					-0.1390				
Goodness of Fit*	0.816	0.823	0.907	0.659	0.991	0.942	0.9686	0.891	0.838
Model	Logistic	Logistic	Logistic	Logistic	Logistic	Logistic	Gaussian	Gaussian	Gaussian
Model P-value	0.0718	0.0275	0.0470	0.0231	1.000	0.484	0.0158	0.00466	<0.00001

* As the deviance of binomial generalized linear model (GLIM) is only a function of the fitted values and not the observed values, the usual R^2 cannot be estimated. Instead, "an index of correlation" or pseudo-R^2 is presented for the logistic model. This is the Spearman rank correlation coefficient between observed and predicted values.

4.2.3 Consensus Model Building for Biological Toxicity Test Responses. Automated model building procedures are generally frowned upon as previously described. Due to the importance of the relationship between biological toxicity test responses and sediment physical/chemical parameters, models are also fitted manually. The summary of hypothesis tests section discusses the models built using this approach rather than the automated approach.

Model building tools include graphical plots of the data, knowledge of sediment physical chemistry, prior experience modeling sediment toxicity test responses, and an understanding of the statistical models being employed. Thus, in some instances an interpretation or model specification may vary depending on the person interpreting the data. All subjective model interpretations are discussed in the model-building procedures that follow.

The model building goal was to determine the most important variables contributing to the observed response, rather than prediction. Therefore goodness of fit was sacrificed for a concise model, but not to the point where highly significant variables were omitted from the model. (The goodness of fit can be perfect if one variable is allowed for each observation in the data set. Unfortunately this model is only a restatement of the data and is not very explanatory!)

The model building procedures used follow those generally prescribed in statistical textbooks[2]. A complete description of the model building procedures used is beyond the scope of this document. Interested readers may consult Collett (1991) and McCullagh and Nelder (1989) for an account of model building in the generalized linear model context. General statistical model building is described in Atkinson (1985).

4.2.3.1 Amphiporeia virginiana. *A. virginiana* survival is well predicted by the model:

$$\text{logit (probability of surviving)} = 1.17027 - 0.00001828032 * \text{total sediment PAHs}$$

where logit, refers to the logit transformation; log (p/(1-p). A pseudo R^2 value of 0.806 as described in the preceding section, is only marginally lower than the R^2 value generated by stepwise methods.

Other possible single parameter models include porewater Hg, and PAHs, sediment-associated, Cu and Pb as variables. An examination of the relationship between *A. virginiana* survival rate and non-contaminant variables shows that total organic carbon and redox potential are strongly correlated with survival rates (r = 0.880 and 0.776, respectively). The correlation between TOC and total sediment PAHs is 0.883 and between redox potential and total sediment PAHs is -0.721. Total PAH concentration is related to sediment TOC, as expected, but the relationship between *A. virginiana* survival and total sediment PAH is stronger than that with TOC.

4.2.3.2 Eohaustorius estuarius. *E. estuarius* survival is almost as well predicted as *A. virginiana* survival. The following model best reflects the data set.

$$\text{logit (probability of surviving)} = 3.034739 - 0.1188341* \text{total porewater PAHs}$$

where logit, refers to the logit transformation; log (p/(1-p). A pseudo R^2 value of 0.746 as described in the preceding section, is lower than the R^2 value generated by stepwise methods.

Other possible single parameter models include porewater Hg and sediment-associated, Cu, Pb, total PAHs, total PCBs, and TOC. Of these,

[2] These include a criterion for the inclusion of variables to the model with a significant increment in explicable variance with a penalty for over-parameterization (following the discussion in the previous paragraph), lack of residual structure, which may indicate a sufficient parameterization, and an understanding of how confounding factors and contaminants interact to best choose which contending variables should be included in the model.

models using porewater PAHs or sediment-associated Pb or Cu produce almost identical fits.
Thus although total porewater PAHs were selected as the variable best describing *E. estuarius* survival, sediment-associated Pb (> TEL at 5 of 6 stations and > PEL at 3 of 6 stations) or Cu (> TEL at 5 of 6 stations, but < PEL at 6 of 6 stations) could equally well have been chosen. Again, the effects of TOC cannot be discounted as it is extremely highly correlated with Pb and Cu (r = 0.955 and 0.982, respectively). Note the similarity of this model with that generated using automatic model building methods. In both models total porewater PAHs are significantly associated with survival. The manual approach chose TOC as a contending variable while the automatic model building procedure included percent silt. The association between TOC and silt explains the apparent discrepancy between these two models.

4.2.3.3 Eohaustorius washingtonianus. Of the initial variables selected to describe *E. washingtonianus* survival, all were sediment-associated metals. This occurs, as *E. washingtonianus* survival is slightly depressed at St. Ann's Harbour relative to Stations 9 and 12. This coincides with a general decrease in sediment-associated As, Cu, Hg, and Pb from Stations 1 to 12, with a small increase in St. Ann's Harbour.

The best fitting single variable model was survival as a function of sediment-associated Cu. An examination of the model residuals shows a disturbing spread at intermediate values of Cu. Given that Cu is highly correlated with PAHs, PCBs, and TOC (r = 0.934, 0.973, and 0.990, respectively), it is possible that Cu correlates well with *E. washingtonianus* survival due to the depression in survival at St. Ann's Harbour but not as well at the intermediate stations, where organic compounds may be having a greater effect than metals. Models are then tested with Cu and either Eh, porewater PAHs, or TOC.

None of these models fit significantly better, nor do they reduce the lack of fit at intermediate values of Cu. Consequently, the following model is adopted:

logit (probability of surviving) = 2.449244 - 0.02648029* sediment-associated Cu.

This model produces a pseudo R^2 value of 0.792 that is lower than the R^2 value generated by stepwise methods. However, the current model uses only one variable to explain the response and is more scientifically defensible. Other single variable models that are not significantly different from the Cu-only model include sediment-associated As, Hg, and Pb as independent variables.

4.2.3.4 Rhepoxynius abronius. *R. abronius* survival is moderately depressed at Station 5 and shows only a very slight dose response. Initial correlation analyses show that the same parameter set responsible for decreased survival in *E. washingtonianus* (sediment-associated metals) is also correlated with *R. abronius* survival. However the correlations are approximately 30–40% lower. The best fitting single-variable model is given by:

logit (probability of surviving) = 1.924936 - 0.008457369* redox potential.

This model fits poorly with the bulk of the model fit ascribable to the intercept. The pseudo R^2 value of 0.349 is much lower than the R^2 value generated by stepwise methods.

4.2.3.5 Dendraster excentricus. *D. excentricus* percent fertilization gradually decreases from the top of the gradient to the Station 12 reference site and is very high in St. Ann's Harbour sediments. The variability of the responses as depicted in Figure 11, is very high at Station 1. The relative magnitude of the within replicate (laboratory replicates) and among replicate (field replicates) variability is tested in Section 4.3. Often an increase in variability is associated with a response to a stressor.

An examination of potential independent variables shows no association with contaminants. *D. excentricus* percent fertilization is strongly correlated with porewater and sediment ammonia, sulphide, moisture, and pH. All of these variables covary (see Table 20).

Table 20 Summary of Correlation Between Potential Independent Variables

	Porewater Ammonia	Sulphide	Percent Moisture	Sediment Ammonia	pH
Porewater Ammonia	1.000	-0.914	-0.969	-0.771	0.679
Sulphide		1.000	0.903	0.750	-0.803
Percent Moisture			1.000	0.896	-0.783
Sediment Ammonia				1.000	-0.840
pH					1.000

All single variable models fit equally badly with none explaining a significant proportion of the observed variability. It should be noted that in Figure 11, the median percent fertilization follows the PAH gradient. Robust methods that trim outlying data points would likely find a significant correlation between percent fertilization and some of the variables that decrease along the gradient such as PAHs. However, this type of modelling would not mask the variability in the response of *D. excentricus* to the sediments collected during this study. This variability would still play a role in pass/fail decisions using traditional statistical tools such as Dunnett's test. Again, robust methods could be employed to deal with some of the outlying data points.

4.2.3.6 Lytechinus pictus. There is some evidence of a reduction in percent fertilization along the gradient, with an increase at the Station 12 reference site, but the responses are quite variable (see H3, for a comparison of within and among replicate variability). Initial plots show little evidence of a predictable dose-response. The dose-response is even less predictable than that of *D. excentricus*. The highest correlations occur with non-contaminant variables (sand, silt, and Eh).

4.2.3.7 Boccardia proboscidea. No dose response is seen for *B. proboscidea* weight, although reduced weight gain occurs at the Station 12 reference site. This data set consists of only four points, as data for Stations 5 and 6 is unavailable. Additionally, due to sediment homogenization procedures, the samples for a station are subsamples or laboratory replicates rather than true or field replicates. The dose response is not predictable. The highest non-trivial correlations occur with non-contaminant variables (sediment ammonia and pH).

4.2.3.8 Polydora cornuta. *P. cornuta* exhibits no dose response along the gradient until St. Ann's Harbour where an increase in growth occurs, similar to that found in the control sediments. It appears that some factor is present in Stations 1, 5, 6, 9, and 12 that is depressing growth and that is not present in the control or St. Ann's Harbour reference station. The available control sediment chemistry for Conrad Beach (Cd = 0.05 µg/g, Pb = 3.03 µg/g, PAHs < 0.01 ng/g, PCBs < 0.01 ng/g, TOC = 0.36%, gravel = 0.1%, sand = 95.2%, and % fines = 4.1%) shows the percent sand is much higher at Conrad Beach than at all of the stations used in this study. Given the absence of similarities between control sediment and St. Ann's Harbour contaminant chemistry, only the correlations between weight change and sediment physical/chemical variables can be examined.

Porewater Cr and Zn are highly correlated with growth in *P. cornuta* as they are virtually undetectable at all stations except St. Ann's Harbour. It is unlikely that porewater Cr and Zn enhance growth although some metals are required as cofactors. Coincidental with the spike in porewater Cr and Zn, is a minimum in pH at this station that is similar to the pH of the control sediment. It is possible that *P. cornuta* exhibits improved growth at neutral pH. As no replicate samples are available, this hypothesis was not tested.

4.2.3.9 Photoluminescent Bacteria. The photoluminescent bacterial assay exhibits a peak in IC50s at the Station 12 reference site, with the IC50 in the St. Ann's Harbour station being

higher than at Stations 1, 5, 6, and 9. A log-transformation reduces the effect of the high IC50 at the Station 12 reference site. The $\log_{10}$ IC50 is modeled as a function of the sediment physical/chemical parameters. The use of the log transformation implies that the difference in response across independent variables is proportional to changes in the independent variable, rather than additive. As an example, if sediment-associated Ni increases by 1 unit, the $\log_{10}$ (IC50) will decrease by 0.1032 units. However, the IC50 will decrease by $10^{-0.1032}$ or 0.7885 units.

The following variables are considered for inclusion in the model based upon correlations: Eh, sediment-associated As, Cr, Cu, Hg, Ni, and Pb, and TOC and porewater PCBs. A series of model trials, results in the following model:

$$\log_{10} (IC50) = 5.837466 - 0.1032447 * \text{sediment-associated Ni.}$$

The R^2 value of 0.908 is higher than the R^2 value generated by stepwise methods and contains fewer parameters. Model residual diagnostics are very good. However, models with sediment-associated Cr, or Eh fit almost as well due to their high correlation with sediment-associated Ni. Note that this model considers sediment-associated Cr as a possible explanatory variable.

4.2.3.10 Summary of Consensus Model Building for Biological Toxicity Test Responses. Table 21 summarizes the major conclusions of the dose-response models. The reader should refer to specific sections for a more detailed interpretation of the modeled dose response.

A strong response to the PAH gradients occurred among the amphipods *A. virginiana* and *E. estuarius*. Other organisms responded to different contaminants such as metals. Other than redox potential, and in one instance pH, non-contaminant effects such as grain sizes were not observed. The TOC was usually a contending explanatory variable due to its high correlation with contaminants.

The exploration of the various models <u>not</u> presented, shows that non-contaminant variables are also associated with the responses, but not as strongly as with contaminants. The correlation between contaminants and non-contaminant factors, particularly TOC and redox potential was often high.

Table 21 Summary of Model Building for Biological Toxicity Test Responses

Response	Comments
A. virginiana	A strong dose response well predicted by total sediment PAHs but also related to TOC.
E. estuarius	A strong dose response well predicted by total porewater PAHs, but also related to TOC and TOC-associated variables.
E. washingtonianus	Good dose response predicted by sediment-associated Cu, possibly due to increase in concentrations at the reference station.
R. abronius	Limited dose response with redox potential weakly correlated with percent survival.
D. excentricus	The dose-response is not predictable. The highest correlations occur with non-contaminant variables.
L. pictus	The dose-response is not predictable. The highest correlations occur with non-contaminant variables.
B. proboscidea	No dose-response is seen for *B. proboscidea* weight gain.
P. cornuta	No dose-response but improved growth at the St. Ann's Harbour station may be due to pH effects.
V. fischeri	The log-transformed dose response is very well predicted by sediment-associated Ni.

The consensus model building approach produces simpler, more interpretable models than the automated model building approach. These simpler models usually do not fit quite as well as the computer-generated models for reasons described at the beginning of Section 4.2.3.

4.2.4 Model Building for Tissue Bioaccumulation. The same model building techniques are employed to investigate which sediment physical chemical variables are correlated with tissue toxicant levels in *M. nasuta*. Only the following tissue contaminants showed evidence of change and were further examined: Cr, Cu, Ni, Pb, Zn, total PAHs, and PCBs. Models summarizing a significant proportion of variability in the data set are highlighted in the row "Model P-value." The table is in the same format as Table 19 and is read in the same manner. The tissue PCB model fits the observed data moderately well with an R^2 value of 0.687 and describes a significant proportion of the total variability in the data set as the model p-value of 0.0415 is less than 0.05.

The automated model building procedure for tissue PAHs chose TOC as the independent variable that best explains tissue levels of PAHs. However an examination of the correlation between TOC and sediment and porewater levels of PAHs found that the correlations were 0.899 and 0.938, respectively. Consequently, a model was fit with porewater PAHs as the only predictor variable. This model has an R^2 of 0.904. The associated F-statistic is only slightly smaller than that of the model in Table 22. There is interest in describing the relationship between environmental contaminants and tissue contaminants, and the contending models are equally acceptable; therefore, the following model for tissue levels of PAHs is adopted:

Tissue PAH = 601.3638 + 407.736* Porewater PAH

This is an example of how automated model building procedures may produce models that do not incorporate a mechanistic understanding of the process(es) involved.

Only the tissue PAH and tissue PCB models describe the variability in the data set to a significant extent. Both of these tissue contaminants are strongly correlated with its porewater concentration.

4.3 H3: Do biological toxicity tests perform consistently within sites?

This hypothesis may be assessed using replicates within stations to examine the variability of a given response at a station. Significant differences in subsample variability among the replicates may indicate an inconsistent test, or micro-scale differences in sediment physical/chemical quality, possibly due to poor homogenization techniques. This test of hypothesis is similar in intent, to the control charts using reference toxicants for biological toxicity tests; namely consistency of the test under conditions that are as similar as possible.

The test is conducted by partitioning each treatment sums of squares into an among-sites sum of squares and among subsample (or laboratory replicate) sums of squares. The within site sums of squares represents the subsampling error. If a test is excessively variable then the subsampling error will be greater than the treatment error. This is tested using the ratio of the two mean square errors. In tests where the response is survival, logistic regression is used to estimate the two error components. The deviances corresponding to each term in the logistic model are assumed to have chi-squared distribution. An F-test for equality of variances is used to test the hypothesis that the two variances are equal.

Table 23 shows that the within sample mean square error for *E. washingtonianus* and *D. excentricus* biological toxicity tests were both greater than the treatment mean square error. The graphic demonstrating percent survival for *E. washingtonianus* (Figure 5) exhibits large variability at Stations 1 and 6, while *D. excentricus* percent fertilization (Figure 11) was quite variable in sediments from Stations 1, 5, and 6.

For this data set, these two tests exhibit more variability in subsamples or laboratory replicates than among true replicates or field replicates. If this is generally the case, then the absence of a "failure" may be due to excessive variability and not due to a true lack of effect. Although not important from a strictly ocean disposal perspective, a large degree of variability may also restrict modeling biological responses as shown in Section 4.2.

The coefficients of variation (see Table 24) for reference toxicant tests are examined using Cu and/or Cd, conducted at the Environment Canada Toxicology Laboratories, in Moncton and Vancouver. The reported coefficients of variation are estimated from the LC50s.

Table 22 Summary of Regression of Tissue Contaminant Levels on Sediment and Porewater Variables

Tissue Contaminant	Cr	Cu	Ni	Pb	Zn	Total PAH	PCB
Intercept	-1664.961	596.1422	-1217.053	3.767	286.750	-1382.230	116.566
Porewater Variables							
Ag							
Ammonia							
Cr							
Hg							
Eh							
Pb							
PCB							0.687
pH							
Salinity	63.532	-16.164	46.809	0.0321			
Sulphide							
Total PAHs							
Zn					-83.520		
Sediment Variables							
Ammonia							
As							
Cd							
Cr							
Cu							
Dry Weight							
Hg							
Ni							
PCB							
Pb							
Total PAHs							
Zn							
Moisture							
TOC						-0.835	
Percent Sand							
Percent Silt							
Percent Clay							
Goodness of Fit R^2	0.573	0.425	0.581	0.559	0.851	0.911	0.687
Model P-value	0.0814	0.160	0.0781	0.0875	0.00879	0.00358	0.0415

Table 23 Summary of Tests of Equality of Variance

Test Species	Degrees of Freedom	P-Value
A. virginiana	14, 84	0.9759354
B. proboscidea	Test Not Possible	
E. estuarius	14, 84	0.5348306
E. washingtonianus	14, 84	0.02955153
D. excentricus	8, 45	0.04456763
L. pictus	8, 45	1
P. cornuta	Test Not Possible	
R. abronius	14, 84	0.09214276
V. fischeri	No Laboratory Replication	

Table 24 Summary of Reference Toxicant Coefficients of Variation

Test Species	Coefficient of Variation (%)	Sample Size
A. virginiana	54.0765	33
B. proboscidea	21.687	6
E. estuarius	56.0	13
E. washingtonianus	40.58	3
D. excentricus @ 10° C	27.57	14
D. excentricus @ 15° C	9.74	9
L. pictus	45.678	13
P. cornuta	16.293	6
R. abronius	43.575	5
V. fischeri	27.19	138

Exposure temperatures of 15° C are currently being used for the *D. excentricus* biological toxicity test. *Eohaustorius washingtonianus* exhibits a large coefficient of variation in reference toxicant tests, but *D. excentricus* (at 15° C) does not. It would be useful to compare the coefficients of variation from different field studies to see if the findings are validated.

