00-54

Environment Canada
Water Science and Technology Directorate

Direction générale des sciences et de la technologie, eau
Environnement Canada

Using Benthic Assessment Techniques to Determine Combined Sewer Overflow and Stormwater Impacts in the Aquatic Ecosystem

By:

Q. Rochfort, L. Grapentine, J. Marsalek....

10-54

MANAGEMENT PERSPECTIVE

This is part of the CLEAN ENVIRONMENT business line, Result: Reduction of impacts of wet-weather pollution on aquatic ecosystems. This study addresses the impacts of wet-weather pollution on receiving water sediments and habitat. It also relates to the issue of sediment quality guidelines and protection of the receiving water environment.

This study showed that although the lowest effect level and severe effect level of the sediment quality guidelines were exceeded for some PAHs and metals, biological effects were not evident. Benthic community structure indicators and toxicity test endpoints did not appear to be related to the observed contaminant levels.

The sites selected in this initial investigation were used to determine the extent of impacts that could be observed due to combined sewer overflows or stormwater discharges. The results from this study indicate that only limited comparisons could be made, owing to the difference in habitats between exposed and control sites. The next step would be to collect samples for benthic community structure and toxicity (where possible) immediately upstream and downstream of the selected stormwater or combined sewer overflow outfall, to provide appropriate reference sites nd determine impacts.

SOMMAIRE À L'INTENTION DE LA DIRECTION

Ceci fait partie du secteur d'activité UN ENVIRONNEMENT SAIN, résultat : « réduction des effets produits sur les écosystèmes aquatiques par la pollution des périodes pluvieuses ». Cette étude porte sur les effets de la pollution des périodes pluvieuses sur les sédiments et l'habitat des eaux réceptrices. Elle examine également la question des directives sur la qualité des sédiments et sur la protection de l'environnement des eaux réceptrices.

Cette étude a démontré que, bien que la concentration minimale avec effet et la concentration d'effet dangereux des directives sur la qualité des sédiments soient dépassées pour certains HAP et certains métaux, il n'y avait pas d'effets biologiques évidents. Les indicateurs de structure de communauté benthique et le résultat des tests de toxicité ne semblaient pas reliés aux concentrations de contaminants observées.

Les sites sélectionnés pour cette étude initiale ont été utilisés pour déterminer la portée des effets des rejets de déversoirs d'orage et d'eau de ruissellement qui ont été observés. Les résultats indiquent qu'on ne peut faire que des comparaisons limitées, en raison de la différence des habitats entre les sites exposés et les sites contrôlés. L'étape suivante consisterait à recueillir des échantillons pour les structures de communautés benthiques et pour les tests de toxicité (là où la chose est possible) immédiatement en amont et en aval du point sélectionné de rejet d'eau de ruissellement ou de déversoir d'orage afin d'obtenir des sites de référence appropriés et de déterminer les effets.

Using Benthic Assessment Techniques to Determine Combined Sewer Overflow and Stormwater Impacts in the Aquatic Ecosystem

Q. Rochfort, L. Grapentine, J. Marsalek, B. Brownlee, T. Reynoldson, S. Thompson, D. Milani, C. Logan

National Water Research Institute - Environment Canada, 867 Lakeshore Rd., Burlington, Ontario L7R 4A6

Abstract

Urban wet-weather sources of pollution such as stormwater and combined sewer overflows (CSOs) can contribute significantly to the contamination of receiving waters, particularly in sediment depositional areas near outfalls. Analyses of sediment chemistry alone are not sufficient to fully assess the effects of these discharges. Toxicity testing and evaluations of benthic invertebrate communities, in conjunction with chemical analyses, provide a more complete characterization. This study assessed relationships among three separate aspects of the benthic environment: sediment chemistry (metals, PAHs and nutrients) and particle size, sediment toxicity (ten endpoints with four benthic taxa) and benthic invertebrate community structure. In this initial survey, ten sites in five different study areas, representing a range of receiving water environments exposed to stormwater and CSO discharges, were sampled in October 1998. Results of analyses indicated that while contaminant (metals and PAHs) concentrations were relatively high in sediments, biological effects were not evident. Toxicity of sediments was low, and altered benthic communities were not detected. Neither toxicity endpoints nor benthic community descriptors were related to sediment contaminant levels. To improve the power of these assessments, future investigations of stormwater and CSO discharge impacts should use "upstream/downstream" sampling designs, and study sites with minimal variability of habitat conditions.

Key Words: Benthic Toxicity, Community Structure, Wet-Weather Discharges

Résumé

Les sources de pollution urbaine en période pluvieuse, telles que les eaux de ruissellement et les déversoirs d'orage, peuvent contribuer de façon significative à la contamination des eaux réceptrices, particulièrement dans les zones de dépôt de sédiments voisines des points d'évacuation. Les analyses chimiques des sédiments ne sont pas suffisantes à elles seules pour évaluer tous les effets de ces déversements. Les tests de toxicité et les évaluations de communautés d'invertébrés benthiques, conjointement avec les analyses chimiques, permettent d'obtenir une caractérisation plus complète. Cette étude a évalué les relations entre trois aspects distincts des environnements benthiques : la composition chimique des sédiments (métaux, HPA et nutriments) et les tailles des particules, la toxicité des sédiments (dix résultats avec quatre taxa benthiques) et la structure des communautés d'invertébrés benthiques. Dans cette enquête initiale, dix sites situés dans cinq zones différentes et représentant une gamme d'environnements d'eaux réceptrices exposées à des rejets d'eau de ruissellement et de déversoir d'orage ont été échantillonnés en octobre 1998. Les résultats des analyses ont indiqué que, bien que les concentrations de contaminants (métaux et HPA) soient relativement élevées dans les sédiments exposés, il n'y avait pas d'effets biologiques apparents. La toxicité des sédiments était faible et aucune communauté benthique altérée n'a été détectée. Ni les paramètres de toxicité recherchés ni des descripteurs de communauté benthique n'étaient reliés au niveau de contamination des sédiments. Pour améliorer la puissance de ces évaluations, les études futures sur les effets des rejets d'eau de ruissellement et de déversoirs d'orage devraient utiliser des échantillonnages « amont/aval » et viser les sites où les conditions d'habitat n'ont qu'une variabilité minimale.

Water Qual. Res. J. Canada, 2000
Volume 35, No. 3, 365-397

Using Benthic Assessment Techniques To Determine Combined Sewer Overflow and Stormwater Impacts in the Aquatic Ecosystem

Q. ROCHFORT, L. GRAPENTINE,* J. MARSALEK, B. BROWNLEE, T. REYNOLDSON, S. THOMPSON, D. MILANI AND C. LOGAN

National Water Research Institute, Environment Canada, 867 Lakeshore Road, Burlington, Ontario L7R 4A6

Urban wet-weather sources of pollution such as stormwater and combined sewer overflows (CSOs) can contribute significantly to the contamination of receiving waters, particularly in sediment depositional areas near outfalls. Analyses of sediment chemistry alone are not sufficient to fully assess the effects of these discharges. Toxicity testing and evaluations of benthic invertebrate communities, in conjunction with chemical analyses, provide a more complete characterization. This study assessed relationships among three separate aspects of the benthic environment: sediment chemistry (metals, PAHs and nutrients) and particle size, sediment toxicity (ten endpoints with four benthic taxa), and benthic invertebrate community structure. In this initial survey, ten sites in five different study areas, representing a range of receiving water environments exposed to stormwater and CSO discharges, were sampled in October 1998. Results of analyses indicated that while contaminant (metals and PAHs) concentrations were relatively high in sediments, biological effects were not evident. Toxicity of sediments was low and altered benthic communities were not detected. Neither toxicity endpoints nor benthic community descriptors were related to sediment contaminant levels. To improve the power of these assessments, future investigations of stormwater and CSO discharge impacts should use "upstream/downstream" sampling designs and study sites with minimal variability of habitat conditions.

Key words: **benthic toxicity, community structure, wet-weather discharges**

Introduction

Urban non-point sources of pollution such as stormwater and combined sewer overflows (CSOs) can contribute significantly to the contamination of receiving waters. In addition to dissolved materials, stormwater and CSO outfalls discharge particulate material which accumulates in depositional areas in the vicinity of the outfall. These sediments may contribute the major load of pollutants to the receiving stream, including adsorbed polycyclic aromatic hydrocarbons (PAHs) and metals (Lee et al. 1997). While these substances may not be immediately bioavail-

* Corresponding author; *Lee.Grapentine@cciw.ca*

able to all aquatic organisms in the receiving stream, benthic communities may become severely impacted over time. The intermittent and highly variable nature of these wet-weather discharges has made the characterization and quantification of contaminant loadings problematic (Seager and Maltby 1989). Analysis of chemical constituents alone was not sufficient to fully define these types of discharges and it has been suggested that aquatic toxicity testing be performed in conjunction with these chemical analyses for a more complete characterization (Marsalek et al. 1999).

Recognizing high variability of toxicity of stormwater and CSOs, it was of interest to explore sediment toxicity using benthic toxicity techniques. The use of benthic toxicity testing in detecting impacts of wet-weather discharges has been limited and changes to benthic community structure due to impacts from these discharges have not been well characterized. Some researchers, however, have found that benthic invertebrate methods were suitable for detecting the impacts of urban wet-weather pollution, as follows.

Maltby et al. (1995) used toxicity identification evaluation (TIE) procedures to isolate the source of sediment toxicity to the amphipod *Gammarus pulex*. This species was present both upstream and downstream of a stormwater outfall from a large motorway (M1, U.K.), but was observed to have much lower population downstream. They determined that hydrocarbons, Cu and Zn were responsible for the observed toxicity to *G. pulex*. Isolation of the hydrocarbons by chromatography fractionation showed that the 2–5 ring PAHs were responsible for most of the toxicity (Boxall and Maltby 1995). Further testing with pure compounds and sediment extracts in laboratory exposures showed that three PAHs accounted for most of the toxicity: pyrene (44.9%), fluoranthene (16%) and phenanthrene (3.5%) (Boxall and Maltby 1997).

