Evaluating methods for assessing sediment quality in a Great Lakes embayment

Judy L. Crane,[1*] Carl Richards,[2**] Dan Breneman,[3] Stephen Lozano,[4†] and Jeffrey A. Schuldt[3‡]

[1]*Minnesota Pollution Control Agency, Environmental Analysis and Outcomes Division, 520 Lafayette Road North, St. Paul, MN 55155-4194, Phone 651 297 4068; Fax 651 297 7709*

[2]*University of Minnesota Duluth, Minnesota Sea Grant College Program, 2305 East Fifth Street, Duluth, MN 55812*

[3]*University of Minnesota Duluth, Natural Resources Research Institute, 5013 Miller Trunk Highway, Duluth, MN 55811*

[4]*U.S. Environmental Protection Agency, Environmental Effects Research Laboratory, Mid-Continent Ecology Division, 6201 Congdon Boulevard, Duluth, MN 55804*

[**]*Current address: U.S. Environmental Protection Agency, Mid-Continent Ecology Division, 6201 Congdon Boulevard, Duluth, MN 55804.*

[†]*Current address: National Oceanic and Atmospheric Administration, Great Lakes Environmental Research Laboratory, 2205 Commonwealth Boulevard, Ann Arbor, MI 48105-2945*

[‡]*Current address: University of Wisconsin Superior, Department of Biology and Earth Sciences, Belknap and Catlin, Superior, WI 54880-4500*

[*]*Corresponding author: E-mail: judy.crane@pca.state.mn.us*

A probability-based, sediment quality assessment was conducted during 1995 in the lower St. Louis River Area of Concern, located in western Lake Superior. A regional application of the intensified sampling grid developed for the United States Environmental Protection Agency's Environmental Monitoring and Assessment Program was used to randomly select 90 sites for measuring the following sediment quality indicators: sediment chemistry, physical parameters, sediment toxicity, and benthic macroinvertebrate community structure. Screening methods were used to assess sediment chemistry and sediment toxicity at all sites, whereas more conventional metrics were used at a subset of sites. In addition, sediment quality data were collected from 20 a priori training sites, 10 in low impact areas and 10 in high impact areas. Mean probable effect concentration quotients were calculated for sediment chemistry variables at each site. As the range of mean probable effect concentration quotients values increased, the incidence of sediment toxicity increased. Benthic data from the training sites were used to establish standard criteria for developing two benthic integrity indices based on multimetric analysis and discriminant function analysis. Based on the training site results, the discriminant function analysis categorized the macroinvertebrate community at all random sites as 45 percent low impact and 55 percent high impact. A multimetric approach categorized 55 percent of the random sites as low impact and 36 percent as high impact. Due to the overlap of 95 percent confidence intervals, the multimetric approach also placed 9 percent of the random sites into an indeterminate category. The incidence of high impact sites appears to be primarily due to physical habitat characteristics. This finding was supported by the sediment quality triad assessment of 52 random sites that indicated alteration of the benthic community at 71 percent of sites was probably not due to chemical contamination.

Keywords: benthic macroinvertebrates, Lake Superior, toxicity tests, chemistry, multimetric index, discriminant function analysis

Aquatic Ecosystem Health & Management, 8(3):323–349, 2005. ISSN: 1463-4988 print / 1539-4077 online
DOI: 10.1080/14634980500242852

Introduction

Many embayments and harbors within the Great Lakes have been impacted by a variety of environmental stressors, resulting in possible impairments due to increased eutrophication, alteration of aquatic and nearshore habitats, expansion of exotic species populations, bioaccumulation of persistent chemicals in fish tissue, and fluctuating water levels. The St. Louis River Area of Concern (AOC) is one of 41 current Great Lakes AOCs in which impaired use criteria, identified by the International Joint Commission (IJC), have been observed (IJC, 1989). Contaminated sediments have contributed to several use impairments in this western Lake Superior AOC, including the issuance of fish advisories, restrictions on dredging, and habitat impairments to bottom-feeding organisms (MPCA, WDNR, 1992, 1995; Crane et al., 2002a). Chemicals of potential concern in portions of the lower St. Louis River AOC include: polycyclic aromatic hydrocarbons (PAHs), polychlorinated biphenyls (PCBs), dioxins and furans, diesel range organics, pesticides (e.g., DDT metabolites, toxaphene), trace metals (copper, lead, mercury, nickel, zinc), and tributyltin (Crane et al., 2002a).

The IJC, through a formal protocol agreement between Canada and the United States, is charged with reviewing the remedial action plans (RAPs) for each Great Lakes AOC. The RAPs are being prepared in a staged approach to evaluate impaired uses, to develop and implement a plan for restoring beneficial uses, and to evaluate the success of any remedial measures that are conducted. A Stage I RAP for the St. Louis River AOC was completed in 1992 (MPCA, WDNR, 1992), and a three-phase strategy to reduce sediment-related impairments was recommended in the Stage II RAP (MPCA, WDNR, 1995).

The United States Environmental Protection Agency's (USEPA) Assessment and Remediation of Contaminated Sediments Program recommends using a weight-of-evidence approach to assess contaminated sediment sites (Fox et al., 1994). The Sediment Quality Triad (Triad) approach is one way in which sediment quality can be characterized through synoptic chemical and physical analyses, whole-sediment toxicity tests, and benthic infauna community surveys (Chapman et al., 1987,1997; Long and Wilson, 1997). The Triad approach is effective in addressing bioavailability and contaminant toxicity (Long and Chapman, 1985; IJC, 1989; Burton and Ingersoll, 1994; Canfield et al., 1994a,b, 1996) and it is considered a comprehensive and ecologically relevant metric for site characterizations. In addition, the Triad approach is valuable for status and trends monitoring of aquatic ecosystems because it reflects highly localized conditions that are relatively static over short periods of time.

Although the Triad approach has been applied to a focused assessment of contaminated sites in the lower St. Louis River AOC (Crane et al., 1997), a large scale Triad study was needed for status and trends monitoring throughout this AOC. The resource status information obtained from status and trends monitoring will help evaluate restoration efforts of impaired aquatic habitats at several sites within the lower St. Louis River AOC, including two Superfund sites. These restoration efforts will also benefit from implementation of the lower St. Louis River Habitat Plan (St. Louis River CAC, 2002).

One approach to determining resource status is through statistical sampling design principles and selection of appropriate status indicators employed in the Great Lakes Environmental Monitoring and Assessment Program (EMAP; USEPA, 1993). This process provides a means for identifying and sampling the appropriate number of sites to acquire the baseline information necessary to measure progress in restoring impaired areas. Combining these design principles with a comprehensive examination of the physical, chemical, and biological characteristics of the sediments should provide sufficient data to make sound environmental decisions.

We previously reported on environmental influences on benthic community structure in the lower St. Louis River AOC as part of a Regional EMAP (R-EMAP) investigation (Breneman et al., 2000). In this paper, we describe other objectives of the R-EMAP project to: determine impaired areas using the Triad approach, provide status and trends data for evaluating sediment quality in the lower estuary, and to provide the first test of EMAP concepts to the Great Lakes Harbors and Embayments resource class. This paper will describe methods used to calculate mean probable effect concentration quotients (PEC-Qs) for total PAHs and/or metals, identify sites having significant acute toxicity, and develop two benthic indices based on multimetric and multivariate analyses. The results from 20 a priori training sites chosen in low and high impact areas will be used to evaluate sediment quality conditions associated with randomized sites throughout the lower St. Louis River AOC.

Materials and methods

Study area

The St. Louis River is a fifth order system that lies on the western most expanse of Lake Superior, bordering Duluth, MN and Superior, WI (Figure 1). The confluence of the St. Louis and Nemadji Rivers with Lake Superior forms a freshwater estuary that covers approximately 4,856 ha. This area is comprised of numerous large bays, peninsulas, and islands. Much of the lower estuary (below the Fond du Lac dam) is characterized as a shallow embayment ($\sim$2 m depth) consisting of over 166 km of shoreline. The lower estuary culminates in the Duluth-Superior Harbor, which is one of the most heavily used ports in the Great Lakes basin. The only habitats within the lower estuary reaching depths greater than 2 m are the St. Louis River natural channel (i.e., non-maintained portion of the St. Louis River channel upstream to Fond du Lac, MN), several commercial boat slips and marinas, deep holes resulting from the excavation of gravel in the harbor, portions of Superior Bay, and designated navigation channels (St. Louis River CAC, 2002). About 114,700 m^3 of sediments are dredged annually from the navigation channels, primarily in the outer harbor (USACE, 1997).

The lower estuary supports a variety of industrial, commercial, residential, and recreational activities. In addition, these areas provide essential habitats for aquatic organisms and aquatic-dependent wildlife species. However, aquatic habitats in some of these areas have been adversely affected by economic development of the St. Louis River over the past 130 years. A number of point and nonpoint sources of nutrients and chemicals of potential concern to the St. Louis River AOC were identified during the Stage I RAP (MPCA, WDNR, 1992). A mixture of elevated sediment contaminants have been observed in several localized areas, especially at the Interlake/Duluth Tar and USS Superfund sites in the inner Duluth

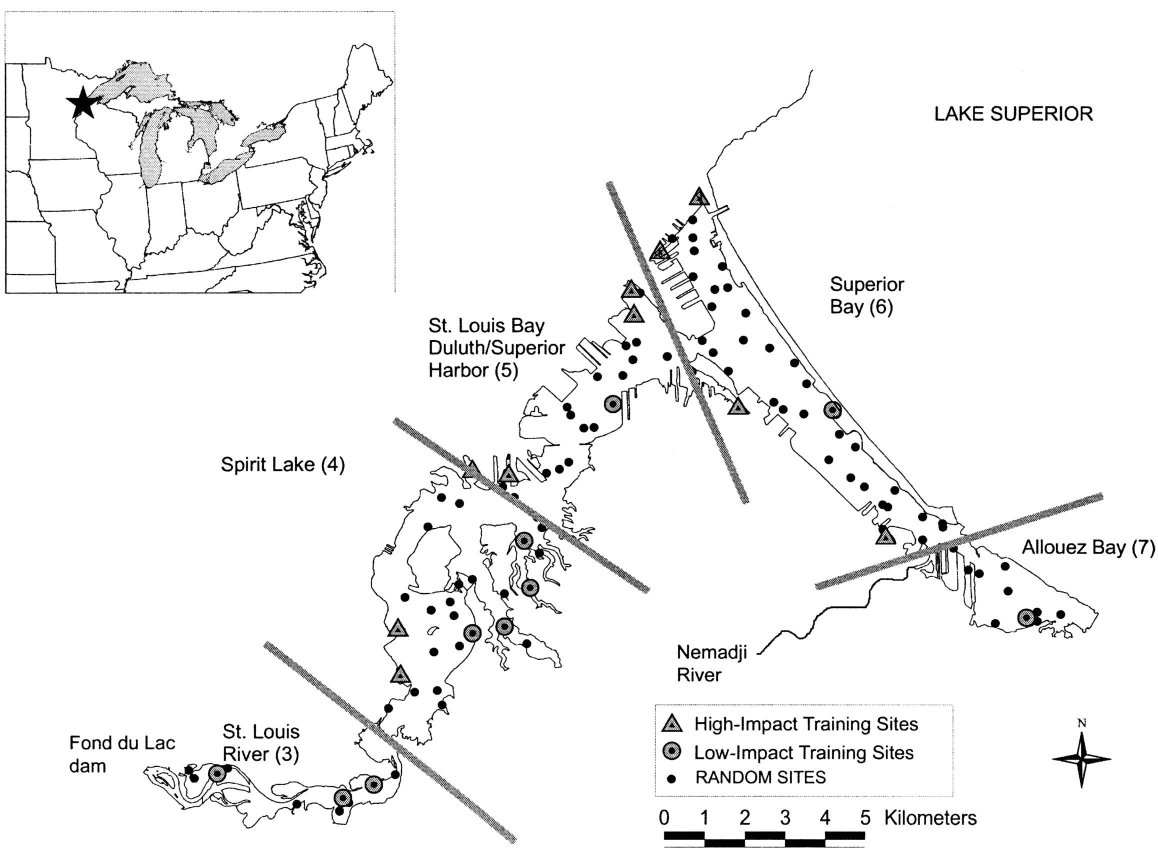

Figure 1. The St. Louis River Area of Concern (AOC) longitudinally delineated into distinct habitat areas based on ecological conditions and anthropogenic modification. Moving from upstream to downstream the areas include; 3) St. Louis River, 4) Spirit Lake, 5) Duluth/Superior Harbor, 6) Superior Bay, and 7) Allouez Bay.

Harbor (Schubauer-Berigan and Crane, 1997; IT Corporation, 1997; URS Corporation, 2002). Other areas with elevated sediment contamination in the Duluth-Superior Harbor include Hog Island Inlet/Newton Creek (Redman and Janisch, 1995; SEH Inc., 2000, 2003), as well as several boat slips, areas adjacent to wastewater treatment plants, and other areas with historical sources of contaminants (Crane et al., 1997; Schubauer-Berigan and Crane, 1997; Crane, 1999; Crane et al., 2002b).

Experimental design and site selection

Our experimental design used a set of a priori selected sites (training sites) to classify a set of random sites that were selected following a randomized sampling design based on the format used in the Great Lakes EMAP system (USEPA, 1993). The randomly selected sites provided a means for generally characterizing condition of the region. The training sites (10 low impact and 10 high impact) were chosen in an attempt to categorize the best and worst case sediment quality conditions existing in the lower estuary. Training sites were selected based upon previous sediment quality studies of the region (MPCA, WDNR, 1992; Schubauer-Berigan and Crane, 1996, 1997; Crane et al., 1997) and upon the best professional judgment of local resource professionals. Much of this professional judgment about local sediment quality and habitat conditions has been documented in the lower St. Louis River Habitat Plan (St. Louis River CAC, 2002).

Random sites were chosen based on the National Oceanic and Atmospheric Administration (NOAA) shipping chart for the Duluth-Superior Harbor which was digitally categorized into shallow (water depth <5.5 m) or channelized (>5.5 m in depth) habitats. A 49-fold enhancement of the Great Lakes EMAP grid was applied to accessible areas, and the number of sites evaluated within each habitat class was based on the total surface area represented by that particular class. Ninety sample site locations were randomly selected from the resulting hexagonal grid, with each site positioned by a latitudinal and longitudinal coordinate. Sixty random sites were located in the shallow habitat class, while 30 random sites were located in the channel habitat of the federal navigation channel. Sites were stratified into shallow and channel due to known influence of this physical feature on sediment conditions.