4.4 H4: Does the suite of biological toxicity tests provide a consistent interpretation of the status of the sediment?

This hypothesis is tested by testing two sub-hypotheses. These are: Do all biological toxicity tests characterize the sample in the same way? and Do biological toxicity tests rank the stations in the same way?

4.4.1 H4a: Do all biological toxicity tests characterize the sample in the same way?

This hypothesis may be tested using concordance analysis where the number of "agreements" between pairs of biological toxicity tests is statistically analyzed. A lack of concordance may indicate that constituents of the battery are providing complementary rather than redundant information, which is the *raison d'être* for a battery, Munawar *et al.* (1992) and Keddy *et al.* (1994).

4.4.1.1 Biological Toxicity Tests Pass/Fail Status Relative to Control Sediments (Table 25).

The biological toxicity tests are generally in agreement when using the response

in control sediments as a point of reference. Stations 1, 5, and 6 show clear evidence of adverse effects while Stations 9, 12, and St. Ann's Harbour indicate some adverse effects. The most sensitive tests or species in descending order are: *D. excentricus*, *L. pictus* = photoluminescent bacteria = *A. virginiana* = *E. washingtonianus*, *E. estuarius*, and *R. abronius*.

The pass/fail data is used to test the null hypothesis H4a: All biological toxicity tests pass or fail stations consistently. Cochran's test is used (Cochran, 1950) and the assumption is made that the stations were randomly chosen from among all possible stations. This is true within the criteria for choosing stations (see Materials and Methods, Site Selection Criteria). Cochran's test statistic of 12.222 on 6 degrees of freedom is associated with a p-value of 0.0572. This suggests that the group of biological toxicity tests passes or fails stations in the same way when the basis for comparison is control sediment.

4.4.1.2 Biological Toxicity Tests Pass/Fail Status Relative to St. Ann's Harbour (Table 26). The biological toxicity tests show less agreement when using the response in St. Ann's

Table 25 Summary of Station Pass/Fail Status* Relative to Control Sediments

Response	Test/Species	Stations						
		1	5	6	9	12	St. Ann's Harbour	Proportion of Sites Failing
Survival	*A. virginiana*	F	F	F				50
	E. estuarius	F						16.7
	E. washingtonianus	F	F	F				50
	R. abronius							0
Reproduction	*D. excentricus*	F	F	F	F	F		83.3
	L. pictus		F		F		F	50
Luminescence	*V. fischeri* (Solid Phase)	F	F	F				50
	Proportion of Tests Failing	71	71	57	29	14	14	

* Interim interpretation criteria from Environment Canada (1996)

Table 26 Summary of Station Pass/Fail Status* Relative to St. Ann's Harbour

Test/Species	Stations					
	1	5	6	9	12	Proportion of Sites Failing
A. virginiana	F	F	F			60
D. excentricus	F	F	F	F	F	100
E. estuarius	F					20
E. washingtonianus	F		F			40
L. pictus						0
V. fischeri (Solid Phase)	F	F	F			60
R. abronius						0
Proportion of Tests Failing	71	43	57	14	14	

* Interim interpretation criteria from Environment Canada (1996)

Harbour sediments as a point of reference than when using control sediments. Stations 1, 5, and 6 show evidence of adverse effects while Stations 9 and 12 may indicate adverse effects. The most sensitive tests or species in descending order are: *D. excentricus*, *A. virginiana* = photoluminescent bacteria, *E. washingtonianus, E. estuarius,* and *R. abronius = L. pictus*.

The pass/fail data is used to test the null hypothesis H4a: All biological toxicity tests pass or fail stations consistently. Cochran's test statistic of 11.636 on 4 degrees of freedom is associated with a p-value of 0.0203. This suggests that the group of biological toxicity tests does not pass or fail stations in the same way.

4.4.1.3 Biological Toxicity Tests Pass/Fail Status Relative to Station 12 Reference Site (Table 27). As before, the biological toxicity tests are generally in agreement and Stations 1, 5, and 6 show clear evidence of adverse effects while Station 9 may be adversely affected. The ordering of the relative sensitivity of species is changed. The most sensitive tests or species in descending order are: *E. washingtonianus = A. virginiana* = photoluminescent bacteria, *L. pictus, R. abronius = E. estuarius* and *D. excentricus*.

The pass/fail data is used to test the null hypothesis H4a: All biological toxicity tests pass or fail stations consistently. Cochran's test statistic of 9.333 on 3 degrees of freedom is associated with a p-value of 0.0252. This suggests that the group of biological toxicity tests does not pass or fail stations in the same way.

As a group, the biological toxicity tests pass or fail stations in the same way only when the basis for comparison is a control sediment. When the pass/fail decision is made relative to a reference sediment, then the group of biological toxicity tests passes/fails sediments differently. This is intuitively satisfying as we know that the response at a control sediment must be "good" or the sediment would not have been chosen as a control sediment. The agreement in pass/fail status when using the group of biological toxicity responses relative to control sediment responses implies that the biological tests are responding adversely to either confounding variables or contaminant effects at the exposure sites. The relationship observed between the amphipods and the PAH gradient (see Section 4.2.3) and other contaminants for those responses which were predictable) suggests that the pass/fail decision is due to the presence of contaminants.

The lack of agreement in pass/fail status when using the group of biological toxicity responses relative to reference sediment responses may be due to the effects of factors present at the reference sites. If some organisms exhibit a negative response in the reference sediments, and other organisms do not, the pass/fail status of sediments (relative to the reference stations) will not be homogeneous among the group of toxicity tests. This was the inference or conclusion made

Table 27 Summary of Station Pass/Fail Status* Relative to Station 12 Reference Site

Test/Species	Stations				
	1	5	6	9	Proportion of Sites Failing
A. virginiana	F	F	F		75
D. excentricus					0
E. estuarius	F				25
E. washingtonianus	F	F	F		75
L. pictus		F		F	50
V. fischeri (Solid Phase)	F	F	F		75
R. abronius		F			25
Proportion of Tests Failing	57	71	43	14	

* Interim interpretation criteria from Environment Canada (1996)

from the three Cochran's tests on the homogeneity of biological toxicity test responses.

Examples of biological toxicity tests performing differently than expected (i.e., non-monotonically) are seen in the percent survival for *E. washingtonianus*, photoluminescent bacteria light inhibition, percent fertilization for *D. excentricus* and *L. pictus* as shown in Section 3.1.

The relative sensitivity of the biological test species changes when reference stations are changed. This observation is on the surface, unsettling. However, when the two species exhibiting reversals in sensitivity, *D. excentricus* and *L. pictus*, are examined, it can be seen that the extreme responses for these species were observed in the control sediments, St. Ann's Harbour or the Station 12 reference site. Thus these stations are acting as pivotal stations, when pass or fail decisions are being made. The choice of station (or pivot) affects the pass/fail decision.

4.4.2 H4b: Do biological toxicity tests rank the stations in the same way? Now that it has been shown that the stations are not passed or failed in the same way using different biological toxicity tests when the basis for comparison is a reference station, it is investigated if the biological toxicity tests at least rank the stations in the same order. The biological toxicity test raw data (i.e., within station responses, not station pass/fail status) from all stations is used, to test the null hypothesis, H4b: Each biological toxicity test ranks the sites equally. The Spearman rank correlation coefficient is used to test this hypothesis. Note that normally when many statistical tests are conducted on a data set, the probability of rejecting the null hypothesis when it should not be rejected increases for the group of comparisons as a whole. Often a correction factor is applied to the α value for a single test. For the following data set, the Bonferroni-adjusted α value would be equal to 0.0025. Even when sites are perfectly correlated, the p-value has a theoretical minimum of 0.0298. Thus the null hypothesis can never be rejected when using the Bonferroni adjustment. The p-values are presented in Table 28 without correction. The table format is "Spearman's ρ, p-value". Highlighted cells indicate the pairs of tests that rank the stations in the same way.

The *D. excentricus*, *R. abronius*, and *L. pictus* biological toxicity tests, rank stations differently than all the other tests. Only four pairs of the tests rank the stations similarly.

The general conclusions reached in this section are:

- The biological toxicity tests tend to fail the stations closest to Muggah Creek, irrespective of choice of "reference" sediment.

- As a group, the biological toxicity tests pass or fail stations in the same way when the basis for comparison is a control sediment.

Table 28 Summary of Multiple Comparisons Among Biological Toxicity Tests

	D. excentricus	*E. estuarius*	*E. washingtonianus*	*L. pictus*	Microtox	*R. abronius*
A. virginiana	0.3714, 0.3711	**0.9429, 0.0409**	0.8857, .0553	-0.08571, 0.7983	**1, 0.0298**	0.6, 0.2013
D. excentricus	-	-0.4286, .3067	-0.6571, 0.1252	-0.08571, 0.7983	-0.3714, 0.3711	0.08571, 0.8983
E. estuarius	-	-	**0.9429, 0.0409**	-0.2571, 0.5229	**0.9429, 0.0409**	0.4857, 0.3067
E. washingtonianus	-	-	-	-0.3142, 0.4433	0.8857, 0.0553	0.8857, 0.0553
L. pictus	-	-	-	-	-0.08571, 0.7983	0.31428, 0.5229
Microtox	-	-	-	-	-	0.6, 0.2013

- The choice of reference station may greatly affect the pass/fail decision rendered by a single biological test.

- There is general concordance between tests using pass/fail criteria and the test ranks. *Rhepoxynius abronius* biological toxicity test is one of the least sensitive tests of those with sufficient data to test, and it tends to rank stations differently from other tests.

- The photoluminescent bacterial test ranks stations similarly to *E. estuarius* and *A. virginiana* and is similar in sensitivity (as determined by proportion of station "failures") to *A. virginiana* and *E. washingtonianus*.

- Amphipods from the genus *Eohaustorius* rank the stations similarly.

4.5 H5: Do the biological toxicity tests indicate an effect when the TELs or PELs are exceeded?

Table 29 summarizes the mean parameter values for those parameters that have interim sediment quality guidelines (ISQGs) established. Values exceeding the probable effect level (PEL) and threshold effect level (TEL) are highlighted, with parameters exceeding the PEL being shaded more darkly.

The table shows that all stations would fail, according to the current values. If the mean PEL quotient, given as:

$$\text{PEL Quotient} = \sum_{i=1}^{n}\left(\frac{concentration_i}{PEL_i}\right)\Big/ n$$

following Long *et al.*, (1998) is used, the probability that stations along the gradient would exhibit toxicity are 56% for Stations 1, 5, and 6, and 24% for Stations 9, 12 and St. Ann's Harbour. Given all the caveats of the PEL Quotient method (assumes additivity, all contaminants were measured, database derived largely from amphipod survival tests, etc.) Table 30 shows a good deal of concurrence between expected toxicity using sediment contaminant levels and biological toxicity tests.

H5: "Do the biological toxicity tests indicate an effect when the guidelines suggest there should be an effect?" is then tested using concordance analysis to test the agreement between the characterization of a sediment using biological toxicity tests and TELs or PELs. The proportion of parameters exceeding the PEL or TEL is compared with the proportion of stations failing the biological toxicity tests.

The data matrix used in estimating the correlations in the first two rows of the summary table is presented in Table 31, as an aid in understanding the rationale behind this analysis. The correlation between the proportions of parameters exceeding the TEL at a given station is estimated (1.00,1.00, 0.95, 0.75, 0.10, and 0.15) with the proportion of biological tests failing a station relative to control sediment (0.71, 0.71, 0.57, 0.29, 0.14, and 0.14). The value of 0.733 in the summary table (Table 31), is the correlation (as measured by Kendall's Tau) between these two vectors.

Statistically, the hypothesis, Ho: "There is no correlation among the proportions using Kendall's Tau" is tested. Note that this differs from the current paradigm for characterizing a site as failing. Tests significant at the 5% level are highlighted in Table 32.

There is a significant correlation between the proportion of failures using biological toxicity tests and TELs and PELs when biological toxicity test inferences are made relative to control sediment. When the biological toxicity test inferences are made relative to the St. Ann's Harbour station, only the relationship with PELs is significant. The correlation between pass/fail status using biological toxicity tests and TELs is highest when biological toxicity tests inferences are made relative to the Station 12 reference site. This correlation is not significant due to the low number of data points (four pairs of pass/fail decisions) available for the comparison.

Table 29 Summary of Sediment Contaminants with PELs and TELs

	Stations						TEL	PEL
Parameter	**1**	**5**	**6**	**9**	**12**	**St. Ann's Harbour**		
Metals (*μ*g/g dry weight)							**µg/g dry**	**µg/g dry**
As*	41.0000	39.3333	32.6667	10.0000	9.6667	15.6667	7.24	41.6
Cd*	1.16667	0.9267	0.4667	0.1467	0.08000	0.2467	0.7	4.2
Cr*	81.2333	86.5667	61.7000	33.9000	24.1000	41.9333	52.3	160
Cu*	101.3333	73.6667	53.7000	22.7000	13.0000	37.3000	18.7	108
Hg*	0.7077	0.4863	0.3330	0.0377	0.0243	0.0550	0.17	0.70
Pb*	285.6667	214.0000	133.3333	32.0000	21.0000	37.0000	30.2	112
Zn*	516.2667	865.6667	281.5667	91.1000	56.2000	84.2333	124	271
PCBs (ng/g dry weight)							**ng/g dry**	**ng/g**
total PCB*	2095.0230	1186.2303	642.7346	-69.1310[3]	-63.5605	-30.8474	21.5	189
PAHs (ng/g dry weight)*							**ng/g dry**	**ng/g dry**
2-Methyl-Naphthalene	1161.3778	710.9085	361.3295	88.1666	39.7952	5.5598	20.2	201
Acenaphthylene	690.1267	372.3373	202.4355	7.2086	1.1304	0.0000	5.87	128
Acenaphthene	419.2858	259.1971	152.4286	27.6189	2.6713	0.8682	6.71	88.9
Fluorene	1636.1570	901.9746	432.1399	50.3778	12.4128	5.1443	21.2	144
Phenanthrene	8839.2643	4929.8272	2427.1508	340.4782	50.5420	13.3471	54.1	86.7
Anthracene	5498.4707	2787.5911	1362.9712	125.8528	41.7342	9.3728	46.9	245
Fluoranthene	13651.301	6054.4574	2950.9182	337.5117	41.2303	31.9409	113	1494
Pyrene	19589.421	6888.9134	3450.7664	289.4181	34.9358	24.5577	153	1398
Benzo*(a)* Anthracene	14229.563	6848.0221	3147.7763	283.8230	24.9308	22.4324	74.8	693
Chrysene	17391.319	8184.1967	3798.1966	312.2847	30.7719	23.0015	108	846
Benzo*(a)*Pyrene	23464.113	8502.3410	3352.9856	193.5271	11.3748	17.8804	88.8	763
Dibenzo*(a,h)* Anthracene	4223.4349	1871.4732	671.2666	44.7249	0.0000	0.0000	6.22	135
Mean PEL Quotient**	14.85907	7.593556	3.619002	0.438296	0.108411	0.107492		

* CCME (1999); ** Long *et al*. (1998).

[3] Negative numbers reflect values below the detection limit. The detection limit is a function of the instrument response, the noise in the chromatogram, and the amount of sample used.

Table 30 Comparison of Mean Quotient PELs and Observed Toxicity

	Station					
	1	**5**	**6**	**9**	**12**	**St. Ann's Harbour**
Probability that sediment is acutely toxic (Long *et al.*, 1998)	56	56	56	24	24	24
Proportion of Toxicity tests Failing relative to control sediments (from Table 25)	71	71	57	29	14	14

Table 31 Sample Data Set

Assessment Tool	Station					
	1	**5**	**6**	**9**	**12**	**St. Ann's Harbour**
Proportion Samples > PEL (from Table 29)	0.80	0.75	0.75	0.05	0.00	0.00
Proportion Samples > TEL (from Table 29)	1.00	1.00	0.95	0.75	0.10	0.15
Proportion Stations "Failing" relative to Control Sediment (from Table 29)	0.71	0.71	0.57	0.29	0.14	0.14

Table 32 Summary of Tests of Correlation Between Proportions of Stations Failing Assessment Tool

Biological Toxicity Tests Relative to:	**Comparison with:**	**Kendall's Tau**	**P-value**
Control Sediment	TEL	0.733	0.0293
	PEL	0.733	0.0266
St. Ann's Harbour	TEL	0.6	0.1184
	PEL	0.8	0.0374
Station 12 Reference Site	TEL	0.833	0.071
	PEL	0.5	0.279

Additionally, the proportion of tests failing the sediments, concurs with the probability that a sediment is highly toxic (see Long *et al.*, 1998 for definition), based upon mean PEL quotients.

4.6 H6: Do the biological toxicity tests indicate an effect when the in-situ *benthic macroinvertebrate community does?*

This hypothesis is tested by examining the data set for patterns. Ordinations of the benthic macroinvertebrate abundances are used to explore structure in the data set (Figure 27). Ordination of averaged raw abundances using the correlation matrix showed that a large proportion of the variability in the data set was due to simple numerical dominance in the case of the first principal component and the presence of specific organisms for the second principal component.