Mulliss et al. (1996) caged the amphipod *Gammarus pulex* and the isopod *Asellus aquaticus* upstream and downstream of a CSO outfall in an urban stream. Mortality and heavy metal tissue concentrations were used to demonstrate that these species were sensitive to impacts from the combined sewer. Principal component analysis was used to associate the observed toxicity with measured physical and chemical parameters. Depressed communities in both species were linked to increased hydraulic flow, high suspended solids, biochemical oxygen demand and aqueous Cu concentrations, but it was found that low *G. pulex* populations were associated with high aqueous concentrations of NH_3, Pb, Zn and Cu, whereas low *A. aquaticus* populations were associated with high aqueous Cd concentrations and high tissue concentrations of Zn, Pb, Cd and Cu. These results confirm that such benthic invertebrate species can (along with supporting chemical analyses) indicate variation in sediment quality within a receiving water system and also show that benthic populations may manifest different responses as a result of impacts from these discharges.

Davis (1997) used a macrobenthic survey and chemical testing of sediments to demonstrate the recovery which had occurred in a river system in Texas. A survey performed 19 years previously showed evidence of severe impacts by sewage treatment plant discharges and other urban non-point source pollution. After remediation, most areas showed good to fair sediment chemistry and healthy benthic communities with habitat that was not physically limiting. This observed recovery was largely attributed to improvements in the sewage treatment systems and management of urban wet-weather discharges. Only a few urban areas remained impacted, where total taxa and mean taxa richness factors were still low. These areas were characterized by higher pesticide and metallic oxide concentrations as well as depressed oxygen levels.

Borchardt and Sperling (1997) developed a screening procedure for urban catchments which identified watersheds where receiving waters were likely to be degraded as a result of urban discharges. This screening procedure was based on laboratory and field measurements of chemical and biological conditions as well as model-based predictions. The study concluded that significant impacts could be expected when greater than 5% of the drainage area was impervious, NH_3 concentrations exceeded 0.1 mg/L, and suspended solids were greater than 50 mg/L; however, the extent of the final impacts was highly associated with the flow and depth in the receiving waters.

Hall et al. (1998) demonstrated that sediment samples taken along transects in Vancouver Harbour could be used to delineate exposure zones around CSOs. The richness and abundance of benthic communities and toxicity of the sediment to two marine benthic species (the mussel *Mytilus edulis* and the amphipod *Rhepoxynius abronius*) were used to characterize the effects of these outfalls. Accompanying chemical studies correlated higher PAH contamination with the toxicity of these sites and this was also found to be associated with the degree of heavy industry in the watershed.

Hatch and Burton (1999) were successful in using both laboratory and in situ toxicity testing to investigate the toxicity of sediments below an urban stormwater outfall. The *Hyalella azteca* survival test appeared to be the most successful and reproducible test, with good correlation between field and laboratory measurements. There was high mortality (up to 90%) of organisms exposed to sediment during discharge conditions in the field, and it was suggested that hydraulic action may release pollutants trapped in the sediments and make them temporarily bioavailable to the benthic organisms.

Both benthic community structure analyses and benthic toxicity testing may be useful tools to apply when assessing the impacts of stormwater and CSO discharges in the receiving waters. While water quality may vary considerably during a storm event, exposure periods are generally brief and many aquatic organisms may not be affected in the short term. Sediment deposited by these discharges has a more long-term effect on the receiving environment. Depositional zones in the receiving streams near discharge outfalls integrate effects over time. Similarly, benthic invertebrate communities are exposed to cumulative discharge effects of sediment as well as water quality. Their sedentary nature, ubiquity, responsiveness to disturbances, ease of sampling and importance to other ecosystem components make benthic invertebrate communities highly relevant in environmental studies (Johnson et al. 1993). Assessing sediment and benthic invertebrate conditions is thus a practical alternative to the continued monitoring of water quality affected by these discharges. As such, benthic methods may be applicable in both short-term toxicity assessment and long-term monitoring of wet-weather discharges.

Study Approach

This study was designed as an initial investigation into the feasibility of using benthic invertebrate methods to assess the quality of sediments in the vicinity of stormwater and CSO outfalls and to determine the biological effects observed in the receiving waters as a result of these discharges. The overall approach involved assessing relationships between three separate aspects of the benthic environment: sediment chemistry (metals, PAHs and nutrients) and particle size, sediment toxicity (tests with four benthic taxa), and benthic invertebrate community structure. The scope of this study was limited to a preliminary characterization of benthic conditions at sites exposed to a broad range of CSO and stormwater discharges. Results will be used to identify locations that are amenable to more detailed future assessments that include sampling from reference sites.

Sediment chemistry and particle size can indicate physico-chemical effects of stormwater and CSO discharges. Sediment toxicity and benthic community structure (BCS) can represent biological responses to the physico-chemical conditions of the sediments. Samples were collected in several locations exposed to stormwater and CSO discharges (rather than upstream and downstream of particular outfalls), in order to maximize the range of discharges and receiving water body types. With sediment contaminant concentrations considered indicative of the degree of exposure to stormwater and CSO discharges (assuming the bioavailable fraction of the total contaminant concentration was the same for all sites), sediment toxicity and BCS were analyzed to determine the concordance between physico-chemical and biological responses to discharges. The performance of several best management practices (BMPs) for stormwater and CSO control (Schueler 1992) was also addressed in this study.

In conjunction with sediment chemistry, two biological methods were used: the toxicity of sediment to four benthic toxicity test organisms (*Hyalella azteca, Chironomus riparius, Hexagenia* spp. and *Tubifex tubifex*) and an examination of benthic invertebrate communities in the vicinity of selected wet-weather discharges via community structure analysis. These were adaptations of methods developed by Environment Canada (Reynoldson and Day 1998) to establish biological guidelines for sediment quality in the Great Lakes and were thought to have potential for assessing the quality of sediment from stormwater and CSOs.

Methods

Study Areas and Sample Locations

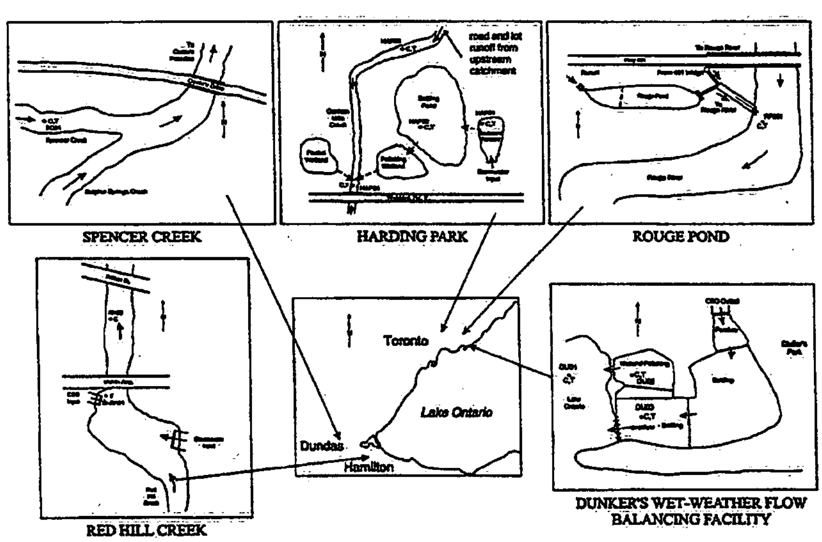
In this initial survey, five different study areas were sampled, representing a range of receiving water types impacted by a variety of stormwater and CSO discharges. All samples were collected in October 1998. Two study areas (one CSO treatment facility and one stormwater pond) were sampled at multiple locations, while the remaining sites were sampled at only one location.

Dunker's wet-weather flow balancing system — Scarborough

The Dunker's wet-weather flow balancing system is a CSO treatment facility designed to protect the near-shore water quality in Lake Ontario. Site DU01 was in a shallow embayment in Lake Ontario near the outfall, site DU02 was a wetland-polishing cell and site DU03 was located in the main settling pond (Fig. 1).

Harding Park — Richmond Hill

This stormwater pond site receives runoff from a largely residential area. Site 1 (HAP01) was located in the sediment forebay, site 2 (HAP02) was located in the main settling pond, site 3 (HAP03) was located in German Mills Creek, upstream of the pond outfall, just above a runoff channel draining from a row of houses (but below drainage outfalls from streets



C *indicates location of community structure sample;* T *indicates location of toxicity sample.*

Fig. 1. Sample collection sites, 1998.

and housing lots) and site 4 (HAP04) was located approximately 10 m downstream of the stormwater pond outfall in German Mills Creek (Fig. 1).

Single-sampling site locations

Sampling site RPO01 was established at the outfall from the Rouge River stormwater pond, located east of Port Union Road and immediately south of Highway 401 (Fig. 1). Site SC01 was sampled on Spencer Creek in Dundas, approximately 50 m upstream of the confluence with Sulphur Springs Creek (Fig. 1). Both above sites were used for sampling toxicity and BCS. Site RH01, which was located just upstream of Melvin Avenue at the CSO discharge, was sampled for toxicity on the Red Hill Creek in Hamilton. A community structure sampling site (RH02) was located between Melvin Avenue and Barton Street, 50 m downstream of the outflow (Fig. 1).

Expected degree of exposure to CSO and stormwater discharges

Based on the relative locations of (a) discharge treatment facilities and outfalls and (b) sampling sites, expected categories of exposure to discharges are shown in Table 1. There was uncertainty about site HAP03 because, while it is upstream of the outfall from the Harding Park treatment facility, it is also downstream of several other stormwater outfalls.

Sample Collection

Not all sites could be sampled in exactly the same way, due to variation in habitat type. As such, different approved methods (Reynoldson et al. 1999b) were employed for sample collection, depending on the conditions existing at the sample sites.

Table 1. Expected degree of exposure to CSO and stormwater discharges

Site	Habitat	CSO/stormwater exposure
DU01	Lake	Far far field
DU02	Pond	Far field
DU03	Pond	Near field
HAP01	Pond	Near field
HAP02	Pond	Far field
HAP03	Stream	Upstream reference/far field
HAP04	Stream	Far far field
RPO01	Stream	Far field
SC01	Stream	Near field
RH01/02	Stream	Near/far field

Sediment quality

Where petit-Ponar samples were taken for benthic community structure, one Ponar was sampled for chemistry. Using a glass dish and teflon spatulas to homogenize the sample, subsamples were obtained for PAHs, metals, particle size and nutrient chemistry.

