

## Exploiting habitat and gear patterns for efficient detection of rare and non-native benthos and fish in Great Lakes coastal ecosystems

Anett S. Trebitz\*, John R. Kelly, Joel C. Hoffman, Gregory S. Peterson and Corlis W. West

*U.S. Environmental Protection Agency, Office of Research and Development, Mid-Continent Ecology Division, 6201 Congdon Boulevard, Duluth MN 55804, USA*

*E-mail: [trebitz.anett@epa.gov](mailto:trebitz.anett@epa.gov) (AST), [kelly.johnr@epa.gov](mailto:kelly.johnr@epa.gov) (JRK), [hoffman.joel@epa.gov](mailto:hoffman.joel@epa.gov) (JCH), [peterson.greg@epa.gov](mailto:peterson.greg@epa.gov) (GSP), [west.corlis@epa.gov](mailto:west.corlis@epa.gov) (CWW)*

*\*Corresponding author*

Received 14 August 2009; accepted in revised form 6 November 2009; published online 19 November 2009

### Abstract

Despite the continued arrival and impacts of non-native aquatic species in the Great Lakes, there is as yet no comprehensive early-detection monitoring program for them. As a step towards implementing such a program, we evaluated strategies for efficient non-native species monitoring based on the ability to detect a diverse set of benthos and fish species currently present in a heavily invaded, spatially complex Great Lakes subsystem. Taxa accumulation analyses confirmed that reliable detection of rare species requires substantial sampling effort but also that there is potential for exploiting patchiness in distributions to increase efficiency. While non-native species monitoring warrants generally comprehensive spatial coverage, it may be possible to identify areas where such taxa are broadly most prevalent (e.g., the lower reaches of our study system) as a way to focus effort. On a finer scale, richness of non-native taxa may vary substantially among stations in close proximity – which in this system was driven by habitat variability rather than distance from potential introduction points. Microhabitats that differ in physical attributes are also likely to differ in species composition and richness. Randomization analyses indicated that some monitoring effort should be directed towards all distinct habitats but that detection rates are maximized by biasing effort towards those habitats or gear yielding the most total, non-native, or rare taxa. For benthic invertebrates, shallow structurally complex (vegetated) habitats yielded the most taxa but shallow open and deep habitats also contributed unique taxa. For fish, fyke-net stations (shallowest habitats) yielded the most taxa, but electrofishing (intermediate-depth) and trawling (deepest) also contributed unique taxa. Our approach to identifying relevant sampling strata and exploiting difference among them to increase the efficiency of early-detection monitoring is applicable to a broad variety of systems.

**Key words:** non-native species, early detection, sampling strategies, optimization, St. Louis River/Duluth-Superior Harbor

### Introduction

To date, the Laurentian Great Lakes have been invaded by almost 190 species of aquatic plants and animals (Mills et al. 1993; Grigorovich et al. 2003a; Ricciardi 2006). Some of these species have been deliberately introduced; many others have been transported into the Great Lakes unintentionally through transoceanic shipping (Grigorovich et al. 2003a; Duggan et al. 2005; Drake and Lodge 2007), recreational boating and fishing (Ludwig and Leitch 1996), and aquaria and horticultural trades (Rixon et al. 2005; Cohen et al. 2007) or have used ship passage

structures to circumvent natural migration barriers. These non-native species have had significant ecological and economic impacts in the Great Lakes (Mills et al. 1993; Vanderploeg et al. 2002; Pimentel et al. 2005), and large sums of money are routinely spent on their control. Even if efforts to slow the arrival and spread of additional species are successful, the increasing global connectivity of human travel and commerce and the lag-times sometimes present between arrival and adverse impacts mean that non-native species are likely to remain a threat for the foreseeable future (Levine and D'Antonio 2003; Ricciardi 2006).

Given the difficulty of controlling non-native species once they become established, detecting their presence while they are still few in number or localized in distribution is key to a successful management response (Meyers et al. 2000; Hulme 2006; Mehta et al. 2007). Unfortunately, achieving some substantial detection probability for new and still rare species is an inherently labor intensive task (Rew et al. 2006; Mehta et al. 2007; Harvey et al. 2009). Time and resource constraints on monitoring programs are inevitable, so information that could be used to make the process more efficient is desirable. One approach would be to target high-risk species and locations as identified through analyses of vector intensity, invader characteristics, and similarity of source and target environments (e.g., Kolar and Lodge 2002; Grigorovich et al. 2003a; Holeck et al. 2004; Herborg et al. 2007). However, with the large number of invaders that have the potential to arrive at any given location in the Great Lakes, a species-by-species approach to monitoring is likely to be inefficient and insufficient. Rather, monitoring ought to be sufficiently comprehensive to detect a potentially large suite of new invaders across a potentially wide variety of locations and habitats. Moreover, as actions are taken to reduce the potential for introductions, a generalized monitoring design should aid in accounting for the effectiveness of such actions. There is, at present, no such comprehensive monitoring program in use for invasion-vulnerable coastal areas of the Great Lakes.

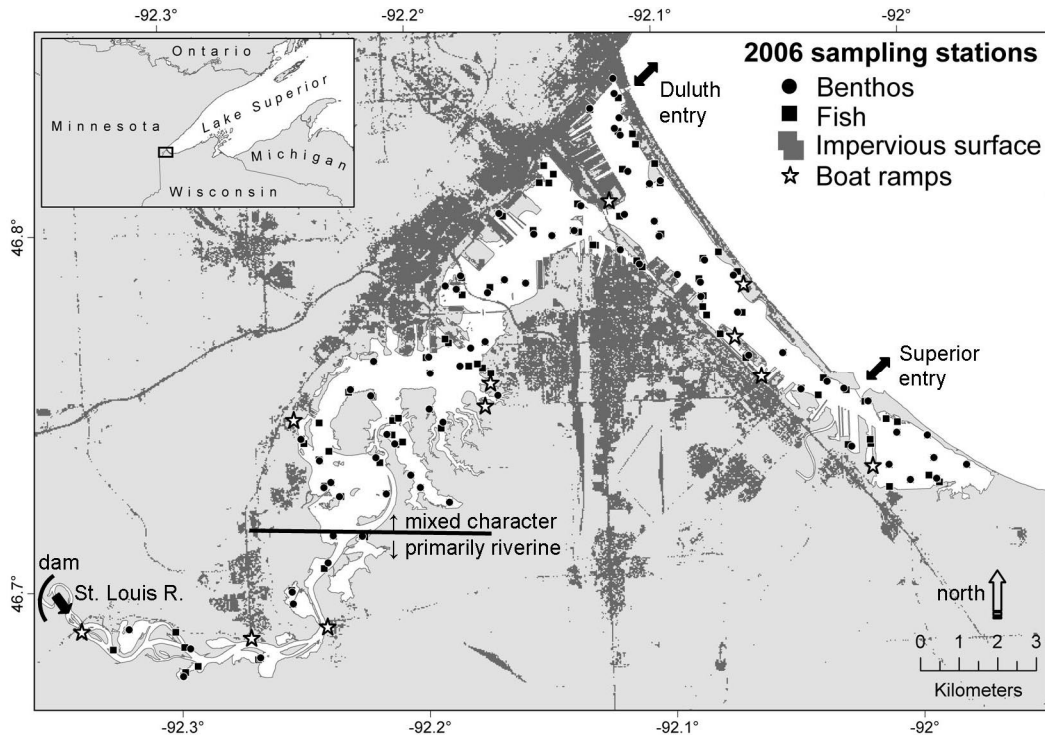
The overall goal of our study was to define an approach for early-detection monitoring for non-native species in general. To this end, we conducted intensive sampling in an invasion-vulnerable Great Lakes sub-system with a variety of sampling designs and gear, so as to amass a comprehensive data set with which to explore monitoring strategies. This paper focus on the question of whether efficiencies could be gained by exploiting patterns in species composition among habitats and sampling gear. We use a combination of endpoints including total richness, non-native richness, and richness of rare taxa, under the premise that strategies for detecting future invaders (with unknown abundances, introduction vectors, life history strategies, and habitat preferences) can be evaluated based on their performance for the diversity of species currently present, regardless of origin. We juxtapose findings for benthic invertebrate assemblages and fish assemblages,

so that similarities and differences among taxonomic groups can be evaluated to attain greater generality. Analytical techniques used include taxa accumulation curves to evaluate sampling sufficiency and efficiency, descriptive analyses of taxa composition patterns and environmental covariates, and randomization analyses to identify optimal allocations of effort among different station types.