The data is well summarized by two components accounting for 82.5% of the total variability in the data set. Plots of station scores show that the first principal component strongly separates Stations 9 and 12 from the other stations. The second principal component strongly separates out the Station 12 reference site and only weakly separates Station 9. Both principal components group Stations 1, 5, 6 and St. Ann's Harbour together. Also, Arenicola Marine (1997) describes the St. Ann's Harbour reference station as being most similar with respect to benthic community structure, to Station 1 (Arenicola Marine, 1997). St. Ann's Harbour was also characterized as one of the two most affected sites, based on benthic community structure

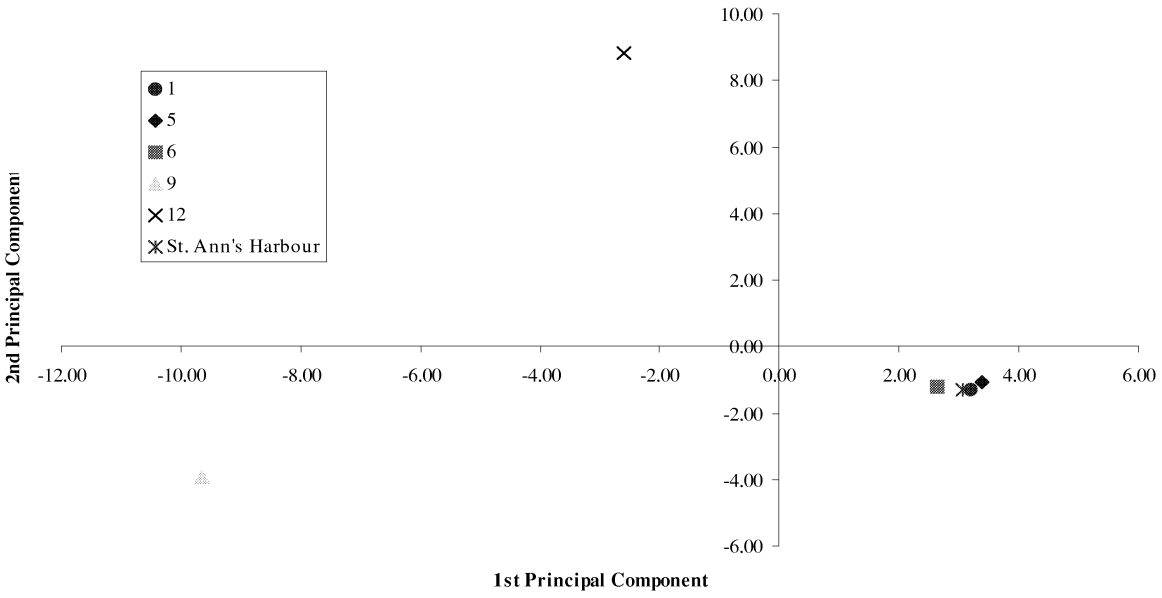


Figure 27 Scores Plot for Benthic Macroinvertebrate Abundances

(Arenicola Marine, 1997). This ordination shows that on the basis of benthic community structure, Stations 1, 5, 6 and St. Ann's Harbour are affected.

The separation of Stations 9 and 12 by the first principal component is due to the numerical abundance of the following taxa: *Nucula delphinodonata*, *Chiridotea tuftsi*, Sabellidae, *Ninoe nigripes*, *Eudorallopsis deformis*, *Diastylis polita*, Nematoda and *Stenothoe minuta*. The separation of the Station 12 reference site by the second component is due to the presence of the following taxa (albeit in small numbers) only at the Station 12 reference site; *Echinarchnius parma*, *Spiophanes bombyx*, Capitellidae, *Phoxocephalus holbolli*, *Gammarus* sp., Copepoda, *Prionospio steenstrupi*, and *Neoleanira tetragona*.

This ordination and observations by Arenicola Marine (1997) led to a dialogue regarding the definition of a reference station. It was decided to retest all hypotheses that used St. Ann's Harbour as a baseline or reference station by using the Station 12 reference site as a reference station and omitting the St. Ann's Harbour station.

The hypothesis Ho: "Do the biological toxicity tests indicate an effect when the *in-situ* benthic macroinvertebrate community does?" is then tested. First, it is determined whether the *in-situ* benthic macroinvertebrate community differs from the benthic community at the reference station. This may done using a permutation test by Clarke (1993) that is a variant of Mantel's test (Mantel, 1967). The ANOSIM method uses the Bray-Curtis similarities between station subsamples, (based upon 4^{th}-root transformed abundance data) to perform an ANOVA-like calculation. Significance tests are generated using permutation distributions.

The ANOSIM procedure showed that all stations differ from either of the reference stations, with p-values of 0.009 90 for all comparisons. Thus using the *in-situ* benthic macroinvertebrate community as a tool for determining the status of sediments, it would be stated that all exposure stations fail. This inference is different from the

ordination which groups Stations 1, 5, 6 and St. Ann's Harbour together, based on raw abundances. The result of the difference is likely the 4th root-transformation that greatly reduces the effects of numerical abundance. The interpretation of the ANOSIM procedure following the ordination is that once the effects of numerical abundance have been reduced, all stations differ from either of the reference stations, likely on the basis of species composition. Undue weight should not be placed upon this finding until the literature corroborates the use of this test with other procedures and data sets.

If the number of times that both biological toxicity tests fail a station (based upon the failure of one acute amphipod test) and benthic communities fail a station is counted, Table 33 can be created.

Fisher's exact test is used to test the null hypothesis that the two row and column variables are independent. The p-value for the both tables is 1. Consequently, it is found that the benthic macroinvertebrate community and the biological toxicity tests do not characterize the sediments in the same way.

4.7 H7: Do the three evaluation tools characterize sediments in the same way?

The degree of agreement in the classification of a sediment or site using the three characterization tools is summarized in Table 34. The stations fail a biological toxicity test if more than 1 test fails or a single amphipod toxicity test fails.

There is perfect agreement between the biological toxicity tests and the sediment PELs, based upon pass/fail status. (Note that the current use of TELs in the ocean disposal context is to trigger a tier 2 assessment, not necessarily prevent dredged material from being disposed of at sea.) The benthic macroinvertebrate community and the TELs concur. However, the ANOSIM procedure used to compare the benthic community macroinvertebrate structure between stations does not agree with the ordination of the benthic macroinvertebrate community structure likely due to the difference in emphasis on numeric abundance. We refrain from making conclusions regarding the concordance of the three assessment tools pending a comparison of the three data sets using the raw data. (See the following final hypothesis.)

4.8 H8: How strongly are the three data sets correlated?

The study design is a gradient design using a sediment quality triad approach. The constituents of the triad are biological toxicity tests, sediment physical/chemistry and *in-situ* benthic macroinvertebrate community structure. The degree of correlation between these data sets is explored using ordination (principal components analysis) and permutation tests (Mantel's test).

Table 33 Concordance Between Station Characterizations using Benthic Community Structure and Biological Toxicity Tests

Biological Toxicity Tests (relative to St. Ann's Harbour)		Benthic Macroinvertebrate Community: Fail	Benthic Macroinvertebrate Community: Pass
	Fail	3	0
	Pass	2	0

Biological Toxicity Tests (relative to Station 12)		Benthic Macroinvertebrate Community: Fail	Benthic Macroinvertebrate Community: Pass
	Fail	3	0
	Pass	1	0

Table 34 Summary of Station Pass/Fail Status* Using All Criteria

	Pass/Fail Status					
	1	**5**	**6**	**9**	**12**	**St. Anne's Harbour**
Biological Toxicity Tests Relative to Control Sediment	F	F	F	P	P	P
Biological Toxicity Tests Relative to St. Ann's Harbour	F	F	F	P	P	NA**
Biological Toxicity Tests Relative to Station 12	F	F	F	P	NA**	P
Benthic Macroinvertebrate Community	F	F	F	F	F***/ NA**	NA**
Sediment TEL	F	F	F	F	F	F
Sediment PEL	F	F	F	P	P	P

* Criteria for biological toxicity test given in Environment Canada (1996).

** By definition, the reference site cannot "fail"; therefore, the pass/fail status of the reference station (either St. Ann's Harbour or station 12) is not in question.

*** Fails when compared to St. Ann's Harbour.

We begin with ordinations of the sediment physical/chemistry and biological toxicity test responses. It was previously shown that individual PAHs are highly correlated with total PAHs, consequently total sediment and porewater PAHs will be used in lieu of individual sediment and porewater PAHs, respectively, for the ordinations (see Figure 28).

The principal components analysis shows that the sediment physical/chemical data set is highly structured. Three principal components describe 80.200% of the total variability in the data set. The first principal component orders the stations in the following order: 1, 5, 6, St. Ann's Harbour, 9 and 12. There is clear evidence of the gradient by the loadings (not shown) on the following 12 variables in order of magnitude: TOC, sediment Cu, Pb, As, Cr, porewater PCBs and PAHs, sediment PAHs, sediment Hg, Ni, total PAHs, and Cd. The St. Ann's Harbour station has intermediate values for these parameters; hence it's position in the ordering due to the first principal component. The other stations are arranged in an order corresponding to the known gradient in PAHs and metals.

The second principal component describing an additional 18.6% of the variability in the data set separates St. Ann's Harbour from the other stations. This is due to loadings (not shown) that contrast porewater Zn, Cr, sulphide, sediment ammonia, moisture and redox potential with porewater pH and ammonia, dry weight, porewater Ag, salinity, and Hg. The separation of the St. Ann's Harbour station from the other stations is due to the low values of pH, porewater ammonia, and dry weight (relative to the other stations). The low pH may have increased the porewater levels of Zn and Cr (Sigg, 1987). Porewater levels of these two metals are the highest in St. Ann's Harbour stations, even though these sediments contain only moderate amounts of sediment-associated Zn and Cr (relative to the other stations).

The third principal component describing an additional 8.01% of the variability in the data set separates Station 9 (not shown) from the other stations due to contrasting loadings on fines (clay and silt) and sand. Station 9 has the largest amounts of clay and silt and the smallest amount of sand relative to the other stations.

Thus we see that the stations are ordered along a concentration gradient, St. Ann's Harbour sediments are moister, (possibly with increased porosity) and have elevated porewater metal levels, possibly due to pH. Station 9 is anomalous with respect to sediment grain size composition containing more clay and less sand than the other stations.

The biological toxicity test responses are ordinated using principal components analysis (PCA). Three components describe 94.600% of the total variability in the data set. The scores plot is shown in Figure 29.

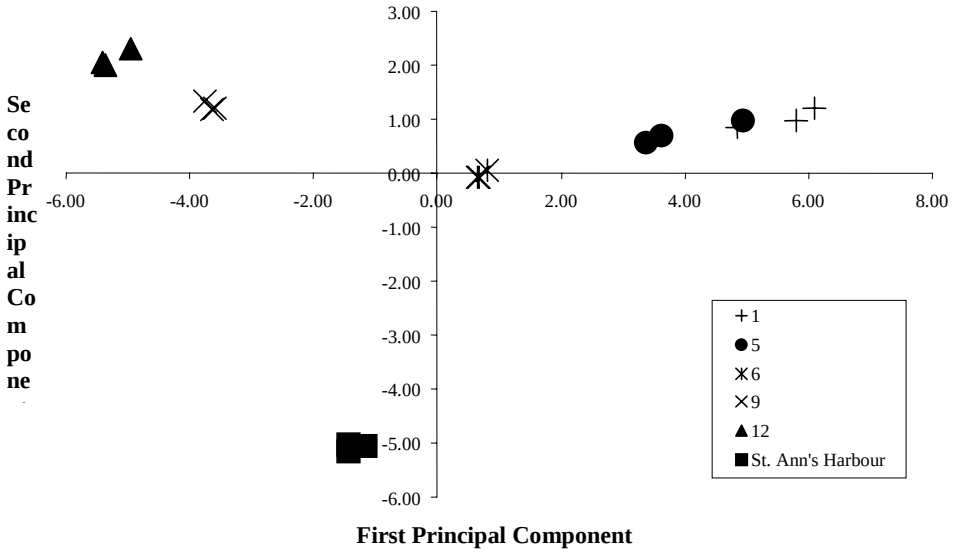


Figure 28 Scores Plot for Sediment Physical/Chemistry Variables

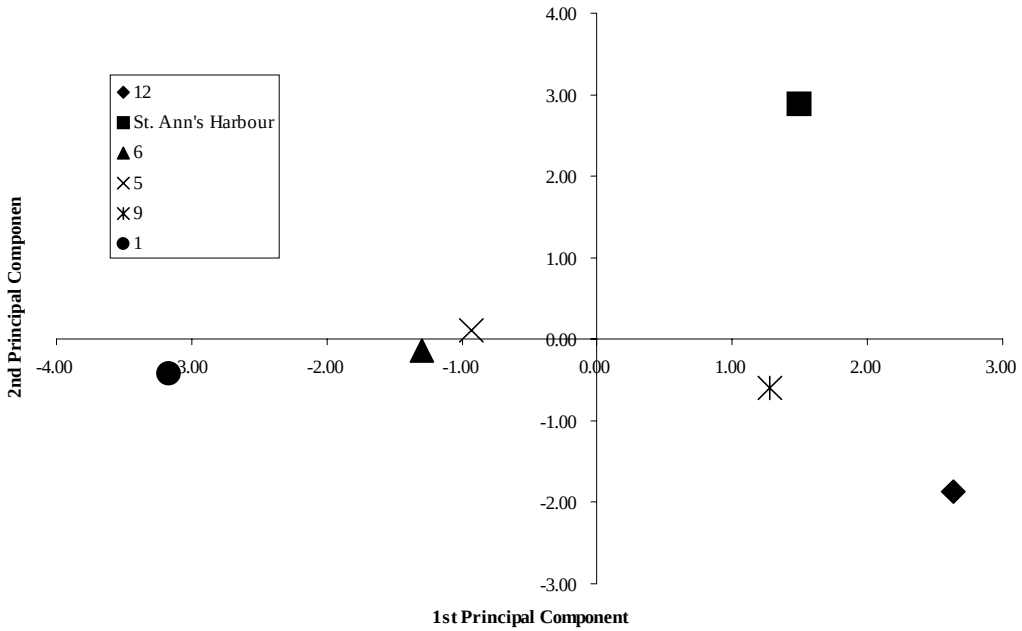


Figure 29 Scores Plot for Biological Test Responses

The first principal component describing 53.2% of the total variability in the data set orders the stations roughly corresponding to the PAH gradient. Stations 5 and 6 are grouped together and there is little difference between Stations 9, St. Ann's Harbour and 12. The ordering of the stations (1, 6, 5, 9, St. Ann's Harbour, and 12) is due to the loadings on the responses from the following biological toxicity tests*: E. washingtonianus*, *E. estuarius*, *A. virginiana*, Microtox® solid phase and *R. abronius*.

The second principal component describing an additional 23.9% of the total variability in the data set, separates the St. Ann's Harbour reference site on the basis of contrasting loadings on the *P. cornuta*, *D. excentricus*, *E. washingtonianus*, *A. virginiana*, and *R. abronius* biological toxicity tests with those from *L. pictus* and Microtox® solid phase. This occurs as the maximal or near maximal response (i.e., highest growth, percent fertilization or percent survival) is found at St. Ann's Harbour for the first group of tests, while a low or intermediate response is found at St. Ann's Harbour for the second group of tests.

The third principal component describing an additional 17.400% of the total variability in the data set weakly separates Station 9 due to contrasting loadings on *E. estuarius*, *A. virginiana* and *E. washingtonianus* biological toxicity tests with those from *R. abronius*, *L. pictus*, Microtox® solid phase, *P. cornuta* and *D. excentricus*. This occurs as the maximal or near maximal survival or fertilization is found at Station 9 for the first group of tests, while a low or intermediate response is found at Station 9 for the second group of tests.

The ordination of the biological toxicity test responses shows that the stations may be ordered along the PAH gradient, the St. Ann's Harbour reference site is separable from the other stations, and that Station 9 is weakly separable from the other stations.

4.8.1 Heuristic Comparison of Ordinations (Table 35). The ordinations show a good deal of concurrence. All three assessment tools group Stations 1, 5 and 6 together. Station 9 is separated on the basis of numerical abundance of specific taxa, contrasts in biological toxicity test responses and sediment grain size differences. The ordinations using biological toxicity test responses and sediment physical chemistry separate St. Ann's Harbour from other stations.

4.8.2 Comparison of the Three Evaluation Tools. The matrices of responses or measurements at a given station, for each of the three evaluation tools are compared using Mantel's test (Mantel, 1967). This is a non-parametric resampling method that tests the null hypothesis, Ho: "There is no association between two distance matrices". The distance matrices are formed by determining the multivariate distance between two stations based on the variables of interest. For example, a matrix of distances based on benthic macroinvertebrate abundance data tells us how "far apart" matrices are based on the abundance of benthic macroinvertebrates. A matrix of distances based on biological toxicity test responses tell us how far apart the same stations are, based on biological toxicity test responses. If benthic abundances and biological toxicity test responses are related to one another, they should produce similar matrices of distances between stations. Mantel's test tests the similarity of the two distance matrices as measured by correlation, using permutations to test the significance of the correlation between the pair of distance matrices being compared. The results of these analyses are summarized in Table 36. Correlations close to either 1 or –1 imply a strong positive or negative correlation, respectively between the two data sets being compared. A small p-value indicates that the correlation is significantly different from 0.

There is no significant correlation between any of the stations compared. However, the p-value for the agreement between sediment physical/chemistry and the biological toxicity tests is moderately low (approaching the generally accepted cutoff of $\alpha = 0.05$), due to some of the predictable dose responses.

Table 35 Summary of Ordinations

Ordination on:	Comment
Benthic Community Structure	First component separates Stations 9 and 12 due to numerical abundance of specific taxa. The second components separates the Station 12 reference site due to the presence of organisms found only at that station. Stations 1, 5, 6, and the St. Ann's Harbour station are grouped together.
Biological Toxicity Tests	Shows gradient, separates St. Ann's Harbour and weakly separates Station 9 due to contrasts in biological toxicity test responses. Stations 1, 5, 6, and 12 are grouped together.
Sediment Physical Chemistry	Shows gradient, separates St. Ann's Harbour due to porewater ammonia, metals and low pH, and separates Station 9 due to sediment grain size differences. Stations 1, 5, 6, and 12 are grouped together.

Table 36 Summary of Mantel's Test Comparisons

Comparison	Correlation	P-Value
Biological Toxicity Tests and Benthic Community Structure	0.133	0.364
Sediment Physical/Chemistry and Benthic Community Structure	-0.205	0.892
Sediment Physical/Chemistry and Biological Toxicity Tests	0.436	0.199

Also, the proportion of tests failing the sediments, concurs with the probability that a sediment is highly toxic (see Long *et al.*, 1998 for definition), based upon mean PEL quotients as shown in Table 30.

4.9 H9: Do the interim biological toxicity test interpretation criteria need to be adjusted to match assessments of sediment quality using benthic community structure or SQGs?

The previous hypothesis tests indicate that the interim biological toxicity test interpretation criteria characterize (i.e., pass/fail) sediments in the same way as the PELs. When the biological toxicity test battery is used to determine the suitability of dredged materials for ocean disposal, the battery will consist of one each, of the following tests: amphipod survival, echinoderm fertilization, bivalve bioaccumulation, and the Microtox® (*V. fischeri*) solid-phase test, (L. Porebski, pers. comm., Environment Canada, Ottawa, ON, 1999). A sediment will fail the biological test battery if two or more tests fail or if one acute toxicity test (amphipod) fails.