Where benthic community structure samples were collected with a kick net, representative samples were obtained for chemical analyses by scooping the sediment directly from the stream bed using a plastic jar.

Benthic sampling

Fine-grain material was required for toxicity test samples (particle size of 250 µm or less). Five 2-L replicates were collected for each site using a petit-Ponar for the toxicity tests. Samples for benthic community structure were obtained using two different collection methods, as the substrate varied considerably at the study areas. Where sediment was muddy, soft and in slow moving water (such as a pond or large stream), sediments were collected by petit-Ponar. Three 2-L replicate samples were collected for community structure. Samples were sieved through a 500-µm mesh bucket in the field prior to preservation with 5% formalin. After 48 h, samples were transferred into 70% ethanol to prevent degradation of mollusc shells (Reynoldson et al. 1999a). Where the water was shallow, fast moving or more rocky (sites HAP03, HAP04, RH02), benthic invertebrates were collected by standardized three-minute kick net sampling.

Sample Analysis

Chemistry

For each site, sediment samples were analyzed for PAHs, particle size distribution, metals, total organic carbon (TOC), total Kjeldahl nitrogen (TKN) and total phosphorus (TP).

PAHs and particle size analyses were performed in-house. Wet PAH samples were ground with anhydrous Na_2SO_4 (to remove moisture), extracted with dichloromethane (DCM) by Accelerated Solvent Extraction (ASE™), and analyzed on a Hewlett-Packard model 5890 Series II gas chromatograph equipped with a HP 5971 mass selective detector (MSD) operating in electron impact mode at 70 eV. Data were acquired in selective ion monitoring (SIM) mode, and the column was a 30 m × 0.25 mm HP-5MS capillary column. Results were reported as ng/g dry weight, and losses were estimated based on internal standard spike recoveries. Particle size was determined using a sedigraph apparatus, and results reported as percent gravel, sand, silt and clay. Mean particle size and particle size of the 25th percentile and 75th percentile were also indicated.

All other analyses were performed by a private (Canadian Association for Environmental Analytical Laboratories [CAEAL] accredited) laboratory following standard methods (APHA 1989), and using appropriate QA/QC methods. Samples for analysis of elements were prepared by nitric acid extraction and analyzed by inductively coupled plasma (ICP). Total nitrogen and total phosphorus were extracted from the sediments using a standard Kjeldahl digestion and analyzed colourimetrically on a spectrophotometer.

Benthic sediment toxicity testing procedures

Sediment samples were stored in the dark at 4°C until used in toxicity tests. Each sediment sample was homogenized and sieved, where possible, through a 250-µm sieve (Nytex®) to remove indigenous macrofauna (Reynoldson et al. 1991, 1994), using a 4:1 ratio of culture water : sediment (2 L culture water:500 mL sediment). The sieved sediment was allowed to settle for a minimum of 24 h, after which the water was decanted and used as the overlying water in the experiment. Culture water was de-chlorinated municipal tap water.

Test methods for all species are detailed in Reynoldson et al. (1999b). Water chemistry variables (pH, dissolved oxygen, temperature and conductivity) were measured and recorded for each replicate test beaker both at the beginning and end of each test. Average total ammonia concentration was also determined for each site by taking a composite water sample from each replicate beaker at the start of the test and at test completion. A laboratory control treatment was included in each test to ensure that the test responses of each species were within acceptable limits. All test beakers were aerated for 7 days prior to and over the course of the test, and water loss due to evaporation was replaced with de-ionised water. Tests were run at 23±1°C in environmental chambers. A photoperiod of 16 h-light:8 h-dark and a light intensity of 500–1000 lux was maintained throughout the tests, with the exception of the *T. tubifex* test, which was run in the dark.

Hyalella azteca 28-day survival and growth test

Each 250-mL test beaker received 50 mL of test sediment and 175 mL of overlying water. Two- to 10-day-old amphipods were randomly added to replicate beakers until 15 organisms were in each chamber. Each chamber was fed 8 mg Nutrafin® fish flakes twice weekly on non-consecutive days. On day 28, the contents of each beaker were rinsed through a 250-µm mesh sieve. Surviving amphipods were counted, dried for 24 h at 60°C, and dry weights recorded.

Chironomus riparius 10-day survival and growth test

Each 250-mL test beaker received 50 mL of test sediment and 175 mL of overlying water. First instar chironomids were added by pipette until the count of 15 per beaker was achieved. Each test chamber was fed 8 mg crushed Nutrafin® fish flakes three times throughout the test on non-consecutive days. On day 10, the contents of each beaker were sieved through a 250-µm sieve. Surviving chironomids were counted, dried for 24 h at 60°C, and dry weights recorded.

Hexagenia spp. 21-day survival and growth test

Each 1-L test chamber received 125 mL of test sediment and 650 mL of overlying water. Ten preweighed (5–8 mg wet weight) mayflies were randomly added to replicate jars. Each test jar was fed 50 mg of a prepared diet (Cerophyll, Nutrafin®, brewers yeast) once weekly. On day 21, the contents of each jar were sieved through a 500-µm sieve. Surviving mayflies were counted, dried for 24 h at 60°C, and dry weights recorded. Initial wet weights were converted to dry weights with the following conversion: initial dry weight = (Initial wet weight + 1.15)/7.35 (Reynoldson and Day 1998). Growth was determined by subtracting the initial dry weight from the final dry weight.

Tubifex tubifex 28-day survival and reproduction test

Each 250-mL test chamber received 100 mL of test sediment and 100 mL of overlying water. Food (80 mg of crushed Nutrafin® per beaker) was added to the test beakers and thoroughly mixed into the sediment. Four sexually mature specimens of *T. tubifex* (identifiable gonads) were randomly added to each replicate beaker. On day 28, the contents of each beaker were rinsed through a 500-µm and 250-µm sieve sequentially. The contents from the two sieves were washed separately into individual gridded plastic dishes for enumeration with a dissecting microscope. The number of surviving adults, the number of full cocoons, the number of empty cocoons, and the number of large offspring were counted from the 500-µm screen. The number of small offspring were counted from the 250-µm mesh screen. Four endpoints were calculated for this test: mean percent of adult survival, the mean number of cocoons per adult, the mean percent of cocoons hatched and the mean number of offspring per adult. If adult survivorship was less than 100%, the number of surviving adults was taken as the average of the actual number of adults surviving and the number of original adults (4). This value was used in the calculation of the number of cocoons per adult and the number of young per adult.

Community structure

Macroinvertebrates were picked from benthic samples and sorted to family level using a low power stereo microscope. Number of individuals per sample for each family were enumerated and recorded. Kick net samples were subsampled (due to the high number of organisms) using a subsampling device (Marchant 1989). The subsampler is a box (35 × 35 × 10 cm) divided into 100 equal cells. The unsorted sample was washed into the box with sufficient water. The box was covered and then shaken to evenly distribute the sample among the cells. Cells were randomly sampled until at least 200 organisms were picked. (All cells started were finished.) The total number of organisms in the sample was estimated by extrapolation to the full 100 cells based on the number of cells counted.

Data Analyses

The data obtained from the sampling procedures fall into three categories of variables: (1) sediment chemistry (7 particle sizes, 3 nutrients, 18 metals, 28 PAH compounds); (2) sediment toxicity (10 endpoints for 4 species); and (3) benthic invertebrate community composition (counts for 32 taxa).

Analyses of these data sets were aimed at (a) characterizing conditions at sampling sites in terms of each of the above categories, and (b) assessing the relationships between sediment chemistry, sediment toxicity and BCS. Site characterizations of sediment chemistry allowed identification of contaminated conditions, based on comparisons to sediment quality guidelines (SQG). Assessing relationships between the physico-chemical and biological data sets indicates the levels of concordance in the patterns of responses to CSO and stormwater discharges.

Although hypothesis testing was restricted, and involved the examination of (a) associations between expected degree of exposure to CSO and stormwater discharges and benthic conditions, and (b) correlations between sediment chemistry, sediment toxicity and BCS, expected outcomes of comparisons under various scenarios can be stated.

If discharges do not contaminate the benthic receiving environment:

- contaminants in sediments should not exceed SGQs *or* should not be related to expected degree of exposure to discharges;
- sediments should not be toxic (or enriching in the case of nutrients);
- biological conditions (sediment toxicity, BCS) should not be related to either sediment contaminant levels or expected degree of exposure to discharges.

If discharges contribute bioavailable contaminants to the benthic receiving environment:

- contaminants in sediments should exceed SGQs *or* should be related to expected degree of exposure to discharges;
- sediments should be toxic (or enriching);
- biological conditions (sediment toxicity, BCS) should be related to either sediment contaminant levels of expected degree of exposure to discharges.

If discharges contribute contaminants that are not bioavailable to the benthic receiving environment:

- contaminants in sediments should exceed SGQs *or* should be related to expected degree of exposure to discharges;
- sediments should not be toxic (or enriching);
- biological conditions (sediment toxicity, BCS) should not be related to either sediment contaminant levels or expected degree of exposure to discharges.

While rigorous statistical testing of these hypotheses was not possible due the nature of the sampling design, these expected outcomes were used as a guide for interpretation of the results.

Characterization of sediment conditions at sites was achieved by univariate and multivariate descriptive statistics and graphs. For the sediment chemistry data, concentrations of metals and PAHs at each site were compared to existing provincial sediment quality guidelines (Persaud et al. 1993) as a means of identifying degree of contamination. To provide a general characterization of sediment chemistry, principal components analysis (PCA) was performed on each of two subgroups of variables: (1) metal, nutrient and grain size variables, and (2) PAH variables. Because the data consist of observations for multiple, possibly interacting and covarying variables, ordination methods such as principal components analysis (PCA) are effective means of "compressing" the information contained in many variables into fewer (often 1–3) *component* variables (Clarke and Green 1988). These principal components (PCs) are orthogonal and uncorrelated. As such, they are superior descriptors of patterns in the data structure and more suitable for further statistical analyses than the original variables. The two sediment chemistry data subsets were therefore reduced to the minimum number of component variables necessary to explain at least 85% of the variability (or information) in the data. These PCs were then used to describe general differences among sites in pairwise plots of their coordinates along the new axes, and for relating to biological conditions of sediments. The first PCA was performed on ln-transformed concentrations of 13 metals (major cations and Cd, which was undetected in all samples, excluded), total organic carbon, total Kjeldahl nitrogen and total phosphorus, and the ln-transformed 25th percentile grain size (calculated from phi-value). The second PCA was performed on ln-transformed concentrations of 28 PAH compounds.