#### *St. Louis River/Duluth-Superior Harbor as case study*

Our study was conducted in the St. Louis River/Duluth-Superior Harbor system formed at the confluence of the river with Lake Superior (Figure 1). The lower portion of the system is flanked by the cities of Duluth, Minnesota and Superior, Wisconsin, and hosts the largest commercial seaport on the Great Lakes. Water in the lower system is relatively cold and clear due to exchange with Lake Superior through two openings in the barrier beach, but urban and industrial development have resulted in widespread shoreline armoring, dredging of shipping channels through shallow sand and mud flats, and loss of once extensive wetlands. In contrast, the more upstream portion of the system is largely undeveloped, contains a greater variety of habitats including sinuous river channels, small islands, shallow flats, and backbay and fringing wetlands, and has warmer and less clear water due to inputs from the St. Louis and several smaller rivers (Figure 1).

The St. Louis River/Duluth-Superior Harbor provides habitat for many resident and migratory fishes and is a popular fishing destination (Lindgren et al. 1997). Potential entry vectors for non-native aquatic species include commercial shipping (ballast water, tank residue, hull fouling), recreational boating and fishing (bait buckets, bilge water, hull and trailer fouling), and general urbanization (land transportation corridors, aquarium dumping). Propagule pressure from the various introduction vectors is apparently high and non-native species are able to prosper in the habitats present, as the system is heavily invaded despite the relative inhospitality of Lake Superior proper to non-native species (Grigorovich et al. 2003b). The St. Louis River/Duluth-Superior Harbor thus provides an excellent case study for examining strategies for the generalized detection of a potentially wide variety of non-native species over a large and spatially complex system.



**Figure 1.** Map of the St. Louis River/Duluth-Superior Harbor system, showing stations sampled for benthos and fish, distribution of natural versus developed (impervious) land, and location of boat ramps. Harbor facilities (shipping docks) are evident as linear indentations and projections along the shoreline. A division between the upstream, primarily riverine and the downstream, more mixed character portion of the system is indicated

## Methods

### *Study design and environmental characterization*

This paper primarily uses data from a comprehensive, system-wide sampling for benthos and fish conducted in 2006, although preliminary sampling was conducted in the lower system in 2005 and additional sampling for fish was conducted in 2007. Sampling in 2006 used two different designs, each of which distributed stations across the system. The first was a spatially-balanced random-probability design (Stevens and Olson 2004) without stratification. The second design targeted sampling locations by entering spatial data on habitat attributes and distances to possible introduction points into a cluster analysis, and manually distributing stations across the resulting cluster map to cover the diversity of areas present. The relative efficacy of these designs will be compared in a separate paper; samples obtained from both were pooled for the analyses described herein.

To aid in analyzing species – environment relationships and targeting sampling areas, publicly available spatial data were used to map several habitat and landscape setting variables on a 100x100 m grid prior to field sampling. Overwater distances from each grid-cell to potential introduction points were computed from data obtained from the States of Minnesota and Wisconsin (boat ramp locations), the St. Louis River Citizens Action Committee (dock and shipyard locations), and National Land Cover Data (road and railroad networks). Distance of each grid-cell from the dam at the upstream end of the system was also calculated. Density of wetlands within a 250 m radius of each grid-cell was computed from the National Wetland Inventory and the Wisconsin Wetland Inventory. Fetch (distance to nearest shore) was characterized for three locally dominant wind directions (west, northeast, and southwest) using the ACOE (1984) method as implemented in an ArcGIS script. Percent fine material (<75  $\mu\text{m}$  size particles) and total organic carbon in the top

sediment layer were interpolated from a Minnesota Pollution Control Agency sediment-quality database. Additional environmental data collected in the field included station depth, near-surface water temperature, turbidity, and dissolved oxygen (multiprobe sensor), sediment density (dry weight as percent of wet weight), and vegetation cover. Plant cover was scored for each of the emergent, floating-leaf, and submerged vegetation zone and summarized as the maximum over all three (0 = no vegetation, 1 = very sparse plants, 2 = substantial vegetation but more than 2× as much open as vegetated area, 3 = vegetated and open area similar, and 4 = vegetated area more than 2× open area).

#### *Benthos and fish sampling and taxonomy*

Field sampling for benthic macroinvertebrates (benthos) and fish sought to characterize the complete assemblage at each station. Sampling took place during late summer, to allow peak development of aquatic vegetation and recruitment of young-of-year fish to the sampling gear. Benthos were collected with a single petite ponar grab sampler (area 236 cm<sup>2</sup>) at each station. Organisms were separated from sediments using a 500 µm screen and preserved in 10 percent buffered formalin for laboratory identification. Fish were collected by one of three types of gear, depending on station depth. Stations in 0-1 m of water were sampled with fyke nets (paired nets set parallel to shore for 24 hr, 48 mm mesh; 0.9 m deep × 1.2 m wide frame). Stations in 1-2 m of water were sampled by daytime electrofishing (10 min. effort, 5000 watt generator, 120 hz, 6-8 amp DC current). Stations >2 m deep were sampled with an otter trawl (5 min tows at ~3 km/hr, 38 mm mesh net with 3 mm cod-end liner). Fish that could not be positively identified in the field were photographed and/or preserved for laboratory identification.

In order to maximize taxa detection rates, we made complete sample counts, identified organisms to the lowest taxonomic level possible, and verified potential non-native species with help of expert taxonomists (and DNA typing in the case of Dreissenid mussels; Grigorovich et al. 2008). However, poor preservation prevented complete identification of some specimens, and immature stages of some taxa cannot be fully keyed out. Our approach to resolving the resulting taxonomic ambiguities (Cuffney et al. 2007) was to assign poorly-identifiable specimens and undifferentiated life stages to whichever fully-

identified taxon was most common at that station or across stations. This distorts relative abundance (only the most common taxa gain individuals), but is more conservative of richness than the alternatives of distributing proportionally among all potential taxa (which can inflate station-level richness) or of collapsing taxonomy to the lowest common denominator (which discards taxonomic information; Cuffney et al. 2007).

For benthos, taxonomy was generally to species or genus and all non-natives were identified to species (at least as adults), but some taxa were only identified to family (due to lack of taxonomic keys or the issues raised above). Our 2005 pilot sampling yielded the first detection of quagga mussels, *Dreissena bugensis*, in Lake Superior (Grigorovich et al. 2008), but since zebra mussels (*D. polymorpha*) far outnumbered quagga mussels among adults, we assigned all juvenile Dreissenids to the zebra mussel category. Fish identifications were to species except that native sunfish (pumpkinseed, *Lepomis gibbosus* and bluegill, *L. macrochirus*) and bullheads (black, *Ameiurus melas* and brown, *A. nebulosis*) were analyzed as genera. These species cannot be reliably distinguished as young-of-year in the field and older specimens were too infrequent to guide species assignments.

#### *Richness-based performance metrics*

Because *detection* is the key issue in monitoring for non-native species, we focused on presence-based (i.e., richness) rather than abundance-based endpoints for performance assessment. The number of non-native taxa detected is obviously of interest, but since not all the non-natives present are necessarily new or rare, the total number of taxa detected is also relevant to evaluating strategies for detection of future invaders, as is the number of rare taxa detected. The meaning of “rare” adopted in this paper is that a taxa has low occurrence frequency, which we measured both as the richness of taxa occurring at ≤ 20% of the stations (Rare-20) and as the richness of taxa occurring at ≤ 5% of the stations (Rare-5) sampled. Taxa contributing to richness in the rare-20 or rare-5 categories were based on the full 2006 data set, so that any given station or set of stations could be compared to this overall pool. Depending on the analyses, richness metrics were computed either for individual stations (to evaluate gear and habitat

effects) or across sets of stations (to evaluate effort allocation strategies).

#### *Taxa accumulation analyses*

We used *rarefaction* curves (statistical expressions of the taxa accumulation pattern over multiple random re-orderings of the sampling sequence) to examine the rate at which taxa were detected and to determine the number of stations to consider in later randomization analyses. We computed sample-based rarefaction curves (after Mao et al. 2005), individual-based rarefaction curves (after Coleman 1981), and the Chao asymptotic richness estimator (Chao 1987) using EstimateS software (Colwell 2006). Sample-based rarefaction accounts for spatial patterns in distributions (entire “blocks” of taxa are acquired at once) while individual-based rarefaction ignores the co-distribution among taxa. The two types of curves eventually converge to the same asymptote for any given data set, but the difference in the *initial* taxa accumulation rate between the two indicates the degree of heterogeneity in the distribution of organisms (Gottelli and Colwell 2001) and by extension, the degree to which sampling efficiency could potentially be gained by exploiting that heterogeneity. The nonparametric Chao estimator is based on the number of taxa captured only once (singletons) or twice (doubletons) and applies even when the rarefaction curve itself lacks a well-defined asymptote. Following Colwell (2006), we report the bias-corrected form of the Chao estimator when its coefficient of variation was  $<0.5$  but its un-corrected form otherwise.