The significant difference, or in the terms of this report, the "failure" of Stations 1, 5, 6 and 9 relative to either of the reference stations based upon the *in-situ* benthic macroinvertebrate communities should not be given undue weight. The ANOSIM procedure is relatively new and may be overly sensitive. Ordination procedures and expert opinion group Stations 1, 5, 6 and St. Ann's Harbour together. This grouping of the reference station with the most highly contaminated sites is contrary to the findings of the biological toxicity tests and sediment contaminant concentrations.

The proportion of biological toxicity tests failing a station is more highly correlated with the proportion of stations failing according to PEL criteria than TEL criteria. Also, the pass/fail status of stations is identical using biological toxicity tests and sediment PELs. Finally, the proportion of tests failing the sediments concurs with the probability that a sediment is highly toxic based upon mean PEL quotients. Empirically and statistically that is seen as a group, the interim biological toxicity test criteria reflect sediment contamination in a meaningful way.

Summary of Hypothesis Tests

At times during the report, a hypothesis test or analysis is repeated when one of the two reference stations is used as a basis for comparison. If changes in inferences resulting from the use of a different reference station occur, the differences are summarized in Table 37 by the inclusion of a separate set of rows for the hypothesis so affected. The results and implications of hypothesis 2 reflect the manual or consensus model building approach rather than the automated model building approach of Section 4.2.2.

Table 37 Summary of Hypothesis Tests

Hypothesis Test	Result	Implication
H1: Homogeneity of confounding factors across station.	Median TOC, percent moisture, sand, silt clay, redox potential, porewater and sediment ammonia and sediment sulphide values are different for at least one of the stations along the gradient.	Biological toxicity test responses and benthic community structure may vary due to confounding effects as well as toxicant effects.
H2a: Effect of confounding factors on biological toxicity test responses.	Of the organisms that reacted with a strong biological response across the gradient none were best predicted by confounding factors. However, TOC was often highly correlated with those variables included in the model and was therefore a contending explanatory variable.	Confounding factors do not affect the observed responses more than the contaminant variables. Confounding factors, particularly TOC are often associated with high levels of contaminants. Confounding factors should continue to be monitored.
H2b: Are the dose responses predictable?	A strong response to the PAH gradients occurred among the amphipods *A. virginiana* and *E. estuarius*. Other organisms responded to different contaminants such as metals. The exploration of the various models not presented, shows that non-contaminant variables are also associated with the responses, but not as strongly as with contaminants. The correlation between contaminants and non-contaminant factors, particularly TOC and redox potential was often high. Non-contaminant effects such as grain size were not observed.	The biological toxicity tests are responding primarily to the contaminants.
H3: Do biological toxicity tests perform consistently within a site?	The within sample mean square error or laboratory replicate variability for *E. washingtonianus* and *D. excentricus* biological toxicity tests were both greater than the treatment mean square error or field replicate variability.	These tests exhibit greater variability within samples than among samples.

Hypothesis Test	Result	Implication
H4: Does the suite of biological toxicity tests provide a consistent interpretation of the status of the sediment?	See following subsections.	The basis for comparison affects the pass/fail decision and consequently the rankings of relative sensitivity. The biological toxicity tests as a group, only pass or fail stations in the same way when the basis for comparison is a control sediment. The choice of the basis for comparison is important in an ocean disposal context.
H4a: All biological toxicity tests pass or fail stations consistently. (St. Ann's Harbour as reference station).	Using pass/fail criteria, *L. pictus* and *R. abronius* were the least sensitive species, *D. excentricus* was the most sensitive. Statistically, the group of biological toxicity does not pass or fail stations in the same way.	The choice of biological toxicity test (and species) influences the pass/fail status of a sediment/station.
H4a: All biological toxicity tests pass or fail stations consistently. (Station 12 as reference station).	Using pass/fail criteria, *D. excentricus* and *R. abronius* were the least sensitive test species while *A. virginiana*, *E. washingtonianus* and *V. fischeri* were the most sensitive organisms. Statistically, the group of biological toxicity tests does not pass or fail stations in the same way.	The choice of biological toxicity test influences the pass/fail status of a sediment/station.
H4a: All biological toxicity tests pass or fail stations consistently. (Control sediment as reference station).	Using pass/fail criteria, *R. abronius* was the least sensitive species while *D. excentricus* was the most sensitive biological toxicity test. Statistically, the group of biological toxicity tests passes or fails stations in the same way.	The choice of biological toxicity test does not influence the pass/fail status of a sediment/station.
H4b: Each biological toxicity test ranks the sites equally.	*R. abronius*, *L. pictus* and *D. excentricus* rank the stations differently from the other biological toxicity tests.	Three biological toxicity tests ranked the stations differently. The most and least sensitive species ranked the stations differently than the other biological toxicity tests.
H5: Do the biological toxicity tests indicate an effect when the TELs or PELs are exceeded?	A comparison of the proportion of biological toxicity tests eliciting a fail response at a station with the proportion of parameters exceeding the PEL or TEL was significant. The degree of correlation between station pass/fail status using biological toxicity tests and chemical values was higher using PEL than TEL.	Biological toxicity test responses are triggered when TEL or PEL values are exceeded. The basis for comparison affects the pass/fail decision and consequently the strength of the relationship between PELs or TELs. The proportion of biological toxicity tests responding, roughly concurs to the proportion expected to respond based upon mean PEL quotients.
H6: Do the biological toxicity tests indicate an effect when the *in-situ* benthic macroinvertebrate community does?	The biological toxicity tests always indicate an effect (based upon the failure of at least one test. The benthic macroinvertebrate community stations always differ from either of the reference stations using permutation methods but not when using ordination methods.	Since all stations "failed" according to both sets of criteria this observation is equivocal. The interpretation of the ANOSIM procedure following the ordination is that once the effects of numerical abundance have been reduced, all stations differ from either of the reference stations, likely on the basis of species composition.

Hypothesis Test	Result	Implication
H7: Do the three evaluation tools characterize sediments in the same way?	The three evaluation tools do not characterize the sediments in the same way. The benthic community fails all stations (see caveats in Section 4.6) as would TELs if employed as pass/fail criteria. There is perfect agreement between pass/fail status using biological toxicity tests and PELs (if PELs are employed as pass/fail criteria).	Sediment chemistry and biological toxicity tests characterize sediments in the same way.
H8: How strongly are the three data sets correlated?	Ordinations using biological toxicity test data and sediment physical chemistry, order the exposure stations along the chemical gradient and separate both Stations 9 and St. Ann's Harbour. The ordination on benthic community structure separates Stations 9 and 12 while grouping Stations 1, 5, 6 and St. Ann's Harbour together. Mantel's tests show no significant correlation between the three data sets using Euclidean distances.	The ordinations on biological toxicity test data and sediment physical chemistry indicate that the biological toxicity tests reflect the sediment chemistry even though this is not corroborated by Mantel's test. The ordination on the *in-situ* benthic community do not order the stations along the known gradient but does group Stations 1, 5, and 6 together. However, the reference station is also grouped with these contaminated sites.
H9: Do the interim biological toxicity test interpretation criteria need to be adjusted to match assessments of sediment quality using benthic community structure or SQGs or ISQGs?	The interim biological toxicity test interpretation criteria reflect a probable effect level.	The current interpretation criteria afford a short-term or lethal level of protection to the environment based upon the responses measured, for the types and concentrations of contaminants encountered. Note: This study used relatively short-term exposures and a small group of taxa. Further investigation would be advisable to clarify the links with ecosystem–level bioaccumulative or sublethal effects.

Discussion

6.1 Performance Evaluation of Biological Toxicity Tests

6.1.1 Acute Survival Tests. The acute survival tests using amphipods show a dose response ranging from moderate to very strong. The responses of *A. virginiana* and *E. estuarius* are related to sediment and porewater PAHs, respectively. The pass/fail comparisons with the control sediments show that the amphipods as a group are less sensitive than echinoderms and *V. fischeri*. When the pass/fail characterizations are made relative to the St. Ann's Harbour reference station, or the Station 12 reference site, the amphipods are of intermediate sensitivity, relative to echinoderms and *V. fischeri*.

Stations are not passed/failed in the same way using the different biological toxicity tests. Note that this evaluation does not include polychaetes, due to the lack of replication. This conclusion holds whether pass/fail decisions are made relative to St Ann's Harbour or the Station 12 reference site. When the biological test responses are used rather than pass/fail status, it is found that *A. virginiana, E. estuarius,* and *E. washingtonianus* rank the stations similarly.

Of the amphipod species tested, the use of *A. virginiana* and *E. estuarius* is recommended based on the general concurrence of pass/fail status, stability of sensitivity to changing reference conditions, and statistically acceptable variability and response to the known organic contaminant gradient in this study.

6.1.2 Sublethal Tests. No dose response was predictable for *D. excentricus, L. pictus, B. proboscidea,* or *P. cornuta*. The log-transformed dose response for *V. fischeri* was well predicted by sediment-associated Ni. When the pass/fail status of sediments relative to control sediments is determined using sublethal tests the echinoderms are the most sensitive biological toxicity tests, with the *V. fischeri* test being of intermediate sensitivity. When the pass/fail test is made relative to the St. Ann's Harbour sediments, the *D. excentricus* fertilization assay remains among the most sensitive tests, while the *L. pictus* biological toxicity test becomes the least sensitive, and relative sensitivity of the *V. fischeri* test remains unchanged. When the pass/fail test is made relative to the Station 12 reference site sediments, the *D. excentricus* fertilization assay is grouped among the least sensitive tests while the other two sublethal tests are of intermediate sensitivity.

Among the sublethal tests, the within sample variability for *D. excentricus* was greater than the variability among samples. This implies that differences between stations or between exposure and control sites may be obscured by the variability within a site.

Of the sublethal tests, the echinoderm tests are the most problematic in this study. Large statistical variability was shown by the *D. excentricus* fertilization assay. The echinoderm pass/fail tests were also sensitive to the choice of reference station. Porebski *et al.* (1998) also found the *D. excentricus* test to be variable and sensitive to the choice of reference station.

The *V. fischeri* pass/fail test remained consistent when reference stations were changed, as the pass/fail decision is not made relative to performance at a reference site. Section 6.6 includes suggestions for modifying interpretation criteria for this test that would affect its consistency.

Polychaete survival was not a function of the gradient and only *P. cornuta* growth exhibited any systematic change along the gradient. However, this change in *P. cornuta* growth, was not predictable. Also, *P. cornuta* did not respond as anticipated to a known metal gradient (Porebski *et al.,* 1998).

Problems with the sublethal toxicity tests based on the results of this study are the general lack of predictable dose responses along a known contaminant gradient (with the exception of the *V. fischeri* toxicity test) and the equivocal pass/fail status of sediments based upon echinoderm toxicity tests.

6.1.3 Bioaccumulation Tests. The bioaccumulation of porewater PAHs and porewater PCBs was highly predictable in *M. nasuta*. The bioaccumulation test using *M. nasuta* is being standardized by the USEPA and is advocated as a bioaccumulation test for the evaluation of dredged materials for disposal at sea (USEPA/USACE, 1998).

The Canadian species *M. balthica* was used to evaluate bioaccumulation along a known metals gradient (Porebski *et al.*, 1998) but the laboratories involved commented that the test was labour intensive due the small size of the organism. *M. nasuta* may provide a better alternative in light of it's larger size, demonstrated performance under an organic contaminant gradient and use by the USEPA in sediment quality assessment which allows a broader basis for comparison. However, *M. nasuta* is not native to the colder waters of Canada. Before a final decision is made to endorse *M. nasuta* over *M. balthica*, comparative sensitivity and variability studies should be undertaken.

6.2 Sediment Physical Chemistry

6.2.1 PELs/TELs and Biological Toxicity Tests. Contaminant levels alone do not fail a station in the ocean disposal context. However, a correlation between guideline levels and biological toxicity test failures provides some degree of assurance that biological effects criteria do not require adjustment. In this study there was a significant correlation between the proportion of failures at a given station using biological toxicity tests and both TELs and PELs when biological toxicity test pass/fail decisions are made relative to control sediment, but only with PELs when St. Ann's Harbour was the basis for biological toxicity test pass/fail decisions.

The proportion of tests failing the sediments (pass/fail decision made relative to control sediment) concurs with the probability that a sediment is highly toxic based upon mean PEL quotients.

6.2.2 Porewater and Sediment Chemistry. The issue of using porewater chemistry in addition to sediment chemistry has been raised. In this study, the bioaccumulation of porewater PAHs and porewater PCBs was more predictable in *M. nasuta* than sediment concentrations of PAHs and PCBs. Of the two strong dose responses observed over the organic contaminant gradient, one was well predicted by sediment-associated PAHs *(A. virginiana* survival) while the other was well predicted by porewater-associated PAHs (*E. estuarius* survival). The porewater biological fertilization tests using echinoderms did not respond in a predictable manner to either the porewater or sediment measured contaminants.

Automated model building methods suggest that porewater variables explain more of the observed variability in the biological toxicity test responses than sediment variables. However, this observation was not confirmed when modelling dose-responses using a hands-on approach.

The work conducted during the course of this study indicates that porewater variables are correlated with some biological test responses. An assessment of the additional benefits that the use of porewater contaminant measurements would bring to the ocean disposal program is beyond the scope of this project. However, it should be noted that other jurisdictions acknowledge the relevance of porewater contaminant concentrations in assessing sediment. The USEPA is developing sediment quality guidelines using an equilibrium partitioning approach (EqP) (USEPA 1992; 1999) for nonionic organic compounds and simultaneously extracted metals (SEM) for metals criteria (Ankley *et al.*, 1996). Both of

these methods are porewater methods in that they address porewater contaminant concentrations indirectly (EqP) or directly (SEM). The inclusion of porewater variables (in some form) would allow for comparison with other sediment assessment methods and may be of use in understanding the results of porewater tests.

6.2.3 Total PAHs versus Individual PAH Measurements. Individual PAHs were highly correlated with total PAHs. This is in part due to the fact that each PAH contributes to the measurement of PAH. However this observation has also been made in investigations of freshwater sediments contaminated with organic compounds (Moran *et al.*, 1997). Swartz (1999) and USEPA (1999) discuss the utility of total PAH measurements as a surrogate for individual PAH measurements. Based on this and other recent studies, the disposal at sea program may wish to continue to rely on a total PAH value for screening sediments.

6.3 Choice of Reference Stations

As a group, the biological toxicity tests do not pass or fail stations in the same way when the basis for comparison is a reference sediment. When the pass/fail decision is made relative to a control sediment, then the group of biological toxicity tests passes/fails sediments in the same way.

The lack of agreement in pass/fail status when using the group of biological toxicity test responses relative to reference sediment responses may be due to the effects of factors present at the reference sites. If some organisms exhibit a negative response in the reference sediments, and other organisms do not, the pass/fail status of sediments (relative to the reference stations) will not be homogeneous among the group of toxicity tests. The practical result in an ocean disposal context may be contradictions within the test battery.

Another consequence of organisms responding differently to reference sediments is the observed change in relative sensitivity of the biological test species. This observation is on the surface, unsettling. However, when the two species exhibiting reversals in sensitivity, *D. excentricus* and *L. pictus*, were examined, the extreme responses for these species were observed in the control sediments, St. Ann's Harbour or the Station 12 reference site. Thus these stations are acting as pivotal stations, when pass or fail decisions are being made. The choice of station (or pivot) affects the pass/fail decision.

The agreement in pass/fail status when using the group of biological toxicity responses relative to control sediment responses implies that the biological tests are responding adversely to either confounding variables or contaminant effects at the exposure sites.

The choice of reference station also affected the strength of the correlation between the proportion of sediment "failures" using biological toxicity tests and proportion of sediment "failures" using TELs or PELs. The strongest correlation between biological toxicity test pass/fail decisions and sediment chemistry pass/fail decisions generally occurred when the control sediment was used as a reference station. There may be some circularity in this observation as PELs and TELs are derived using the geometric mean from percentiles of a biological toxicity test effects and no effects database (EC, 1995b). The circularity of the observation arises as the definition of "effect" and "no effect" may be made relative to a reference or control sediment specific to each experiment. The determination of effect or no effect for each of the individual data sets comprising the effects/no effects database was made by the contributing author, and may have been made relative to a reference or control sediment.

A reference station should be chosen in the context of the experimental/study goals. For this study, the reference station was chosen to match known confounding factors, such as sediment grain sizes and station depths in Sydney harbour in order to explore issues arising when interpreting the pass/fail status of sediments using three assessment tools. Although most

potential confounding factors did vary statistically among stations, the confounding factors were not primarily responsible for predictable biological toxicity test responses. Thus, the observed differences in mean values of confounding factors were statistically significant, but not ecologically significant.

This observation implies that confining confounding factors to class intervals or ranges, as is done for grain size criteria in amphipod biological toxicity tests (see Table 6) is sufficient to ensure comparability of biological toxicity test results between exposure and reference stations. Class intervals or ranges for a given confounding factor within which a biological test would not significantly vary could be used to better select reference sites. The key confounding factors for which class size restrictions should be developed will likely be species-specific.

When choosing a reference site for an ocean disposal permit application, the USEPA (1991) suggests that the grain sizes at the reference station(s) be as similar as is practical to the grain sizes of the dredged material, and that the reference station reflects conditions prevailing at the disposal site, before disposal. This concurs in principle with the suggestion to only compare sediments when known confounding factors fall within the same class interval.

Although the reference site in this study was chosen with due diligence, it did not reflect the reference condition for the *in-situ* benthic macroinvertebrate community but did for the biological toxicity tests being evaluated (with the exception of *L. pictus*). Given that the goals of the study were to explore issues arising when interpreting the pass/fail status of sediments using three assessment tools, the choice of St. Ann's Harbour as a reference station should stand. However, the discrepancy between responses observed at different levels of biological organization (i.e., single organism toxicity tests and benthic macroinvertebrate community structure) suggests that critical elements in reference site selection were not identified in advance.