It is important to note that PCA is used here for data reduction and transformation, rather than for hypothesis testing. The component variables produced by PCA allow testing or evaluation of all ordinated variables jointly and reduce the need to reconcile the often conflicting conclusions obtained from a series of single variable analyses. PCA is generally suitable for normally distributed data with the number of samples greater than the number of variables. The method can, however, be robust to departures from these conditions (Green 1979; Green and Montagna 1996). Although the data sets here have more variables than samples and, in the case of the benthic invertebrate data set, are not normally distributed, ordinations by more suitable methods (principal coordinate analyses [Legendre and Legendre 1983], nonmetric multidimensional scaling [Belbin 1991]) produce similar results. Therefore, for the purpose of characterizing patterns among samples based on multiple variables, followed by simple or informal tests of hypotheses, PCA can be used effectively.

Sediment toxicity results were evaluated by comparing endpoint means for each of the 10 sampling sites with criteria derived from tests with uncontaminated reference sediment from the Great Lakes (Reynoldson and Day 1998). The criteria are based on measurements of the 10 toxicity endpoints for 179–220 reference sediments collected from the nearshores of the Great Lakes over a 3-year period. From these reference data, means and standard deviations (SDs) were calculated for each endpoint and used to define three toxicity categories (Table 2):

Table 2. Criteria for determination of toxicity for nearshore sediments of the Great Lakes (Reynoldson and Day 1998)

Species	No toxicity	Potential toxicity	Confirmed toxicity
H. azteca			
Survival (%)	>67.0	67.0–57.1	<57.1
Growth (mg)	0.75–0.23	0.22–0.10	<0.10
C. riparius			
Survival (%)	>67.7	67.7–58.8	<58.8
Growth (mg)	0.49–0.21	0.20–0.14	<0.14
Hexagenia spp.			
Survival (%)	>85.5	85.4–80.3	<80.3
Growth (mg)	5.04–0.97	0.96–0.0	—
T. tubifex			
Survival (%)	>88.9	88.9–84.2	<84.2
Number of cocoons/adult	12.4–7.2	7.1–5.9	<5.9
Percent hatched	78.1–38.1	38.0–28.1	<28.1
Number of young/adult	46.3–9.9	9.8–0.8	<0.8

- *toxic,* where response is less than the mean minus (3 × SD);
- *potentially toxic,* where response is between the mean minus (2 × SD) and mean minus (3 × SD); and
- *non-toxic,* where response is greater than mean minus (2 × SD).

For the growth and reproduction endpoints, an upper limit for the non-toxic category was set at mean plus (2 × SD). Where responses were in excess of this limit (i.e., growth or reproduction significantly higher than average), sediments were grouped in a fourth category: *enriched.*

Benthic invertebrate communities of the sampling sites were described by multivariate methods. For the sediment chemistry data set, PCA was used to produce several (3) descriptor component variables to characterize BCS. Because benthic invertebrates were collected by two sampling methods — petit-Ponar grab at seven sites and kick net at three sites — the data are not commensurate and cannot be pooled. PCA was therefore performed only with data from seven Ponar stations. Analyses were performed on ln (x+1)-transformed data, where x = number of individuals per taxon (generally family). Station BCS was represented in pairwise plots of the coordinates along the PCs.

Relationships between sediment chemistry descriptors, toxicity test endpoints and BCS descriptors were assessed by correlation analysis and bivariate plots. Sediment chemistry descriptors included PCs from the two ordinations and a grain size variable — ln (25th percentile of particle size) — as a means of assessing the influence of grain size alone. The toxicity endpoints were the 10 measured responses. The BCS descriptors were the first three PCs from the ordinations of the Ponar sample data subset, plus total abundance and taxa richness for each of the Ponar and kick net sample data subsets. Separate correlations were calculated for each of the Ponar and kick samples.

Results

Sediment Chemistry and Grain Size

Concentrations of metals and nutrients and proportions of grain size classes in sediment collected from the study sites are shown in Table 3. Results of analyses for 28 PAH compounds are shown in Table 4. Differences in grain size between samples are marked. The 25th percentile size (based on phi units), which provides a good characterization of the general grain sizes of samples, ranged from 12 μm for the Harding Park settling pond (HAP02) to 3138 μm for the Harding Park stream site (HAP04) below the inflow from the settling pond. The Lake Ontario sample (DU01) was distinct in being the sandiest. Ranges in nutrient concentrations were also high — ratios of maximum : minimum concentrations were 45, 71 and 6 for total Kjeldahl nitrogen, total phosphorus and total organic carbon, respectively.

Included in Tables 3 and 4 are existing Ontario sediment quality guidelines (Persaud et al. 1993). Samples from most of the sites had concentrations of metals and PAHs in excess of the "lowest effect level". "Severe effect level" (SEL) limits were exceeded only for PAHs in the Red Hill Creek sample (RH01). Maximum exceedances were less than 2 × SEL. The PAH values for SEL quoted in Table 4 are expressed in terms of total organic carbon content in the PAH sample, and therefore where the SEL was exceeded, adjusted sample PAH values based on TOC have been included in a footnote.

Multivariate characterizations of sediment chemistry and grain size were achieved by PCA ordinations. Results of the first ordination, involving sediment grain size, metal and nutrient variables, are shown in Fig. 2, which represents the samples based on their scores (i.e., coordinates) along the first two PCs. The plot shows two things: (1) relative variable levels for each sample, as indicated by the position of the data points along each axis, and (2) similarities of samples to each other, based on the closeness of data points. PC scores are interpretable in terms of the original variables by identifying which of these variables are most related to each of the components. These are indicated in the PC plot by annotation along each axis. The first PC (PC1) accounts for 58% of the total variation among samples. It is inversely related to concentrations of all metals and nutrients, and directly related to grain size. Therefore, the lower the score for PC1 (i.e., the further to the left in the plot), the more metal-contaminated, nutrient-enriched and finer the sediment in the sample. The second

Table 3. Sediment chemistry analysis for sample sites[a]

Variable		SQG[b]		Site									
		LEL	SEL	DU01	DU02	DU03	HAP01	HAP02	HAP03	HAP04	RH01	RPO01	SC01
Gravel	%			3.0	0	0	0	0	0	33.3	20.3	0.4	0
Sand	%			93.3	31.8	15.3	6.6	2.7	71.3	60.9	69.8	64.8	4.4
Silt+clay	%			3.7	68.2	84.7	93.4	97.3	28.8	5.8	9.9	34.8	95.6
P size 25%	µm			175	72	33	17	12	338	3138	1814	126	22
TKN	µg/g	550	4800	48	199	1030	2180	1370	716	97	1010	719	2040
TOC	%	1	10	0.04	0.71	1.29	2.64	1.67	2.40	0.42	0.36	0.60	2.84
TP	µg/g	600	2000	547	598	698	726	775	533	290	1750	636	917
Al	%			0.23	0.46	0.79	1.02	1.30	0.33	0.20	1.04	0.38	0.92
As	µg/g	6	33	11	5	12	2.5	1.3	7	2.5	15	2.5	14
Cd	µg/g			< 1	< 1	< 1	< 1	< 1	2	< 1	< 1	< 1	< 1
Co	µg/g			2	3	4	1	5	0.5	0.5	5	0.5	7
Cr	µg/g	26	110	6	12	19	25	24	16	7	35	11	18
Cu	µg/g	16	110	6	13	26	55	37	66	18	75	11	41
Fe	%	2	4	0.88	1.02	1.47	1.52	2.05	1.13	0.74	2.00	0.89	2.22
Hg	µg/g	0.2	2	0.026	0.054	0.221	0.165	0.064	0.054	0.032	0.056	0.032	0.058
Mn	µg/g	460	1100	179	295	364	481	484	232	238	766	261	895
Ni	µg/g	16	75	7	12	19	22	24	12	14	21	9	19
Pb	µg/g	31	250	13	17	26	42	24	30	29	24	15	43
Ti	µg/g			273	401	490	356	527	194	107	167	290	188
V	µg/g			17	18	25	25	34	15	9	27	16	23
Zn	µg/g	120	820	34	38	77	219	117	124	64	180	41	313

[a] Shaded area: exceeds lowest effect level.
[b] SQG: sediment quality guidelines; LEL: lowest effect level; SEL: severe effect level (Persaud et al. 1993).