To aid in summarizing results, site locations within the St. Louis River AOC were also delineated, upstream to downstream, into five areas (Figure 1). As described in the following sections, not all of the metrics for these indicators were measured at all sites. Instead, screening methods were used to develop a short-list of sites for running comprehensive sediment toxicity tests and chemical analyses.

Sediment sampling

All sediment samples were collected during a 30 d period, beginning in June 1995 as described by Breneman et al. (2000). In brief, composited sediment samples of the upper 5 cm depth interval were split for sediment chemistry analyses, physical measurements, and sediment toxicity tests. Field replicates for sediment chemistry, particle size, and sediment toxicity were collected at approximately every 10 sites, for a total of 11 replicates (8 random sites and 3 training sites). Benthic macroinvertebrates were sampled at the same time in triplicate using a 0.023 m^2 petite Ponar grab sampler (Breneman et al., 2000). The storage and processing of sediment samples are described in Breneman et al. (2000). A Global Positioning System was used to determine latitude and longitude coordinates (Breneman et al., 2000). Water depth and the depth of soft sediments were also measured at each site (Breneman et al., 2000). Sediment samples were collected from all 20 training sites for sediment chemistry, particle size, sediment toxicity, and benthic macroinvertebrates. From the original 90 random sample locations, sediment chemistry and Microtox® samples were obtained and analyzed for 87 sites, including 58 shallow and 29 channelized habitats. Cohesive sediment samples could not be obtained from three random sample sites because the substrates were composed of unconsolidated sand and gravel. Extra sediment samples were also collected at all 87 sites for possible use in 10 d sediment toxicity tests. Benthic macroinvertebrate samples were collected from 89 random sites.

Sediment chemistry and physical parameters

The procedures for analyzing sediment samples for sediment chemistry and physical parameters were described in detail in Breneman et al. (2000). In brief, simultaneously extractable metals (SEM, sum of Cd, Cu, Ni, Pb, Zn concentrations in μmol g^{-1}), acid volatile sulfide (AVS), mercury, percent total organic carbon (TOC), KCl-extractable ammonia, screening PAHs (as measured using a fluorescence procedure; Peterson et al., 2002), and particle size (six size classes) were measured on sediment from all random and training sites. The following PAH compounds were measured by GC/MS-SIM on 42 random and 9 training site samples: acenaphthene, acenaphthylene, anthracene, benz(a)anthracene, benzo(b)fluoranthene, benzo(k)-fluoranthene, benzo(g,h,i)perylene, benzo(a)pyrene, chrysene, dibenz(a,h)anthracene, fluoranthene, fluorene, indeno(1,2,3-cd)pyrene, naphthalene, perylene, phenanthrene, and pyrene. The analysis of PCBs, by the immunoassay method, was dropped from this study because a comparison of the results of PCBs measured by GC/ECD in an earlier contaminated sediment survey of the Duluth-Superior Harbor (Schubauer-Berigan and Crane, 1997) indicated that approximately 90% of the sites sampled were below the detection limit of the immunoassay method.

Low molecular weight (LMW) PAHs were calculated as the sum of the dry weight concentrations of acenaphthene, acenaphthylene, anthracene, fluorine, naphthalene, and phenanthrene (2-methylnaphthalene, the other LMW PAH, was not measured in this study). High molecular weight (HMW) PAHs were calculated as the sum of the dry weight concentrations of benz(a)anthracene, benzo(a)pyrene, chrysene, dibenz(a,h)anthracene, fluoranthene, and pyrene. Total PAHs were calculated as both the sum of the LMW and HMW PAHs ($TPAH_{12}$) and as the sum of the 17 PAHs measured in this study ($TPAH_{17}$).

A statistical summary of the dry weight concentrations of individual PAHs and $TPAHs_{17}$ was completed by calculating the arithmetic mean and standard deviation of random samples in each habitat class. In addition, the range of low and high values was determined for each habitat class. A t-test ($\alpha = 0.05$) was conducted on the mean $TPAH_{17}$ concentrations in the shallow and channelized habitats using SigmaStat® Version 3.0. Ratios of phenanthrene/anthracene and fluoranthene/pyrene were calculated for each sample, and the arithmetic mean and standard deviation were determined for the random samples in each habitat class. The percentage composition of individual PAHs was calculated by normalizing the individual PAH concentrations by the $TPAH_{17}$ value for each habitat class, and multiplying the result by 100. Summary statistics of the arithmetic mean, standard deviation, and range were determined for both habitat classes.

Chemical and physical parameter data from all random sites were pooled together for conducting linear regression analyses using SigmaPlot® 8.0. Ninety-five percent confidence intervals and prediction intervals were plotted, but no data points exceeding these intervals were excluded from the regression analyses.

Mean probable effect concentration quotients (PEC-Qs)

A subset of the sediment chemistry data were compared to Level II sediment quality targets (SQTs) adopted for use in the St. Louis River AOC (Crane et al., 2000, 2002a). The Level II SQTs are intended to identify contaminant concentrations above which harmful effects on sediment-dwelling organisms are likely to be frequently (i.e., $\geq$75%) or always observed. Most of the Level II SQTs were composed of consensus-based probable effect concentrations (PECs; MacDonald et al., 2000). The PEC values were determined to be reliable if the predictive ability was $\geq$75% for $\geq$20 samples (MacDonald et al., 2000). The predictive ability of the numerical SQTs was evaluated using the matching sediment chemistry and toxicity data set for the St. Louis River AOC as described in Crane et al. (2000, 2002a). This evaluation involved determination of the incidence of toxicity to amphipods (*Hyalella azteca*) and midges (*Chironomus dilutus,* formally known as *C. tentans*) within five ranges of Level II SQT quotients (i.e., mean PEC quotients, PEC-Qs). The incidence of toxicity was determined based on the results of 10 d toxicity tests with amphipods (endpoints: survival and growth) and 10 d toxicity tests with midges (endpoints: survival and growth). For both toxicity tests, the incidence of toxicity increased as the mean PEC-Q ranges increased. Thus, the predictive ability of the Level II SQTs (through the use of mean PEC-Qs) was improved when the SQTs for the various chemicals of potential concern were used together to classify sediments from the St. Louis River AOC (Crane et al., 2000, 2002a).

For this study, an evaluation was conducted to determine the incidence of toxicity and incidence of high impact sites (based on macroinvertebrate indices) within the following ranges of mean PEC-Qs: $\leq$0.1, >0.1 to $\leq$0.5, >0.5 to $\leq$1.0, >1.0 to $\leq$5.0, and >5.0. These ranges were analogous to the mean PEC-Q ranges used by Ingersoll et al. (2001) and Crane et al. (2000, 2002a) to evaluate the predictive ability of freshwater sediment quality guidelines and SQTs. In this evaluation, mean PEC-Qs were calculated using the methods that were recommended by Ingersoll et al. (2001) and outlined in Crane et al. (2000, 2002a). In brief, mean PEC-Qs were calculated as follows:

$$\text{PEC-Q} = \frac{\text{chemical concentration (dry wt.)}}{\text{corresponding PEC value}}$$

$$\text{mean PEC-Q} = (\text{mean PEC-Q}_{\text{metals}} + \text{PEC-Q}_{\text{Total PAHs}} + \text{PEC-Q}_{\text{Total PCBs}})/n$$

where n = number of classes of chemicals for which sediment chemistry data were available (i.e., 1 to 3).

Only the metals (i.e., arsenic, cadmium, chromium, copper, lead, nickel, and zinc) for which reliable Level II SQTs were available were used to calculate mean PEC-Qs (Crane et al., 2000, 2002a). As a conservative estimate, SEM concentrations were assumed to be equivalent to total metal concentrations for Cd, Cu, Ni, Pb and Zn. This assumption was based on evaluating a sediment quality data set for the Mississippi River (Canfield et al., 1998) in which both SEM and total metals were measured. For each sample, the ratios of the SEM and total concentration for each of the five metals were calculated (D. MacDonald, pers. comm., 2004). The average ratio was then calculated for each metal. The results of these analyses showed that the ratios for the various metals averaged about 70 to 75%, which was considered within the range of analytical variability. Thereafter, SEM and total metals were treated as functionally equivalent in a North America-wide database of matching sediment chemistry and toxicity data (i.e., the SEDTOX database; (Field et al., 2002; MacDonald et al., 2002)).

The PEC for total PAHs was used in the calculation to avoid double counting the individual PAH PEC values (MacDonald et al., 2000). Since PCBs were not measured in this study, the mean PEC-Q values were based on two classes of chemicals.

Descriptive statistics of mean PEC-Q values for the random sites (shallow and channelized habitats) and training sites (low impact and high impact) were conducted using SigmaStat® Version 2.0. A Kruskal-Wallis one way analysis of variance (ANOVA) on ranks of median mean PEC-Qs was conducted on the four groups of sites since the normality test failed ($p < 0.001$) for conducting an ANOVA on mean PEC-Qs for these groups. A pairwise multiple comparison (Dunn's Method) was used to isolate the group or groups that differed from the others.

Toxicity testing

Microtox® tests: Two Microtox® acute toxicity screening assays were conducted on all random and training site sediment samples: 1) the solid-phase test, and 2) a 90% pore water screening test. Both the solid-phase and pore water assays measure the reduction in bioluminescence of the marine bacteria *Photobacterium phosphoreum*. For the solid-phase test, methods followed those described in Brouwer et al. (1990) and Tung et al. (1990). Bacteria were directly exposed to an aqueous sediment suspension for 20 min., after which the *P. phosphoreum* were separated from the sediment by a filter column. The filtrate was then analyzed using the Microtox® toxicity meter (model M500). Data from the test were expressed as an EC_{50}, which is defined as the effective concentration of a sample that causes a 50% reduction in luminescence relative to the control. Samples producing an EC_{50} less than 0.5% (±0.05%) sediment were considered toxic. The 90% pore water screening test followed the methods described in the product manual (Microbics Corp., 1994). Pore water was prepared by centrifugation and was frozen until time of analysis. Tests were read at 5 and 15 min. A reduction in light output of 20% (±0.5%) or greater relative to the controls was considered a toxic response. The criteria chosen for toxic responses in both tests were based on technical discussions with staff from Microbics Corp.

Ten-day sediment toxicity tests: The results of the solid-phase and pore water Microtox® tests were used to screen sediment samples for 10 d sediment toxicity tests. Sediment toxicity tests were run on a proportion of the samples that displayed a toxic response in both Microtox® tests (100%), in only the solid-phase test (74%), and in only the pore water test (71%). To reduce the chance of missing toxic samples due to false negative results, sediment toxicity tests were run on 41% of sites displaying nontoxic responses in both Microtox® tests. In all, ten-day sediment toxicity tests were conducted on fourteen training sites (all in shallow areas) and fifty-two random sites (36 shallow and 16 channelized habitat sites). The test organisms were *H. azteca* (7 to 14 d old) and *C. dilutus* (third instar or younger, with at least 50% of the organisms at third instar). Survival endpoints were compared to organisms similarly exposed to control sediment collected from West Bearskin Lake (Cook County, MN; Ankley et al., 1994a,b). The test organisms were exposed to sediment samples in a portable, mini-flow system (Benoit et al., 1993; USEPA, 1994) using modified procedures of USEPA (1994) and Crane et al. (1997). The test apparatus consisted of 300-ml glass beakers held in a glass box supplied with water from an acrylic plastic headbox. The test set-up could accommodate a batch of replicates for the control and up to 11 test sediments for each of the *H. azteca* and *C. dilutus* tests, which were run concurrently. In order to conduct the 66 sediment toxicity tests for this study, seven batches of tests were run.

For each batch of toxicity tests, sediments were homogenized by hand, and about 100 ml of sediment was added to a 300-ml test beaker. Each sediment test was set up with four replicates of *H. azteca* and four replicates of *C. dilutus*. Approximately 170 ml of aerated, artesian well water was added to each beaker, and the sediments were allowed to settle from two hr to overnight. The test period began when ten individual organisms were randomly added to each beaker. Tests were conducted at 23°C on a 16L:8D photoperiod. Each day, two liters of aerated, artesian well water were exchanged through each glass box via the headbox delivery system. *Hyalella azteca* were fed 1 ml of yeast-cerophyll-trout chow (YTC) each day, whereas *C. dilutus* were fed 1 ml of Tetrafin daily. Temperature, pH, and dissolved oxygen were measured in each treatment daily in the overlying water of one replicate beaker. Similarly, conductivity was measured on days 0 and 10 for the last two batches of toxicity tests. At test termination, test sediment was sieved through a 425 μm mesh screen and all organisms removed. Organisms not accounted for were presumed to be dead, and percent survival was determined. The removed *C. dilutus* for each treatment were placed in aluminum weighing pans, dried at approximately 100°C for at least 4 h, desiccated to room temperature, and weighed.

A 96 h reference toxicant (NaCl) test was run in conjunction with each batch of toxicity tests to determine the condition and sensitivity of the *H. azteca* culture (USEPA, 1994). Comparable reference toxicant (NaCl) tests, using *C. dilutus*, were run with the last two batches of toxicity tests.

The 10 d survival data were statistically analyzed using TOXSTAT (Gulley and WEST, Inc., 1994). All survival data were expressed as a proportion and transformed using an arc sine-square root transformation prior to analysis. The Shapiro-Wilk's test for normality and Bartlett's test for homogeneity of variance were run on the transformed data. Next, an Analysis of Variance (ANOVA) was conducted. Statistical guidelines given in USEPA (1994) and Gilbert (1987) were used to select the most appropriate parametric or nonparametric one-tailed statistical test. A sample was considered toxic when mean percent survival was significantly lower ($p = 0.05$) than the corresponding response in the control sediment. Due to a quality assurance/quality control issue, the growth data were not analyzed. The LC_{50} values for the reference toxicant tests were determined by the Trimmed Spearman-Karbar Method.