#### *Descriptive analyses*

Descriptive analyses were used to examine how patterns in fish or benthos composition among stations were related to environmental or sampling attributes, and to identify appropriate categorizations of continuous environmental data (e.g., water depth) in support of randomization analyses. We examined composition patterns using both whole-assembly analyses and summary richness metrics. Richness metrics were related to continuous station descriptors (sediment character, water quality, fetch, distance to introduction points) via scatter plots and Pearson correlation analyses, and to categorical station descriptors (gear type, vegetation cover, depth categories) via box plots and ANOVA (Tukey HSD test). Because

vegetation only occurred at stations  $<3$  m deep, a depth stratification at 3 m was used prior to other habitat association analyses. Assemblage patterns were summarized by non-metric multi-dimensional scaling ordination using Bray-Curtis similarity (Primer software; Clarke and Warwick 2001) computed from presence/absence data so as to focus on patterns of occurrence rather than abundance. Assemblage patterns were related to categorical station descriptors via coding of points in ordinations and via ANOSIM (analysis of similarity) tests of group structure (a non-parametric analogue of ANOVA; Clarke and Warwick 2001).

#### *Randomization analyses*

In contrast to descriptive analyses which examined taxa composition at the station level, randomization analyses examined taxa aggregated over sets of stations, in order to identify the most effective allocation of some fixed sampling effort among different combinations of gear or habitat types. Randomizations were based on 20-station sets, as representing a point along the taxa accumulation curves where a large portion (~75 %) of the total taxa had been obtained yet there remained substantial potential for increasing sampling efficiency (separation between sample-based and individual-based rarefaction curves). We worked with triads of categories (trawl, fyke, and electrofishing gear for fish; deep, shallow-vegetated, and shallow-open habitat for benthos) which could be depicted as combinations in a two-dimensional triangular space, and generated 10 random draws for each of the 66 combinations possible under even-number allocation of categories. For example, to generate the various mixes of fishing gear, we randomly selected 0, 2, 4, 6, ... 18, or 20 trawl records and for each of these added 0, 2, ... 20 fyke records and 0, 2, ... 20 electrofishing records as needed to achieve a total of 20 stations. The process of randomizing station combinations and cumulating richness metrics across them was automated with custom programs written in C++ (Boorland International, Inc.) and SYSTAT (SPSS, Inc.) software. Results were depicted as contour plots using richness-endpoints averaged across all 10 draws for each station combination as the underlying data grid.

To check that our findings concerning optimal station combinations also apply to larger sample sizes, we ran additional randomization tests for

fish using 40 or 60 stations per draw. In order to obtain enough stations for each gear type, we added data from a 2007 sampling in the St. Louis River/Duluth-Superior Harbor system (same field protocols and target-zone design) to the 2006 data already described. Station combinations were constructed and evaluated exactly as for the 20 station draws, except that station types were allocated in sets of 4 or 6 instead of 2 to maintain a constant overall number of combinations analyzed (i.e., 66). Analyses with larger sample sizes were not possible for benthic invertebrates because we did not sample for them in 2007 and the pilot data collected in 2005 did not cover a comparable area.

## Results

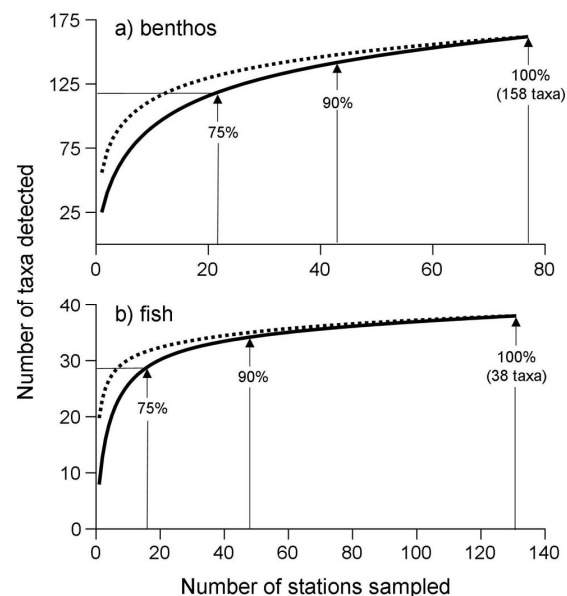
### *Patterns of richness and rareness*

Petite ponar grab samples from 77 stations in the St. Louis River/Duluth-Superior Harbor system yielded 158 benthic invertebrate taxa. The 16 non-native benthic species (Table 1) differed widely in occurrence frequency and length of time since first detection, and these two factors were not necessarily related (e.g., one newly detected species had >20% occurrence, while some long present species had  $\leq 2\%$  occurrence). Of all benthic taxa, 114 (72%) were found at  $\leq 20\%$  of stations (rare-20, Table 2) and included thirteen non-native taxa, and 59 (37%) were found at  $\leq 5\%$  of stations (rare-5, Table 2) and included nine non-native taxa. No single habitat type yielded all the taxa or all the non-native or rare taxa, although shallow vegetated habitats yielded the highest percentage of the total for all four endpoints (Table 2). The most taxa-rich single station (52 taxa) yielded less than one-third of the pool of non-native or rare benthos (non-native = 5, rare-5 = 7, rare20 = 24) and the most taxa-poor stations yielded <10 taxa each.

We collected 38 taxa of fish across the 131 stations sampled (Table 2). Our collection included all 9 previously known summer non-native fish species (Table 1; non-native salmonids migrate into the system in spring and fall – Lindgren et al. 1997). As with the benthos, some non-native fish species were widely distributed while others occurred at only a few stations (Table 1). Ten fish taxa (26% of the total) including two non-native species were found at  $\leq 5\%$  of stations, and 23 taxa (61%) including five non-native species were found at  $\leq 20\%$  of

stations (Table 2). No single gear collected all fish taxa or all rare taxa, although fyke nets did collect all non-native species (Table 2). The most taxa-rich single station (17 taxa) yielded less than one-fourth of the rare fish (rare-5 = 2, rare-20 = 5) and only two-thirds of the non-native species, while the most taxa-poor stations yielded <5 taxa each.

Taxa accumulation curves (rarefaction curves) for both benthos and fish continued to rise slowly over a large number of stations rather than leveling off quickly (Figure 2), indicating the presence of a number of patchily distributed taxa for which reliable detection would require a substantial overall sampling effort. For benthos, the number of taxa accumulated over all stations actually sampled was well below the 197 taxa predicted to be in system by the Chao asymptotic richness estimator. For fish, the final number of taxa accumulated was close to the predicted number of 40. The accumulation curves also show that a more modest number of stations detected most taxa (~75% of taxa by 20 stations



**Figure 2.** Taxa accumulation curves for a) benthos, and b) fish showing the cumulative number of taxa detected as a function of the number of stations sampled. Solid lines are sample-based rarefaction curves (accounting for actual distribution of taxa among stations) and dotted lines are individual-based rarefaction curves (as if taxa were randomly distributed). The number of stations by which 75% or 90% of the final (100%) richness was obtained is indicated with arrows. Note the difference in the initial taxa acquisition rate between the two kinds of curves; it takes about twice as many stations to collect 75% of the taxa (where curves intersect the thin horizontal line) under sample-based rarefaction as under individual-based rarefaction

**Table 1.** Non-native benthos and fish taxa detected in 2006, in order of decreasing percent occurrence across the stations sampled. Approximate times since detection in the St. Louis River/Duluth-Superior Harbor system are relative to our 2006 sampling effort (some “newly detected” taxa were also found in 2005 pilot sampling). Non-native but of North American origin means native to areas outside the upper Great Lakes. Times since first detection were compiled from Grigorovich et al. 2003b, the U.S. Geological Survey nonindigenous aquatic species website (<http://nas.er.usgs.gov>), and the Minnesota Sea Grant aquatic invasive species website (<http://www.seagrant.umn.edu/ais/index>)

Non-native species	Origin	% Occurrence	Time since detection
<b>Benthos</b>			
zebra mussel, <i>Dreissena polymorpha</i>	Eurasia	36	~15 years
peaclam, <i>Pisidium henslowanum</i>	Eurasia	22	newly detected
amphipod, <i>Gammarus fasciatus</i>	North America	21	~5 years
oligochaete, <i>Potamothrix vejdvskyi</i>	Eurasia	15	<5 years
oligochaete, <i>Potamothrix moldaviensis</i>	Eurasia	8	~5 years
peaclam, <i>Pisidium moitessierianum</i>	Eurasia	5	~5 years
European valve snail, <i>Valvata piscinalis</i>	Eurasia	4	~20 years
amphipod, <i>Echinogammarus ischnus</i>	Eurasia	3	newly detected
quagga mussel, <i>Dreissena bugensis</i>	Eurasia	2	newly detected
amphipod, <i>Gammarus tigrinus</i>	North America	2	~20 years
New Zealand mudsnail, <i>Potamopyrgus antipodarum</i>	New Zealand	2	newly detected
oligochaete, <i>Paranais frici</i>	Eurasia	2	newly detected
oligochaete, <i>Pristina acuminata</i>	Eurasia	2	newly detected
oligochaete, <i>Ripistes parasita</i>	Eurasia	2	~5 years
Asian clam, <i>Corbicula fluminea</i>	Eurasia	1	~5 years
greater European peaclam, <i>Pisidium amnicum</i>	Eurasia	1	~20 years
<b>Fish</b>			
Eurasian ruffe, <i>Gymnocephalus cernuus</i>	Eurasia	44	~20 years
round goby, <i>Neogobius melanostomus</i>	Eurasia	41	~10 years
white perch, <i>Morone americana</i>	North America	37	~20 years
tubenose goby, <i>Proterorhinus marmoratus</i>	Eurasia	32	~5 years
common carp, <i>Cyprinus carpio</i>	Eurasia	15	>100 years
rainbow smelt, <i>Osmerus mordax</i>	North America	8	>50 years
alewife, <i>Alosa pseudoharengus</i>	North America	5	>50 years
brook silverside, <i>Labidesthes sicculus</i>	North America	4	<5 years
threespine stickleback, <i>Gasterosteus aculeatus</i>	North America	2	~20 years