The two reference stations were chosen based on similarity of known confounding factors with exposure sites (see Section 2.1). Upon chemical analysis, the St. Ann's harbour station exceeded Cu and Pb TELs while the Station 12 reference site exceeded TELs for 2-Methyl-Naphthalene and As. Similarly, Porebski *et al.* (1998) found that the reference site for Belledune harbour exceeded at least one TEL. These two observations highlight the practical difficulties in choosing a reference station. It is unlikely that many reference stations could be found that did not exceed at least one TEL.

The importance of reference site selection should not be under-emphasized, especially in the case of ocean disposal permit application, as this is essentially an experiment with two "treatments", namely a reference sediment and a sediment being considered for ocean disposal. Good experimental design suggests equal experimental effort be expended upon both "treatments" in the absence of *a priori* knowledge. Thus an equal number of reference and exposure sites represents the best allocation of sampling effort. Note that multiple samples from <u>one</u> reference site or area, do not constitute replicates. These samples are best described as pseudoreplicates (Hurlbert, 1984), and have been decried as not being representative of natural variability (Underwood, 1991).

The use of multiple reference stations to determine potential contaminant effects is endorsed by the USEPA (1991) in an ocean disposal context and Environment Canada (1998b) in the context of assessing potential pulp and paper mill environmental effects.

Some problems do exist if multiple reference sites are used as a basis for comparing biological toxicity test results. These include restrictions on choice of reference site to avoid biasing a permitting application, the availability of one, let alone several suitable reference sites, the presence of unsuspected contaminants and the additional costs incurred.

6.4 Reference Stations versus Control Stations

The basis of comparison affected: 1) the pass/fail decision for a sediment; 2) the consistency of interpretation of the biological toxicity tests; and 3) the degree of correlation between biological toxicity test pass/fail assessments and sediment contaminants.

1) The choice of reference station affected the pass/fail decision rendered by a single biological test.

2) As a group, the biological toxicity tests pass or fail stations in the same way when the basis for comparison is a control sediment but not when the basis for comparison are reference sediments.

3) The biological pass/fail assessments concur with those made using sediment chemistry values when the biological pass/fail assessments are made relative to control sediment. When reference sediment is used as the basis for a making pass/fail decision with a biological test, the biological test pass/fail assessments are less strongly correlated with the sediment assessments made using contaminant criteria. This occurs as the proportion of biological tests failing the stations along the gradient does not decrease monotonically when the basis for biological pass/fail decisions are the reference stations. The number of failures triggered by TEL/PEL values does decrease monotonically along the gradient.

6.5 Comparison of Three Sediment Characterization Methods

Ordinations show that all three assessment tools group Stations 1, 5, and 6 together. Station 9 is separated from the other stations on the basis of numerical abundance of specific taxa, contrasts in biological toxicity test responses, and sediment grain size differences. The ordinations using biological toxicity test responses and sediment physical chemistry also separate St. Ann's Harbour from other stations.

This finding concurs with the observed correlation between the pass/fail characterizations using biological toxicity tests and TELs or PELs and the observed bioaccumulation of organic compounds in bivalves. The lack of correlation between the health of the *in-situ* benthic macroinvertebrate community and sediment contamination may be more a consequence of an effect at the St. Ann's Harbour reference station than a lack of effect, lower in the gradient. Figure 18 shows that the benthic macroinvertebrate community richness increases almost monotonically until the St. Ann's Harbour reference station is reached. The failure of the benthic community at this single station is enough to mask this trend when comparing ordinations and may in part be responsible for the lack of significance when using Mantel's test.

6.5.1 Interpretation Criteria. Sediment contaminant TELs and PELs do concur with biological toxicity test results, with PELs being more highly correlated with biological toxicity test responses than TELs (note that in an ocean disposal context a station does not fail on the basis of sediment chemistry alone). Empirically and statistically it can be seen that as a group, the current interim biological toxicity test criteria reflect sediment contamination in a meaningful way.

The current interpretation criterion for the bivalve bioaccumulation test is a statistically significant difference between reference or control sites and exposure sites. The pass/fail status of the test was determined, but should be interpreted cautiously due to a lack of replication. The test did respond well to the organic contaminant gradient. Models exist demonstrating the link between sediment organic contaminants (PCBs and DDTs) and bioaccumulation in *M. nasuta* (Boese *et al.*, 1997). Guidelines based upon risks to human health are available for some contaminants (USEPA/USACE, 1998). Health Canada or

other regulatory agencies such as those responsible for tissue consumption advisories may be able to suggest alternatives representing Canadian concerns. Alternatively, interpretation criteria could be based on a lack of significance, **and** ecologically relevant differences in bioaccumulation between reference site(s) and material being considered for ocean disposal. This follows the paradigm used for echinoid and amphipod pass/fail decisions.

6.5.2 Pass/Fail Decisions. Pass/fail decisions using echinoids and amphipods are implemented by applying two decision rules (EC, 1996):

1) A statistically significant difference must exist between the mean response at an exposure site and a mean response at a reference or control site.

2) A pre-specified, absolute decrease in response (the interim interpretation criteria) between the exposure and reference or control site must be exceeded.

The first decision rule does not provide an explicit probability value, but it is generally accepted that statistical significance occurs at an α value[4] of 5%. This α value is split into two halves because the alternative hypothesis being considered is one of **differences** between the mean response. Thus statistical significance may occur if the mean reference/control site response is much larger, **or** much smaller than the mean exposure site response. Since we are always looking at a decrease in response, our working α value is really only 2.5%. Then, when the second decision rule is sequentially applied, this working α value is not only not maintained, but is further reduced. The final pass/fail decision has a much lower α value than 5%.

[4] An α value of 5% means that there is a 5% chance of rejecting the null hypothesis when it should not be rejected. In the context of the decision rule, it would be said that there is a statistically significant difference between reference/control and exposure sites when there is really no difference.

The result of this decreased α value is that it is more difficult to fail a station than when the generally accepted α value of 5% is maintained. This procedure affords a lesser degree of protection to the environment than one where the generally accepted overall α value of 5% is maintained.

It is suggested that expert knowledge be combined in the form of the interim interpretation criteria and statistical objectivity when assessing pass/fail status of sediment by a redefinition of alternative hypotheses. This would maintain the overall pass/fail decision α value of 5%, and consequently the degree of protection afforded to the environment.

6.6 Further Work

The following suggestions are provided to improve the ocean disposal permit granting program by reducing the effects of confounding variables and improving the interpretation of the biological toxicity test battery results.

- Class sizes or tolerance ranges for known confounding variables should be established for all biological toxicity tests such that comparisons of sediments falling within a class are valid. The class size interval should correspond to the most restrictive comparison made within a given class or be species-specific. Also, the list of key confounding factors for which tolerance ranges are currently being developed should be reviewed for completeness. This list should remain species-specific. The establishment of allowable class sizes for valid comparison with, and selection of, reference stations follows the intent of the criteria for valid use of the amphipod biological toxicity tests (EC, 1998a) and should also be considered.

- Address the issue of multiple versus single reference sites and/or control sites as a basis for comparison. As described in Section 6.3, multiple reference sites provide a better understanding of the reference condition, but

may be difficult to find and will incur greater costs. The choice of reference sites changes the pass/fail decision (see Section 4.4.1 H4a: do all biological toxicity tests characterize the sample in the same way?) and therefore criteria for the selection of reference sites should be developed, possibly as previously described. Alternatively, control sediments could form the basis for making pass/fail decisions. This would obviate the problems with reference sites but may reduce the ecological relevance of the comparison.

- As porewater variables were often associated with the biological toxicity test responses, a series of spiked-sediment biological toxicity tests may be conducted to explore the relative strengths of the relationship between sediment and porewater-associated contaminants. This should be coupled with a literature review to explore the relative merits of porewater, sediment, sediment-water interface and possibly elutriate contaminant concentrations, as Phase 1 screening tools. This should also include a cost/benefit analysis of adding other measurements of sediment contamination to the Phase I screening tools.

- It is suggested that expert knowledge be combined in the form of the interim interpretation criteria and statistical objectivity when assessing pass/fail status of sediment by a redefinition of alternative hypotheses. This would maintain the overall pass/fail decision α value of 5%, and consequently the degree of protection afforded to the environment.

- Explore the consequences of comparing the change in bioluminescence of *V. fischeri* exposed to a given sediment relative to a control or reference sediment in the same way that amphipod, echinoderm, and other test responses are compared.

- Determine interpretation criteria for the bivalve bioaccumulation test as discussed in Section 6.5.1. This test responded well to the organic contaminant gradient but currently, only a statistically significant difference between reference or control sites and exposure sites is used to make pass/fail decisions.

- Assess the comparability of bioaccumulation tests using both *M. balthica* and *M. nasuta* to validate the substitution of a non-native species.

- Further examine the within-test variability of echinoids and also the relative responses to contaminants versus non-contaminants to further validate the use of this test as a regulatory test.

- Explore the use of robust methods when making pass/fail or other regulatory decisions with organisms such as *D. excentricus* that routinely produce outlying data points. Newer tests provide an alternative to traditional statistical tools such as Dunnett's test.

- A re-examination of the Polychaete species data may be warranted once the method development has been finalized.

References

Allen, H.E., G. Fu, and, B. Deng. 1993. Analysis of acid-volatile sulfide (AVS) and simultaneously extracted metals (SEM) for the estimation of potential toxicity in aquatic sediments. *Env. Tox. Chem., 12:*1441–1453.

Ankley, G. T., D. M. Di Toro, D. J. Hansen, and W. J. Berry. 1996. Technical basis and proposal for deriving sediment quality criteria for metals. *Env. Tox. Chem., 15:*2056–2066.

Arenicola Marine. 1997. Sample collection and analysis for benthic community structure: Sydney Harbour, NS, July 1997.

Atkinson, A. C. *Plots, transformations, and regressions.* Oxford Science Publications, Oxford, England, 282 pp.

BEAK Consultants Ltd. 1996. Development of interpretive guidance for biological tests using chemical gradient studies. Phase 1: Confirmation study. report prepared for Environment Canada, Marine Environment Division, Ottawa, ON.

Boese, Bruce L., Henry Lee II, and S. Echols. 1997. Evaluation of a first-order model for the prediction of the bioaccumulation of PVBs and DDTs from sediment into the marine deposit-feeding clam, *Macoma nasuta*. *Environmental Toxicology and Chemistry, 16:*1545–1552.

British Columbia Ministry of Environment, Lands and Parks. 1994. Sand dollar sperm inhibition bioassay manual. Unpublished report.

CCME (Canadian Council of Ministers of the Environment) 1999. Canadian Sediment Quality Guidelines. CCME, Winnipeg, Manitoba.

(CEPA) Canadian Environmental Protection Act. 1999. Statutes of Canada, Chapter 33. Federal Government of Canada, Ottawa, ON, Canada.

Chapman, P.M. 1992. Sediment Quality Triad Approach. In: USEPA, Sediment Classification methods compendium. 10-1-10-18, Office of Science and Technology, Washington, DC.

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

Cochran, W.G. 1950. The comparison of percentages in matched samples. *Biometrika, 37:*256–266.

Collett, D. 1991. *Modelling binary data*. Chapman and Hall, Toronto, Ontario, 369 pp.

Devitt, S. 1997. Field activity report for collection of marine sediments from Sydney Harbour. Prepared for Ms. L. Porebski, Environment Canada, Marine Environment Division, Ottawa, Ontario.

Edwards, D. and J. J. Berry. 1987. The efficiency of simulation-based multiple comparisons. *Biometrics, 43:*913–928.

EC (Environment Canada). 1992. Analytical methods for organochlorinated pesticides, polychlorinated biphenyls, polynuclear aromatic hydrocarbons, chlorinated benzenes, and organophosphorous pesticides in water (liquid-liquid extraction, GC determination), Atlantic Region Environmental Quality Laboratories, Moncton, NB.

———. 1992a. Biological test method: Acute test for sediment toxicity using marine or estuarine amphipods. Report EPS 1/RM/26. Environment Canada, Environmental Protection, Ottawa, Ontario.

———. 1992b. Biological test method: Fertilization assay using echinoids (sea urchin and sand dollars). Report EPS 1/RM/27,

Environment Canada, Conservation and Protection, Ottawa, Ontario.

———. 1992c. Biological test method: Toxicity testing using luminescent bacteria. Report EPS 1/RM/24, Environment Canada, Conservation and Protection, Ottawa, Ontario.

———. 1994. Interim sediment quality assessment values. Environment Canada, Ecosystem Conservation Directorate, Soil and Sediment Quality Section, Guidelines Division. Ottawa, Ontario.

———. 1995a. "Biological Test Method: Test for Growth and Survival in Sediment using Marine or Estuarine Polychaete worms", Environment Canada, Conservation and Protection, Ottawa, Ontario, September 1995. (Draft Manuscript).

———. 1995b. Protocol for the Derivation of Canadian Sediment Quality Guidelines for the Protection of Aquatic Life. Report CCME EPC-98E, Environment Canada, Ottawa, Ontario.

———. 1995c. Users guide to the application form for ocean disposal. Report EPS 1/MA/1, Environment Canada, Marine environment Division, Ottawa, Ontario.

———. 1996. 1996 National Compendium Monitoring at Ocean Disposal Sites. Environment Canada, Ottawa, Ontario.

———. 1998a. Reference Method for Determining Acute Lethality of Sediment to Marine or Estuarine Amphipods. Report EPS1/RM/35, prepared for Method Development and Application Section, Environmental Technology Advancement Center, Environment Canada, Ottawa.

———. 1997. Polynuclear aromatic hydrocarbons and total arochlors in sediments (GC/MSD), Atlantic Region Monitoring and Evaluation Branch, Moncton, NB.

———. 1998b. Pulp and Paper EEM Technical Guidance Document, Environment Canada, Ottawa, Ontario.

Fennell, M. and G. van Aggelen. 1997. Biological Assessment of Sediments Collected from Sydney Harbour, Nova Scotia: Site of the 1997 Pollution Gradient Study Conducted by Environment Canada's National Disposal at Sea Program. (Unpublished Report).

Finney, D.J. 1971. *Probit Analysis.* 3rd edition. Cambridge University Press, Cambridge, Massachusetts.

Hurlbert, S. H. 1971. The nonconcept of species diversity: a critique and alternate parameters. *Ecology, 52:*577–586.

Hurlbert, S.H. 1984. Pseudoreplication and the design of ecological field experiments. *Ecol. Monog., 54:*187–211.

Jackman, P. and K.G. Doe. 1997. Toxicity of Sediments from Sydney Harbour, N.S. - Results of the 1997 Pollution Gradient Study. (Unpublished Report), Environment Canada, Atlantic Region, Moncton, NB.

Keddy, C., J. C. Greene, and M. A. Bonnell. 1994. A review of whole organism bioassays for assessing the quality of soil, freshwater sediment, and freshwater in Canada. Scientific Series 198. pp125, Environment Canada, Ecosystem Conservation Directorate, Ottawa, Ontario.

Long, E.R., L. J. Field, and D. D. MacDonald. 1998. Predicting toxicity in marine sediments with numerical sediment quality guidelines. *Env. Tox. Chem., 17(4):*714–727.

Mantel, N. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Res., 27:*209–220.

Margalef, R. 1958. Information theory in ecology. *Gen. Syste., 3:*36–71.

Matheson, R. A. F., G. L. Trider, W. R. Ernst, K. G. Hamilton, and P. A. Hennigar. 1983. Investigation of polynuclear aromatic hydrocarbon contamination of Sydney Harbour, N.S., Environment Canada, Environmental Protection Service, Dartmouth, NS.

McIntosh, R.P. 1967. An index of diversity and the relation of certain concepts to diversity. *Ecology, 48:*392–404.

McCullagh, P. and J. A. Nelder. 1989. *Generalized linear models.* 511pp. Chapman and Hall, Toronto, Ontario.

Microbics Corporation. 1992. Microtox® Update Manual. Carlsbad, CA.

Moran, T., D. Adams, K. Walker, and B. Zajdlik. 1997. Lambton Industrial Society 1994/1995 St. Clair River Sediment Program Defining Spatial Extent and Environmental Conditions. prepared for the Lambton Industrial Society by Pollutech Enviroquatics Limited, Point Edward, Ontario.

Munawar, M., I. F. Munawar, P. Ross, and R. Dermott. 1992. Exploring aquatic ecosystem health: A multi-trophic and an ecosystemic approach. *J. Aquatic. Ecosys. Health, 1:*237–252.

Norberg-King, T.J.. 1993. A linear interpolation method for sublethal toxicity: The inhibition concentration (ICp) approach (Version 2.0). Tech. Report 03-93 of National Effluent Toxicity Assessment Center, July 1993. USEPA, Environ.Res. Lab.-Duluth, Duluth, Minnesota.

PESC (Pacific Environmental Science Centre). 1999a. Metals Analysis in Soils and Sediments. Version 5.0, Issue Date: September, 1999, PESC, North Vancouver, BC.

———. 1999b. Metals Analysis in Water. Version: 3.3, Issue Date: September 1, 1999, PESC, North Vancouver, BC.

———. 1999c. Mercury in Water Analysis. Version 3.7, Issue Date: September, 1999, PESC, North Vancouver, BC.

Porebski, L. M. , K. G. Doe, B. A. Zajdlik, D. Lee, P. Pocklington, G. Atkinson, and J. Osborne. 1998. Interpretive guidance for bioassays using pollution gradient studies – Belledune, New Brunswick. EPS Technical Report WM-20, Environment Canada, Marine Environment Division, Ottawa, Ontario.

Porebski, L. M., K. G. Doe, B. A. Zajdlik, D. Lee, P. Pocklington, and J. Osborne. 1999. Evaluating techniques for a tiered testing approach to dredged sediment assessment – A study over a metal concentration gradient. *Env. Tox. Chem., 18:*2600–2610.

Schiewe, M.H., E.G. Hawk, D.I. Actor and M.M. Krahn. 1985. Use of bacterial bioluminescence assay to assess toxicity of contaminated marine sediments. *Can. J. Fish. Aquat. Sci. 42:*1244–1248.

Shannon, C.E. and W. Weaver. 1949. *The Mathematical Theory of Communication.* The University of Illinois Press, Urbana, IL.

Sigg, L. 1987. Surface chemical aspects of the distribution and fate of metal ions in lakes. pp. 319–349, In: *Aquatic surface chemistry: Chemical processes at the particle-water interface.* Stumm, W. (ed.) , John Wiley and Sons, New York.

Simpson, E. H. 1949. Measurement of diversity. *Nature, 163:*688.