Benthic macroinvertebrates

Benthic habitats in the St. Louis River AOC were examined by evaluating macroinvertebrate abundance at 89 randomized sites. Benthic community structure and functional assemblages associated with the St. Louis River AOC were compared to the 10 best available (low impact) sites, as well as 10 impaired-use (high impact) sites occurring throughout the harbor. Experimental design, collection techniques for all sample types, and laboratory processing procedures are described by Breneman et al. (2000). In summary, benthic samples were collected in triplicate using a petite Ponar dredge. Macroinvertebrates were sorted by hand, subjected to QA/QC protocols, identified to genus whenever possible, and total numbers/sample determined.

Benthic invertebrate data were used to generate a set of 21 biological metrics that described the macroinvertebrate community (Table 1). Biological metrics used to develop an overall score reflect the trophic status, functional feeding group distribution, behavioral attributes, and taxonomic classifications associated with the benthic community. Information to assign invertebrate community characteristics was taken from ecological texts (Hilsenhoff, 1981; Wiederholm, 1983; Brinkhurst, 1986; Thorp and Covich, 1991; Merritt and Cummins, 1996). Benthic sediment samples from a single site at the mouth of the Superior entry in Allouez Bay were abandoned due to an inability to capture sediment particles. Invertebrate data from two sites that lacked corresponding sediment chemistry and toxicity data were also excluded from further analysis.

Table 1. Biological metrics created from the benthic macroinvertebrates collected in the lower St. Louis River AOC.

Biological metrics	
Abundance values (no. m^{-2})	Taxonomic richness per site
By Taxonomic Group	By Mechanistic Behavior
Amphipoda	Burrowers
Chironomidae (Diptera)	Climbers
Diptera	Clingers
Ephemeroptera	Sprawlers
Isopoda	
Mollusca	By Habitat Preference
Oligochaeta	Obligate Depositional Taxa
Polychaeta	
Trichoptera	Obligate Erosional Taxa
	Mean Number of Taxa
By Trophic Group	
Carnivores	
Detritivores	
Herbivores	
Omnivores	
Total Macroinvertebrates	

Two methods were used to classify random site sediment conditions based on macroinvertebrate community comparisons with training site results. The first method utilized a modification of a multi-metric approach described by Gerritsen (1995). Metrics have been used extensively with fish and benthic communities (Barbour et al., 1995) and are encouraged by the US EPA as a means of developing biological indicators for evaluating aquatic system health and sediment quality (Barbour et al., 1999). The second method utilized a multivariate procedure (discriminant function analysis; SAS, 1988) to classify the random sites based on the community structure associated with the best available conditions as described by Reynoldson et al. (1997). For this study, the calibration procedure for establishing reference conditions was modified. The technique of using the best available conditions from among all sites sampled was simply replaced with a separate training site data set. This technique was used to quantitatively assign an upper boundary from low-impact sites and a lower boundary from high-impact sediments within the harbor.

Multimetric biotic index

The 21 biological community metrics (Table 1) were subjected to t-tests to determine whether significant ($p \leq 0.1$) differences existed between the 10 low impact and 10 high impact sites. Metrics were first transformed to meet equality of variance assumptions, and a significance level of $p = 0.1$ was used to include as many metrics as possible in the final score. Only metrics that distinguished between low-impact and high-impact sites (Figure 2) were retained for further index development.

Biotic scores were constructed by assigning individual rankings for each random site based on a metrics distribution within the boundaries set by the training site data. Specifically, random site data were plotted among the means and 95% confidence intervals (CI) from both the low-impact and high-impact site results. Rankings were assigned as follows; 1) metric values greater than the lower 95% CI from the low-impact sites were assigned a score of '5,' 2) metric values less than the upper 95% CI for high-impact sites were

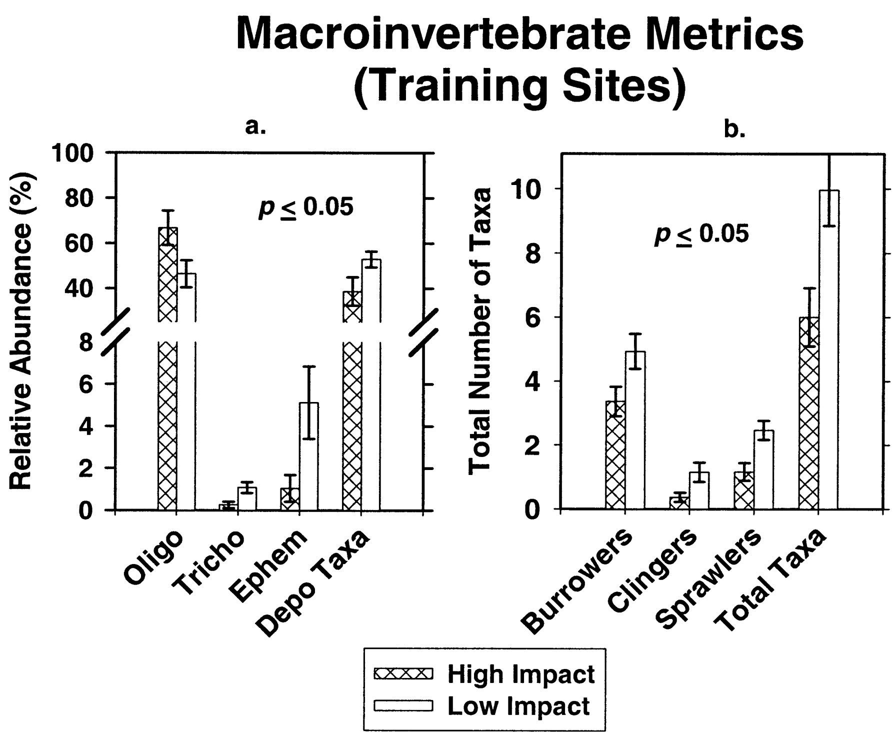

Figure 2. Benthic macroinvertebrate variables used to create biological index scores for sample sites throughout the St. Louis River Area of Concern (AOC). Variables expressed are means ± one standard error based on results of Student's t-Tests ($p < 0.1$). a) Includes the relative abundance (as percent of total abundance) of Oligo = Oligochaeta, Tricho = Trichoptera, Ephem = Ephemeroptera, Depo Taxa = taxa associated with depositional sediments. b) Represents the total number of burrowing taxa (burrowers), clinging taxa (clingers), sprawling taxa (sprawlers), and total taxa between high and low impacted sites.

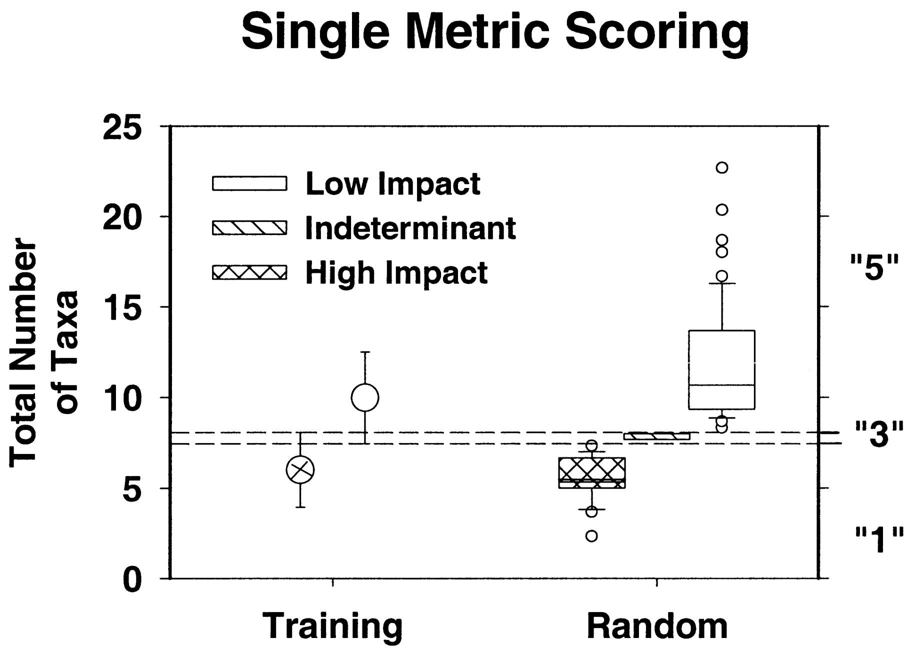

Figure 3. Ranking method used to generate scores ('5,' '3,' or '1')/site for each metric included in the overall biotic index. Figure represents one (e.g., total aquatic taxa) of the eight metrics used to create a total score/sample site. Box plots represent ranked score distribution plus or minus the 95% confidence interval. Designated categories of low impact, indeterminate, and high impact are assigned to each random site based on comparisons to the training site values.

assigned a score of '1,' and 3) when the lower 95% CI for the low-impact sites overlapped with the upper 95% CI for the high-impact sites for a particular metric, a third category was created (Figure 3). A random site value plotted between those intervals was assigned a score of '3.' Ranked scores for eight metrics, those that originally showed a significant difference between low and high-impact sites, were then summed to obtain an overall boundary for determining distribution of random site scores (Figure 4).

Total biotic scores for each random site were then plotted against the means and 95% CI established from the total biotic scores of the low and high-impact sites. Random sites then received a categorical classification based on the same method used to develop a single metric. Total score summaries used categorical names (low-impact, indeterminate, high-impact) instead of a '5,' '3,' or '1' score (Figure 4).

Results

Sediment chemistry and physical parameters

Summary statistics for sediment chemistry parameters (i.e., predicted and measured PAHs, [SEM]-[AVS], TOC, KCl-extractable ammonia) and physical variables (i.e., water depth, soft sediment depth, particle size ranges, median particle diameter) were grouped in Breneman et al. (2000) according to longitudinal location and habitat class within the St. Louis River AOC. An evaluation of a fluorometric screening method for predicting total PAH concentrations in sediment samples from this study was provided in Peterson et al. (2002). No upstream to downstream trends in sediment chemical concentrations were evident (Breneman et al., 2000). However, shallow habitats had the highest mean concentrations for predicted and measured total PAHs and % TOC, as well as negative [SEM]-[AVS] values (Breneman et al., 2000). Channelized habitats contained the highest KCl-extractable ammonia concentrations and, on average, near zero [SEM]-[AVS] values (Breneman et al., 2000). Some upstream to downstream trends were evident in physical variables among the site locations as described in Breneman et al. (2000).

The distribution of individual PAHs, as well as the sum of the 17 PAHs ($TPAH_{17}$) measured in this study, is provided in Table 2 for random samples from shallow and channelized habitats. There was no significant difference ($\alpha = 0.05$) between mean $TPAH_{17}$ concentrations in shallow and channelized habitats. Less variability in individual PAH concentrations was observed for channelized sites than for shallow sites. Based on

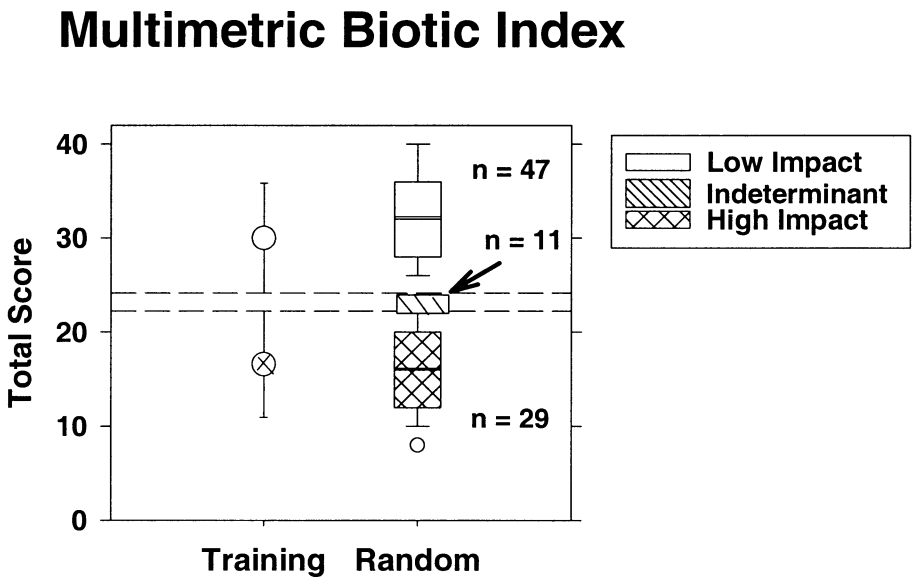

Figure 4. Total biotic index score for each site based on the sum of eight individual metric scores used in the ranking process. Single plots represent mean scores for high-impact and low-impact training sites, plus or minus the 95% confidence interval. Box plots represent the total metric score distribution and designated categories (low impact, indeterminate, and high impact) for random sites based on training site values.

the percentage composition of PAH compounds, perylene displayed the greatest standard deviation in shallow (21%) and channelized (10%) sites, whereas acenaphthene displayed the least variability (0.27% and 0.30%, respectively) in shallow and channelized sites (Table 2). Perylene, fluoranthene, and pyrene comprised the top three PAHs in terms of mean percentage composition in shallow and channelized habitat sites (Table 2). Acenaphthene contributed the lowest mean percentage of PAHs (0.34% and 0.57%, respectively) in shallow and channelized sites (Table 2). The mean phenanthrene/anthracene ratios were 3.2 (SD = 1.4) and 3.4 (SD = 0.61), respectively, in random samples from shallow and channelized habitats. The mean fluoranthene/pyrene ratios were 1.1 (SD = 0.14) and 1.2 (SD = 0.13), respectively, in random samples from shallow and channelized habitats. The LMW and HMW PAHs comprised 28.3% and 71.7%, respectively, of the sum of the 12 LMW and HMW PAHs ($TPAH_{12}$) measured in random samples from shallow and channelized habitats.

The results of linear regression analyses of chemical and physical parameters measured in random samples from shallow and channelized habitats are provided in Table 3 in descending order of their coefficient of determination (r^2) values. The strongest correlations were generally found for SEM metals. In particular, SEM Zn accounted for 87% of the variance in SEM lead ($r^2 = 0.874$), whereas SEM copper accounted for 84% of the variance in SEM Ni ($r^2 = 0.841$; Table 3). The LMW PAHs accounted for 73% of the variance in HMW PAHs (Table 3). The strongest regression relationship for the particle size class of silt (<63 μm) was with SEM Ni ($r^2 = 0.688$; Table 3). Regression relationships of chemical and physical parameters with TOC were weak ($r^2 < 0.31$), and regression relationships of chemical and physical parameters with $TPAH_{12}$ were weaker ($r^2 < 0.22$; Table 3). Dry weight concentrations of $TPAH_{12}$ were particularly independent of TOC, SEM Cu and silt (<63 μm; Table 3).