**Table 2.** Cumulative values for benthic invertebrate and fish richness endpoints (total richness, richness of taxa with ≤20% overall occurrence, richness of taxa with ≤5% overall occurrence, and richness of non-native taxa) over all stations within a type category. The richness achieved over stations of one type as a percentage of that over all stations is given in parentheses. The division among deep and shallow stations is at 3 m, and open versus vegetated stations have cover scores of 0-1 versus 2-4

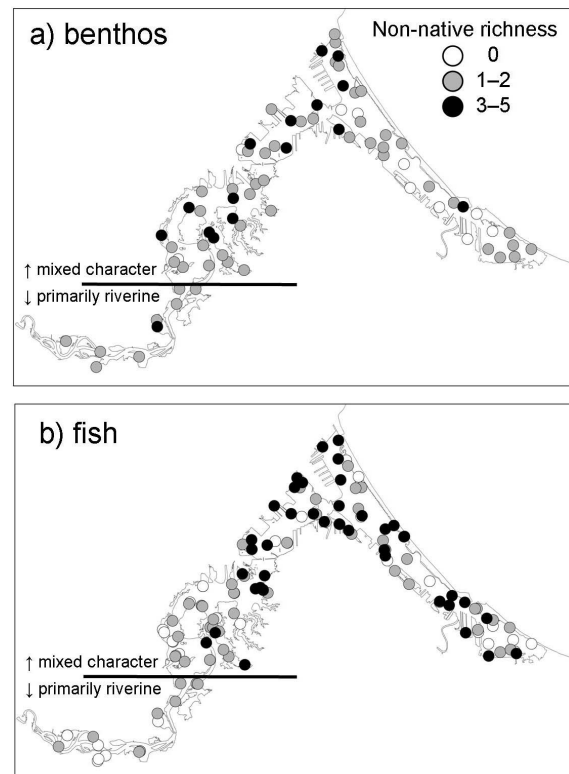
	Station type	N	Richness endpoint			
			Total	Rare-20	Rare-5	Non-native
Benthos	deep	23	91 (58%)	47 (41%)	13 (22%)	10 (64%)
	shallow-open	28	105 (66%)	61 (54%)	18 (31%)	8 (50%)
	shallow-vegetated	26	137 (87%)	94 (82%)	42 (71%)	11 (69%)
	all	77	158	114	59	16
Fish	fyke-nets	39	33 (87%)	18 (78%)	6 (60%)	9 (100%)
	electrofishing	51	33 (87%)	19 (83%)	7 (70%)	8 (89%)
	trawls	41	28 (74%)	14 (61%)	3 (30%)	7 (78%)
	all	131	38	23	10	9

and ~85% by 40 stations; Figure 2). The initial separation between the sample-based and the individual-based rarefaction curves (Figure 2) indicates considerable spatial heterogeneity in the distribution of benthos and fish across the system. This means that there is potential for accelerating the rate at which taxa are accumulated (increasing the steepness of the initial rise in the curves) if the most unique or species-rich sites could be predicted and targeted.

#### *Distribution across space, gear, and habitat*

Benthos composition appeared to be structured both along an upstream – downstream gradient and by water depth and vegetation cover. Total richness and richness of rare taxa generally decreased with increasing distance from the upstream end of the study system (correlation  $r \approx 0.35$ ). Stations having three or more non-native species were much more frequent downstream of the transition from the primarily riverine to the more mixed-character portion of the system, and the highest non-native richness generally occurred along the heavily developed north-west portion of the system (Figure 3). On a finer spatial scale, however, stations quite close together could differ substantially in non-native benthos richness (Figure 3) and also in total or rare richness (not shown). As a consequence of this fine-scale variability, non-native richness was only weakly related to distance from potential introduction points (roads and railroads, harbor facilities, boat ramps) in the downstream portion of the system (all  $r < 0.25$ ), even though these introduction points were considerably more prevalent in the lower than the upper system (Figure 1).

Water depth alone was not sufficient to explain patterns in benthos ordinations or in richness among stations, but richness of non-native and rare benthos was higher at shallow stations having substantial vegetation cover than at stations having little or no vegetation (Figure 4). We therefore constructed a three-category classification using a combination of water depth and vegetation (shallow-open:  $< 3$  m and cover score 0-1; shallow-vegetated:  $< 3$  m and cover score 2-4; deep:  $> 3$  m and no vegetation). Shallow-vegetated stations had higher values of total, rare-20, and rare-5 richness than either shallow-open or deep stations, while non-native richness was similar at shallow-vegetated and



**Figure 3.** Map of non-native richness across the St. Louis River/Duluth-Superior Harbor system for a) benthos, and b) fish. Horizontal lines indicate a division between the upstream, primarily riverine and downstream, mixed character portion of the system below which non-native taxa appear to be more prevalent. Refer to Figure 1 for additional map reference points

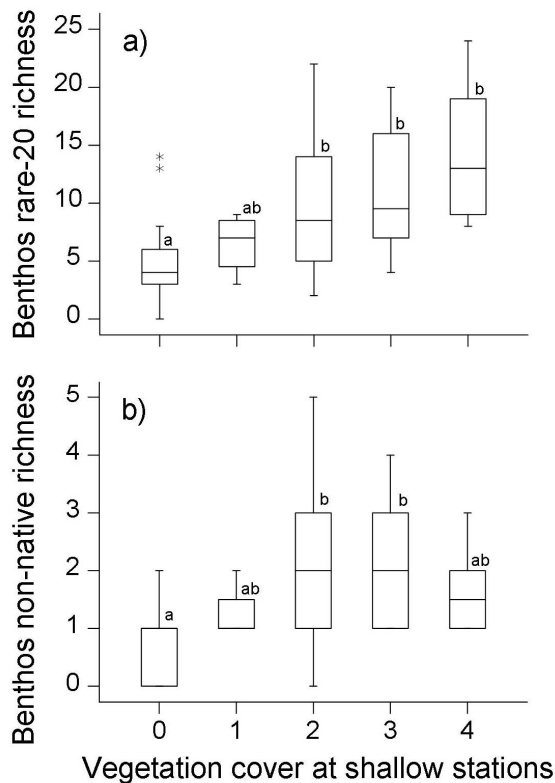
deep stations and higher than at shallow-open stations (Table 3). Benthos assemblages differed significantly among depth- and vegetation-based categories (ordination in Figure 5, ANOSIM  $r = 0.22$ ; all ANOSIM results reported are significant at  $p < 0.001$ ). Because shallow-vegetated stations had warmer water than shallow-open or deep stations, benthic richness endpoints also were related to water temperature. However, benthic richness was not related to other water quality metrics (turbidity, dissolved oxygen), nor to fetch or sediment composition (percent fine material, sediment density, total organic carbon).