Stebbing, A. R. D., V. Dethlefsen, and M. Carr. 1992. Biological effects of contaminants in the North Sea. Results of the ICES/IOC Bremerhaven Workshop. Marine Ecology Progress Series, Volume 91.

Stephan, C.E. 1977. "Methods for calculating an LC50" , p. 65-84, In: *Aquatic Toxicology and Hazard Evaluation.* ASTM STP 634, Mayer, F.L. and J.L. Hamelink (eds.) American Society for Testing and Materials, Philadelphia, PA.

Swartz, R.C., 1999. Consensus sediment quality guidelines for polyaromatic hydrocarbon mixtures. *SETAC,18:*780–787.

Underwood, A.J. 1991. Beyond BACI: Experimental designs for detecting human environmental impacts on temporal variations in natural populations. *Australian Journal of Marine and Freshwater Research, 42:*569–587.

USEPA (United States Environmental Protection Agency), 1991. Evaluation of Dredged Material Proposed for Ocean Disposal Testing Manual, EPA 503/8-91/001. February 1991, USEPA, Washington, DC.

———. 1992. An SAB Report: Review of Sediment Criteria Development Methodology for Non-Ionic Organics. Sediment Quality Subcommittee of the Ecological Processes and Effects Committee. EPA-SAB-EPEC-93-002, USEPA, Washington, DC.

———. 1993. Guidance manual bedded sediment bioaccumulation tests. Report EPA/600/R-93/183. September 1993, USEPA, Washington, DC.

———. 1999. In press. Methods for the derivation of site-specific equilibrium partitioning sediment guidelines (ESGs) for the protection of benthic organisms. USEPA Washington, DC. In Press.

USEPA/USACE (United States Environmental Protection Agency/United States Army Corps of Engineers). 1998. Evaluation of dredged material proposed for discharge in waters of the U.S.—testing manual. EPA 823-B-98-004. February 1998.

Wilhm, J.L. 1970. Range of diversity index in benthic macroinvertebrate populations. *J. Wat. Pollut. Control Fed., 42:*221–224.

Field Data

Table A-1 Sample Location Coordinates

Sample Site Number	Preferred Sample Coordinates (1996)	Garman 75™ Sample Site Coordinates	Trimble™ Sample Site Coordinates
1	46 º 09' 09.5" 60 º 13' 28.0"	46 º 09' 09.7" 60 º 12' 26.8"	46 º 09' 09.1" 60 º 12' 27.1"
5	46 º 10' 12.1" 60 º 12' 18.8"	46 º 10' 10.3" 60 º 12' 13.8"	46 º 10' 12.5" 60 º 12' 14.6"
6	46 º 10' 32.2" 60 º 12' 26.5"	46 º 10' 31.1" 60 º 11' 26.2"	46 º 10' 34.0" 60 º 12' 25.0"
9	46º 12’ 24.0” 60º 13’ 14.4”	46 º 12' 25.9" 60 º 13' 15.2"	46 º 12' 25.8" 60 º 13' 17.3"
12	46 º 15' 07.0" 60 º 10' 23.5"	46 º 15' 03.6" 60 º 10' 30.3"	46 º 15' 03.8" 60 º 10' 30.0"
St. Ann’s Harbour Reference Site	Not Previously Used	46 º 15' 21.0" 60 º 33' 35.6"	46 º 15' 22.6" 60 º 33' 40.1"

Sample coordinates use North American Datum (1983). The column “preferred sample coordinates” refers to coordinates from the initial site screening survey undertaken in the fall of 1996.

Table A-2 Field Data

Project Number S970602

Location	Sydney Harbour		
Sample Number	Station #1	**Replicate Number**	A
Sample Identifier	See Random #	**Date:**	11-Jul-97
Sample Medium	Marine sediment	**Time:**	1045 am
Position			
Sample North	46 ° 09' 09.5"		
Sample West	60 ° 13' 28.0"	**Weather**	Sunny, westerly breeze 5 - 10 km
Garman **GPS North**	46 ° 09' 09.7"	**Tide**	Falling
GPS West	60 ° 12' 26.8"		
Trimble **Boat North**	46 ° 09' 09.1"	**Personnel Present**	SD, EG, LP, PT,
Boat West	60 ° 12' 27.1"		AM, JS, FP

Water Quality

Parameter	**Units**	**Depth (m)** Surf	3	6	9	12	Bottom 14.3
Temperature	°C	13.8	12.2	11.1	8.7	8.3	7.4
Conductivity	µS/cm	39.8	39.7	40.6	40.9	40.9	41.5
Salinity	‰	25.2	25.5	25.5	26.5	26.5	26.6
pH		7.96	8.00	7.98	7.94	7.91	7.85
Dissolved Oxygen	mg/L	8.85	9.66	9.4	9.69	9.42	9.63
ORP	mv	381	380	379	380	381	382

Sediment Quality

Parameter			
Temperature	°C	7	
ORP	mV	-7	
pH		7.1	

Sediment Description: Black anoxic mud, soupy in consistancy, some gravel and wood material in sediment.

Odour: Metallic smell

Benthic Invertebrates: Few visible

Other Observations:

Project Number S970602

Location Sydney Harbour

Sample Number Station #6 **Replicate Number** A

Sample Identifier See Random # **Date:** 10-Jul-97

Sample Medium Marine sediment **Time:** 1740 pm

Position

Sample North 46 ° 10' 32.2"

Sample West 60 ° 12' 26.5" **Weather** Overcast, foggy

Garman **GPS North** 46 ° 10' 31.1" **Tide** falling

GPS West 60 ° 11' 26.2"

Trimble **Boat North** **Personnel Present** SD, EG, LP. PT

Boat West AM, JS, FP

Water Quality

Parameter	Units	Depth (m)					
		Surf	3	6	9	12	Bottom 18
Temperature	°C	12.9	12.8	10.2	9.4	8.2	6.9
Conductivity	µS/cm	39.7	39.4	40.6	40.6	41.1	41.9
Salinity	‰	25.1	25.2	26.1	26.2	26.4	26.8
pH		8.02	8.02	7.98	7.99	7.94	7.89
Dissolved Oxygen	mg/L	9.84	9.35	9.35	9.55	9.18	9
ORP	mv	333	338	347	349	353	345

Sediment Quality

Parameter			
Temperature	°C	5	
ORP	mV	-5	
pH		7.15	

Sediment Description: Black mud with a sand/ grit base

Odour Mild H_2S

Benthic Invertebrates Polycheates

Other Observations

Project Number S970602

Location Sydney Harbour

Sample Number Station #9 **Replicate Number** A

Sample Identifier See Random # **Date:** 10-Jul-97

Sample Medium Marine sediment **Time:** 1430 pm

Position
Sample North 46 ° 12' 24.0"
Sample West 60 ° 13' 14.4" **Weather** Overcast, rain squalls

Garman **GPS North** 46 ° 12' 25.9" **Tide** High
GPS West 60 ° 13' 15.2"

Trimble **Boat North** **Personnel Present** SD, EG, LP. PT
Boat West AM, JS, FP

Water Quality

Parameter	**Units**	**Depth (m)** Surf	3	6	9	12	Bottom 10.5
Temperature	°C	15.8	14.9	12.7	10.9		10.6
Conductivity	µS/cm	38.2	39.5	39.8	40.3		41.2
Salinity	‰	24.9	24.7	26	26.4		26.4
pH		8.04	8.04	8.02	8.00		7.99
Dissolved Oxygen	mg/L	9.11	8.88	9.22	9.48		9.75
ORP	mv	334	340	348	352		355

Sediment Quality

Parameter			
Temperature	°C	10	
ORP	mV	-28	
pH		7.5	

Sediment Description: Black mud with some sand, soupy.

Odour H_2S odour
Benthic Invertebrates
Other Observations

Project Number S970602

Location Sydney Harbour

Sample Number Station #12 **Replicate Number** A

Sample Identifier See Random # **Date:** 11-Jul-97

Sample Medium Marine sediment **Time:** 1356 pm

Position

Sample North 46 ° 15' 07.0"

Sample West 60 ° 10' 23.5" **Weather** sunny, westerly breeze

Garman **GPS North** 46 ° 15' 03.6" **Tide** Falling

GPS West 60 ° 10' 30.3"

Trimble **Boat North** 46 ° 15' 03.8" **Personnel Present** SD, EG, LP, PT

Boat West 60 ° 10' 30.0" AM, JS, FP

Water Quality

Parameter	Units	Depth (m)					
		Surf	3	6	9	12	Bottom 16.9
Temperature	°C	14.6	14.5	14.2	12.5	11.6	9.39
Conductivity	µS/cm	39.9	39.9	40.1	40.7	41.1	41.5
Salinity	‰	25.4	25.5	25.6	26.1	26.4	26.7
pH		8.00	8.02	8.00	8.00	7.99	7.98
Dissolved Oxygen	mg/L	9.28	9.02	8.92	9.17	9.22	9.6
ORP	mv	370	370	370	372	373	374

Sediment Quality

Parameter			
Temperature	°C	9.00	
ORP	mV	-18	
pH		7.60	

Sediment Description: Fine to medium grained sand, brown.

Odour None

Benthic Invertebrates Wide variety of polycheates

Other Observations

Project Number S970602

Location St. Annes Bay

Sample Number Reference **Replicate Number** A

Sample Identifier See Random # **Date:** 12-Jul-97

Sample Medium Marine sediment **Time:** 1030 am

Position
Sample North
Sample West **Weather** Sunny, westerly wind 10 - 15 km

Garman **GPS North** 46 ° 15' 21.0" **Tide** Falling
GPS West 60 ° 33' 35.6"

Trimble **Boat North** **Personnel Present** SD, EG, LP, PT
Boat West AM, JS, FP

Water Quality

Parameter	Units	Depth (m) Surf	3	6	9	12	Bottom 15.0
Temperature	°C	14.2	14.0	13.5	10.0	9.3	9.1
Conductivity	µS/cm	37.9	39.2	38.4	39.2	39.4	39.6
Salinity	‰	24.5	25	24.9	25.1	25.4	25.6
pH		7.92	7.94	7.95	7.94	7.93	7.92
Dissolved Oxygen	mg/L	9.09	9.25	9.33	9.96	10.13	10.03
ORP	mv	394	392	391	392	393	398

Sediment Quality

Parameter			
Temperature	°C	7	
ORP	mV	-22	
pH		7.46	

Sediment Description: Black, grey, smooth silt and clay

Odour None
Benthic Invertebrates Few, polycheates.
Other Observations

Table A-3 Blind Sample Numbers

Station	Polychaetes	Bioaccumulation	Replicate 1	Replicate 2	Replicate 3	Replicate 4	Replicate 5
Number		Pore water,		metals, PAH/PCB (SED), TOC, particle size			benthos
		(PAH/PCBS		ammonia/sulphide, archives, amphipod, Microtox, echinoid, benthos			
		metals), AVS					
1	**66**	46	23	80	75	8	5
5	57	31	34	71	40	65	56
9	92	96	87	2	84	88	70
6	36	9	73	100	11	77	53
Ref 1 (old stn 12)	**76**	35	6	33	48	94	90
St Ann	**68**	72	15	38	24	49	89

Raw Data for Toxicity Tests and Bioaccumulation

East Coast Amphipod Biological Toxicity Test Results

Table B-1 Results of 10-Day Test with *Amphiporeia virginiana*

Sample ID	Sample No.	Percent Survival					Mean Percent Survival	SD
		Rep #1	Rep #2	Rep #3	Rep #4	Rep #5		
19R1- Martinique	97AT001351	90	100	80	80	55	81	16.73
19R2- Martinique	97AT001351	95	70	85	85	90	85	9.35
19R3- Martinique	97AT001351	90	90	75	85	80	84	6.52
2	97AT001321	65	90	85	55	65	72	14.83
6	97AT001322	60	85	85	75	70	75	10.61
11	97AT001323	50	45	55	55	35	48	8.37
15	97AT001324	90	60	70	55	70	69	13.42
23	97AT001325	0	0	5	0	15	4	6.52
24	97AT001326	75	90	85	75	75	80	7.07
33	97AT001327	90	70	75	70	75	76	8.22
34	97AT001328	60	55	60	60	60	59	2.24
38	97AT001329	85	90	70	80	85	82	7.58
40	97AT001330	55	40	30	60	40	45	12.25
48	97AT001331	90	85	90	95	75	87	7.58
71	97AT001332	65	50	65	40	40	52	12.55
73	97AT001333	45	70	60	60	55	58	9.08
75	97AT001334	5	0	0	0	0	1	2.24
80	97AT001335	5	0	0	15	0	4	6.52
84	97AT001336	60	85	85	70	90	78	12.55
87	97AT001337	80	70	65	70	75	72	5.70
100	97AT001338	50	60	50	55	50	53	4.47

Table B-2 Results of 10-Day Test with *Rhepoxynius abronius*

Sample ID	Sample No.	Percent Survival					Mean Percent Survival	SD
		Rep #1	Rep #2	Rep #3	Rep #4	Rep #5		
1R1- Whidby	97AT001351	100	100	95	100	100	99	2.24
1R2- Whidby	97AT001351	100	100	100	100	100	100	0.00
1R3- Whidby	97AT001351	100	95	100	100	100	99	2.24
2	97AT001321	90	100	75	75	75	83	11.51
6*	97AT001322	92.5	92.5	95	100	100	96	3.79
11	97AT001323	95	75	70	70	80	78	10.37
15	97AT001324	85	100	100	95	80	92	9.08
23	97AT001325	95	85	95	70	95	88	10.95
24	97AT001326	95	85	85	100	85	90	7.07
33	97AT001327	90	95	90	95	95	93	2.74
34	97AT001328	70	70	70	60	60	66	5.48
38	97AT001329	95	100	100	95	85	95	6.12
40	97AT001330	65	70	80	60	90	73	12.04
48	97AT001331	95	95	100	85	100	95	6.12
71	97AT001332	95	70	80	60	90	79	14.32
73	97AT001333	70	80	75	70	70	73	4.47
75	97AT001334	100	80	80	70	90	84	11.40
80	97AT001335	90	95	80	80	85	86	6.52
84	97AT001336	75	70	70	75	65	71	4.18
87	97AT001337	65	95	85	85	95	85	12.25
100	97AT001338	100	90	95	95	90	94	4.18

*40 animals were added per jar to all 5 replicates by mistake.

Table B-3 Survival Results for the 28-Day Bioaccumulation Test with *Macoma nasuta*

Sample ID	Sample No.	Total tissue weight (g)	Percent Survival			Mean Percent Survival/ Rep	SD/ Rep	Mean Percent Survival/ Sample
			Rep #1	Rep #2	Rep #3			
INITIAL-1		13.12						
INITIAL-2		14.05						
INITIAL-3		12.14						
9-1	97AT001339	14.80	100	100	100	100.00	0.00	82.23
9-2	97AT001339	14.78	100	66.7	100	88.90	19.23	
9-3	97AT001339	9.69	100	66.7	66.7	77.80	19.23	
9-4	97AT001339	10.46	100	100	66.7	88.90	19.23	
9-5	97AT001339	8.50	66.7	0	100	55.57	50.92	
31-1	97AT001340	12.54	100	100	66.7	88.90	19.23	97.78
31-2	97AT001340	17.20	100	100	100	100.00	0.00	
31-3	97AT001340	15.18	100	100	100	100.00	0.00	
31-4	97AT001340	15.90	100	100	100	100.00	0.00	
31-5	97AT001340	13.96	100	100	100	100.00	0.00	
35-1	97AT001341	7.00	0	66.7	66.7	44.47	38.51	64.45
35-2	97AT001341	9.74	100	0	100	66.67	57.74	
35-3	97AT001341	15.01	100	66.7	100	88.90	19.23	
35-4	97AT001341	3.51	0	0	100	33.33	57.74	
35-5	97AT001341	12.25	100	100	66.7	88.90	19.23	
39-1 Control	97AT001345	12.71	100	100	100	100.00	0.00	88.89
39-2 Control	97AT001345	12.55	100	66.7	100	88.90	19.23	
39-3 Control	97AT001345	13.21	100	66.7	100	88.90	19.23	
39-4 Control	97AT001345	15.64	100	100	100	100.00	0.00	
39-5 Control	97AT001345	10.32	66.7	33.3	100	66.67	33.35	
46-1	97AT001342	11.57	100	100	100	100.00	0.00	85.56
46-2	97AT001342	13.23	100	100	100	100.00	0.00	
46-3	97AT001342	8.90	50	100	100	83.33	28.87	
46-4	97AT001342	7.63	100	100	0	66.67	57.74	
46-5	97AT001342	10.09	100	66.7	66.7	77.80	19.23	
72-1	97AT001343	9.28	100	100	66.7	88.90	19.23	86.67
72-2	97AT001343	9.17	33.3	100	66.7	66.67	33.35	
72-3	97AT001343	11.19	100	66.7	100	88.90	19.23	
72-4	97AT001343	13.98	100	100	100	100.00	0.00	
72-5	97AT001343	12.78	100	66.7	100	88.90	19.23	
96-1	97AT001344	3.14	0	0	66.7	22.23	38.51	28.89
96-2	97AT001344	9.85	33.3	66.7	100	66.67	33.35	
96-3	97AT001344	3.24	0	66.7	0	22.23	38.51	
96-4	97AT001344	4.36	100	0	0	33.33	57.74	
96-5	97AT001344	0	0	0	0	0.00	0.00	

West Coast Amphipod Biological Toxicity Test Results –
Table B-4 Results of 10-day Sediment Assays using *Eohaustorius estuarius*