Mean PEC-Qs

Mean PEC-Qs were calculated for each random and training site (Appendix A). The mean PEC-Q value was based solely on the mean PEC-Q for reliable metals at 45 random and 11 training sites. At the other sites (i.e., 42 random and 9 training sites), the mean PEC-Qs were based on two classes of chemicals (i.e., metals and $TPAH_{12}$). The mean PEC-Q values for two high-impact training sites (i.e., sites 44 and 102) appeared to be too low (i.e., 0.29 and 0.0717, respectively) for these known contaminated areas (i.e., Hog Island Inlet and USS Superfund site, respectively). Individual PAH compounds were not measured at these sites in 1995. When these sites were resampled during June 1996, and

Table 2. Distribution of PAH compounds in random sites from shallow and channelized habitats within the lower St. Louis River AOC.

	PAHs																
Statistical summary	Nap	Acy	Ace	Fle	Phe	Ant	Fla	Pyr	Baa	Cry	Bfa(b)	Bfa(k)	Bap	Per	Idp	Dba	Bgp	$TPAH_{17}$
Chemical Concentration (mg kg[1])																		
Shallow Habitat, $n = 31$																		
Mean	0.67	0.05	0.05	0.15	0.78	0.32	1.25	1.03	0.71	0.72	0.65	0.62	0.74	0.49	0.46	0.10	0.41	9.21
SD	1.9	0.10	0.12	0.30	1.3	0.61	1.9	1.4	1.1	0.97	0.88	0.85	1.1	0.41	0.61	0.14	0.49	13.3
Range: Low	0.0005	0.0005	0.00005	0.0005	0.0037	0.0014	0.0062	0.0068	0.0035	0.0056	0.0049	0.002	0.004	0.0035	0.002	0.0005	0.0005	0.049
Range: High	10.00	0.47	0.65	1.46	6.06	2.63	7.51	5.41	4.38	4.07	3.61	3.50	4.69	1.57	2.37	0.58	1.92	54.0
Channelized Habitat, $n = 11$																		
Mean	0.042	0.022	0.007	0.022	0.19	0.061	0.28	0.22	0.13	0.14	0.11	0.11	0.12	0.11	0.086	0.031	0.073	1.76
SD	0.094	0.034	0.008	0.027	0.29	0.095	0.43	0.32	0.20	0.20	0.15	0.16	0.18	0.13	0.13	0.045	0.098	2.44
Range: Low	0.0025	0.0005	0.0007	0.0019	0.0104	0.0033	0.0139	0.0146	0.0091	0.0103	0.0096	0.0035	0.0045	0.028	0.0026	0.002	0.003	0.158
Range: High	0.322	0.105	0.030	0.092	0.953	0.304	1.39	1.03	0.603	0.556	0.406	0.462	0.522	0.464	0.352	0.148	0.296	7.29
Percentage Composition (%)																		
Shallow Habitat, $n = 31$																		
Mean	0.67	0.49	0.34	1.3	7.2	2.4	11.3	10.2	6.3	7.3	6.9	5.9	6.8	17.6	4.8	1.2	4.6	100
SD	6.1	0.28	0.27	0.76	3.6	0.99	3.3	2.9	2.4	2.5	2.1	2.4	2.5	20.8	1.9	0.75	1.7	
Range: Low	0.006	0.094	0.042	0.30	1.7	0.39	1.7	2.0	0.82	0.97	0.80	0.63	0.63	1.6	0.40	0.06	0.63	
Range: High	20.6	1.3	1.2	4.1	18.2	4.9	15.3	14.8	9.7	11.4	10.0	10.3	9.6	88.3	8.9	3.7	8.8	
Channelized Habitat, $n = 11$																		
Mean	2.4	1.3	0.57	1.3	10.0	2.9	13.4	11.4	6.6	7.6	7.0	6.3	6.0	12.3	4.7	1.9	4.5	100
SD	1.9	2.4	0.30	0.36	4.1	0.85	3.7	2.1	1.1	0.98	1.4	1.8	1.7	9.9	2.0	1.4	1.6	
Range: Low	0.16	0.15	0.11	0.70	4.7	1.9	8.3	8.2	4.6	6.0	5.4	2.2	2.8	1.9	1.6	0.57	1.9	
Range: High	5.5	8.3	1.2	1.8	18.5	4.2	19.9	14.5	8.3	9.0	10.0	9.3	8.2	36.7	8.7	5.8	7.8	

n = number of samples; SD = standard deviation. Codes for PAHs: Nap = Naphthalene; Acy = Acenaphthylene; Ace = Acenaphthene; Fle = Fluorene; Phe = Phenanthrene; Ant = Anthracene; Fla = Fluoranthene; Pyr = Pyrene; Baa = Benz(a)anthracene; Cry = Chrysene; Bfa(b) = Benzo(b)fluoranthene; Bfa(k) = Benzo(k)fluoranthene; Bap = Benzo(a)pyrene; Per = Perylene; Idp = Indeno(1,2,3-cd)pyrene; Dba = Dibenz(a,h)anthracene; Bgp = Benzo(g,h,i)perylene; $TPAH_{17}$ = sum of 17 PAHs measured in this study.

Table 3. Results of regression analyses of chemical and physical parameters measured in random sites from the lower St. Louis River AOC.

Y variable	Intercept (b_0)	Slope (b_1)	X variable	r^2 value	N
SEM Pb	−0.73	0.26	SEM Zn	0.874	87
SEM Ni	1.6	0.42	SEM Cu	0.841	87
Hg	−0.038	3	SEM Zn	0.827	87
Hg	−0.014	0.011	SEM Pb	0.792	87
HMW PAHs	1.4	1.4	LMW PAHs	0.73	42
SEM Ni	1.3	0.096	silt (<63 μm)	0.688	87
SEM Ni	1.3	7.2	SEM Cd	0.66	87
SEM Pb	0.16	1.3	SEM Cu	0.636	87
SEM Zn	−2.41	91	SEM Cd	0.635	87
SEM Zn	−2.6	10	SEM Ni	0.608	87
SEM Cd	0.17	0.01	silt (<63 μm)	0.598	87
SEM Zn	10	4.5	SEM Cu	0.574	87
SEM Cu	1.7	0.19	silt (<63 μm)	0.546	87
Hg	−0.047	0.28	SEM Cd	0.534	87
SEM Pb	−1.4	2.6	SEM Ni	0.534	87
SEM Pb	−0.93	23	SEM Cd	0.527	87
SEM Cu	1.8	14	SEM Cd	0.517	87
SEM Zn	6.8	1	silt (<63 μm)	0.486	87
Hg	−1.9	0.013	SEM Cu	0.439	87
Hg	−0.027	0.027	SEM Ni	0.411	87
SEM Pb	2.4	0.24	silt (<63 μm)	0.348	87
Hg	−1.2	2.8	silt (<63 μm)	0.327	87
SEM Cd	0.5	0.059	TOC	0.304	87
$TPAH_{12}$	1.1	0.073	SEM Zn	0.217	42
SEM Cu	8.5	0.9	TOC	0.186	87
SEM Ni	5	0.41	TOC	0.186	87
$TPAH_{12}$	2.1	0.21	SEM Pb	0.15	42
Mean PAH PEC-Q	0.046	1.6	Mean metal PEC-Q	0.149	42
$TPAH_{12}$	2.9	17	Hg	0.117	42
SEM Zn	48	4.2	TOC	0.116	87
SEM Pb	11	1.1	TOC	0.107	87
$TPAH_{12}$	1.6	5.6	SEM Cd	0.0949	42
$TPAH_{12}$	1.5	0.68	SEM Ni	0.0889	42
TOC	1.7	0.034	silt (<63 μm)	0.0792	87
Hg	0.11	0.011	TOC	0.0756	87
$TPAH_{12}$	3.4	0.51	TOC	0.0649	42
$TPAH_{12}$	3	0.21	SEM Cu	0.0511	42
$TPAH_{12}$	3	0.051	silt (<63 μm)	0.0395	42

Regression Relationships are in the form of $Y = b_0 + b_1(X)$, with variables defined above; N = number of samples; HMW = high molecular weight PAHs; LMW = low molecular weight PAHs; PEC-Q = probable effect concentration quotient; SEM = simultaneously extractable metals; TOC = total organic carbon; $TPAH_{12}$ = sum of 12 LMW and HMW PAHs.

both SEM metals and PAH compounds were analyzed, the mean PEC-Qs increased to 0.407 and 26.2 for sites 44 and 102, respectively. Since sediment contamination can be very heterogeneous at these sites, the 1996 mean PEC-Q values should be considered only as an approximation of the 1995 mean PEC-Q values. Sites 44 and 102 were further excluded from additional data analysis involving mean PEC-Qs.

The data set of mean PEC-Q values for the random shallow habitat sites and for high-impact training

Table 4. Summary statistics of mean PEC-Q values for random sites and training sites in the lower St. Louis River AOC.

	Mean PEC-Qs			
	Shallow random sites	Channelized random sites	Low impact training sites	High impact training sites
Mean	0.18	0.094	0.091	3.9
SD	0.19	0.060	0.071	9.1
Median	0.15	0.077	0.077	0.67
Minimum	0.0095	0.013	0.010	0.33
Maximum	1.0	0.21	0.25	26.4
Skewness	2.6	0.39	1.2	2.8
Kurtosis	9.1	−1.1	1.6	8.0

SD = standard deviation.

sites displayed significant positive skewing, as well as a significant kurtosis problem (as determined from the standard errors of skewness and kurtosis; Table 4). Therefore, it is more appropriate to compare the median values of these groups rather than mean values (Brown, 1997). A statistically significant difference ($p = <0.001$) was observed among the median values of groups. The high-impact training sites were statistically different ($p < 0.05$) from the low impact training sites, random channelized habitat sites, and random shallow habitat sites. The other pairwise comparisons were not significant.

Based on the random sampling design utilized in this study, one new area of high (i.e., mean PEC-Q >0.6) chemical contamination was observed outside the boundaries of the St. Louis River Interlake/Duluth Tar Superfund site at sites 71 and 73 (Appendix A). As a result, site 73 has been included in the remediation zone for this Superfund site. In addition, sites 69 and 70, further downstream from this Superfund site, showed a moderate (i.e., mean PEC-Q >0.1 to 0.5) level of contamination due primarily to PAHs (Appendix A). All other random sites had mean PEC-Qs <0.4 (Appendix A), indicating a low to moderate level of contamination.

Toxicity testing

Microtox® tests: The solid-phase Microtox® tests identified 31% of random sites and 55% of training sites (i.e., 20% of low-impact sites and 90% of high-impact sites) as toxic. Similarly, the pore water Microtox® tests identified 29% of random sites and 50% of training sites (i.e., 40% of low-impact sites and 60% of high-impact sites) as toxic. The concordance between toxic responses observed in both Microtox® tests was less in the random (8%) and low-impact training sites (10%) than in the high-impact training sites (60%). The individual sample results for random and training sites are provided in Appendix A.

Ten-day sediment toxicity tests: The pH ranges were acceptable in all of the 10 d sediment toxicity tests. Dissolved oxygen concentrations fell below 40% saturation for at least 1 day in 11% of the *H. azteca* and 49% of the *C. dilutus* toxicity tests. Temperature was less than the recommended range of 23 ±1°C (US EPA, 1994) for 96% of the 10 d sediment toxicity tests for at least 1 day of the tests. The mean water temperature was less than the recommended range of 23 ±1°C for 67% of the tests. Conductivity measurements made on the last two batches of toxicity tests were within acceptable limits (USEPA, 1994). The 96 h *H. azteca* reference toxicant (NaCl) tests were all acceptable with control survival exceeding 90%, whereas one of the *C. dilutus* reference toxicant (NaCl) tests failed the control survival requirement of ≥90% (USEPA, 1994).

With one exception, all batches of 10 d *H. azteca* toxicity tests had acceptable control survival (i.e., ≥80%), and all batches of 10 d *C. dilutus* tests had acceptable control survival (i.e., ≥70%). For the failed batch of *H. azteca* tests, seven of the eleven test samples had mean survival values exceeding 80%. The corresponding reference toxicant control survival of these organisms was also acceptable (i.e., ≥90%). Therefore, the *H. azteca* culture appeared to be healthy, although the data could not be analyzed statistically due to control failure (i.e., 78% survival). A qualitative data examination of this batch of toxicity tests indicated that no test sediments appeared to be toxic.

None of the 52 random samples in shallow and channelized habitats exhibited significant toxicity to *H. azteca* survival (Appendix A). For *C. dilutus*, two random sediment samples (sites 13 and 43) were significantly toxic at $p = 0.05$ (Appendix A). Site 13 was located in a channel east of the Blatnik Bridge, adjacent to a cement facility. Site 43 was located at the Hog Island Inlet/Newton Creek area in Superior, WI; this area is contaminated with diesel range organics (i.e., mid-range petroleum products such as diesel or fuel oil), PAHs, oil and grease, lead, chromium, and mercury (Redman and Janisch, 1995; SEH Inc., 2000, 2003). The specific cause of toxicity could not be determined for either site.

Toxicity tests were conducted on 14 training site sediment samples (Appendix A). For *H. azteca*, sites 56, 72, 77, and 102 had a significant reduction in mean survival compared to the corresponding control at $p = 0.05$. For *C. dilutus*, sites 44, 56, 72, 99, and 102 had significant toxicity; all of these sites, except 44, caused complete mortality of *C. dilutus*. All of the training sites which caused significant acute toxicity, except site 77 (Kimballs Bay), were in a priori high-impact areas.

Benthic macroinvertebrate community classification

Multimetric analysis: Eight of the 21 macroinvertebrate metrics provided significant ($p \leq 0.1$) differences between high-impact and low-impact training sites (Figure 2). These metrics were used to create an overall biotic score. The metrics included percent abundance values from Ephemeroptera, Oligochaeta, Trichoptera, and obligate depositional taxa. Oligochaeta were the most numerous taxa in both habitat types. Ephemeroptera, Trichoptera, and obligate depositional taxa were observed in low densities at high-impact training sites, where Oligochaeta numbers were abundant. Since Oligochaeta have been reported to have high tolerance to contaminants (Breneman and Pontasch, 1994), the Oligochaeta metric was scored to reflect those conditions. An increase in Oligochaeta abundance produced lower ranked scores per site. The remaining four significant metrics reflected aspects of species richness with respect to three common behavioral categories (burrowers, clingers, crawlers) and total taxa richness. All taxa metrics had lower values in high-impact training sites.