Fish composition was primarily structured by the two closely-related factors of water depth and sampling gear, although upstream – downstream patterns and relationships to vegetation were evident as well. Stations having three or more non-native fish species occurred much more frequently in the downstream, mixed-character portion of the system than in the



**Table 3.** Average values for benthos and fish richness endpoints over individual stations within a type. Station types with different letters have significantly different average richness (ANOVA a posteriori Tukey HSD test,  $p \leq 0.05$ ). Variable definitions are as in Table 2

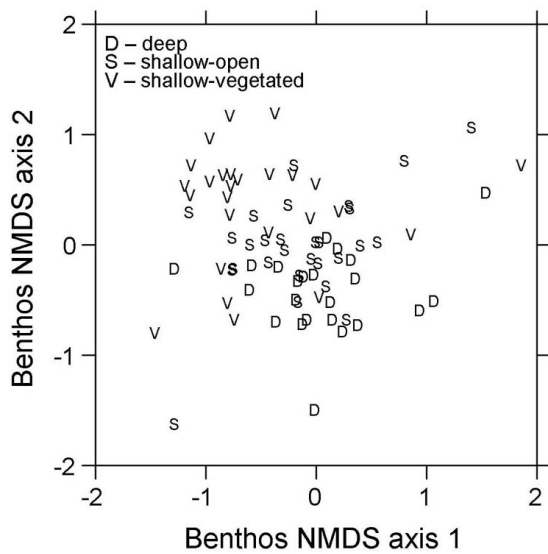
	Station type	N	Richness endpoint			
			Total	Rare-20	Rare-5	Non-native
Benthos	deep	23	21.3 a	4.8 a	0.7 a	2.0 a
	shallow-open	28	24.4 a	5.0 a	0.7 a	1.0 b
	shallow-vegetated	26	31.9 b	11.3 b	2.3 b	2.1 a
Fish	fyke-nets	39	10.4 a	2.3 a	0.3 a	2.2 a
	electrofishing	51	6.6 b	1.5 b	0.2 a	0.7 b
	trawls	41	7.3 b	1.4 b	0.1 a	2.4 a



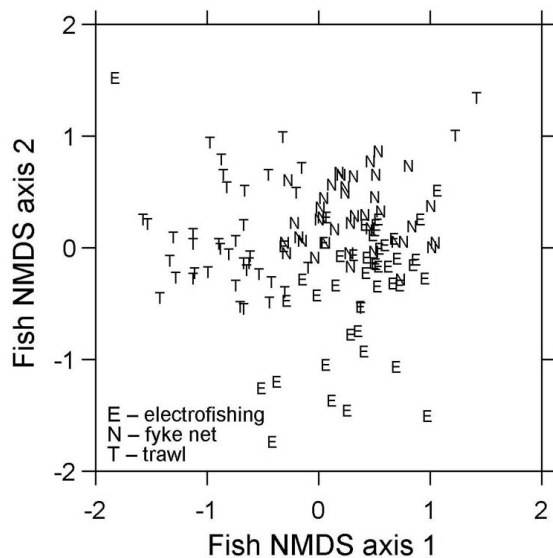
**Figure 4.** Box plots showing richness of a) rare benthos (<20% overall occurrence) and b) non-native benthos in relation to vegetation cover score at shallow (<3 m) stations. Lines across boxes are medians, box ends are quartiles, whiskers show ranges, and asterisks are outliers. Vegetation cover groups with different letters had significantly different mean richness (Tukey HSD test,  $p \leq 0.05$ )

upstream, primarily riverine portion of the system, but like for benthos, there was considerable among-station variability at a finer spatial scale (Figure 3). Non-native fish richness was only very weakly related to distance from boat ramps, harbor facilities, or roads and railroads within the downstream portion of the system (all  $r < 0.20$ ). Total and rare fish richness did not show any upstream – downstream pattern, nor were they obviously related to distance to potential introduction points.

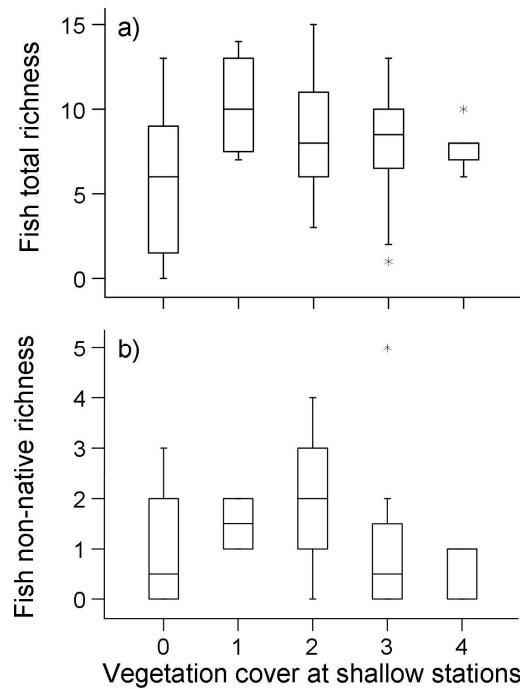
On average, total and rare-20 fish richness were higher at fyke-net stations (0-1 m deep) than at either electrofishing (1-2 m) or trawl (>2 m) stations (Table 3). Non-native richness was similar at fyke-net and trawl stations but lower at electrofishing stations (Table 3). Rare-5 fish richness at individual stations was very low and did not differ among station types (Table 3). Gear explained a good bit of the assemblage patterns (overall ANOSIM on gear  $r = 0.40$ ), with trawls differing substantially from the other two, while fyke-net and electrofishing assemblages overlapped somewhat (Figure 6). The best shallow versus deep partition in fish assemblage patterns was at 3 m, even though a partition at 2 m better matched the division between electrofishing and trawl stations (pairwise ANOSIM on depth categories  $r = 0.27$  vs.  $r = 0.21$ ). Deep-station assemblages differed more strongly from shallow-vegetated than from shallow-open stations (pairwise ANOSIM on habitat categories  $r = 0.39$  vs.  $0.18$ ). Total fish richness appeared to be lower at shallow stations lacking vegetation than those having any amount of vegetation, but the difference was not statistically significant (Figure 7). Non-native fish richness was unrelated to vegetation (Figure 7), but was generally



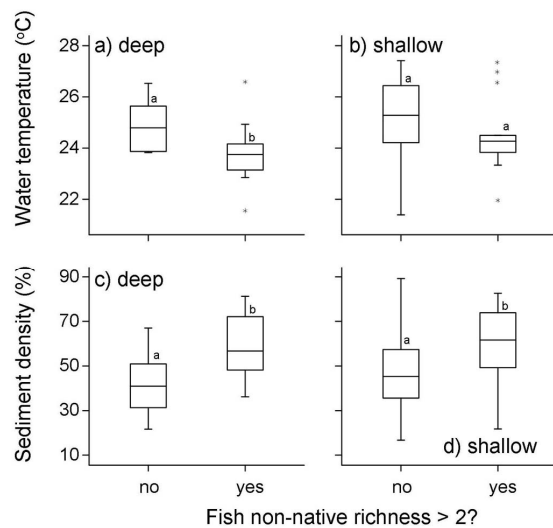
**Figure 5.** Plot of first two axes from a three-dimensional NMDS ordination of benthos presence/absence data across all stations sampled (stress = 0.16). Letters denote station depth and vegetation cover categories (shallow is <3 m, vegetated is cover >1)



**Figure 6.** Plot of first two axes from a three-dimensional NMDS ordination of fish presence/absence data (stress = 0.15). Letters denote sampling gear used. The ordination excludes 4 electrofishing stations at which no fish were caught and one outlier station that yielded only a single fish



**Figure 7.** Box plots showing richness of a) all fish, and b) non-native fish in relation to vegetation cover at shallow (<3 m) stations. Box layout as in Figure 4. Vegetation cover groups did not have significantly different mean richness (Tukey HSD test,  $p$  was not  $\leq 0.05$ ).



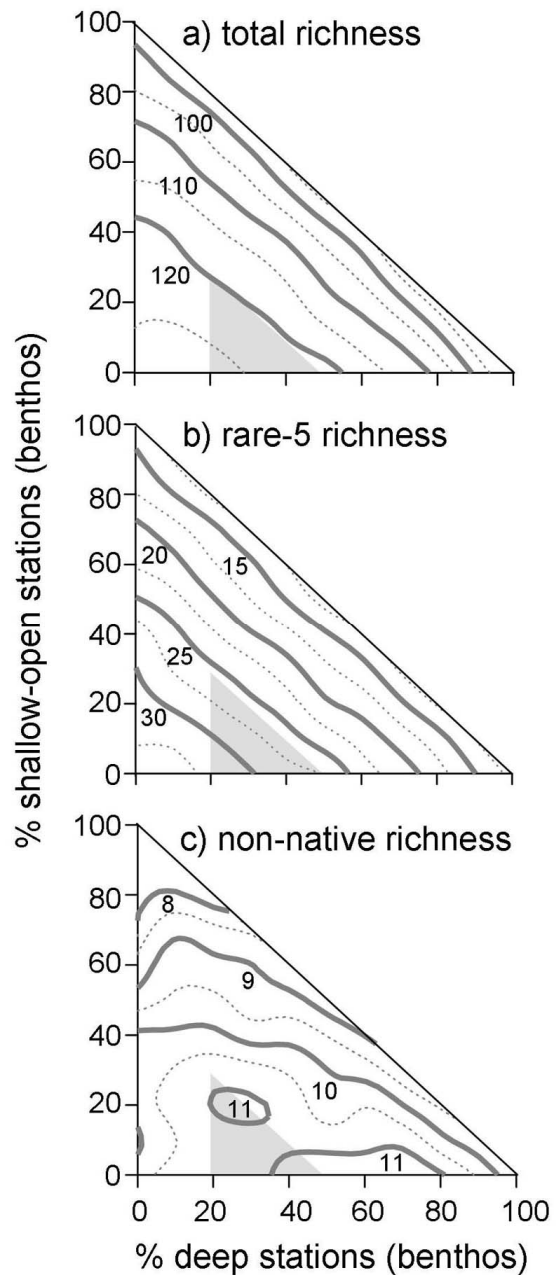
**Figure 8.** Box plots showing differences in water temperature (top row) and sediment density (dry weight as % of wet weight; bottom row) among deep (>3 m, left column) or shallow (<3 m, right column) stations having high or low non-native fish richness. Box layout as in Figure 4. Non-native richness groups with different letters had significantly different mean sediment density or water temperature (Tukey HSD test,  $p \leq 0.05$ )

higher with colder water and increasingly sandy sediments (Figure 8). Total and rare fish richness were not related to water quality or substrate at either shallow or deep stations.