Table 4. Concentrations (ng/g) of PAHs in sediment collected from CSO/Stormwater study sites[a]

Compound	LEL (ng/g)	SEL (ng/g OC)	DU01	DU02	DU03	HAP01	HAP02	HAP03	HAP04	RH01	RP001	SC01
Naphthalene			4.67	11.4	28.5	42.3	22.0	28.2	75.6	494	13.6	25.2
Naphthalene, 2-methyl			4.10	7.51	17.9	54.3	24.9	26.5	73.7	199	6.67	22.3
Naphthalene, 1-methyl			2.52	4.03	9.66	31.7	13.8	16.5	40.6	132	3.68	11.8
Naphthalene, 2,7-dimethyl			6.32	7.97	19.7	66.3	186	28.7	20.8	67.8	6.33	29.9
Naphthalene, 1,3-dimethyl			8.38	8.11	16.2	33.0	13.0	14.4	22.6	61.4	4.23	11.9
Naphthalene, 1,5-dimethyl			4.13	4.52	6.78	10.5	10.3	4.20	6.52	14.3	1.67	4.10
Acenaphthylene			0.25	3.06	12.1	22.2	12.0	12.1	12.4	203	2.64	16.5
Naphthalene, 1,2-dimethyl			1.67	2.25	4.15	11.4	5.72	5.21	8.73	18.4	0.44	3.12
Acenaphthene			2.38	8.45	25.0	83.4	33.3	55.5	69.0	336	11.5	15.5
Naphthalene, 2,3,5-trimethyl			23.6	14.6	25.2	33.4	19.2	22.6	22.9	33.3	1.00	7.31
Fluorene	190	160000	3.03	12.8	36.8	159	85.3	74.8	88.4	742[d]	13.5	38.3
Dibenzothiophene			2.59	8.50	24.4	140	73.3	44.2	74.1	348	10.4	25.0
Phenanthrene	560	950000	24.6	119	457	2360	1210	645	1290	6310[e]	219	357
Anthracene	220	370000	1.30	18.8	65.2	182	78.9	113	184	902	18.3	165
Anthracene, 2-methyl			0.60	3.57	12.9	32.7	14.7	19.5	33.9	119	2.42	20.8
Phenanthrene, 1-methyl			13.9	11.2	26.9	132	69.4	35.4	48.1	134	7.58	25.1
Fluoranthene	750	1020000	16.9	185	746	5290	2370	1170	2500	5030[f]	463	874
Pyrene	490	850000	27.5	191	515	3760	1820	889	1860	3130[g]	323	663
Benzo[a]anthracene	320	1480000	1.97	72.0	323	1180	511	382	818	1450	176	313
Chrysene+triphenylene	340[b]	460000	16.8	92.5	406	2440	1220	497	1070	1470	259	501

(continued)

Table 4. *(concluded)*

Benzo[b]fluoranthene			3.63	90.4	406	1800	876	316	766	1200	280	445
Benzo[k]fluoranthene	240	1340000	2.78	37.2	169	886	409	164	440	626	125	225
Benzo[e]pyrene			5.66	57.7	246	1120	547	217	520	768	172	318
Benzo[a]pyrene	370	1440000	0.65	61.5	277	1260	532	270	678	1320	222	353
Perylene			108	145	223	296	119	110	317	372	120	122
Indeno[1,2,3-cd]pyrene	200	320000	1.66	113	450	796	365	141	362	706	393	277
Dibenzo[a,h]anthracene	60	130000	0.00	11.6	47.7	122	54.0	25.8	64.3	124	33.8	46.3
Benzo[ghi]perylene	170	320000	5.38	81.6	281	742	337	132	326	563	272	268
TOTAL[c]	4000		110	1385	4878	23082	11032	5461	11790	26879	3159	5185
Total organic carbon (%)			0.04	0.71	1.29	2.64	1.67	2.4	0.42	0.36	0.6	2.84

[a] Shaded area: exceeds lowest effect level; inverse printing: exceeds severe effect level.
[b] Guideline for chrysene.
[c] Total for SQG is based on sum of 16 U.S. EPA compounds; for site concentrations, total is sum of 15 of the 16 compounds used for SQG + triphenylene (Persaud et al. 1993).
[d] 206000 ng/g OC.
[e] 1750000 ng/g OC.
[f] 1400000 ng/g OC.
[g] 870000 ng/g OC.

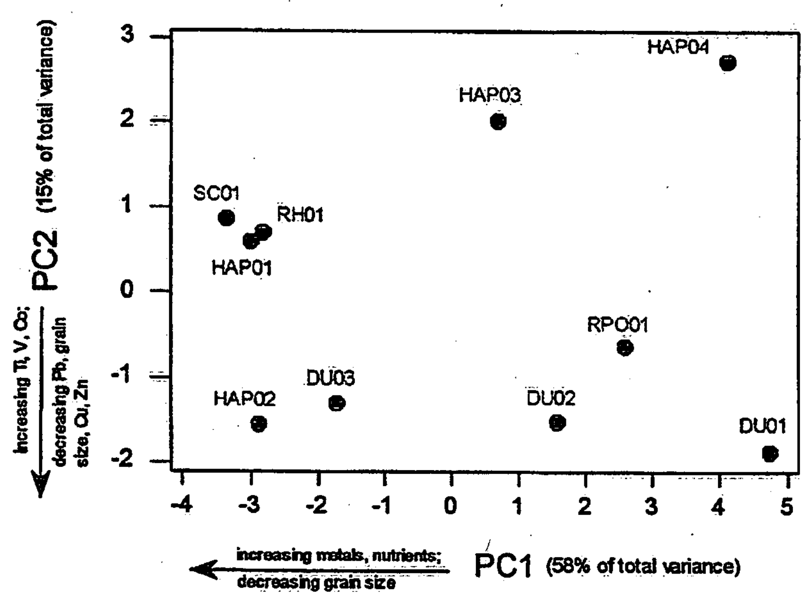


Fig. 2. Characterization of sediment grain size and metal and nutrient concentrations in samples from CSO/stormwater study sites, 1998. Samples are represented by scores for the first two principal components (PCs) from a PCA on ln-transformed data. Interpretation of PCs is shown by annotation of each axis with a labelled arrow indicating gradients for the original variables that are most correlated with the PC.

PC (PC2) accounts for 15% of the total variation, and is inversely related to concentrations of Ti, V and Co, and directly related to Pb, Cu and Zn concentrations and grain size. Thus, samples shown near the top of Fig. 2 are low in Ti, V and Co content, high in Pb, Cu and Zn content, and coarse grained relative to samples plotted further from the top. The third PC (not shown in a plot), accounts for 13% of the total variation. It is mainly related to As and total phosphorus (directly) and total organic carbon and grain size (inversely). While Fig. 2 and the PC scores indicate sample similarities, lack of replicate sampling of sites precludes determination of the significance of differences. However, the main purpose of the PCA was to convert multiple measurements of conditions of sediment metals, nutrients and grain size into a smaller number of variables for comparison to measurements of sediment toxicity and benthic community structure.

The second PCA resulted in a more readily interpretable pattern of sample plots. The first PC explains 87% of the total among sample variation, and was inversely proportional to total concentration of PAH compounds. In fact, PC1 is strongly related to total ln-transformed PAH concentrations ($r = -0.996$, $P < 0.001$). The second PC, accounting for 8% total variable, is related to the proportions of light- and heavy-weight

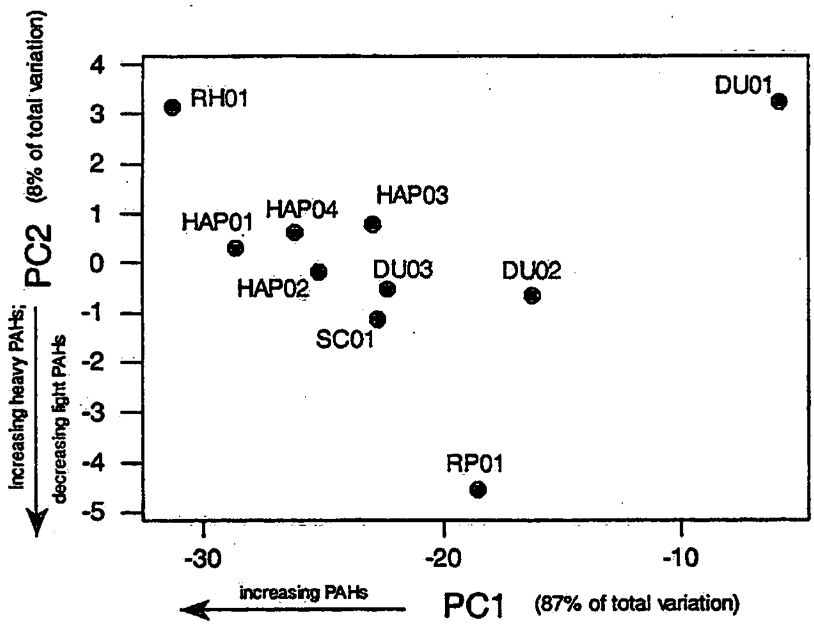


Fig. 3. Characterization of sediment PAH concentrations in samples from CSO/stormwater study sites, 1998. Samples are represented by scores for the first two principal components (PCs) from a PCA on ln-transformed data. Interpretation of PCs is shown by annotation of each axis with a labelled arrow indicating gradients for the original variables.

PAHs: inversely proportional to compounds lighter than fluoranthene and directly proportional to all other measured compounds (except perylene). In Fig. 3, the further to the left in the plot of PC2 versus PC1 scores, the higher the concentration of PAHs. Because (a) "total PAH concentration" is a more comprehensible descriptor of sediment PAH contamination than PC1, (b) total PAH is highly correlated to PC1, and (c) PC1 accounts for a high proportion of the variability among samples, "total PAHs" alone was used to characterize sediments rather than PC1 from the ordination.

Benthic Toxicity Tests

There were many indigenous species of worms present in sites HAP03, HAP04 and RH01. This created a problem for the enumeration of the *T. tubifex* reproduction test and necessitated removal of some replicates. Also, the presence of leeches in certain sediments necessitated the removal of replicates. In general, the tests did not show strong evidence of toxicity at these sites. Ammonia levels (a common factor in toxicity) were undetectable or very low at all sites. The coefficients of variation (CV) were determined for all test sites for each of the 10 endpoints to evaluate the precision of the test. In general, the mean coefficients of variation were good for field replicated samples, ranging from 2.96 to 47.10. Highest variability was noted in the *T. tubifex* reproduction endpoint (number of young per adult) which was typical for these tests and did not affect test accuracy.

Dunker's wet-weather flow balancing system — Scarborough

Only site DU01 showed evidence of confirmed toxicity, with reduced *Hyalella* survival and potential toxicity to *Hyalella* growth (Table 5). Sites DU02 and DU03 showed evidence of enrichment based on *Hexagenia* spp. growth, which exceeded the upper limit of the non-toxic category. DU02 also showed a negative response (reduction in young production) for *T. tubifex*. Variability in test responses (as observed by the CVs) was higher for sites having moderate to high toxicity.