Total multimetric scores at training sites ranged from 8 to 38, with 40 as a possible total score (Figure 4). High-impact training sites had significantly lower values (t-test, $p \leq 0.05$) than low-impact training sites. The random sites had total multimetric scores similar to the same range of values as observed in the training site data set. Most random sites (55%) were classified as low impact by the multimetric approach (Appendix A). Thirty-six percent of the random sites were classified as high impact, and 9% were classified as indeterminate (Appendix A). A larger portion (48%) of the random channelized habitat sites were classified as high impact than the random shallow habitat sites (25%).

Discriminant function analysis: Backwards selection techniques retained 16 of 18 benthic metrics in the discriminant function analysis for the training sites. The two metrics eliminated from the model included Diptera abundance and those taxa associated with only depositional habitats. Using the retained metrics, there was a significant difference between high-impact and low-impact training sites (MANOVA $F = 113.17$, $p = 0.0012$). The 16 metric discriminant function classified the training sites perfectly (cross validated error rate = 0%).

A larger proportion of the random sites were classified as high impact with the discriminant analysis technique. The discriminant function classified 48 (55%) of the random sites as high impact and 39 (45%) as low impact (Appendix A). A larger proportion (69%) of the random channelized sites were classified as high impact than were the random shallow sites (48%). All random sites were classified with greater than 98% probability; therefore, no sites were classified into the indeterminate category.

Comparison of techniques: The multimetric technique classified the random sites as follows: 36% high impact, 55% low impact and 9% as indeterminate. In comparison, 55% of random sites were classified as high impact using the discriminant function analysis and 45% of sites as low impact. Of the 31 random sites classified as high impact by the multimetric technique, 22 sites (71%) were similarly classified as high impact in the discriminant function analysis (Appendix A). In addition, the multimetric method classified 48 random sites as low impact that included 28 sites (58%) similarly classified by the discriminant function procedure.

Classifications were also compared with respect to the location of sites along the length of the St. Louis River estuary (Figure 1). Both methods generally classified more sites in the upper estuary as low impact and a larger portion of the lower estuary sites as high impact (Figure 5). This dichotomy was strongest, however, with the multimetric method which classified very few of the upper sites (4%) as high impact. In contrast, over 40% of the sites in the upper estuary were classified as high impact by the multivariate procedure. The discriminant function analysis, however, classified over 50% of the sites as high impact in all regions. Both techniques identified an increase in the number of sites classified as

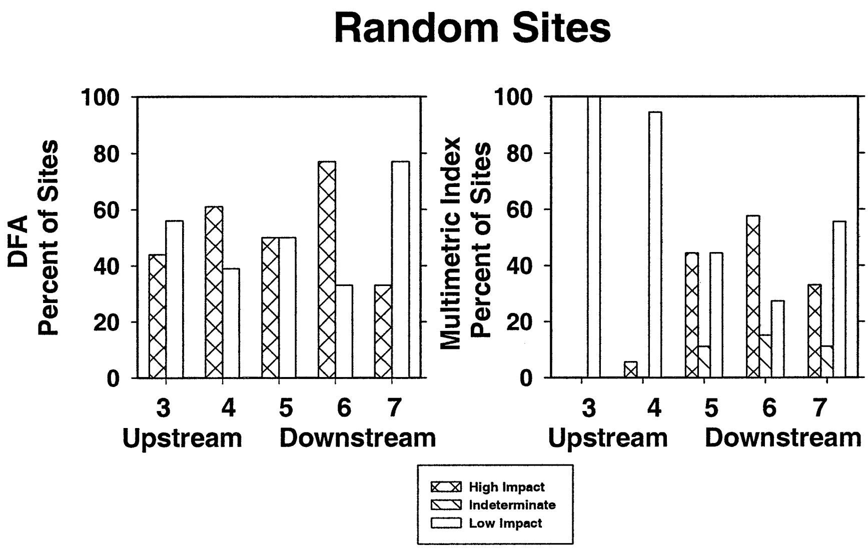

Figure 5. Distribution of the multimetric and multivariate site categories (Discriminant function analysis (DFA) and multi-metric biotic index (MBI)) by region within the St. Louis River Area of Concern (AOC). Sample areas are designated in the schematic diagram on Figure 1. Habitat areas include; 3) St. Louis River, 4) Spirit Lake, 5) Duluth/Superior Harbor, 6) Superior Bay, and 7) Allouez Bay.

low impact in the Allouez Bay region (Figure 1). This sheltered bay encompasses the area east of the Nemadji River entrance into Superior Bay and the Superior entrance to Lake Superior. Allouez Bay is an example of a pulse-stable wetland community; the seiche causes pulses of water and sediment to move in and out of the bay, helping to prevent the wetlands from filling in with sediment (St. Louis River CAC, 2002). This bay includes some of the highest quality remaining wetlands in the lower estuary (St. Louis River CAC, 2002), and it has been shown to possess a similar benthic community structure as the upper reaches of the St. Louis River AOC (Breneman et al., 2000).

There were also differences in percent classification when comparing random samples collected from shallow and channelized habitats. Shallow habitats received 71% and 51% low impact classifications by the multimetric and discriminant function analysis techniques, respectively. Channelized habitats were classified as high impact at 62% and 69% of the sites by the multimetric and discriminant function analysis techniques, respectively. Based on the multimetric category, 7% of shallow habitat sites and 14% of channelized habitat sites were classified as indeterminate.

Comparison of mean PEC-Q, toxicity, and benthic indicators

Mean PEC-Qs have been found to be predictive of sediment toxicity in the St. Louis River AOC, based on a compilation of matching sediment chemistry and toxicity data that have been incorporated into a Microsoft™ Access database (Crane et al., 2002a). The predictive ability of six ranges of mean PEC-Qs for the Microtox®, amphipod, and midge toxicity tests from this study are shown in Table 5. This table included data from both random and training sites. In addition, the incidence of high-impact sites, determined by the two benthic community classification techniques, was assessed for the same range of mean PEC-Qs (Table 5). These assessments were most appropriate for the two lowest mean PEC-Q ranges ($\leq$0.1 and $>$0.1 to $\leq$0.5) because the minimum data requirements (20 samples; CCME, 1995) were met. Any comparisons of the next three higher mean PEC-Q ranges should be made with caution due to the small number of samples in each range. These three ranges were also compressed into one mean PEC-Q range of $>$0.5 to provide a greater number of samples for evaluating qualitative trends in the incidence of toxicity and incidence of high-impact sites.

For all toxicity tests, the incidence of toxicity increased as the mean PEC-Q range increased from $\leq$0.1, $>$0.1 to $\leq$0.5, and $>$0.5 (Table 5). Pore water and solid-phase Microtox® tests, when considered separately, showed a higher incidence of toxicity than the amphipod and midge toxicity tests. When the Microtox® tests were considered together, the incidence of toxicity in tests that were both toxic was lower and was within 4% of the incidence of toxicity for the amphipod and midge

Table 5. Incidence of effects for mean PEC-Q ranges as determined using sediment quality data collected from random and training sites in the lower St. Louis River AOC (number of samples in parentheses). Sites 44 (Hog Island Inlet site) and 102 (USS Superfund site) were excluded from the incidence of effect calculations for specific mean PEC-Q ranges due to incomplete sediment chemistry data (i.e., no PAH data) for these known contaminated areas. Thus, the mean PEC-Q values were predicted to be too low for these high impact training sites.

	Incidence of toxicity (%)					Incidence of high impact sites (%) based on benthic community classification techniques		
Mean PEC-Q range	Microtox® solid phase EC_{50}	Microtox® pore water reduction	Both Microtox® tests with toxic endpoints	10 d *H. azteca* survival	10 d *C. dilutus* growth or survival	Multimetric classification	DFA category	Both benthic indices with high impact sites
≤0.1	10.6% (47)	26% (46)	0% (46)	0% (22)	3.7% (27)	40.4% (47)	44.7% (47)	27.6% (47)
>0.1 to ≤0.5	50% (50)	34% (50)	18% (50)	3.8% (26)	6.4% (31)	28% (50)	58% (50)	20% (50)
>0.5 to ≤1.0	100% (4)	50% (4)	50% (4)	50% (2)	50% (2)	75% (4)	75% (4)	50% (4)
>1.0 to ≤5.0	33.3% (3)	33.3% (3)	0% (3)	0% (2)	0% (3)	66.7% (3)	66.7% (3)	66.7% (3)
>5.0	100% (1)	100% (1)	100% (1)	100% (1)	100% (1)	0% (1)	100% (1)	0% (1)
>0.5	*75% (8)*	*50% (8)*	*37.5% (8)*	*40% (5)*	*33.3% (6)*	*62.5% (8)*	*75% (8)*	*50% (8)*

toxicity tests at mean PEC-Q ranges of ≤ 0.1 and >0.5 and within 14% at the mean PEC-Q range of >0.1 to ≤ 0.5. Of the 14 random and training sites that had significant toxic responses in both Microtox® tests, 50% of these sites were also toxic in either one or both of the corresponding amphipod and midge toxicity tests.

For the discriminant function analysis technique, the incidence of high-impact sites increased as the mean PEC-Q range increased from ≤ 0.1, >0.1 to ≤ 0.5, and >0.5 (Table 5). The multimetric classification showed a higher incidence of high-impact sites at mean PEC-Qs of ≤ 0.1 than at >0.1 to ≤ 0.5, followed by an increase in high impact sites at mean PEC-Qs >0.5. When both benthic indices were considered together, the incidence of high-impact sites in both indices decreased for most mean PEC-Q ranges compared to the individual indices.

Discussion

This study provided a good comparison of sediment assessment techniques within a Great Lakes Harbor and Embayment resource class and also provided a means for assessing the status of this resource utilizing the US EPA's EMAP statistical sampling concepts. The various assessment techniques used in this study provided some generalities in trends and results but also showed distinct differences in the nature of their response to environmental conditions. These differences are useful for assessing condition and evaluating their potential for future trend and assessment monitoring.

Estuarine benthic fauna can provide a reliable and sensitive indicator of disturbance from chemical stressors at exposure levels that are well below those associated with a similar incidence of sediment toxicity in laboratory survival tests with single marine species (Hyland et al., 1999, 2003). In the St. Louis River AOC, the macroinvertebrate community appeared to be more sensitive to physical disturbances of the region than the acute toxicity tests at low mean PEC-Qs. Breneman et al. (2000) found that the strongest environmental factors influencing macrobenthos in the St. Louis River AOC were physical parameters. These physical variables were water depth and site distance from the headwaters (Breneman et al., 2000). When considering both benthic indices together, the classification of more high-impact sites in channelized habitats (48%) than shallow habitats (14%) and greater proportion of high-impact sites in the lower regions of the harbor with commercial activity are consistent with those observations. The navigation channels in the Duluth-Superior Harbor must be periodically dredged to maintain a 27-foot depth. Prop wash from ship traffic frequently resuspends sediments in the channels. In addition, other factors such as wave-induced resuspension of sediments and ice scour contribute to physical disturbances of sediments in the harbor area. Benthic macroinvertebrates may be the best indicator of physical disturbance for restoration efforts in the Duluth-Superior Harbor.

Both macrobenthic classification techniques indicated a higher proportion of random sites were classified as high impact in the lower sections of the AOC. However, the multimetric classification technique exhibited this trend most often. The multimetric technique classified 0 and 5.6% of sites in the upper two sections of the study region as high impact (Spirit Lake area and upper St. Louis River sections, Figure 5) and a much larger proportion (44% to 58%) of sites in the lower sections (downstream from Spirit Lake including the harbor and Allouez Bay) as high impact. The discriminant function analysis technique classified over 50% of sites as high impact in all areas but with some increase (50 to 77%) in the lowermost sections. In terms of their abilities to correctly classify the training sites, the discriminant analysis function technique correctly classified 100%, while the multimetric technique classified 70% as correct.

Discrepancies in the two techniques undoubtedly lie in the inherent differences in the manner in which macrobenthic community data were manipulated. The multimetric technique eliminated individual metrics that were unable to discriminate between the high impact and low impact training sites. The remaining metrics were equally weighted when aggregating to a site score. This technique in effect emphasizes the greatest differences between the high-impact and low-impact sites, but at the same time allows less influence by individual metrics. Thus, more subtle, but less distinct differences among the training sites that may have been present in the variables are eliminated from analysis. The less rigorous criterion for variable inclusion in the discriminant function analysis technique, and the ability of this technique to weight the importance of individual metrics based on their ability to differentiate among the training sites, resulted in a more precise classification of the training sites. However, due to its inherent sensitivity, the discriminant function analysis technique is more sensitive to the a priori selection of training sites than the more robust multimetric method. Without extensive sediment quality databases to work from, future studies in Great Lakes embayments may utilize multimetric approaches as an effective means for establishing sediment condition. However, the definition of training sites needs to be clearly delineated. The eight macroinvertebrate metrics utilized in this study may provide robust indicators in other Great Lakes embayments. However, their utility should be evaluated in other areas.

Polycyclic aromatic hydrocarbons are widespread chemicals of potential concern in the lower St. Louis River estuary. The percentage compositions of LMW and HMW PAHs in the random site samples of this study were similar to site-specific observations in both Minnesota Slip and Slip C in the Duluth Harbor (Crane et al., 2002b). The mean phenanthrene/anthracene and fluoranthene/pyrene ratios were indicative of pyrogenic (combustion) sources of PAHs to the estuary (i.e., phenanthrene/anthracene ratio <10 and fluoranthene/pyrene ratio >1.0; Budzinski et al., 1997). The high percentage composition of perylene in some sediment samples may be due to other sources. The presence of perylene in sediments has often been attributed to early diagenesis of organic matter (Wakeham et al., 1980a; Venkatesan, 1988). Recent radiocarbon results indicate that both natural (diagenetic) and anthropogenic (fossil fuel combustion) sources of perylene were evident in two marine sediments (Reddy et al., 2002). The high mean fluoranthene and pyrene concentrations in the random samples from shallow and channelized habitats were consistent with these compounds being the most abundant individual combustion PAHs observed in sediments from the Washington coastal region (Prahl and Carpenter, 1984) and throughout the world (Laflamme and Hites, 1978; Wakeham et al., 1980b).