*Optimal allocation of effort among gear and habitats*

For benthos, we focused on depth and vegetation cover in effort allocation analyses, because these habitat variables were significant in structuring composition among sampling locations. The optimal allocation of effort among stations depended on the richness endpoint examined. Total richness and rare-5 richness were maximized in randomization analyses when the effort was allocated entirely to shallow-vegetated stations and declined steadily as the number of either shallow-open or deep stations increased (Figure 9a, 9b; patterns in rare-20 richness, not plotted, were the same as for total richness). However, non-native richness was maximized by allocating roughly equal effort to shallow-vegetated and deep stations and almost no effort to shallow-open stations (the contour “ridge” along the x-axis in Figure 9c), or by allocating roughly half the effort to shallow vegetated stations and the rest equally among open and deep stations (the contour “hilltop” in Figure 9c). Allocating at least 50% of the effort to shallow-vegetated stations, 20 to 50% effort to deep stations, and only 0 to 30% to shallow-open stations (triangular shaded areas in Figure 9) achieved fairly high values over all richness endpoints simultaneously.

For fish, effort allocation analyses focused on gear types (and associated depth strata) because of the differences in fish composition among them. Again, the optimal allocation of effort among stations depended on the richness endpoint. Total richness was maximized when roughly half the effort was fyke-net stations with the rest allocated equally among trawling and electrofishing (the contour “hilltop” in Figure 10a; patterns in rare-20 richness, not plotted, were the same as for total richness). Richness of rare-5 taxa increased as the number of trawl stations decreased, and was higher when fyke-netting and electrofishing were mixed than when either was used alone (Figure 10b). Richness of non-native taxa generally increased as the effort allocated to fyke-netting increased, but including a few trawl stations yielded higher non-native richness than fyke-nets alone (Figure 10c). Results across multiple stations of any one type



**Figure 9.** Contour plot of a) total, b) rare-5, and c) non-native benthos richness accumulated over 20 station-sets as a function of the mix of habitats sampled. The data contoured are averages over 10 random draws for each station combination and the axes and upper right diagonal bound the possible station-type combinations, so percent shallow-vegetated stations equals 100 minus the percent of shallow-open (y-axis value) and deep (x-axis value) stations. As an aid to integrating information across the figure panels, the small shaded triangles highlight a region where all three richness endpoints are simultaneously fairly high (its vertices lie at 20/30/50%, 20/0/80%, and 50/0/50% mixtures of deep/shallow-open/shallow-vegetated stations)

differed from results at single stations; for example individual trawl stations had average non-native richness as high as individual fyke stations (Table 3), but groups of fyke stations yielded higher non-native richness than groups of trawl stations (Figure 10b). Allocating 45 to 85% of the effort to fyke-nets, 10 to 30% to electrofishing, and only 5 to 25% to trawl stations (square shaded areas in Figure 10) achieved fairly high values of all richness endpoints simultaneously.

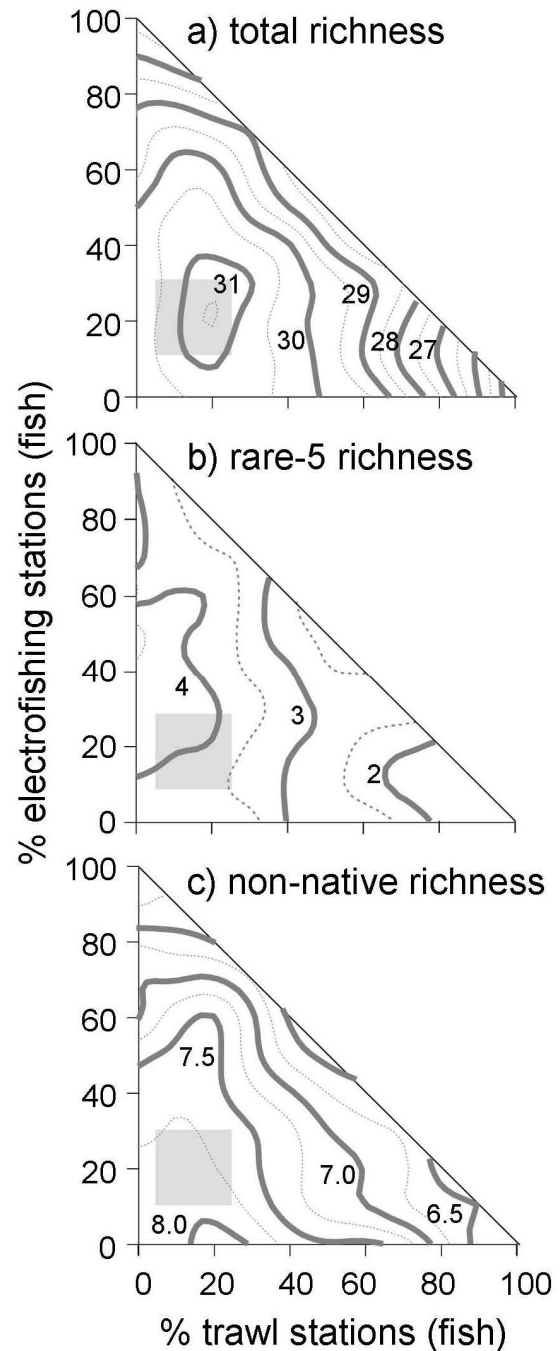
Results from randomization tests using 40 or 60 stations per draw confirmed that our findings concerning optimal station combinations also apply to larger sample sizes. Including more stations yielded more fish taxa per richness endpoint, but the relative performance among various station-type combinations remained consistent with larger sample sizes (Figure 11).

## Discussion

Monitoring designs inevitably must balance desired detection probabilities against available resources and other logistical constraints. Our analyses of data from a comprehensive sampling of a spatially-complex, invasion-vulnerable Great Lakes subsystem show that there are monitoring efficiencies to be gained through knowledge of how taxa detection patterns vary among habitats and sampling gear. Clearly, some of the details concerning benthic invertebrate and fish distributions and the sampling design recommendations that spring from these are specific to the study system. More broadly, however, our approach and our finding that covariances in species distributions and environmental attributes can be expected and capitalized upon are applicable to a variety of systems in which monitoring for non-native species is of interest. The consistencies between two quite different taxonomic groups, benthos and fish, lends confidence in the generality of the patterns observed.

### *Distributions across space, gear, and habitat*

Since detection is accomplished with finding a single specimen and the goal of early detection is finding non-native species while they are still rare, we focused on presence rather than abundance, and examined performance based on detection of rare species as well as non-native species. Some of the non-native benthos and fish

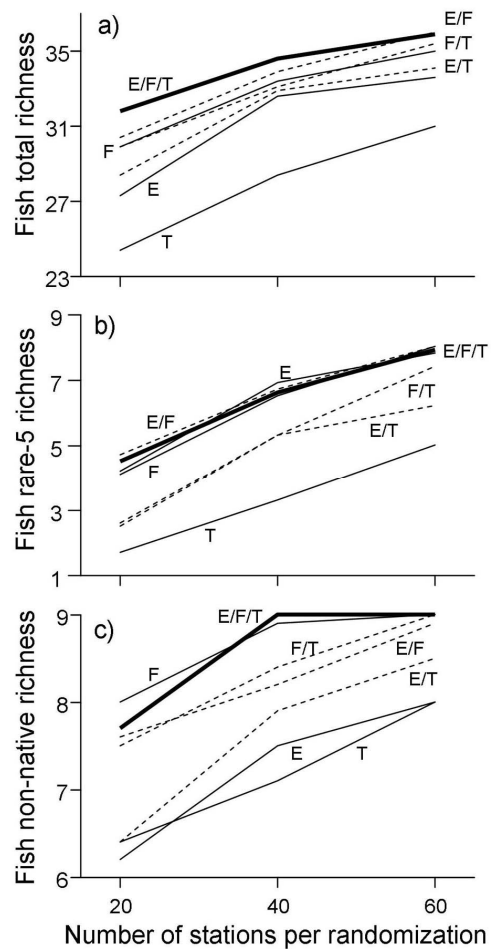


**Figure 10.** Contour plots of a) total, b) rare-5, and c) non-native fish richness accumulated over 20 station-sets as a function of the mix of fishing gears employed. Plot layout as in Figure 9; percent fyke-net stations equals 100 minus the percent electrofishing (y-axis value) and trawl (x-axis value) stations. As an aid to integrating information across the figure panels, the small shaded squares highlight a region where all three richness endpoints are simultaneously fairly high (its corners lie at 5/30/65%, 25/30/45%, 5/10/85%, and 25/10/65% mixtures of trawl/electrofishing/fyke-net stations)

are far from rare in the system (e.g., round goby – Bergstrom et al. 2008, zebra mussel – Grigorovich et al. 2008), so detecting these is an insufficient test of monitoring effectiveness. For fish, any given subset of stations actually yielded a higher percentage of the pool of non-native species than of rare species, although the converse applied for benthos. More than one definition of “rare” was informative in our analyses, depending on the scale over which patterns were evaluated. At individual stations, there were typically so few rare-5 taxa that patterns could not be discriminated, but richness of rare-20 taxa differentiated station types for both fish and benthos. Aggregated over multiple stations, patterns in rare-20 richness were essentially the same as for total richness, but patterns in rare-5 richness provided additional insight on sampling strategies.