Harding Park — Richmond Hill

Only the tests performed on stream sites (HAP03 and HAP04) at Harding Park showed any form of toxic response. Two of the four species (*H. azteca* and *T. tubifex*) showed toxicity at the downstream HAP04 site in the sediment toxicity tests (Table 5). This site also had a high percentage of sand and gravel similar to DU01 (Table 3). The receiving stream above the stormwater pond (HAP03), which was exposed to untreated stormwater discharges, showed some evidence of enrichment, with *Hexagenia* spp. growth above the upper non-toxic limit, while at the same time it showed potential toxicity to *H. azteca* and confirmed and potential toxicity in *T. tubifex* (Table 5). The sediment forebay (HAP01) and the main settling pond (HAP02) did not show any evidence of toxicity.

Single-sampling site locations

The toxicity test results from single-sampling site locations were used to provide information about sites which may be generally impacted by multiple sources upstream of the site as well as local point source discharges. There was enrichment in sites SC01, RPO01 and RH01 evidenced by the *Hexagenia* spp. growth data, and RH01 also showed enrichment for *C. riparius* growth (Table 5), which corresponded to elevated nutrient levels (Table 3). Only the Rouge Pond outfall (RPO01) site showed toxicity in the *T. tubifex* test (in both the number of cocoons hatching and number of young produced) at the same time as showing evidence of enrichment in *Hexagenia* spp..

Benthic Invertebrate Communities

A total of 32 benthic invertebrate taxa (primarily families) were identified from 10 samples of sediment collected from the study sites (Table 6). Overall, oligochaete worms (naidids and tubificids) and chironomids were the dominant taxa. Number of taxa per sample ranged from only 2

Table 5. Summary of ten endpoints from toxicity testing data from all sites[a]

Site Code	DU01	DU02	DU03	HAP01	HAP02	HAP03	HAP04	RPO01	SC01	RH01
H. azteca										
Survival (%)	50.7	70.7	89.3	85.3	80.0	58.7	33.3	82.7	94.7	81.3
Growth (mg)	0.18	0.30	0.39	0.36	0.40	0.32	0.37	0.54	0.36	0.38
C. riparius										
Survival (%)	93.3	88.0	81.3	80.0	90.7	96.7	80.0	80.0	74.7	97.3
Growth (mg)	0.30	0.34	0.36	0.33	0.33	0.43	0.44	0.34	0.25	(0.63)
Hexagenia spp.										
Survival (%)	90.0	100	100	100	96.0	96.7	100	96.0	100	98.0
Growth (mg)	1.90	(7.30)	(5.08)	2.78	4.83	(5.39)	1.62	(7.01)	(7.56)	(5.45)
T. tubifex										
Survival (%)	100	100	100	100	100	100	87.5	100	100	100
Cocoons (No.)	7.7	11.2	10.3	11.0	10.3	11.5	5.6	10.6	10.7	10.1
Hatch (%)	74.2	33.8	58.2	57.2	58.8	15.4	68.8	20.4	40.5	60.8
Young/Adult (No.)	11.2	2.5	20.6	22.5	23.8	5.5	6.6	1.6	9.8	33.8

[a] Regular font: non-toxic; inverse printing: toxic; shaded area: potentially toxic; brackets: enriched.

Table 6. Summary of mean abundance (number per sample) of families in the invertebrate communities at sample sites

Site	DU01	DU02	DU03	HAP01	HAP02	HAP03	HAP04	RPO01	SC01	RH02
Habitat	Lake	Pond	Pond	Pond	Pond	Stream	Stream	Stream	Stream	Stream
Collection method	Ponar	Ponar	Ponar	Ponar	Ponar	Kick-net	Kick-net	Ponar	Ponar	Kick-net
Taxon										
Asellidae						25				300
Baetidae					0.6					
Chaoboridae		0.2	0.6		4.4					
Ceratopogonidae								0.4	4.2	25
Chironomidae	2.6	52.4	13	30.4	3	575	475	120.6	66.6	1750
Cladocera family		0.2	0.4					9		25
Copepoda family		0.8	0.4	4.2		700	175			300
Corixidae		0.8								
Corydalidae				0.2						
Enchytraeidae						25	25			
Elmidae								7.2	0.6	25
Empididae										25
Erpobdellidae						25	125		0.6	
Heptageniidae								0.2		
Hydridae				1.2						
Lebertiidae			0.2							
Limnesiidae							25			

(continued)

Table 6. *(concluded)*

Leptohyphidae								0.2		
Lymnaeidae									0.4	
Macrobiotidae								0.2		
Muscidae				0.2						
Naididae		23.8	16.8	2		750	975	7.2	0.8	2500
Nematoda family		1		63.8		200		0.6		
Ostracoda family				0.8			25	0.6		
Physidae			0.6	0.4			150		0.6	150
Planorbidae				2.2						
Sminthuridae				0.2						
Spongillidae			0.8							
Stratiomyidae				0.2						
Tabanidae								0.2		
Tubificidae	4.8	94.4	18.4	104.4	3.2	3650	3775	512	151.6	325
Turbellaria family		0.2						0.6	3.6	
Total abundance	**7.4**	**173.8**	**51.2**	**210.2**	**11.2**	**5950**	**5750**	**659**	**229**	**5425**
Richness	**2**	**9**	**9**	**13**	**4**	**8**	**9**	**13**	**9**	**10**

in the Lake Ontario embayment (DU01) to 13 in Harding Park (HAP01) and Rouge Pond (RPO01). Because samples were collected by two different methods, data from all stations are not comparable. Therefore, separate analyses were performed on each subset of stations.

Ponar samples

The three Dunker's samples (DU01, DU02, DU03), two of the Harding Pond samples (HAP01, HAP02) and samples from Rouge Pond (RPO01) and Spencer Creek (SC01) ranged in total number of individuals per sample (based on mean of three samples per site) from 7.4 at DU01 to 659 at RPO01. PCA, involving 28 taxa, resulted in 90% of the total variation explained by the first three PCs. PC1 was inversely related to abundance of most taxa, and especially to tubificids, chironomids and naidids. It was also negatively correlated with taxa richness ($r = -0.917$, $P = 0.004$) and ln (total abundance) ($r = -0.989$, $P < 0.001$). PC2 was inversely related to nematode abundance. PC3 (13% of variance explained) was inversely related to naidid abundance. Fig. 4 shows the seven Ponar samples represented by PC1 and PC2. The sample from RPO01 was shown to have the most individuals and taxa, whereas samples from DU01 and HAP02 had the least. The sample from site HAP01 was distinct from the other

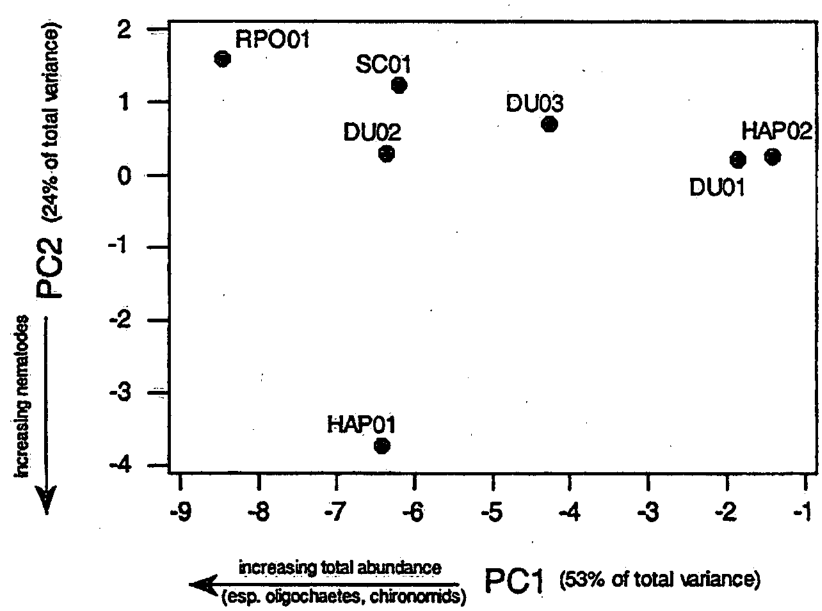


Fig. 4. Benthic invertebrate community structure of Ponar samples from CSO/stormwater study sites, 1998. Samples are represented by scores for the first two principal components (PCs) from a PCA on ln-transformed taxon counts. Taxa showing most important relationships to PCs are indicated in annotations of axes.

samples based on PC2, largely due to its comparatively high count of nematodes.

Kick net samples

Taxa richness and total abundance in the three stream kick net samples were less variable than those from the standing water sites, ranging from 8 to 10 taxa per sample and 5425 to 5950 individuals per sample, respectively. Kick sample numbers are not, however, directly comparable to the Ponar sample numbers. Due to small sample size, PCA was not performed. As for the Ponar-sampled benthos, the dominant taxa were tubificids, naidids and chironomids. Copepods were also numerically important.

Relationships Between Sediment Quality and Biological Conditions

For each of the three sets of data on sediment conditions, variability among samples has been characterized by a selected number of descriptor variables. To assess the degree to which physico-chemical conditions were predictive of sediment toxicity and BCS, pairwise plots and correlations between variables from predictor and response data sets were examined. For sediment quality, descriptor variables include the first three PCs from the ordination of metal, nutrient and grain size data (labelled "chemPC1 – 3"), and total PAH concentration. To determine if grain size alone explained biological conditions of samples, "ln-transformed 25th percentile grain size" was also compared to biological variables. The biological descriptors were (a) the 10 toxicity test endpoints for the sediment toxicity data set, and (b) the first three PCs for the ordination of the benthic invertebrate data (Ponar-sampled data only). In addition to the BCS PCs, taxa richness and total abundance were also assessed because, though redundant and not as informative as PCs, they are commonly measured descriptors and more comprehensible.