The lack of correlations between $TPAH_{12}$ and either TOC or silt (<63 μm) in the random site sediments may be due to multiple sources of PAHs in the lower St. Louis River watershed and to a broad range (0.02–19.2%) of TOC values in the lower estuary. In contrast, a site-specific study at Slip C in the Duluth Harbor (Crane, 1999) demonstrated strong correlations between total PAHs and TOC, as well as silt (53–2 μm). Mercury and SEM metals accounted for less than 22% of the variance in $TPAH_{12}$ in the random site samples of this study, whereas mercury and lead have been found to account for 77 and 88%, respectively, of the variance in total PAHs in Slip C (Crane, 1999). Thus, the broad, random sampling strategy of this R-EMAP study may mask out chemical concentration trends at small, contaminated sites that have limited sources of contaminants.

Although the EMAP sampling design is unbiased with respect to space, it may be biased with respect to risk (Suter, 2001). Sampling randomly over space tends to emphasize the importance of wide-spread stressors such as introduced species and land use and minimize potentially severe stressors such as point source effluents, waste dumping, gravel dredging, mine drainage, or failures of mine or animal waste lagoons (Suter, 2001). A number of known contaminated sediment areas in the Duluth-Superior Harbor were missed in the selection of random sites, indicating a combination of random and purposive sampling should be conducted in future status and trends monitoring to characterize the full range of aquatic ecological risks in the harbor and the rest of the St. Louis River AOC. A mixed strategy of random and purposive sampling will also allow state agencies, responsible parties, Natural Resource Trustees (including federal partners), the St. Louis River Citizens Action Committee (CAC), and other stakeholders to more effectively track changes in sediment quality after remediation actions and habitat restoration efforts have been implemented at sites requiring remediation.

Survival was the principal metric for the short-term amphipod and midge toxicity tests. The chronic 28 to 42 d *H. azteca* toxicity test is more sensitive than either the 10 d amphipod or midge tests (Ingersoll et al., 2001), and its use would reduce the potential for false negatives at low mean PEC-Qs. Because longer-term toxicity tests provide the most effective mean of discriminating among moderately contaminated sediment samples and because in situ benthic macroinvertebrates are exposed to contaminated sediments for an extended period, it would be prudent in future R-EMAP studies to evaluate sediment toxicity using the 28 to 42 d *H. azteca* test (endpoints: survival and growth) in sediments with mean PEC-Qs <5.0. However, it would be more cost-effective to utilize acute toxicity tests to characterize the toxicity of highly contaminated sediments (i.e., mean PEC-Qs ≥ 5.0).

Although the combined use of the pore water and solid phase Microtox® tests provided a better indicator of acute sediment toxicity than the individual Microtox® tests, it still greatly overestimated the occurrence of acute toxicity in concurrent random samples (Appendix A). Only 14% of the random sites in which toxic responses were observed in both Microtox® tests also exhibited corresponding toxicity in either the 10 d *H. azteca* or *C. dilutus* tests. However, the combination of both Microtox® tests was a strong indicator of sediment toxicity in the high impact training sites where toxic responses in both Microtox® tests corresponded to 83% of these sites also exhibiting toxicity in either acute toxicity test (Appendix A). In particular, sites 56, 72, and 102 displayed toxicity in both Microtox®

tests and both acute toxicity tests. Site 56 was located near the outfall of the Western Lake Superior Sanitary District, site 72 was located within the St. Louis River Interlake/Duluth Tar Superfund site, and site 102 was located within the USS Superfund site. Significant acute toxicity to *C. dilutus* has been observed at each of these contaminated sites during previous sediment investigations (Crane et al., 1997; Schubauer-Berigan and Crane, 1997).

Benton et al. (1995) noted that extreme caution should be taken in interpreting the results of solid-phase Microtox® tests when testing sediments of varying composition or that may be differentially contaminated or contain a suite of contaminants. Pardos et al. (1999) demonstrated the role of elemental sulfur in giving false positive toxicity results in the Microtox® bioassay. Although elemental sulfur was not measured in sediments from the St. Louis River AOC, sulfide may be oxidized to elemental sulfur by bacteria in sediments. The most labile fraction of sulfides in sediments is represented by AVS, and AVS is associated with the more soluble iron and manganese monosulfides (Di Toro et al., 1990). Acid volatile sulfide was detected in 78% of the random sites and 95% of the training sites. Due to the potential for a majority of these sites to form elemental sulfur, interpretation of the pore water and solid phase Microtox® tests is difficult.

A more complete assessment of sediment quality in the lower St. Louis River AOC was attained using a weight-of-evidence approach of several sediment quality indicators. A sediment quality triad assessment was conducted for 52 random sites that included all three components of the Triad (i.e., sediment chemistry, 10 d sediment toxicity tests with midges and amphipods, and benthic macroinvertebrate community structure). Possible conclusions of the triad results were based on Chapman (1992). For 21% of these sites, there was strong evidence for the absence of pollution-induced degradation (Table 6). For 71% of sites, alteration of the benthic macroinvertebrate community was probably not due to toxic chemical contamination (Table 6). Each of the following scenarios applied to about 2% of sites: contaminants were not bioavailable, unmeasured toxic chemicals caused degradation, chemicals were either not bioavailable or alteration of the benthic macroinvertebrate community was not due to toxic chemicals, and unmeasured chemicals or conditions existed that had the potential to cause degradation. These results were consistent with Breneman et al.'s (2000) finding that physical habitat features best explained variability in benthic macroinvertebrate community structure in the lower St. Louis River AOC based on multivariate redundancy analysis on 13 environmental parameters.

The potential for sediment-associated use impairments to aquatic life (e.g., fish) at the random Triad sites was as follows: highly unlikely (21%), unlikely (73%), and likely (6%; Table 6). Channelized habitat sites were more unlikely (94%) to cause impairments to aquatic life than shallow habitat sites (64%). These findings are consistent with field observations in the lower St. Louis River AOC. Native fish populations have rebounded since water quality in the estuary began to improve in the late 1970s (St. Louis River CAC, 2002). Approximately 45 native species have been documented by state agencies in the lower St. Louis River. Piscivorous species such as yellow perch, white bass, muskie, walleye and northern pike contribute to a popular sport fishery in this area. Although the Triad approach does not indicate a high degree of impairments to aquatic life, this approach does not consider bioaccumulation. Fish consumption guidelines for human health are currently in effect for a number of fish species in the lower estuary, primarily due to mercury contamination (MDH, 2002). In addition, chronic impacts to aquatic life were not assessed in this study.

We recommend that the Triad approach be used in future R-EMAP surveys of the lower St. Louis River AOC. Furthermore, we recommend measuring conventional metals instead of SEM and AVS for making comparisons to the Level I and Level II SQTs (Crane et al., 2000, 2002a). Long et al. (1998) noted that relative to SEM:AVS concentrations, sediment guidelines based upon dry weight-normalized concentrations were equally or slightly more accurate in predicting both nontoxic and toxic results in laboratory tests. When SEM exceeds AVS by a factor of 5 (on a molar basis), a higher incidence of toxicity (80 to 90%) has been observed in freshwater and saltwater sediment amphipod tests (USEPA, 1997). This situation did not occur in the St. Louis River sediments (Appendix A). Most sites that were toxic to amphipods and/or midges had a greater concentration of AVS compared to SEM metals, implying that SEM metals were probably not available for uptake by benthic biota. Morse and Rickard (2004) noted that AVS is metastable and as the AVS minerals dissolve, any metals sequestered with them also dissolve and the final trace metal toxicity is determined by their behavior with respect to pyrite rather than AVS. Thus, the toxicity of these samples appears to be due to other chemicals such as PAHs. Future R-EMAP surveys should also take into consideration the potential for phototoxicity of PAHs (Monson et al.,

Table 6. Sediment quality triad results in the lower St. Louis River AOC for sites with matching sediment chemistry, toxicity, and benthic community data.

Number of sites	Habitat class	Sediment chemistry[a]	Sediment toxicity[b]	Benthic community impact[c]	Potential for sediment-associated use impairments: aquatic life[d]	Possible conclusions[e]
Random sites						
11	Shallow	−	−	Low	Highly unlikely	Strong evidence for absence of pollution-induced degradation
1	Shallow	+	−	Low	Unlikely	Contaminants are not bioavailable
22	Shallow	−	−	High	Unlikely	Alteration is probably not due to toxic chemical contamination
1	Shallow	+	−	High	Likely	Chemicals are not bioavailable or alteration is not due to toxic chemicals
1	Shallow	−	+	High	Likely	Unmeasured toxic chemicals are causing degradation
15	Channel	−	−	High	Unlikely	Alteration is probably not due to toxic chemical contamination
1	Channel	−	+	Low	Likely	Unmeasured chemicals or conditions exist that have the potential to cause degradation
Training sites						
Low impact						
5	Shallow	−	−	Low	Highly unlikely	Strong evidence for absence of pollution−induced degradation
1	Shallow	−	+	High	Likely	Unmeasured toxic chemicals are causing degradation
High impact						
1	Shallow	−	−	High	Unlikely	Alteration is probably not due to toxic chemical contamination
2	Shallow	+	−	High	Likely	Chemicals are not bioavailable or alteration is not due to toxic chemicals
3[f]	Shallow	−	+	High	Likely	Unmeasured toxic chemicals are causing degradation
2	Shallow	+	+	High	Highly likely	Strong evidence for pollution-induced degradation

[a] (−) = mean PEC-Q <0.6; (+) = mean PEC-Q ≥0.6.
[b] (−) = non-toxic response in both 10-d sediment toxicity tests; (+) = toxic response in either 10-d sediment toxicity test.
[c] low impact or indeterminate benthic community based on both benthic classification procedures (i.e., multimetric and DFA); high impact benthic community designated by either the multimetric or DFA category.
[d] Based on contingency table given in Crane et al. (2000).
[e] Based on Chapman (1992).
[f] PAH compounds were not analyzed in sediment samples collected from sites 44 (Hog Island Inlet) and 102 (USS Superfund site). This represents a data gap in the determination of mean PEC-Qs for these high impact training sites.

1995; Diamond et al., 2003) in shallow areas and that sequestration in sediments affects the bioavailability of hydrophobic organic chemicals to benthic organisms (Kraaij et al., 2002; Ghosh et al., 2003).

This study has demonstrated that the US EPA's EMAP concepts can be effectively applied to the Great Lakes Harbors and Embayments resource class. Additional sediment quality indicators that could be applied at other Great Lakes sites are given in Crane et al. (2000). Although R-EMAP studies are resource intensive, they are vital for tracking the health of harbors and embayments over time. This type of information is especially important for reaching decisions on whether sediment-related impairments at Great Lakes AOCs have been alleviated enough to contribute to the delisting of AOCs.

Acknowledgments

We wish to thank the many participants of this study. Harold Wiegner, Scot Beebe, Patti King, Carol Hubbard, and Jerry Flom, all of the Minnesota Pollution Control Agency (MPCA), provided field sampling assistance and conducted the 10 d sediment toxicity tests. James Beaumaster (MPCA) provided graphical support. Field assistance and benthic sample processing was conducted by Jacque Alexander, John Juaire, Barb Peichel, Maartje Thjssen, and Aaron Undeland at the University of Minnesota Duluth (UMD) Natural Resources Research Institute (NRRI), Duluth, MN. In addition, Sherry Bertelsen, Tim Corry, Jill Scharold, and Jo Thompson, USEPA Duluth, MN, assisted in processing benthic samples. Sediment chemistry analyses were conducted by John Ameel, Rich Axler, Sarah Kohlbry, Keith Lodge, Irene Moser, and Greg Peterson at the Analytical Chemistry Laboratory and Trace Organic Laboratory at UMD-NRRI. Microtox® analyses were conducted by Greg Peterson and Chris Owen at UMD-NRRI. Particle size analyses were done by Keith Lodge at UMD-NRRI. This research was largely supported by Cooperative Agreement CR823497 between the MPCA and US EPA and Grant Contract Number 47002 between the MPCA and UMD-NRRI. This research was partially supported by the Minnesota Sea Grant College Program (MSGCP), and this paper is journal No. JR 505 of the MSGCP. This paper is journal No. 368 of the Center for Water and the Environment of the NRRI. The views expressed herein are those of the authors and do not necessarily reflect the views of the MPCA, UMD-NRRI, Minnesota Sea Grant, or the USEPA.

References

Ankley, G. T., Benoit, D. A., Balogh, J. C., Reynoldson, T. B., Day, K. E., Hoke, R. A., 1994a. Evaluation of potential confounding factors in sediment toxicity tests with three freshwater benthic invertebrates. Environ. Toxicol. Chem. 13, 627–635.

Ankley, G. T., Call, D. J., Cox, J. S., Kahl, M. D., Hoke, R. A., Kosian, P. A., 1994b. Organic carbon partitioning as a basis for predicting the toxicity of chlorpyrifos in sediments. Environ. Toxicol. Chem. 13, 621–626.

Barbour, M. T., Stribling, J. B., Karr, J. R., 1995. Multimetric approach for establishing biocriteria and measuring biological condition. In: W. S. Davis and T. P. Simon (Eds.), *Biological Assessment and Criteria: Tools for Water Resource Planning and Decision Making*, pp. 63–76. Lewis Publishers, Ann Arbor, MI.

Barbour, M. T., Gerritsen, J., Snyder, B. D., Stribling, J. B., 1999. Rapid bioassessment protocols for use in streams and wadeable rivers: Periphyton, benthic macroinvertebrates and fish, Second Edition. Office of Water, USEPA, Washington, DC. EPA 841-B-99-002.

Benoit, D. A., Phipps, G., Ankley, G. T., 1993. A sediment testing intermittent renewal system for the automated renewal of overlying water in toxicity tests with contaminated sediments. Water Res. 27, 1403–1412.

Benton, M. J., Malott, M. L., Knight, S. S., Cooper, C. M., Benson, W. H., 1995. Influence of sediment composition on apparent toxicity in a solid-phase test using bioluminescent bacteria. Environ. Toxicol. Chem. 14, 411–414.