The relationship between effort expended and the number and rarity of species detected is well known in ecology (Rosenzweig 1995). As might be expected for a complex ecotone system (Willis and Magnuson 2000), benthos and fish assemblages in the St. Louis River/Duluth-Superior Harbor are fairly taxa rich, and many of the taxa occur at only a few locations. This pattern of richness and patchiness is reflected in accumulation curves that continue to rise slowly over many samples, confirming the need to commit to a considerable total sampling effort to detect rare taxa. Sampling for benthos might need to cover substantially more than the 77 stations analyzed here for the accumulation curve to approach an asymptote, yet even so, our sampling detected several non-native species not formerly known from the system. The accumulation curves also suggest that there is considerable potential for increasing the efficiency of sampling. The initially much faster rate of taxa accumulation in individual-based than sample-based rarefaction curves for both benthos and fish indicates that the distribution of organisms is heterogeneous, with some stations tending to contribute relatively many new taxa to the overall pool when encountered early in the sampling sequence, while other stations contribute no new information. The key question for increasing sampling efficiency is whether that heterogeneity can be effectively described using easily collected attributes of the sampling locations and thus be exploited in a monitoring design.

Our analyses focused on location, habitat type, distance to introduction vectors, and sampling



**Figure 11.** Plot showing average fish richness endpoints for selected station combinations obtained from randomizations over 20-, 40-, or 60-station sets. Thin solid lines are 100% electrofishing, fyke-net, or trawl stations (labeled E, F, or T respectively), dashed lines are 50/50% combinations of any two gear types, and the heavy solid “E/F/T” line is a combination of 20% electrofishing, 60% fyke-nets, and 20% trawling. Relative to the triangular contoured area in Figure 10, these station combinations represent points at each vertex, points mid-way along each side, and one within the shaded box

gear as predictors for fish and benthos distributions, and identified two distinct spatial scales of variation. On a broad spatial scale, some areas contributed relatively little information about the non-native species present and might therefore warrant reduced sampling effort. Non-native richness was consistently low for both benthos and fish in the upstream, primarily riverine part of the system, and often much higher in the downstream, mixed-character portions. The area spanning the highest incidences of non-native taxa was more centralized for benthos and more

dispersed for fish, perhaps reflecting differences in motility and dominant introduction vectors between these groups. This broad pattern in non-native richness was not simply a corollary of overall richness, as total benthos richness increased in the upstream direction (consistent with Swanson 1999; Breneman et al. 2000) while total fish richness showed the opposite pattern, and the correlation between total and non-native richness was weak for both fish and benthos. The low incidence of non-native species in the upstream-most portion of the system might be interpreted as reflecting the generally larger distance from introduction points. However, within the downstream portion of the system, non-native taxa richness did not increase with proximity to shipping facilities, land transportation corridors, or boat-ramps (in fact, these areas might represent poor benthos habitat due to industrial contamination; Edsall et al. 2004). Alternatively, low non-native richness in the upstream-most area might be a function of its riverine morphology (Figure 1), with a generally stronger current (a barrier to up-stream dispersal of some taxa) and little sand and mud flat habitat (a habitat type that might be preferred by some taxa). Most of the lower system appears to be readily accessible to even passively dispersing species (perhaps aided by seiche induced flow reversals and wind-induced horizontal transport; Stortz and Sydor 1980). The considerable fine-scale variability in the number of non-native taxa detected coupled with a lack of relationship to introduction points suggests that their distributions largely reflect environmental sorting rather than vector-driven patterns.

On a local scale, habitat attributes offer clear potential as stratification factors in sampling designs. Benthos and fish assemblages appeared to be structured by a combination of water depth and vegetation cover, which is consistent with the literature for other systems (e.g., Voigts 1976; Brinkhurst 2002; Smokorowski and Pratt 2007) and with an earlier study of benthos in the St. Louis River/Duluth-Superior Harbor (Breneman et al. 2000). For fish, depth effects on composition are confounded with gear effects, as we deployed whatever gear was appropriate for the station. Trawling is difficult in shallow water while fyke nets are only effective in shallow water, and electrofishing requires areas deep enough to maneuver a small boat but shallow enough for the current field to intersect the bottom. However the depth that best differentiated fish assemblage groups in ordinations was

the same as for benthos and corresponded with the depth limit for aquatic vegetation rather than the depth separating fishing gear types. Because invertebrates were collected from sediments rather than from plants themselves, we had expected relationships to vegetation to be stronger for fish than benthos, but the opposite was true. This may have been due to a better match in spatial scale of the vegetation data to the petite ponar than to the fishing gear (vegetation was characterized over a 5 m radius circle, while the area covered by the fishing gear was much larger). We observed congregation of some non-native species (e.g., Asian clam) in the unique thermal habitat around a sewage treatment plant outfall (which, unlike the rest of the system, remains ice-free in winter). However, the increased incidence of non-native fishes at stations having colder water probably reflects proximity to Lake Superior as source for lake-migrant species such as alewife. The diversity of patterns observed cautions that sampling should not simply be limited to habitats presumed to be preferred by non-native species.

#### *Implications for sampling design*

The inevitable differences in species composition obtained with different sampling gear are often seen as impeding the ability to compare among data sets. However collectively, a combination of complementary gear can more thoroughly characterize a complex system than any single gear would accomplish alone (Magnuson et al. 1994, Jackson and Harvey 1997, Turner and Trexler 1997). A positive relationship generally exists between habitat heterogeneity and species diversity (e.g., Voigts 1976, Benson and Magnuson 1992, Palmer and Poff 1997), and describing an entire system in a way that captures diversity across habitats and locations is precisely what is desired of effective non-native species monitoring. Such monitoring has a different goal than many studies in ecology, in that it emphasizes qualitative rather than the quantitative endpoints (detection, not abundance) and seeks to embrace rather than control for variability. A random design with sufficient station density is one way of achieving coverage of space, but with some system-specific knowledge, sampling strategies can be refined through stratification or deliberately unequal allocation across habitat and gear types so that species are detected more efficiently and monitoring becomes less resource-intensive.

In our study, stations with some particular attributes yielded more rare or non-native taxa than others, but no single gear or habitat yielded all the rare or non-native taxa. The presence of unique taxa at all station types but in different frequencies suggests that species detection rates would be maximized when all significant gear or habitat categories are sampled but with the effort biased towards the ones yielding the most species. For benthos, shallow vegetated stations yielded substantially more total and rare taxa than shallow-open or deep stations, but our randomization analyses suggested that rates of detection for non-native taxa are maximized by allocating roughly half the effort to shallow-vegetated stations and the other half to some combination of deep and shallow-open stations. This finding is consistent with an earlier study of benthos in the St. Louis River/Duluth-Superior Harbor system by Swanson (1999), who found the most taxa in shallow flats (she did not distinguish vegetated from open), but that natural channels and dredged areas also contributed unique taxa. For fish, fyke-nets (shallowest stations) were the most effective single gear, but rare taxa were equally well detected across electrofishing stations (intermediate depth), and detection rates across all richness endpoints were maximized by allocating at least half the effort to fyke nets and distributing the rest roughly equally among electrofishing and trawling.