12 - 22 Aug 1997								
		Replicates						
Treatment		**A**	**B**	**C**	**D**	**E**	**Mean**	**SD**
#29 - Control - Pachena Bay								
	% survival	100	100	100	100	100	100	0.0
	% at surface	0	0	0	0	0	0	0.0
#29 - Control - Pachena Bay								
	% survival	100	100	100	100	100	100	0.0
	% at surface	0	0	0	0	0	0	0.0
#29 - Control - Pachena Bay								
	% survival	100	100	100	100	100	100	0.0
	% at surface	0	0	0	0	0	0	0.0
			Combined Mean for Control Survival				*100*	*0.0*
#2	% survival	100	100	100	95	100	99	2.2
	% at surface	0	0	0	0	0	0	0.0
#6	% survival	100	100	95	90	95	96	4.2
	% at surface	0	0	0	0	0	0	0.0
#11	% survival	90	90	85	85	95	89	4.2
	% at surface	0	10	0	0	0	2	4.5
#15	% survival	100	100	100	100	95	99	2.2
	% at surface	20	10	10	10	15	13	4.5
#23	% survival	80	65	60	65	65	67	7.6
	% at surface	0	5	5	0	5	3	2.7
#24	% survival	95	90	90	95	100	94	4.2
	% at surface	0	0	0	0	0	0	0.0
#33	% survival	100	100	100	95	100	99	2.2
	% at surface	0	0	0	0	0	0	0.0
#34	% survival	90	90	100	85	85	90	6.1
	% at surface	0	0	5	0	0	1	2.2
#38	% survival	100	100	100	100	85	97	6.7
	% at surface	0	0	0	0	0	0	0.0
#40	% survival	80	90	85	95	80	86	6.5
	% at surface	0	0	0	10	10	4	5.5
#48	% survival	95	100	100	90	100	97	4.5
	% at surface	0	0	0	0	0	0	0.0
#71	% survival	70	90	70	85	90	81	10.2
	% at surface	0	0	0	0	0	0	0.0
#73	% survival	90	90	100	70	65	83	14.8
	% at surface	0	10	5	0	0	3	4.5
#75	% survival	75	60	60	75	35	61	16.4
	% at surface	30	10	10	0	0	10	12.2
#80	% survival	100	55	45	70	70	68	20.8
	% at surface	0	0	10	10	0	4	5.5
#84	% survival	95	90	100	100	95	96	4.2
	% at surface	0	0	0	0	0	0	0.0
#87	% survival	85	100	85	100	100	94	8.2
	% at surface	0	0	0	0	0	0	0.0
#100	% survival	90	90	75	80	60	79	12.4
	% at surface	0	0	5	0	0	1	2.2
#80 frozen	% survival	50	75	45	50	65	57	12.5
	% at surface	25	20	15	5	10	15	7.9

Table B-5 Results of 10-day Sediment Assays using *Eohaustorius washingtonianus*

29 Jul - 08 Aug 1997									
					Replicates				
Treatment			**A**	**B**	**C**	**D**	**E**	**Mean**	**SD**
#14 - Control - Esquimalt Lagoon									
	% survival		100	95	100	85	95	95	6.1
	% at surface		0	0	0	0	0	0	0.0
#14 - Control - Esquimalt Lagoon									
	% survival		95	95	100	95	100	97	2.7
	% at surface		0	0	0	0	0	0	0.0
#14 - Control - Esquimalt Lagoon									
	% survival		100	95	100	100	100	99	2.2
	% at surface		0	0	0	0	0	0	0.0
				Combined Mean for Control Survival				*97*	*4.1*
#2	% survival		95	95	85	85	95	91	5.5
	% at surface		0	10	0	0	0	2	4.5
#6	% survival		100	80	85	95	100	92	9.1
	% at surface		0	0	0	0	0	0	0.0
#11	% survival		65	35	45	70	85	60	20.0
	% at surface		0	5	5	5	25	8	9.7
#15	% survival		85	80	65	80	75	77	7.6
	% at surface		20	10	10	10	15	13	4.5
#23	% survival		80	55	50	40	40	53	16.4
	% at surface		10	10	5	0	0	5	5.0
#24	% survival		75	85	95	95	75	85	10.0
	% at surface		0	10	10	0	5	5	5.0
#33	% survival		100	85	95	85	100	93	7.6
	% at surface		0	0	0	0	0	0	0.0
#34	% survival		55	70	80	65	50	64	11.9
	% at surface		5	10	5	0	0	4	4.2
#38	% survival		75	85	85	90	90	85	6.1
	% at surface		0	0	0	0	15	3	6.7
#40	% survival		45	65	65	65	55	59	8.9
	% at surface		0	15	5	5	0	5	6.1
#48	% survival		100	95	100	100	100	99	2.2
	% at surface		0	0	0	0	0	0	0.0
#71	% survival		65	70	75	85	75	74	7.4
	% at surface		5	0	5	5	15	6	5.5
#73	% survival		75	45	75	75	10	56	28.8
	% at surface		0	0	5	10	0	3	4.5
#75	% survival		45	50	55	60	60	54	6.5
	% at surface		5	0	0	5	10	4	4.2
#80**	% survival		0	55	30	25	70	36	27.2
	% at surface		0	0	5	0	5	2	2.7
#84	% survival		95	85	85	95	90	90	5.0
	% at surface		5	0	10	0	5	4	4.2
#87	% survival		95	90	75	85	75	84	8.9
	% at surface		5	5	0	0	0	2	2.7
#100	% survival		45	45	50	55	70	53	10.4
	% at surface		0	0	0	0	15	3	6.7
#80 frozen	% survival		45	25	25	35	35	33	8.4
	% at surface		0	0	5	0	5	2	2.7
** sample seemed to immobilize amphipods; some difficulty telling live from dead									

East Coast Sea Urchin Biological Toxicity Test Results-

Table B-6 Results of Fertilization Inhibition Test with *Lytechinus pictus*

Sample #	Sample ID	IC50 (95% Confidence limits) (%)	IC25 (95% Confidence limits) (%)
97AT001321	2	78.5 (63.7–97.1)	52.9 (43.4–65.8)
97AT001322	6	> 100	>100
97AT001323	11	> 100	> 100
97AT001324	15	7.07 (6.01–9.56)	3.54 (2.99–4.82)
97AT001325	23	6.04 (5.66–6.72)	3.03 (2.83–3.35)
97AT001326	24	35.2 (30.6–41.1)	15.1 (8.67–20.5)
97AT001327	33	> 100	> 100
97AT001328	34	> 100	> 100
97AT001329	38	8.97 (7.51–10.5)	4.46 (3.67–5.24)
97AT001330	40	57.7 (45.7–76.8)	6.96 (5.72–8.75)
97AT001331	48	26.6 (-6.0–55.0)	5.09 (4.06–6.44)
97AT001332	71	6.17 (5.70–6.67)	3.08 (2.87–3.33)
97AT001333	73	> 100	> 100
97AT001334	75	> 100	> 100
97AT001335	80	> 100	> 100
97AT001336	84	7.36 (7.07–7.72)	3.68 (3.55–3.83)
97AT001337	87	40.6 (33.4–47.7)	9.82 (7.06–17.4)
97AT001338	100	28.2 (No CL)	7.92 (6.12–10.2)

West Coast Sea Urchin Biological Toxicity Test Results-

Table B-7 Echinoid Fertilization Inhibition Test using the Eccentric Sand Dollar Test

Date: Aug 7/97							
Site	**Sample Number**	**Percent Fertilization in Replicates**			**Mean**	**Percent Fertilization**	**SD**
		at 100% Concentration				**after Abbott's**	
		A	**B**	**C**		**Correction**	
Control	_____	94	97	91	94	100	3.00
2	974418-1	4	4	3	4	4	0.58
6	974418-2	3	0	0	1	1	1.73
11	974418-3	70	75	59	68	72	8.19
15	974418-4	97	93	96	95	100	2.08
23	974418-5	76	84	78	79	84	4.16
24	974418-6	95	97	94	95	101	1.53
33	974418-7	0	2	1	1	1	1.00
34	974418-8	30	41	32	34	37	5.86
38	974418-9	97	94	93	95	100	2.08
Control	_____	96	93	92	94	100	2.08
40	974418-10	29	28	23	27	4	3.21
48	974418-11	12	15	9	12	1	3.00
71	974418-12	82	86	89	86	72	3.51
73	974418-13	25	28	24	26	101	2.08
75	974418-14	28	40	31	33	84	6.24
80	974418-15	87	91	80	86	101	5.57
84	974418-16	23	22	25	23	1	1.53
87	974418-17	0	2	1	1	37	1.00
100	974418-18	88	80	74	81	101	7.02
80	974418-26	88	85	88	87	0	1.73
(frozen)							

East Coast Polychaete Biological Toxicity Test Results-

Table B-8 Results of 14-Day Test with *Boccardia proboscidea*

Sample ID	Sample Number	Wt/worm (mg)	Mean Wt/Treat. (mg)	SD	Survival (%)	Mean Percent Survival/ treat.	SD
Initial		0.626	0.64	0.0857	N/A	N/A	N/A
		0.558			N/A	N/A	N/A
		0.615			N/A	N/A	N/A
		0.760			N/A	N/A	N/A
4	97AT001353	1.233	1.401	0.231	100	93.75	12.50
Whitty's		1.176			100		
Beach		1.647			75		
		1.546			100		
35	97AT001322	1.826	1.409	0.344	100	87.5	14.43
		0.988			75		
		1.361			100		
		1.460			75		
66	97AT001348	1.319	1.551	0.370	100	100	0
		2.091			100		
		1.301			100		
		1.491			100		
68	97AT001349	1.694	1.583	0.247	100	93.75	12.5
		1.558			100		
		1.251			75		
		1.828			100		
92	97AT001337	1.530	1.557	0.161	100	87.5	14.43
		1.792			75		
		1.441			100		
		1.465			75		

Table B-9 Results of 14-Day Test with *Polydora cornuta*

Sample ID	Sample Number	Wt/worm (mg)	Mean Wt/Treat. (mg)	SD	Survival (%)	Mean Percent Survival/ treat.	SD
Initial		0.109	0.14	0.104	N/A	N/A	N/A
		0.088			N/A	N/A	N/A
		0.09			N/A	N/A	N/A
		0.326			N/A	N/A	N/A
		0.088			N/A	N/A	N/A
6	97AT001322	0.875	0.952	0.336	80	72	30.33
		1.541			80		
		0.85			100		
		0.788			80		
		0.705			20		
32	97AT001352	1.9225	2.113	0.829	100	92	10.95
Conrad's		3.1925			80		
Beach		2.63			100		
		1.029			100		
		1.789			80		
36	97AT001346	1.278	0.953	0.194	100	84	16.73
		0.868			80		
		0.857			60		
		0.975			80		
		0.788			100		
57	97AT001347	1.299	1.026	0.171	100	96	8.94
		0.922			100		
		1.038			80		
		1.02			100		
		0.849			100		
66	97AT001348	0.41	0.356	0.1	80	80	20
		1.85			60		
		0.42			100		
		0.348			100		
		0.415			60		
68	97AT001349	1.288	2.241	1.609	80	96	8.94
		1.08			100		
		2.339			100		
		1.507			100		
		4.989			100		
92	97AT001337	0.5008	0.466	0.314	100	44	35.78
		0.898			40		
		0.225			40		
		0.243			40		
		N/A			0		

West Coast Polychaete Analyses

Table B-10 *Boccardia proboscidea* in 14-Day Survival and Growth Tests

Test period: 7 - 21 August 1997								
Test organism age: 20 - 23d old, collected 15 - 18 July 1997								
		Replicates					**Mean**	**SD**
Treatment	**parameter**	**A**	**B**	**C**	**D**	**E**		
#4 - Control	% survival	100	100	100	40	60	80	28
	ttl 14d biomass (mg)	3.45	3.77	3.01	0.96	1.91	2.62	1.16
	dry wt/worm (mg)	0.69	0.75	0.60	0.48	0.64	0.63	0.10
	growth rate (mg/d)	0.043	0.047	0.036	0.028	0.039	0.038	0.007
	growth increase	7.3	8.0	6.4	5.1	6.8	6.7	1.1
	initial wt (mg/worm)	0.094						
#6 (Stn. 12)	% survival	80	60	60	40	20	52	23
	ttl 14d biomass (mg)	1.91	2.33	1.38	0.15	0.05	1.16	1.03
	dry wt/worm (mg)	0.48	0.78	0.46	0.08	0.05	0.37	0.31
	growth rate (mg/d)	0.027	0.049	0.026	-0.001	-0.003	0.020	0.022
	growth increase	5.1	8.2	4.9	0.8	0.5	3.9	3.2
	initial wt (mg/worm)	0.094						
#66 (Stn. 1)	% survival	40	60	60	40	60	52	11
	ttl 14d biomass (mg)	2.26	1.02	1.13	2.12	1.26	1.56	0.59
	dry wt/worm (mg)	1.13	0.34	0.38	1.06	0.42	0.67	0.39
	growth rate (mg/d)	0.074	0.018	0.020	0.069	0.023	0.041	0.028
	growth increase	12.0	3.6	4.0	11.3	4.5	7.1	4.2
	initial wt (mg/worm)	0.094						
#68 (St. Ann's)	% survival	0	40	60	0	100	40	42
	ttl 14d biomass (mg)	n/a	0.99	1.72	n/a	2.06	1.59	0.55
	dry wt/worm (mg)	n/a	0.50	0.57	n/a	0.41	0.49	0.08
	growth rate (mg/d)	n/a	0.029	0.034	n/a	0.023	0.029	0.006
	growth increase	n/a	5.3	6.1	n/a	4.4	5.2	0.9
	initial wt (mg/worm)	0.094						

Table B-11 ***Polydora cornuta*** **in 14-Day Survival and Growth Tests**

Test period: 7 - 21 August 1997								
Test organism age: 20-d old, collected 18 July 1997								
		Replicates					**Mean**	**SD**
Treatment	**parameter**	**A**	**B**	**C**	**D**	**E**		
#32 - Control	Percent survival	100	100	100	100	100	100	0
	Total 14-d biomass (mg)	6.63	13.09	8.30	8.45	5.84	8.46	2.81
	dry wt/worm (mg)	1.33	2.62	1.66	1.69	1.17	1.69	0.56
	growth rate (mg/d)	0.090	0.182	0.114	0.116	0.078	0.116	0.040
	growth increase	19.1	37.7	23.9	24.4	16.8	24.4	8.1
	initial wt (mg/worm)	0.069						
#6 (Stn. 12)	Percent survival	100	100	100	100	100	100	0
	total 14-d biomass (mg)	4.97	7.01	7.80	7.74	6.15	6.73	1.19
	dry wt/worm (mg)	0.99	1.40	1.56	1.55	1.23	1.35	0.24
	growth rate (mg/d)	0.066	0.095	0.106	0.106	0.083	0.091	0.017
	growth increase	14.3	20.2	22.5	22.3	17.7	19.4	3.4
	initial wt (mg/worm)	0.069						
#66 (Stn. 1)	Percent survival	40	80	100	60	100	76	26
	total 14-d biomass (mg)	0.11	1.52	1.21	1.42	1.27	1.11	0.57
	dry wt/worm (mg)	0.06	0.38	0.24	0.47	0.25	0.28	0.16
	growth rate (mg/d)	-0.001	0.022	0.012	0.029	0.013	0.015	0.011
	growth increase	0.8	5.5	3.5	6.8	3.7	4.0	2.3
	initial wt (mg/worm)	0.069						
#68 (St. Ann's)	Percent survival	100	100	100	80	80	92	11
	total 14-d biomass (mg)	2.87	6.87	4.47	5.53	1.63	4.27	2.08
	dry wt/worm (mg)	0.57	1.37	0.89	1.38	0.41	0.93	0.45
	growth rate (mg/d)	0.036	0.093	0.059	0.094	0.024	0.061	0.032
	growth increase	8.3	19.8	12.9	19.9	5.9	13.3	6.5
	initial wt (mg/worm)	0.069						

Bacterial Photoluminescence Results

Table B-12 Bacterial Photoluminscence Results

Site	Sample number	IC50 Wet not moisture corrected (%)	IC50 Dry moisture corrected (%)
2	974418-1	0.151	0.11
6	974418-2	1.07	0.83
11	974418-3	0.0256	0.0083
15	974418-4	0.690	0.23
23	974418-5	0.0230	0.0081
24	974418-6	0.410	0.14
33	974418-7	1.73	1.33
34	974418-8	0.0221	0.0081
38	974418-9	0.462	0.15
40	974418-10	0.0353	0.0140
48	974418-11	2.32	1.80
71	974418-12	0.0349	0.0147
73	974418-13	0.0496	0.0224
75	974418-14	0.0296	0.0107
80	974418-15	0.0255	0.0103
84	974418-16	0.153	0.11
87	974418-17	0.115	0.0828
100	974418-18	0.0291	0.0127
80 (frozen)	974418-26	0.0224	0.0088
HS-6 REF	_____	0.0278	_____

Bioaccumulation Results for *Macoma nasuta*

Mean metal concentrations are presented in Table 3 Summary of *Macoma nasuta* Mean Tissue Metal Levels (μg/g dry weight).