Correlations between sediment quality descriptors and toxicological responses are shown in Table 7. Among the 50 correlations calculated, only 5 were significant at the $\alpha = 0.05$ level. Of those, none indicate a relationship of increasing sediment toxicity with increasing contaminant concentration. Two endpoints — survival of *H. azteca* and number of tubificid young produced per adult — were negatively correlated with "chemPC1", indicating that survival and reproduction were highest in the most metal-contaminated, nutrient-rich, fine-grained sediment. Which of the sediment variables associated with "chemPC1" was/were responsible cannot be determined from the data. However, several variables are candidates, based on high correlations with *H. azteca* survival. These include the nutrients total nitrogen and total phosphorus, and several metals (Al, Cr, Fe, Mn, V), all of whose correlations for ln-transformed concentrations were between 0.725 and 0.843 ($P = 0.018$ to 0.002). Elevated metal concentrations are not known to increase survival of *H. azteca*; therefore, a nutrient effect is the best hypothesis. Growth of *H. azteca*, though, was not

Table 7. Pearson correlations between sediment descriptors and toxicological responses[a]

Sediment descriptor[b]		Toxicological response									
		H. azteca		*C. riparius*		*Hexagenia* spp.		*T. tubifex*			
		Survival	Growth	Survival	Growth	Survival	Growth	Survival	# Cocoons	% Hatched	# Young
ln (PS25%)	*r*	-0.678	-0.073	0.362	0.754	-0.079	-0.307	-0.604	-0.577	0.122	-0.081
	P	0.031	0.842	0.304	0.012	0.828	0.389	0.064	0.081	0.738	0.823
chemPC1	*r*	-0.803	-0.271	0.104	-0.067	-0.474	-0.367	-0.464	-0.611	0.011	-0.680
	P	0.005	0.448	0.775	0.854	0.166	0.297	0.177	0.061	0.975	0.030
chemPC2	*r*	-0.350	0.133	-0.123	0.415	0.417	-0.191	-0.594	-0.243	-0.142	-0.100
	P	0.321	0.713	0.734	0.233	0.231	0.596	0.070	0.498	0.696	0.783
chemPC3	*r*	0.014	0.312	-0.401	-0.480	0.265	-0.105	-0.111	0.181	-0.266	-0.305
	P	0.969	0.381	0.251	0.160	0.460	0.773	0.760	0.617	0.457	0.392
Total PAH	*r*	0.259	0.442	-0.060	0.567	0.655	-0.007	-0.212	0.138	0.082	0.601
	P	0.470	0.201	0.869	0.087	0.040	0.985	0.556	0.704	0.821	0.066

[a] Regular font: not significant; inverse printing: significant ($p<0.05$).
[b] As defined in the text.

correlated to nutrient levels. Grain size is another possible factor affecting *H. azteca* survival because ln-transformed 25th percentile grain size was significantly related to survival of *H. azteca* ($P < 0.031$) and growth of *C. riparius*, accounting for 46 and 57%, respectively, of the variation in the toxicity endpoint. Survival of *Hexagenia* spp. was correlated to PAH concentrations ($P = 0.040$), but as for chemPC1, the "healthiest" response was associated with highest contaminant levels. In addition, an outlying data point (corresponding to the Lake Ontario sample) highly influenced the slope and significance of the relationship for these variables.

Benthic community descriptors showed fewer significant relationships to sediment quality descriptors than did the toxicity endpoints (Table 8). The only significant correlation ($P = 0.013$) involved the taxa richness for the kick net samples and total PAHs. Besides being based on only three samples, the direction of the correlation suggested higher richness with greater PAH contamination. Therefore, these results provide no evidence that difference in BCS is associated with increase in sediment contaminant concentration.

Discussion

This study was designed to be a preliminary assessment of the benthic conditions in a series of locations exposed to CSO and stormwater discharges varying in magnitude and type. The purpose was to provide a characterization of physico-chemical and biological conditions of samples collected from the study sites, and to illustrate an approach for assessing benthic effects of CSO/stormwater discharges involving multiple lines of evidence.

Assessments based on multiple lines of evidence are superior to those based on single lines for several reasons. First, although physical and chemical attributes of receiving environments (such as water flow regime or contaminant concentrations in sediments) are easier to measure than the comparatively more complex and variable biological ones (such as toxicological responses or benthic community structure), it is in terms of the biological conditions that ecosystem "damage" is commonly recognized (Beanlands and Duinker 1983; McIntyre 1984). Second, biological conditions represent integrated responses to all exposure pathways. For example, benthic communities are exposed not only to particles deposited as sediment in their habitat, but also to suspended particles and dissolved substances that are not typically characterized in benthic assessments. Third, multiple lines of evidence allow confirmation of patterns of effects, as well as determination of potential causes (which can be verified in further experimental studies).

Effects of Wet-Weather Discharges

The approach for this study was based on the assumptions that contaminant concentrations would indicate degree of exposure to wet-weather

Table 8. Pearson correlations between sediment descriptors and benthic community descriptors[a]

Sediment descriptor[b]		Benthic invertebrate community descriptor						
		Ponar samples (n = 7)					Kick samples (n = 3)	
		Richness	ln (total abundance)	benth C1	benth C2	benth C3	Richness	ln (total abundance)
ln (PS25%)	*r*	-0.121	0.026	-0.158	0.393	0.179	-0.613	0.724
	P	0.796	0.955	0.735	0.384	0.700	0.580	0.485
chemPC1	*r*	-0.268	-0.140	0.008	0.315	0.191	0.627	-0.505
	P	0.561	0.764	0.986	0.491	0.682	0.568	0.663
chemPC2	*r*	0.644	0.650	-0.586	-0.330	-0.583	0.748	-0.641
	P	0.118	0.114	0.167	0.469	0.170	0.462	0.557
chemPC3	*r*	0.543	0.289	-0.239	-0.538	0.291	0.944	-0.884
	P	0.207	0.530	0.606	0.212	0.526	0.215	0.310
Total PAH	*r*	0.501	0.308	-0.192	-0.465	-0.071	1.000	0.992
	P	0.252	0.502	0.680	0.293	0.880	0.013	0.082

[a] Regular font: not significant; inverse printing: significant ($p<0.05$).
[b] As defined in the text.

discharges, and that benthic biological effects would reflect the extent of the chemical impacts. It was the intent of this initial investigation to characterize sediment sampled from selected sites in terms of three categories: chemistry and grain size, toxicity, and benthic community structure. These results were then statistically compared to determine concordance of responses and identify *possible* causes.

Sediment chemistry

Sites for this study were selected based on the potential influence of a number of wet-weather discharge types, and the chemistry data for sediments showed a range of contamination. Based on exceedance of sediment quality guidelines for Ontario (Persaud et al. 1993), samples from several sites — SC01, RH01, HAP01, HAP02, DU03, HAP04 — were contaminated by a variety of metals and/or PAH compounds, whereas samples for other sites — DU01, DU02, RPO01 — were relatively uncontaminated (Tables 3 and 4, Fig. 2 and 3). At both the Dunker's and Harding Park wet-weather discharge treatment facilities (Fig. 1), sediment samples collected furthest "upstream" in the treatment chain (primary settling areas) were higher in contaminant concentrations compared to samples collected further "downstream" (e.g., DU03 versus DU02 versus DU01; HAP01 and HAP02 versus HAP03 versus HAP04 [metals only]). In addition, the Rouge River sample (RPO01), collected downstream of a sedimentation basin, was comparatively low in contaminants, whereas the Spencer Creek and Red Hill samples, collected from sites exposed to untreated wet-weather discharges, were among the most contaminated. These results suggest that treatment of CSO or stormwater discharges using sedimentation basins may reduce contaminant concentrations in sediments of the receiving environment. Conclusive evidence though, can only be obtained by a more complete assessment of sites upstream of these outfalls.

Sediment toxicity

In contrast to the pattern of sediment contamination among samples, toxicity did not correspond to expected degree of exposure to untreated wet-weather discharge. Greatest overall toxicity was observed in samples from sites DU01 and HAP04, which were from sites furthest "downstream" in their discharge treatment systems. As well, samples from Spencer and Red Hill creeks, which were exposed to untreated discharges, showed the lowest toxicity (Table 5), although it is possible that the sediment from these discharges may have been carried further downstream than the locations sampled. Toxicity of samples from the other sites showed minor differences. Overall, toxicity of samples was not high, and differences among sites were not apparent due to the variability of replicates.

It is possible that toxicity observed at site DU01 may be attributable to the physical nature of the sediment and, as such, particle size may be an important attribute to consider when interpreting test results. Thus, the toxicity evident for *H. azteca* at site DU01 (Lake Ontario) may be related to the high percentage of sand (and also low TOC) observed in this sediment (Table 3), rather than the presence of toxic substances. On the other hand, sandy and low-TOC sediments have lower contaminant binding capacity than fine grained, high-TOC sediments (Campbell and Tessier 1989). Therefore, available contaminants (such as As, which exceeded the SQG: lowest effect level) could be more bioavailable than in samples from other sites. It is also possible that an unmeasured contaminant is related to low *H. azteca* survival in the sample.

Benthic community structure

Benthic communities can be altered in several ways: in their taxonomic composition, in the abundance of taxa relative to each other, in the number of taxa present (richness), and in total abundance of all taxa. Decreases in the latter two attributes are the clearest signs of detrimental effects. By these measures, communities from sites exposed to untreated discharges (DU03, SC01, HAP01) were not negatively affected compared to other Ponar-sampled sites exposed to lesser amounts of untreated discharges (DU01, HAP02; Table 6, Fig. 4). Due to differences in the habitat sampled (pond, lake, stream sediment, stream riffle) and sampling method (Ponar grab, kick net), the ability to detect trends in BCS was low. This limitation and its rectification are discussed below.

Integration of Results

The multiple line of evidence approach used for this assessment of wet-weather discharges was expected to show (a) degree of contamination of sediments exposed to discharges, and (b) the extent to which contaminant concentrations explain sediment toxicity and BCS. Correlations between either toxicity test endpoints or benthic community descriptors and sediment quality descriptors were mostly non-significant and did not suggest any association between negative biological conditions and high contaminant concentrations. These results were surprising because contaminant levels for many metals and PAHs in samples from several sites exceeded provincial sediment quality guidelines ("lowest effect level" for most instances; "severe effect level" for four PAH compounds at one site). Potential explanations for the apparent lack of biological response to elevated contaminant concentrations include:

- contaminants were not bioavailable (not measured in this study);
- biota were resistant to contaminant concentrations in the samples collected;
- effects of other confounding factors (e.g., grain size, habitat type, nutrient levels) obscured contaminant effects;
- biota were not exposed to conditions characterized by sediments collected for physico-chemical analyses; and
- low sample sizes, resulting in low power to detect relationships.