Breneman, D., Pontasch, K., 1994. Stream microcosm toxicity tests: predicting the effects of fenvalerate on riffle insect communities. Environ. Toxicol. Chem. 13, 381–387.

Breneman, D., Richards, C., Lozano, S., 2000. Environmental influences on benthic community structure in a Great Lakes embayment. J. Great Lakes Res. 26, 287–304.

Brinkhurst, R. O., 1986. Guide to the freshwater aquatic microdrile oligochaetes of North America. Can. Spec. Publ. Fish. Aquat. Sci. 84,1– 259.

Brouwer, H., Murphy, T., McArdle, L., 1990. A sediment-contact bioassay with *Photobacterium phosphoreum*. Environ. Toxicol. Chem. 9, 1353–1358.

Brown, J. D., 1997. Skewness and kurtosis. Shiken: JALT Testing & Evaluation. SIG Newsletter. 1, 16–18.

Budzinski, H., Raoux, C., Baumard, P., Bellocq, J., Garrigues, P., 1995. In: Grimalt, J., Dorronsorro, C. (Eds.), *Organic Geochemistry: Developments and Applications to Energy, Climate, Environment and Human History*, pp. 616–618. A.I.G.O.A.: Donostia: San Sebastian, Spain.

Burton, G. A., Jr., Ingersoll, C. G., 1994. Evaluation of sediment toxicity. In: PTI Environmental Services (Eds.), Assessment Guidance Document, pp. 86–130. Great Lakes National Program Office, USEPA, Chicago, IL. EPA 905-B94-002.

Canfield, T., La Point, T., Swift, M., 1994a. Assessment of benthic invertebrate community structure. In: PTI Environmental Services (Eds.), Assessment Guidance Document, pp.131–175. Great Lakes National Program Office, USEPA, Chicago, IL. EPA 905-B94-002.

Canfield, T. J., Kemble, N. E., Brumbaugh, W. G., Dwyer, F. J., Ingersoll, C. G., Fairchild, J. F., 1994b. Use of benthic invertebrate community structure and the sediment quality triad to evaluate metal-contaminated sediment in the upper Clark Fork River, MT. Environ. Toxicol. Chem. 13, 1999–2012.

Canfield, T. J., Dwyer, F. J., Fairchild, J. F., Haverland, P. S., Ingersoll, C. G., Kemble, N. E., Mount, D. R., La Point, T. W., Burton, G. A., Swift, M. C., 1996. Assessing contamination in Great Lakes sediments using benthic invertebrate communities and the sediment quality triad approach. J. Great Lakes Res. 22, 565–583.

Canfield, T. J., Brunson, E. L., Dwyer, F. J., Ingersoll, C. G., Kemble, N. E., 1998. Assessing sediments from the upper Mississippi River navigational pools using benthic invertebrate evaluations and the sediment quality triad approach. Arch. Environ. Contam. Toxicol. 35, 202–212.

CCME, 1995. Protocol for the derivation of Canadian sediment quality guidelines for the protection of aquatic life. Task Group on Water Quality Guidelines, Canadian Council of Ministers of the Environment, Ottawa, ON.

Chapman, P. M., 1992. Sediment quality triad approach. In: USEPA Sediment Oversight Technical Committee (Eds.), Sediment Classification Methods Compendium, pp. 10–1 to 10–18. Office of Water, USEPA, Washington, DC. EPA 823-R-92-006.

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

Chapman, P. M., Anderson, B., Carr, S., Engle, V., Green, R., Hameedi, J., Harmon, M., Haverland, P., Hyland, F., Ingersoll, C., Long, E., Rodgers, J., Salazar, M., Sibley, P. K., Smith, P. J., Swartz, R. C., Thompson, B., Windom, H., 1997. General guidelines for using the sediment quality triad. Mar. Pollut. Bull. 34, 368–372.

Crane, J. L., 1999. Assessment of contaminated sediments in Slip C, Duluth Harbor, Minnesota. Great Lakes National Program Office, USEPA, Chicago, IL. EPA-905-R-99-007. (http://www.pca.state.mn.us/water/sediments/slipc.pdf)

Crane, J. L., Schubauer-Berigan, M., Schmude, K., 1997. Sediment assessment of hotspot areas in the Duluth/Superior Harbor. Great Lakes National Program Office, USEPA, Chicago, IL. EPA-905-R97-020. (http://www.pca.state.mn.us/water/sediments/94mudpuppy.pdf)

Crane, J. L., MacDonald, D. D., Ingersoll, C. G., Smorong, D. E., Lindskoog, R. A., Severn, C. G., Berger, T. A., Field, L. J., 2000. Development of a framework for evaluating numerical sediment quality targets and sediment contamination in the St. Louis River Area of Concern. Great Lakes National Program Office, USEPA, Chicago, IL. EPA-905-R-00-008. (http://www.pca.state.mn.us/water/sediments/sqt-slraoc.pdf)

Crane, J. L., MacDonald, D. D., Ingersoll, C. G., Smorong, D. E., Lindskoog, R. A., Severn, C. G., Berger, T. A., Field, L. J., 2002a. Evaluation of numerical sediment quality targets for the St. Louis River Area of Concern. Arch. Environ. Contam. Toxicol. 43,1–10.

Crane, J. L., Smorong, D. E., Pillard, D. A., MacDonald, D. D., 2002b. Sediment remediation scoping project in Minnesota Slip, Duluth Harbor. Great Lakes National Program Office, USEPA, Chicago, IL. EPA-905-R-02-002. (http://www.pca.state.mn.us/water/sediments/mnslipreport.pdf)

Diamond, S. A., Milroy, N. J., Mattson, V. R., Heinis, L. J., Mount, D. R., 2003. Photoactivated toxicity in amphipods collected from polycyclic aromatic hydrocarbon-contaminated sites. Environ. Toxicol. Chem. 22, 2752–2760.

Di Toro, D. M., Mahony, J. D., Hansen, D. J., Scott, K. J., Hicks, M. B., Mayr, S. M., Redmond, M. S., 1990. Toxicity of cadmium in sediments: The role of acid volatile sulfide. Environ. Toxicol. Chem. 9, 1487–1502.

Field, L. J., MacDonald, D. D., Norton, S. B., Ingersoll, C. G., Severn, C. G., Smorong, D., Lindskoog, R., 2002. Predicting amphipod toxicity from sediment chemistry using logistic regression models. Environ. Toxicol. Chem. 21, 1993–2005.

Fox, R., Landrum, P., McCrone, L., 1994. Introduction. In: PTI Environmental Services (Eds.), Assessment Guidance Document, pp. 1–9. Great Lakes National Program Office, USEPA, Chicago, IL. EPA 905-B94-002.

Gerritsen, J., 1995. Additive biological indices for resource management. J. N. Am. Benthol. Soc. 14, 451–457.

Ghosh, U., Zimmerman, J. R., Luthy, R. G., 2003. PCB and PAH speciation among particle types in contaminated harbor sediments and effects on PAH bioavailability. Environ. Sci. Technol. 37, 2209–2217.

Gilbert, G. O., 1987. *Statistical Methods in Environmental Pollution Monitoring*. Van Nostrand Reinhold, New York, NY.

Gulley, D. D., WEST, Inc., 1994. TOXSTAT 3.4. WEST, Inc., Cheyenne, WY.

Hilsenhoff, W. L., 1981. Aquatic Insects of Wisconsin. Keys to Wisconsin Genera and Notes on Biology, Distribution, and Species. Natural History Council, University of Wisconsin-Madison, Madison, WI.

Hyland, J. L., Van Dolah, R. F., Snoots, T. R., 1999. Predicting stress in benthic communities of southeastern U.S. estuaries in relation to chemical contamination of sediments. Environ. Toxicol. Chem. 18, 2557–2564.

Hyland, J. L., Balthis, W. L., Engle, V. D., Long, E. R., Paul, J. F., Summers, J. K., Van Dolah, R. F., 2003. Incidence of stress in benthic communities along the U.S. Atlantic and Gulf of Mexico coasts within different ranges of sediment contamination from chemical mixtures. Environ. Monitoring Assess. 81, 149–161.

IJC (International Joint Commission), 1989. Great Lakes water quality agreement of 1978 (as amended by Protocol signed November 18, 1987). IJC, Windsor, ON.

Ingersoll, C. G., MacDonald, D. D., Wang, N., Crane, J. L., Field, L. J., Haverland, P. S., Kemble, N. E., Lindskoog, R. A., Severn, C., Smorong, D. E., 2001. Predictions of sediment toxicity using consensus-based freshwater sediment quality guidelines. Arch. Environ. Contam. Toxicol. 41, 8–21.

IT Corporation, 1997. Remedial investigation data report, sediment operable unit, St. Louis River/Interlake/Duluth Tar site. Volume 1 of 5. International Technology Corp., St. Paul, MN.

Kraaij, R., Seinen, W., Tolls, J., Cornelissen, G., Belfroid, A. C., 2002. Direct evidence of sequestration in sediments affecting the bioavailability of hydrophobic organic chemicals to benthic deposit-feeders. Environ. Sci. Technol. 36, 3525–3529.

Laflamme, R. E., Hites, R. A., 1978. The global distribution of polycyclic aromatic hydrocarbons in recent sediments. Geochim. Cosmochim. Acta. 42, 289–303.

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

Long, E. R., Wilson, C. J., 1997. On the identification of toxic hot spots using measures of the sediment quality triad. Mar. Pollut. Bull. 34, 373–374.

Long, E. R., MacDonald, D. D., Cubbage, J. C., Ingersoll, C. G., 1998. Predicting the toxicity of sediment-associated trace metals with simultaneously extracted trace metal:acid volatile sulfide concentrations and dry weight-normalized concentrations: A critical comparison. Environ. Toxicol. Chem. 17, 972–974.

MacDonald, D. D., Ingersoll, C. G., Berger, T. A., 2000. Development and evaluation of consensus-based sediment quality guidelines for freshwater ecosystems. Arch. Environ. Contam. Toxicol. 39, 20–31.

MacDonald, D. D., Ingersoll, C. G., Smorong, D. E., Lindskoog, R. A., Sparks, D. W., Smith, J. R., Simon, T. P., Hanacek, M. A., 2002. An assessment of injury to sediments and sediment-dwelling organisms in the Grand Calumet River and Indiana Harbor Area of Concern. Arch. Environ. Cont. Toxicol. 43, 141–155.

MDH (Minnesota Department of Health), 2002. Site-specific meal advice for tested lakes and rivers. Environmental Health Division, MDH, St. Paul, MN.

Merritt, R. W., Cummins, K. W. (Eds.), 1996. *An Introduction to the Aquatic Insects of North America*, 3rd ed. Kendall/Hunt Publishing Co., Dubuque, IA.

Microbics Corp., 1994. Microtox7 M500 manual: A toxicity testing handbook. Vol. 3. Microbics Corporation, Carlsbad, CA.

Monson, P. D., Ankley, G. T., Kosian, P. A., 1995. Phototoxic response of *Lumbriculus variegatus* to sediments contaminated by polycyclic aromatic hydrocarbons. Environ. Toxicol. Chem. 14, 891–894.

Morse, J. W., Rickard, D., 2004. Chemical dynamics of sedimentary acid volatile sulfide. Environ. Sci. Technol. 38, 131A–136A.

MPCA (Minnesota Pollution Control Agency), WDNR (Wisconsin Department of Natural Resources), 1992. The St. Louis River system remedial action plan. Stage one. MPCA, St. Paul, MN and WDNR, Madison, WI. (http://www.stlouisriver.org/rap.html)

MPCA, WDNR, 1995. The St. Louis River system remedial action plan. Progress report. MPCA, St. Paul, MN and WDNR, Madison, WI. (http://www.stlouisriver.org/stage2/stage2_report.pdf)

Pardos, M., Benninghoff, C., Thomas, R. L., Khim-Heang, S., 1999. Confirmation of elemental sulfur toxicity in the Microtox® assay during organic extracts assessment of freshwater sediments. Environ. Toxicol. Chem. 18, 188–193.

Peterson, G. S., Axler, R. P., Lodge, K. B., Schuldt, J. A., Crane, J. L., 2002. Evaluation of a fluorometric screening method for predicting total PAH concentrations in contaminated sediments. Environ. Monitoring Assess. 78, 111–129.

Prahl, F. G., Carpenter, R., 1984. Hydrocarbons in Washington coastal sediments. Estuarine Coastal Shelf Sci. 18, 703–720.

Reddy, C. M., Pearson, A., Xu, L., McNichol, A. P., Benner Jr., B. A., Wise, S. A., Klouda, G. A., Currie, L. A., Eglinton, T. I., 2002. Radiocarbon as a tool to apportion the sources of polycyclic aromatic hydrocarbons and black carbon in environmental samples. Environ. Sci. Technol. 36, 1774–1782.

Redman, S., Janisch, T., 1995. Newton Creek system sediment contamination site characterization report. Sediment Management and Remediation Techniques Program, WDNR, Madison, WI. PUBL-WR-433-95.

SAS, 1988. SAS User's Guide: Statistics. SAS Institute Inc., Cary, NC.

Schubauer-Berigan, M., Crane, J. L., 1996. Preliminary contaminant assessment of the Thomson, Forbay, and Fond du Lac Reservoirs. Water Quality Division, MPCA, St. Paul, MN. (http://www.pca.state.mn.us/water/sediments/cloquet.pdf)

Schubauer-Berigan, M., Crane, J. L., 1997. Survey of sediment quality in the Duluth/Superior Harbor: 1993 sample results. Great Lakes National Program Office, USEPA, Chicago, IL. EPA 905-R97-005. (http://www.pca.state.mn.us/water/sediments/93mud.pdf)

SEH Inc., 2000. Site investigation report. Newton Creek segments B and C, Superior, Wisconsin. Short Elliott Hendrickson Inc., Chippewa Falls, WI. SEH No. WIDNR9905.00.

SEH Inc., 2003. Remedial investigation report. Newton Creek, Douglas County, Wisconsin. Short Elliott Hendrickson Inc., Chippewa Falls, WI. SEH No. WIDNR9905.02.

St. Louis River CAC, 2002. Lower St. Louis River habitat plan. St. Louis River Citizens Action Committee, Duluth, MN. (http://www.stlouisriver.org/habitatplan/habitatplan.html)

Suter, G. W. II, 2001. Applicability of indicator monitoring to ecological risk assessment. Ecol. Indicat. 1, 101–112.