Table B-13 Summary of Total PAH Tissue Levels (ng/g dry tissue) in *Macoma nasuta*

Site	Subsample	Response
control	1	118.47
control	2	127.97
control	3	135.72
control	4	109.59
control	5	114.37
1	1	9223.23
1	2	9135.14
1	3	6538.85
1	4	6423.35
1	5	9037.66
5	1	7050.04
5	2	6677.48
5	3	6656.63
5	4	4812.54
5	5	7063.65
6	1	4448.33
6	2	4641.10
6	3	4841.02
6	4	3239.62
6	5	4606.60
9	1	401.91
9	2	288.13
9	3	396.18
9	4	333.96
12	1	194.10
12	2	147.72
12	3	143.26
12	4	283.13
12	5	147.32
St. Ann’s Harbour	1	199.40
St. Ann’s Harbour	2	209.29
St. Ann’s Harbour	3	226.08
St. Ann’s Harbour	4	162.48
St. Ann’s Harbour	5	148.74

Table B-14 Summary of Total PCB Tissue Levels (ng/g dry tissue) in *Macoma nasuta*

Site	Subsample	Response
control	1	0.48
control	2	0.48
control	3	0.48
control	4	0.48
control	5	0.48
1	1	159.5971619
1	2	133.5544726
1	3	95.96944106
1	4	84.19546897
1	5	134.1281444
5	1	434
5	2	385
5	3	412
5	4	371
5	5	443
6	1	233
6	2	240
6	3	272
6	4	217
6	5	175
9	1	0.48
9	2	0.48
9	3	0.48
9	4	0.48
12	1	19
12	2	17
12	3	15
12	4	0.48
12	5	17
St. Ann’s Harbour	1	0.48
St. Ann’s Harbour	2	0.48
St. Ann’s Harbour	3	0.48
St. Ann’s Harbour	4	0.48
St. Ann’s Harbour	5	0.48

Benthic Macroinvertebrate Community Raw Data-Table B-15 Mean Number of Benthic Macroinvertebrate Taxa

	Station					
Site	1	5	6	9	12	St. Ann's Harbour
Taxon						
Aricidea suecica	0	0	0.2	0	0.8	0
Capitella capitata	0	0	0	0	0	0.2
Capitellidae	0	0	0	0	1.4	0
Cirratulidae	0	0	0	0.4	0.6	0
Cossura longicirrata	0	0	0	0.4	0	0
Eteone longa	0	0	1	29.4	0.6	0
Euchone incolor	0	0	0	0	0.4	0
Exogone hebes	0	0	0.2	0	0.2	0
Harmothoe imbricata	0.2	0	0	0.8	0	0
Neoleanira tetragona	0	0	0	0	0.2	0
Nephtys ciliata	8.6	21.6	13.2	21.4	10.4	21.8
Ninoe nigripes	0	0.2	0.8	39.2	14	0.2
Pherusa plumosa	0	0	0.2	2	0.2	0
Pholoe minuta	0	0	0	0.4	1.4	0.2
Phyllodoce mucosa	0	0	0	1.4	0.2	0
Polydora quadrilobata	1.6	44.8	113	2522.4	213	0
Prionospio steenstrupi	0	0	0	0	0.6	0
Sabellidae	0	0	0	1.6	0.6	0
Scoloplos armiger	0	0	0	0	0	0
Spiophanes bombyx	0	0	0	0	11.6	0
Tharyx marioni	0	0	0	0	0	0
Marine Oligochaete	0	2	0	0	0.6	0
CRUSTACEANS	0	0	0.2	0	0	0
Anonyx sarsi	0	0	0	5.2	0	0.2
Chiridotea tuftsi	0	0	0	2.2	1.2	0
Copepoda	0	0	0	0	0.4	0
Diastylis polita	0	0.2	0	80.6	7.2	0
Edotea triloba	0	0	0	25.4	0.4	0
Eudorallopsis deformis	0	0	0.2	19.4	0.6	0
Gammarus sp.	0	0	0	0	0.8	0
Leptochirus pinguis	0	0.2	0.4	0	0	0
Orchomenella pinguis	0	0	0	24.8	0	0
Phoxocephalus holbolli	0	0	0	0	0.2	0
Stenothoe minuta	0	0	0	106	0.4	0
Cylichna gouldi	0	0	0	0.4	0	0.4
Cerastoderma pinulatum	0	0	0	2	0	0
Ilynassa trivittatus	0	0	0	1	12.6	0
Macoma balthica	0	0	0	0.2	0.4	0.2
Macoma tenta	0	0	0.2	1.8	4.2	0
Margarites groenlandica	0	0	0	0	0.4	0
Mya truncata	0	0	0	0.8	0	0
Nucula delphinodonta	0	0	0	5.4	13.6	0
Yoldia limatula	0	0	0	2.4	0	0
Edwardsia sp.	0	0.8	0.2	3.8	0.2	0.2
Echinarachnius parma	0	0	0	0	0.2	0
NEMERTEA	0.4	1.4	1.6	10	0.2	1
NEMATODA	0	0	1.2	67.4	67	0
Sagitta elegans	0	0.2	0	0	0	0

Rank Correlations Data

Table C-1 Summary of Rank Correlations among Sediment Variables with |r| ≥ 0.90

Variable 1	Variable 2	Rank Correlation
Acenaphthene	Acenaphthylene	0.9803
Acenaphthene	Anthracene	0.9873
Acenaphthene	Benzo*(a)*Anthracene	0.9798
Acenaphthene	Benzo*(a)*Pyrene	0.9497
Acenaphthene	Benzo*(b)*Fluoranthene	0.9434
Acenaphthene	Benzo*(e)*Pyrene	0.9619
Acenaphthene	Benzo*(g,h,i)*Perylene	0.9510
Acenaphthene	Benzo*(k)*Fluoranthene	0.9550
Acenaphthene	Cd	0.9464
Acenaphthene	Chrysene	0.9781
Acenaphthene	Cu	0.9307
Acenaphthene	Dibenzo*(a,h)*Anthracene	0.9616
Acenaphthene	Fluoranthene	0.9716
Acenaphthene	Fluorene	0.9882
Acenaphthene	Indeno(1,2,3-cd)Pyrene	0.9420
Acenaphthene	Naphthalene	0.9582
Acenaphthene	PCB	0.9513
Acenaphthene	Pb	0.9677
Acenaphthene	Perylene	0.9529
Acenaphthene	Phenanthrene	0.9893
Acenaphthene	Pyrene	0.9531
Acenaphthene	1-Methyl-Naphthalene	0.9942
Acenaphthene	1-Methyl-Phenanthrene	0.9781
Acenaphthene	2,6-Di-Methyl-Naphthalene	0.9936
Acenaphthene	2-Methyl-Naphthalene	0.9922
Acenaphthylene	Anthracene	0.9936
Acenaphthylene	Benzo*(a)*Anthracene	0.9911
Acenaphthylene	Benzo*(a)*Pyrene	0.9769
Acenaphthylene	Benzo*(b)*Fluoranthene	0.9725
Acenaphthylene	Benzo*(e)*Pyrene	0.9867
Acenaphthylene	Benzo*(g,h,i)*Perylene	0.9776
Acenaphthylene	Benzo(*k)*Fluoranthene	0.9798
Acenaphthylene	Cd	0.9743
Acenaphthylene	Chrysene	0.9918
Acenaphthylene	Cu	0.9487
Acenaphthylene	Dibenzo*(a,h)*Anthracene	0.9855
Acenaphthylene	Fluoranthene	0.9898
Acenaphthylene	Fluorene	0.9897
Acenaphthylene	Indeno(1,2,3-cd)Pyrene	0.9713
Acenaphthylene	Naphthalene	0.9816
Acenaphthylene	PCB	0.9745
Acenaphthylene	Pb	0.9647
Acenaphthylene	Perylene	0.9794
Acenaphthylene	Phenanthrene	0.9948
Acenaphthylene	Pyrene	0.9769

Variable 1	Variable 2	Rank Correlation
Acenaphthylene	1-Methyl-Naphthalene	0.9875
Acenaphthylene	1-Methyl-Phenanthrene	0.9860
Acenaphthylene	2,6-Di-Methyl-Naphthalene	0.9904
Acenaphthylene	2-Methyl-Naphthalene	0.9906
Anthracene	Benzo*(a)*Anthracene	0.9981
Anthracene	Benzo*(a)*Pyrene	0.9857
Anthracene	Benzo*(b)*Fluoranthene	0.9819
Anthracene	Benzo*(e)*Pyrene	0.9927
Anthracene	Benzo*(g,h,i)*Perylene	0.9859
Anthracene	Benzo*(k)*Fluoranthene	0.9869
Anthracene	Cd	0.9563
Anthracene	Chrysene	0.9982
Anthracene	Cu	0.9361
Anthracene	Dibenzo*(a,h)*Anthracene	0.9910
Anthracene	Fluoranthene	0.9960
Anthracene	Fluorene	0.9945
Anthracene	Indeno(1,2,3-cd)Pyrene	0.9802
Anthracene	Naphthalene	0.9898
Anthracene	PCB	0.9610
Anthracene	Pb	0.9575
Anthracene	Perylene	0.9881
Anthracene	Phenanthrene	0.9981
Anthracene	Pyrene	0.9856
Anthracene	1-Methyl-Naphthalene	0.9878
Anthracene	1-Methyl-Phenanthrene	0.9855
Anthracene	2,6-Di-Methyl-Naphthalene	0.9913
Anthracene	2-Methyl-Naphthalene	0.9929
As	Cd	0.9074
As	Cu	0.9254
As	Pb	0.9392
Benzo*(a)*Anthracene	Benzo*(a)*Pyrene	0.9895
Benzo*(a)*Anthracene	Benzo*(b)*Fluoranthene	0.9864
Benzo*(a)*Anthracene	Benzo*(e)*Pyrene	0.9948
Benzo*(a)*Anthracene	Benzo*(g,h,i)*Perylene	0.9907
Benzo*(a)*Anthracene	Benzo*(k)*Fluoranthene	0.9917
Benzo*(a)*Anthracene	Cd	0.9546
Benzo*(a)*Anthracene	Cu	0.9390
Benzo*(a)*Anthracene	Dibenzo*(a,h)*Anthracene	0.9953
Benzo*(a)*Anthracene	Fluoranthene	0.9983
Benzo*(a)*Anthracene	Fluorene	0.9944
Benzo*(a)*Anthracene	Indeno(1,2,3-cd)Pyrene	0.9862
Benzo*(a)*Anthracene	Naphthalene	0.9899
Benzo*(a)*Anthracene	PCB	0.9536
Benzo*(a)*Anthracene	Pb	0.9577
Benzo*(a)*Anthracene	Perylene	0.9918
Benzo*(a)*Anthracene	Phenanthrene	0.9953
Benzo*(a)*Anthracene	Pyrene	0.9878
Benzo*(a)*Anthracene	1-Methyl-Naphthalene	0.9800
Benzo*(a)*Anthracene	1-Methyl-Phenanthrene	0.9856
Benzo*(a)*Anthracene	2,6-Di-Methyl-Naphthalene	0.9849
Benzo*(a)*Anthracene	2-Methyl-Naphthalene	0.9857

Variable 1	Variable 2	Rank Correlation
Benzo*(a)*Pyrene	Benzo*(e)*Pyrene	0.9977
Benzo*(a)*Pyrene	Benzo*(k)*Fluoranthene	0.9974
Benzo*(a)*Pyrene	Cd	0.9251
Benzo*(a)*Pyrene	Chrysene	0.9915
Benzo*(a)*Pyrene	Cu	0.9052
Benzo*(a)*Pyrene	Dibenzo*(a,h)*Anthracene	0.9956
Benzo*(a)*Pyrene	Fluoranthene	0.9938
Benzo*(a)*Pyrene	Fluorene	0.9745

Table C-2 Summary of Rank Correlations Among Porewater Variables with |r| ≥ 0.90

Variable 1	Variable 2	Rank Correlation
1-Methyl-Naphthalene	1-Methyl-Phenanthrene	-0.9337
1-Methyl-Naphthalene	2-Methyl-Naphthalene	0.9860
1-Methyl-Naphthalene	Fluorene	0.9643
1-Methyl-Naphthalene	PCB	0.9373
2,3,5-Tri-Methyl-Naphthalene	Acenaphthylene	0.9919
2,6-Di-Methyl-Naphthalene	Acenaphthene	0.9664
2,6-Di-Methyl-Naphthalene	Salinity	-0.9052
2-Methyl-Naphthalene	Anthracene	0.9453
2-Methyl-Naphthalene	Fluorene	0.9948
2-Methyl-Naphthalene	Indeno(1,2,3-cd)Pyrene	0.9008
2-Methyl-Naphthalene	Phenanthrene	0.9544
Acenaphthene	Salinity	-0.9353
Acenaphthylene	PCB	0.9330
Anthracene	Benzo*(a)*Anthracene	0.9358
Anthracene	Benzo*(a)*Pyrene	0.9583
Anthracene	Benzo*(b)*Fluoranthene	0.9642
Anthracene	Benzo*(e)*Pyrene	0.9621
Anthracene	Benzo*(g,h,i)*Perylene	0.9766
Anthracene	Benzo*(k)*Fluoranthene	0.9413
Anthracene	Chrysene	0.9667
Anthracene	Dibenzo*(a,h)*Anthracene	0.9391
Anthracene	Fluoranthene	0.9241
Anthracene	Fluorene	0.9686
Anthracene	Indeno(1,2,3-cd)Pyrene	0.9755
Anthracene	Phenanthrene	0.9943
Anthracene	Pyrene	0.9408
Anthracene	Eh	-0.9110
Benzo*(a)*Anthracene	Benzo*(a)*Pyrene	0.9897
Benzo*(a)*Anthracene	Benzo*(b)*Fluoranthene	0.9877
Benzo*(a)*Anthracene	Benzo*(e)*Pyrene	0.9833
Benzo*(a)*Anthracene	Benzo*(g,h,i)*Perylene	0.9822
Benzo*(a)*Anthracene	Benzo*(k)*Fluoranthene	0.9944
Benzo*(a)*Anthracene	Chrysene	0.9904
Benzo*(a)*Anthracene	Dibenzo*(a,h)*Anthracene	0.9976
Benzo*(a)*Anthracene	Fluoranthene	0.9988
Benzo*(a)*Anthracene	Indeno(1,2,3-cd)Pyrene	0.9820
Benzo*(a)*Anthracene	PCB	0.9117
Benzo*(a)*Anthracene	Perylene	0.9700
Benzo*(a)*Anthracene	Phenanthrene	0.9493
Benzo*(a)*Anthracene	Pyrene	0.9947
Benzo*(a)*Pyrene	Benzo*(e)*Pyrene	0.9985

Variable 1	Variable 2	Rank Correlation
Benzo*(a)*Pyrene	Benzo*(g,h,i)*Perylene	0.9959
Benzo*(a)*Pyrene	Benzo*(k)*Fluoranthene	0.9978
Benzo*(a)*Pyrene	Dibenzo*(a,h)*Anthracene	0.9952
Benzo*(a)*Pyrene	Fluoranthene	0.9867
Benzo*(a)*Pyrene	Fluorene	0.9195
Benzo*(a)*Pyrene	Indeno(1,2,3-cd)Pyrene	0.9975
Benzo*(a)*Pyrene	PCB	0.9831
Benzo*(a)*Pyrene	Perylene	0.9569
Benzo*(a)*Pyrene	Phenanthrene	0.9683
Benzo*(a)*Pyrene	Pyrene	0.9977
Benzo*(b)*Fluoranthene	Benzo*(e)*Pyrene	0.9989
Benzo*(b)*Fluoranthene	Benzo*(g,h,i)*Perylene	0.9974
Benzo*(b)*Fluoranthene	Benzo*(k)*Fluoranthene	0.9962
Benzo*(b)*Fluoranthene	Dibenzo*(a,h)*Anthracene	0.9933
Benzo*(b)*Fluoranthene	Fluoranthene	0.9842
Benzo*(b)*Fluoranthene	Fluorene	0.9255
Benzo*(b)*Fluoranthene	Hg	0.9940
Benzo*(b)*Fluoranthene	Indeno(1,2,3-cd)Pyrene	0.9988
Benzo*(b)*Fluoranthene	PCB	0.9941
Benzo*(b)*Fluoranthene	Perylene	0.9509
Benzo*(b)*Fluoranthene	Phenanthrene	0.9731
Benzo*(b)*Fluoranthene	Pyrene	0.9960
Benzo*(e)*Pyrene	Benzo*(g,h,i)*Perylene	0.9978
Benzo*(e)*Pyrene	Benzo*(k)*Fluoranthene	0.9949
Benzo*(e)*Pyrene	Chrysene	0.9981
Benzo*(e)*Pyrene	Dibenzo*(a,h)*Anthracene	0.9906
Benzo*(e)*Pyrene	Fluoranthene	0.9812
Benzo*(e)*Pyrene	Fluorene	0.9152
Benzo*(e)*Pyrene	Hg	0.9816
Benzo*(e)*Pyrene	Indeno(1,2,3-cd)Pyrene	0.9979
Benzo*(e)*Pyrene	Perylene	0.9492
Benzo*(e)*Pyrene	Phenanthrene	0.9685
Benzo*(e)*Pyrene	Pyrene	0.9934
Benzo*(g,h,i)*Perylene	Benzo*(k)*Fluoranthene	0.9904
Benzo*(g,h,i)*Perylene	Chrysene	0.9981
Benzo*(g,h,i)*Perylene	Dibenzo*(a,h)*Anthracene	0.9876
Benzo*(g,h,i)*Perylene	Fluoranthene	0.9785
Benzo*(g,h,i)*Perylene	Fluorene	0.9285
Benzo*(g,h,i)*Perylene	Hg	0.9603
Benzo*(g,h,i)*Perylene	Indeno(1,2,3-cd)Pyrene	0.9989
Benzo*(g,h,i)*Perylene	Perylene	0.9384
Benzo*(g,h,i)*Perylene	Phenanthrene	0.9794
Benzo(g,h,i)Perylene	Pyrene	0.9888
Benzo*(k)*Fluoranthene	Chrysene	0.9962
Benzo*(k)*Fluoranthene	Dibenzo*(a,h)*Anthracene	0.9989
Benzo*(k)*Fluoranthene	Fluoranthene	0.9938
Benzo*(k)*Fluoranthene	Indeno(1,2,3-cd)Pyrene	0.9912
Benzo*(k)*Fluoranthene	PCB	0.9474
Benzo*(k)*Fluoranthene	Perylene	0.9724
Benzo*(k)*Fluoranthene	Phenanthrene	0.9525
Chrysene	Dibenzo*(a,h)*Anthracene	0.9945
Chrysene	Fluoranthene	0.9870
Chrysene	Fluorene	0.9242
Chrysene	Hg	0.9929
Chrysene	Indeno(1,2,3-cd)Pyrene	0.9986
Chrysene	PCB	0.9893

Variable 1	Variable 2	Rank Correlation
Chrysene	Perylene	0.9518
Chrysene	Phenanthrene	0.9744
Chrysene	Pyrene	0.9958
Cr	Zn	0.9899
Cr	pH	-0.9837
Dibenzo*(a,h)*Anthracene	Fluoranthene	0.9970
Dibenzo*(a,h)*Anthracene	Indeno(1,2,3-cd)Pyrene	0.9878
Dibenzo*(a,h)*Anthracene	PCB	0.9305
Dibenzo*(a,h)*Anthracene	Perylene	0.9765
Dibenzo*(a,h)*Anthracene	Phenanthrene	0.9498
Dibenzo*(a,h)*Anthracene	Pyrene	0.9987
Fluoranthene	Indeno(1,2,3-cd)Pyrene	0.9770
Fluoranthene	Perylene	0.9765
Fluoranthene	Phenanthrene	0.9366
Fluoranthene	Pyrene	0.9931
Fluorene	Indeno(1,2,3-cd)Pyrene	0.9381
Fluorene	Phenanthrene	0.9763
Fluorene	Pyrene	0.9024
Hg	Indeno(1,2,3-cd)Pyrene	0.9702
Hg	Naphthalene	0.9504
Indeno(1,2,3-cd)Pyrene	Perylene	0.9375
Indeno(1,2,3-cd)Pyrene	Phenanthrene	0.9817
Indeno(1,2,3-cd)Pyrene	Pyrene	0.9908
PCB	Pyrene	0.9455
PCB	Eh	-0.9187
Perylene	Pyrene	0.9712
Phenanthrene	Pyrene	0.9541
Zn	pH	-0.9530
Ammonia.p	Sulphide	-0.9142