While it is not possible to identify which of these explanations were responsible, some appear to be more important than others. Sediment heterogeneity was likely a major factor affecting the biological assessments. Two conditions, grain size and nutrient levels (especially TOC), are known to influence toxicity endpoints and affect BCS. Data for two endpoints in this study (*H. azteca* survival, *C. riparius* growth) were significantly correlated with grain size (Table 7). Benthic communities can differ substantially between standing and running water habitats and between substrate types (Thorp and Covich 1991). Thus, the high variability in these conditions in the sediment could overshadow or interact with contaminant effects to obscure the detection of ecotoxicological responses by the biota. Coupled with low sample sizes and lack of information on sites unexposed to wet-weather discharges (i.e., field references), the statistical power of this initial assessment was low. The issue of contaminant bioavailability could be addressed by analyzing concentrations of contaminants in tissues of biota. That could also allow better characterization of exposure to contaminants and determination of whether biota are resistant to the sediment conditions. In this study, it is not likely that biota used in the toxicity tests were exposed to different sediments than those analyzed for contaminants because they were sampled from the same locations with the same apparatus. Benthic communities sampled by kick net, however, could be exposed to quite different conditions than those in the sediment. Kick net samples include invertebrates living above sediments attached to rocks, plants and other debris. The area from which kick net samples were obtained was also much larger than that sampled by Ponar grab. Therefore, those invertebrates in the kick net sample that are exposed to sediments could be from patches of sediment different than the sediment collected for chemical analyses.

Refinement of Study Design

Superior assessments of benthic conditions would be achieved if (a) heterogeneity of sampling conditions was minimized for all factors except exposure to discharges, (b) reference sites were included in the sampling design, (c) sample size was increased, (d) tissue concentrations of contaminants were measured to estimate bioavailability of contaminants in sediments, and (e) size and land use of the drainage areas for the CSO and stormwater outfalls were determined.

Future work should incorporate the use of upstream reference sites appropriate to the outfall area being investigated.

Conclusions

These data provide a preliminary assessment of sediment quality in the vicinity of wet-weather discharge outfalls. In general there was little evidence of confirmed toxicity as a result of these discharges based on both laboratory testing of sediment and community structure analysis. The main effect of these discharges appeared to be enrichment. There was some value in using benthic toxicity testing; however, in this case, field data were equivocal because of the lack of suitable control or reference sites.

The Chironomidae, Tubificidae and Naididae represented the most abundant families in these sediments (with the exception of HAP02). Chironomids and tubificids have been known to dominate in areas of higher organic pollution (Brinkhurst and Gelder 1991). The overall assessment was that these sites had been impacted by the wet-weather pollution, and that these impacts manifested themselves as deviations from the normal population structure with strong tendencies towards pollutant-tolerant species (Johnson et al. 1993). Further testing and research may help improve definitions for impacted sites.

Acknowledgments

We would like to acknowledge the advice and financial support of Environment Canada's Great Lakes 2000 Cleanup Fund. We would also like to thank Gordia MacInnis for her support in the PAH analyses.

References

American Public Health Association (APHA). 1989. Standard methods for the examination of water and wastewater, 17th ed. APHA, Washington, DC.

Beanlands GE, Duinker PN. 1983. An ecological framework for environmental impact assessment in Canada. Institute for Resource and Environmental Studies, Dalhousie University, Halifax, NS.

Belbin L. 1991. Semi-strong hybrid scaling, a new ordination algorithm. J. Vegetation Sci. 2:491–496.

Borchardt D, Sperling F. 1997. Urban stormwater discharges: ecological effects on receiving waters and consequences for technical measures. Water Sci. Technol. 36:173–178.

Boxall ABA, Maltby L. 1995. The characterization and toxicity of sediment contaminated with road runoff. Water Res. 29(9):2043-2050.

Boxall ABA, Maltby L. 1997. The effects of motorway runoff on freshwater ecosystems: 3. Toxicant confirmation. Arch. Environ. Contam. Toxicol. 33:9-16.

Brinkhurst RO, Gelder SR. 1991. Annelida: Oligochaeta and Branchiobdellida, p. 401–435. *In* Thorp JH, Covich AP (ed.), Ecology and classification of North American freshwater invertebrates. Academic Press, Inc., San Diego, CA.

Campbell PGC, Tessier A. 1989. Geochemistry and bioavailability of trace metals in sediments, p. 125–150 *In* Boudou A, Ribeyre F (ed.), Aquatic toxicology: Fundamental concepts and methodologies, Vol. 1. CRC Press, Boca Raton, FL.

Clarke KR, Green RH. 1988. Statistical design and analysis for a biological effects study. Marine Ecology Progress Series 46:213–226.

Davis JR. 1997. Revitalization of a northcentral Texas river, as indicated by benthic macroinvertebrate communities. Hydrobiologia 346:95–117.

Green RH. 1979. Sampling design and statistical methods for environmental biologists. John Wiley & Sons, New York.

Green RH, Montagna P. 1996. Implications for monitoring: study designs and

interpretation of results. Can. J. Fish. Aquat. Sci. 53:2629–2636.
Hall K, McCallum D, Lee K, MacDonald R. 1998. Characterization and aquatic impacts of combined sewer overflows in greater Vancouver, British Columbia. Water Sci. Technol. 38:9–14.
Hatch AC, Burton GA. 1999. Sediment toxicity and stormwater runoff in a contaminated receiving system: consideration of different bioassays in the laboratory and field. Chemosphere 39:1001–1017.
Johnson RK, Wiederholm T, Rosenberg DM. 1993. Freshwater biomonitoring using individual organisms, populations, and species assemblages of benthic macroinvertebrates, p. 40–158. *In* Rosenberg DM, Resh VH (ed.), Freshwater biomonitoring and benthic macroinvertebrates. Chapman and Hall, New York.
Lee P-K, Touray J-C, Ballif P, Ildefonse J-P. 1997. Heavy metal contamination of settling particles in a retention pond along the A-71 motorway in Sologne, France. Sci. Tot. Env. 201:1–15.
Legendre L, Legendre P. 1983. Numerical ecology. Elsevier, New York.
Maltby L, Boxall ABA, Forrow DM, Calow P, Betton CI. 1995. The effects of motorway runoff on freshwater ecosystems: 2. Identifying major toxicants. Environ. Toxicol. Chem. 14(6):1093-1101.
Marchant R. 1989. A subsampler for samples of benthic invertebrates. Bull. Australian Soc. Limnol. 12:49–52.
Marsalek J, Rochfort Q, Mayer T, Servos M, Dutka B, Brownlee B. 1999. Applications of toxicity testing in controlling urban wet-weather pollution: advantages and limitations. Urban Water 1:91–104.
McIntyre AD. 1984. What happened to biological effects monitoring? Mar. Pollut. Bull. 15: 391–392.
Mulliss RM, Revitt DM, Shutes RB. 1996. A statistical approach for the assessment of the toxic influences on *Gammarus pulex* (amphipoda) and *Asellus aquaticus* (isopoda) exposed to urban aquatic discharges. Water Res. 30:1237–1243.
Persaud D, Jaagumagi R, Hayton A. 1993. Guidelines for the protection and management of aquatic sediment quality in Ontario. Ontario Ministry of the Environment and Energy. August 1993.
Reynoldson TB, Thompson SP, Bamsey JL. 1991. A sediment bioassay using the tubificid oligochaete worm *Tubifex tubifex*. Environ. Toxicol. and Chem. 10:1061–1072.
Reynoldson TB, Day KE, Clarke C, Milani D. 1994. Effect of indigenous animals on chronic end points in freshwater sediment toxicity tests. Environ. Toxicol. and Chem. 13:973–977.
Reynoldson TB, Day KE. 1998. Biological guidelines for the assessment of sediment quality in the Laurentian great lakes. NWRI Contribution No. 98-232. National Water Research Institute, Burlington, ON.
Reynoldson TB, Logan C, Pascoe T, Thompson SP. 1999a. Methods manual: I. The creation of reference-condition data bases for benthic invertebrate community structure. NWRI Contribution No. 99-211. National Water Research Institute, Burlington, ON.
Reynoldson TB, Logan C, Milani D, Pascoe T, Thompson SP. 1999b. Methods manual: II. Sediment toxicity testing, field and laboratory methods and data management. NWRI Contribution No. 99-212. National Water Research Institute, Burlington, ON.
Schueler TR, Kumble PA, Heraty MA (Anacostia Restoration Team). 1992. A current assessment of urban best management practices, techniques for

reducing non-point source pollution in the coastal zone. Metropolitan Washington Council of Governments, Washington, DC.

Seager J, Maltby L. 1989. Assessing the impact of episodic pollution. Hydrobiologia 188/189:633–640.

SWAMP. 1999. Performance assessment of Richmond Hill's Harding park stormwater retrofit pond. Stormwater Assessment Monitoring and Performance Program (SWAMP), in press.

Thorp JH, Covich AP. 1991. An overview of freshwater habitats, p. 17–36. *In* Thorp JH, Covich AP (ed.), Ecology and classification of North American freshwater invertebrates. Academic Press, Inc., San Diego, CA.

Environment Canada Environnement Canada

Canada

Canada Centre for Inland Waters
P.O. Box 5050
867 Lakeshore Road
Burlington, Ontario
L7R 4A6 Canada

National Hydrology Research Centre
11 Innovation Boulevard
Saskatoon, Saskatchewan
S7N 3H5 Canada

St. Lawrence Centre
105 McGill Street
Montreal, Quebec
H2Y 2E7 Canada

Place Vincent Massey
351 St. Joseph Boulevard
Gatineau, Quebec
K1A 0H3 Canada

Centre canadien des eaux intérieures
Case postale 5050
867, chemin Lakeshore
Burlington (Ontario)
L7R 4A6 Canada

Centre national de recherche en hydrologie
11, boul. Innovation
Saskatoon (Saskatchewan)
S7N 3H5 Canada

Centre Saint-Laurent
105, rue McGill
Montréal (Québec)
H2Y 2E7 Canada

Place Vincent-Massey
351 boul. St-Joseph
Gatineau (Québec)
K1A 0H3 Canada