Thorp, J. H., Covich, A. P. (Eds.), 1991. *Ecology and Classification of North American Freshwater Invertebrates*. Academic Press, Inc., San Diego, CA.

Tung, K. K., Scheibner, G., Miller, T., Bulich, A. A., 1990. A new method for testing soil and sediment samples. SETAC 11th Annual Meeting Abstract Book. Society of Environmental Toxicology and Chemistry, Pensacola, FL.

URS Corporation, 2002. Former Duluth works LIF investigation report. URS Corporation, Minneapolis, MN.

USACE (United States Army Corps of Engineers), 1997. Phase II report. Draft dredged material management plan study and environmental impact statement. Duluth-Superior Harbor, Minnesota and Wisconsin. USACE, Detroit District, Detroit, MI.

USEPA (United States Environmental Protection Agency), 1993. Regional Environmental Monitoring and Assessment Program (R-EMAP). Office of Research and Development and Office of Science, Planning and Regulatory Evaluation, USEPA, Cincinnati, OH. EPA/625/R-93/012.

USEPA, 1994. Methods for measuring the toxicity and bioaccumulation of sediment-associated contaminants with freshwater invertebrates. Office of Research and Development, USEPA, Duluth, MN. EPA/600/R-94/024.

USEPA, 1997. The incidence and severity of sediment contamination in surface waters of the United States. Volume 1: National sediment quality survey. Office of Science and Technology, USEPA, Washington, DC. EPA 823-R-97-006.

Venkatesan, M. I., 1988. Occurrence and possible sources of perylene in marine sediments. A review. Mar. Chem. 25, 1–28.

Wakeham, S. G., Schaffner, C., Giger, W., 1980a. Polycyclic aromatic hydrocarbons in recent lake sediments—II. Compounds derived from biogenic precursors during early diagenesis. Geochim. Cosmochim. Acta. 44, 415–430.

Wakeham, S. G., Schaffner, C., Giger, W., 1980b. Polycyclic aromatic hydrocarbons in recent lake sediments—I. Compounds having anthropogenic origins. Geochim. Cosmochim. Acta. 44, 403–413.

Wiederholm, T. (Ed.), 1983. Chironomidae of the Holarctic region–Keys and diagnoses. Vol. 1. Larvae. Ent. Scand. Suppl. 19, 1–457.

Appendix A. Compilation of mean PEC-Q values, sediment chemistry, toxicity test results, and benthological classifications of random and training sites in the lower St. Louis River AOC.

					Microtox®		Mean survival (%)		Benthic classification	
Site number	Habitat Class	Mean PEC-Q	TOC (%)	[SEM-AVS] (μmol g^{-1}) (dry wt.)	Solid phase (EC_{50}[a])	Pore water (% reduction[b])	*H. azteca*	*C. dilutus*	Multimetric classification	DFA category
Random sites										
3	1	0.068	1.65	−0.097	1	9			Low Impact	Low Impact
4	1	0.0134	0.20	−0.504	4.9	**25**	86	85	Low Impact	Low Impact
5	1	0.309	4.05	−0.09	**0.4**	8	67	70	High Impact	High Impact
6	1	0.302	3.76	1.12	**0.3**	11	44	58	Indeterminate	High Impact
7	1	0.0129	0.14	0.136	17.8	**67**	58	75	High Impact	Low Impact
8	1	0.0256	0.14	0.2	9.1	**45**	64	82	Low Impact	High Impact
11	1	0.00948	0.08	0.088	>20	**40**	86	90	High Impact	High Impact
14	1	0.393	3.47	−1.02	**0.5**	**20**	94	70	Low Impact	High Impact
15	1	0.348	2.13	1.41	**0.5**	**30**	86	82	High Impact	High Impact
20	1	0.252	4.11	0.635	**0.1**	2	85	78	High Impact	Low Impact
21	1	0.217	4.26	0.32	0.6	5			Low Impact	Low Impact
22	1	0.202	4.41	1.38	**0.3**	3	92	68	Indeterminate	High Impact
26	1	0.157	3.69	0.281	**0.4**	0	94	75	Low Impact	High Impact
33	1	0.305	4.47	−0.845	0.9	**33**	85	68	Low Impact	Low Impact
35	1	0.0253	1.02	0.371	0.8	**30**	90	80	Low Impact	Low Impact
36	1	0.169	4.78	1.01	0.9	8			Indeterminate	Low Impact
37	1	0.185	5.85	−0.078	**0.4**	0			High Impact	Low Impact
38	1	0.0112	0.59	0.091	4	0	85[c]	80	High Impact	High Impact
39	1	0.186	20.20	−0.331	2.3	4			Low Impact	Low Impact
40	1	0.093	19.23	−0.981	1	0			Low Impact	Low Impact
41	1	0.100	9.90	−0.092	2.2	0			Low Impact	Low Impact
43	1	0.329	4.54	−25.5	**0.1**	**30**	69	5[d]	High Impact	High Impact
53	1	0.053	4.18	−0.362	4.8	0			High Impact	Low Impact
54	1	0.177	2.88	−0.786	**0.5**	0	75[c]	60	Low Impact	Low Impact
64	1	0.034	0.19	0.519	8.6	0			Indeterminate	Low Impact
65	1	0.024	0.02	0.432	>20	No data			Low Impact	LowImpact

66	1	0.0177	0.12	0.323	>20	3	95	78	High Impact	High Impact
69	1	0.396	4.36	−1.4	1	0	68	85	Low Impact	High Impact
70	1	0.43	13.98	−7.84	**0.2**	**30**	56	82	High Impact	HighImpact
71	1	0.915	4.13	−1.77	**0.3**	**20**	90	72	High Impact	Low Impact
73	1	1.01	3.07	−2.57	0.7	**32**	75[c]	55	Low Impact	Low Impact
74	1	0.164	4.10	0.781	1.3	**32**			Low Impact	High Impact
75	1	0.019	0.55	0.271	6.3	**33**			Low Impact	Low Impact
76	1	0.0467	1.82	0.267	3.6	0	70	75	Low Impact	High Impact
79	1	0.075	2.82	1.20	**0.3**	7	70	72	Low Impact	Low Impact
81	1	0.188	3.30	0.764	**0.3**	0			Low Impact	Low Impact
82	1	0.152	3.32	−0.051	0.7	8	55	88	Low Impact	Low Impact
83	1	0.164	4.32	−3.18	**0.4**	4			Low Impact	High Impact
84	1	0.0936	4.84	−21.7	**0.1**	18	92	90	Low Impact	Low Impact
85	1	0.0366	0.86	0.326	14.2	3	90	72	Low Impact	Low Impact
87	1	0.0463	0.68	0.21	2.8	**21**	80[c]	55	Low Impact	High Impact
88	1	0.113	3.14	−1.84	**0.5**	**22**	90[c]	70	Low Impact	High Impact
89	1	0.158	3.73	−1.65	**0.4**	11			Low Impact	Low Impact
92	1	0.080	2.35	−1.93	1.1	4			Low Impact	Low Impact
93	1	0.0433	0.61	−0.339	5.4	16	98	65	High Impact	High Impact
94	1	0.211	10.68	−13.7	**0.3**	15	72[c]	70	Low Impact	Low Impact
95	1	0.0837	2.19	−0.216	4.2	**27**	98	75	Low Impact	High Impact
96	1	0.333	17.49	−30.1	1.5	0	75	72	Low Impact	High Impact
97	1	0.200	9.56	−4.41	0.9	0			Low Impact	Low Impact
98	1	0.135	5.34	0.667	1	19	100	75	Low Impact	Low Impact
100	1	0.031	0.73	−0.541	1.3	9			Low Impact	High Impact
101	1	0.117	3.42	−2.84	1.4	**24**			Low Impact	High Impact
103	1	0.131	3.94	−0.676	1.1	**20**			Low Impact	High Impact
104	1	0.149	3.48	−1.17	0.6	**31**			Low Impact	High Impact
105	1	0.255	13.81	−0.741	3.9	0	62	98	Low Impact	High Impact

(Continued on next page)

Appendix A. Compilation of mean PEC-Q values, sediment chemistry, toxicity test results, and benthological classifications of random and training sites in the lower St. Louis River AOC. *(Continued)*

Site number	Habitat Class	Mean PEC-Q	TOC (%)	[SEM-AVS] (μmol g^{-1}) (dry wt.)	Microtox® Solid phase (EC_{50}[a])	Microtox® Pore water (% reduction[b])	Mean survival (%) *H. azteca*	Mean survival (%) *C. dilutus*	Benthic classification Multimetric classification	Benthic classification DFA category
107	1	0.0403	5.77	−1.60	6.7	6	75	75	Low Impact	High Impact
110	1	0.156	3.55	−0.187	0.8	8			Low Impact	High Impact
111	1	0.362	3.98	−2.49	**0.4**	9	100	72	Low Impact	Low Impact
1	2	0.047	0.44	0.396	24.3	0			Low Impact	Low Impact
9	2	0.041	0.24	0.123	>20	8			High Impact	High Impact
10	2	0.023	0.40	0.142	>20	**40**	84	86	High Impact	Low Impact
12	2	0.171	2.67	−0.078	**0.4**	**31**	91	60	Low Impact	High Impact
13	2	0.0589	0.81	0.019	2.1	**27**	89	18[d]	Low Impact	Low Impact
16	2	0.162	3.64	−2.07	0.9	10			Indeterminate	High Impact
17	2	0.016	0.16	0.236	13.9	0			High Impact	High Impact
18	2	0.0615	1.85	−0.038	0.8	15	78	90	High Impact	High Impact
23	2	0.206	3.38	−0.343	**0.3**	19	88	35	High Impact	High Impact
24	2	0.138	0.24	0.194	>20	15	72	52	High Impact	Low Impact
25	2	0.034	0.76	0.225	8.4	8			High Impact	High Impact
27	2	0.126	2.09	0.711	>20	0			Indeterminate	High Impact
28	2	0.138	2.34	0.049	**0.3**	16	98	80	Low Impact	High Impact
29	2	0.182	3.07	−0.487	**0.2**	0	95	70	High Impact	High Impact
30	2	0.113	1.94	−0.028	0.6	**40**	98	62	High Impact	High Impact
31	2	0.0769	1.35	0.845	1	0	92	70	Indeterminate	High Impact
32	2	0.063	1.42	0.575	0.7	**42**	100	60	High Impact	High Impact
34	2	0.013	0.30	0.132	1.4	0			High Impact	High Impact
45	2	0.059	1.40	0.453	**0.5**	0			Indeterminate	High Impact
46	2	0.0768	1.10	0.507	0.8	0	90	62	High Impact	Low Impact
47	2	0.015	0.16	0.095	1	0	85[c]	90	High Impact	High Impact
48	2	0.176	2.93	0.519	**0.2**	14	88[c]	80	High Impact	High Impact
49	2	0.025	0.52	0.232	9.1	13			High Impact	Low Impact
57	2	0.0648	1.55	−1.22	**0.5**	0	88[c]	65	High Impact	High Impact
59	2	0.067	2.55	0.171	1.5	0			High Impact	High Impact
60	2	0.157	2.96	0.126	0.9	0			Low Impact	Low Impact

61	2	0.2	4.02	−0.374	0.6	0	90	90	High Impact	High Impact
62	2	0.106	2.34	−0.432	0.9	0			Low Impact	Low Impact
68	2	0.105	1.94	−0.774	0.9	0			Low Impact	Low Impact
Training sites										
Low impact										
42	1	0.0105	0.35	0.002	11.2	**26**	78[c]	55	Low Impact	Low Impact
52	1	0.0924	3.20	−1.26	0.6	0	100	62	Low Impact	Low Impact
63	1	0.029	0.28	0.313	23.6	0			High Impact	Low Impact
77	1	0.248	4.24	−1.06	**0.1**	**53**	41[d]	70	High Impact	Low Impact
78	1	0.14	3.20	0.579	1.1	**46**	82[c]	50	Low Impact	Low Impact
80	1	0.0734	2.59	1.03	**0.3**	7	98	65	Low Impact	Low Impact
86	1	0.0481	0.56	−0.344	2.1	12	98	68	Low Impact	Low Impact
90	1	0.043	1.37	−1.07	1.2	2			Indeterminate	Low Impact
91	1	0.080	2.49	−1.73	1.1	0			Low Impact	Low Impact
106	1	0.147	3.07	−1.29	1.9	**37**			Low Impact	Low Impact
High impact										
2	1	1.41	4.70	4.25	2.6	13	88	65	High Impact	High Impact
44	1	0.29	3.99	−71.3	**0.2**	**67**	75	10[d]	High Impact	High Impact
51	1	1.1	5.59	−4.34	**0.2**	4	95	72	High Impact	High Impact
55	1	0.328	2.63	−1.10	**0.3**	**34**	75	68	Low Impact	High Impact
56	1	0.748	14.19	−23.4	**0.2**	**68**	12[d]	0[d]	High Impact	High Impact
72	1	26.4	8.26	−29.7	**0.1**	**100**	0[d]	0[d]	Indeterminate	High Impact
99	1	0.445	0.82	4.49	**0.1**	**26**	59	0[d]	High Impact	High Impact
102	1	0.0717	8.40	−2.87	**0.2**	**82**	3[d]	0[d]	High Impact	High Impact
108	1	0.503	5.96	−9.97	**0**	8			High Impact	High Impact
109	2	0.595	2.73	2.79	**0.3**	9			Low Impact	High Impact

Habitat Class 1 = Shallow habitat; Habitat Class 2 = channelized habitat; PEC-Q = probable effect concentration quotient; TOC = total organic carbon; SEM = simultaneously extractable metal; AVS = acid volatile sulfide; DFA = discriminant function analysis; EC_{50} = effective concentration that caused a 50% reduction in luminescence relative to the control.

[a] EC50 as percent dry sediment; toxic response (shown in bold) defined as EC50 < 0.5 (± 0.05)%.

[b] Percent light reduction relative to control; toxic response (shown in bold) defined as >20 (± 0.5)%.

[c] Control failure (i.e., 78% survival); therefore, the toxicity test failed quality assurance requirements.

[d] Mean survival significantly less than control survival at $p = 0.05$.