Species rarefaction curves make it clear that sampling only twenty stations is insufficient for reliably detecting many of the rarer species. We chose 20-station sets for our randomization analyses as representing a point on the accumulation curves where a substantial portion (~75 %) of the total taxa had been obtained yet the effects of varying station combinations could still be discerned (i.e., sample-based and individual-based rarefaction curves had not yet converged) and because we had enough stations in each category for meaningful randomization. By drawing on fish data from additional years, we were able to conduct trials with larger station sets that yielded higher richness overall, but the relative performance among various station-type combinations remained consistent across sample sizes. For future sampling efforts, we therefore recommend allocating effort among station types in the ratios suggested by our analyses of 20-station sets, but over at least 40 stations, a level of effort that obtained at least 85% of the total fish and benthos taxa. Fish sampling at 40 stations can be accomplished with two weeks in

the field by a 3-4 person crew (varies by gear) and some minor laboratory follow-up (10-20 hrs total to verify voucher specimens). Forty petite ponar benthos samples can be collected and field-processed by a 4-person crew in less than a week, but the laboratory effort would be much larger (typically 20-30 person hrs per sample for complete picks and counts, which could be reduced by scanning for as-yet unseen species rather than enumerating all). While this sampling effort is not trivial, it is within the capability of the typical resource management agency, and could be accomplished concurrent with sampling for other goals (e.g., fish species surveys, biotic condition assessments). The sampling effort required to obtain some given proportion of the species pool might well be smaller in a system of less size or habitat complexity (Angermeier and Schlosser 1989) but should be verified with actual assessment.

A general strategy for detecting rare species is to allocate samples widely in space, rather than intensify effort in a small area or over time (Mackenzie and Royle 2005, Rew et al. 2006, Harvey et al. 2009). Our study confirms the need to allocate some effort to all unique habitats, but with the important additional finding that efficiencies can be gained by exploiting the relationship between species distributions and habitat attributes to bias effort towards those locations yielding the greatest number of taxa. Patterns of richness and uniqueness are likely to differ somewhat among ecosystems (i.e., the exact sampling recommendations for our study system may not transfer to each new site), but the basic message that one can expect such patterns, characterize them, and take advantage of them in monitoring for invasive species is broadly applicable. For example, the extent of aquatic vegetation (or the littoral zone) varies among systems, but since this is substantially determined by light penetration (as are other important habitat features such as thermal structure), photic zone depth might be a broadly useful stratification factor. Certainly, other considerations beyond those emphasized here can inform sampling design, including a desire to expend some effort monitoring perceived high-risk invasion points. However distributions in non-native species driven by proximity to introductions points are likely to dissipate over time, whereas distributions driven by habitat are likely to persist. It may be necessary to adapt the monitoring design over time if species compositions or environmental attributes shift, but a

basic strategy of spending some effort covering the space should help assure such a need is recognized and implemented.

Our findings are analogous to those of Smith and Jones (2008) concerning optimal allocation of sampling effort among stream sizes, in which the number of fish species detected was highest when sampling focused on the most species-rich third-order streams but also included some second- and first-order streams. More generally, our findings for how best to detect species mirror the consensus in the conservation literature for how best to protect species, namely by including both biodiversity hotspots and biologically poorer but unique areas in conservation efforts (e.g., Chong and Stohlgren 2007). As pointed out by Peterson and Rabeni (1995), sampling designs which are not balanced across the categories of interest (space, gear, habitat, etc.) somewhat increase the complexity and decrease the power of statistical analyses, but more than make up for these concerns by increased sampling efficiency and decreased cost.

### Acknowledgements

We appreciate the extensive field assistance provided by Charlie Butterworth, Timothy Corry, Brian Sederberg, and Jon Van Alstine. Trawling was conducted in collaboration with Gary Czapinski (U.S. Fish and Wildlife Service). Anthony Olson assisted with the probabilistic sampling design. GIS support was provided by Jesse Adams, Roger Meyer, Tatiana Nawrocki, and Matthew Starry (CSC Corporation). Processing and identification of invertebrate samples was done by Wilson Environmental Laboratories. Identification of difficult specimens could not have been accomplished without the taxonomic expertise of Igor Grigorovich (Wilson Environmental Labs) and Brett Nagle (University of Minnesota Bell Museum). Thanks also to Theodore Angradi, Henry Lee, Hugh MacIsaac, and two anonymous reviewers for helpful manuscript reviews. This research was approved for publication after review by EPA's National Health and Environmental Effects Research Laboratory, but approval does not signify that the contents reflect the views of the Agency, nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

### References

- ACOE (1984) U.S. Army Corps of Engineers. Shore Protection Manual, Vol. 1. Coastal Engineering Research Center, Vicksburg, Maryland, USA
- Angermeier PL, Schlosser IJ (1989) Species - area relationships for stream fishes. *Ecology* 70: 1450-1462, doi:10.2307/1938204
- Benson BJ, Magnuson JJ (1992) Spatial heterogeneity of littoral fish assemblages in lakes: relation to species diversity and habitat structure. *Canadian Journal of Fisheries and Aquatic Sciences* 49: 1493-1500, doi:10.1139/f92-165
- Bergstrom MA, Evrard LM, Mensinger AF (2008) Distribution, abundance, and range of the round goby, *Apollina melanostoma*, in the Duluth-Superior harbor and St. Louis river estuary, 1998-2004. *Journal of Great Lakes Research* 34: 535-543, doi:10.3394/0380-1330(2008)34[535:DAAROT]2.0.CO;2
- Breneman D, Richards C, Lozano S (2000) Environmental influences on benthic community structure in a Great Lakes embayment. *Journal of Great Lakes Research* 26: 287-304
- Brinkurst RO (2002) *The Benthos of Lakes*. Blackburn, Caldwell, New Jersey, USA
- Chao A (1987) Estimating the population size for capture-recapture data with unequal catchability. *Biometrics* 43: 783-791, doi:10.2307/2531532
- Chong GW, Stohlgren TJ (2007) Species - area curves indicate the importance of habitats' contributions to regional biodiversity. *Ecological Indicators* 7: 387-395, doi:10.1016/j.ecolind.2006.03.002
- Clarke KR, Warwick RM (2001) Change in marine communities: an approach to statistical analysis and interpretation. PRIMER-E Corp., Plymouth, UK
- Cohen J, Mirotnick N, Leung B (2007) Thousands introduced annually: the aquarium pathway and non-indigenous plants in the St. Lawrence Seaway. *Frontiers in Ecology and the Environment* 5: 528-532, doi:10.1890/060137
- Coleman BD (1981) On random placement and species area relations. *Mathematical Biosciences* 54: 191-215, doi:10.1016/0025-5564(81)90086-9
- Colwell RK (2006) EstimateS: Statistical estimation of species richness and shared species subsamples. Version 8.0 User's guide and application published at <http://purl.oclc.org/estimates>
- Cuffney TF, Bilger MD, Haigler AM (2007) Ambiguous taxa: effects on the characterization and interpretation of invertebrate assemblages. *Journal of the North American Benthological Society* 26: 286-307, doi:10.1899/0887-3593(2007)26[286:ATEOTC]2.0.CO;2
- Drake JM, Lodge DM (2007) Hull fouling is a risk factor for intercontinental species exchange in aquatic ecosystems. *Aquatic Invasions* 2: 121-131, doi:10.3391/ai.2007.2.2.7
- Duggan IC, van Overdijk CDA, Bailey SA, Jenkins PT, Limén H, MacIsaac HJ (2005) Invertebrates associated with residual ballast water and sediments of cargo-carrying ships entering the Great Lakes. *Canadian Journal of Fisheries and Aquatic Sciences* 62: 2463-2474, doi:10.1139/f05-160
- Edsall TA, Gorman OT, Evrard LM (2004) Burrowing mayflies as indicators of ecosystem health: status of populations in two western Lake Superior embayments. *Aquatic Ecosystem Health and Management* 7: 507-513 doi:10.1080/14634980490513346
- Gotelli NJ, Colwell RK (2001) Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. *Ecology Letters* 4: 379-391, doi:10.1046/j.1461-0248.2001.00230.x



- Grigorovich IA, Colautti RI, Mills EL, Holeck K, Ballert AG, MacIsaac HJ (2003a) Ballast-mediated animal introductions in the Laurentian Great Lakes: retrospective and prospective analyses. *Canadian Journal of Fisheries and Aquatic Sciences* 60: 740-756, doi:10.1139/f03-053
- Grigorovich IA, Kornushin AV, Gray DK, Duggan IC, Colautti RI, MacIsaac HJ (2003b) Lake Superior: an invasion coldspot? *Hydrobiologia* 499: 191-210, doi:10.1023/A:1026335300403
- Grigorovich IA, Kelly JR, Darling JA, West CW (2008) The quagga mussel invades the Lake Superior basin. *Journal of Great Lakes Research* 34: 342-250, doi:10.3394/0380-1330(2008)34[342:TQMITL]2.0.CO;2
- Harvey CT, Qureshi SA, MacIsaac HJ (2009) Detection of a colonizing aquatic nonindigenous species. *Diversity and Distributions* 15: 429-437, doi:10.1111/j.1472-4642.2008.00550.x
- Herborg L-M, Jerde CL, Lodge DM, Ruiz GM, MacIsaac HJ (2007) Predicting invasion risk using measures of introduction effort and environmental niche models. *Ecological Applications* 17:663-674, doi:10.1890/06-0239
- Holeck KT, Mills EL, MacIsaac HJ, Dochoda MR, Colautti RI, Ricciardi A (2004) Bridging troubled waters: biological invasions, transoceanic shipping, and the Laurentian Great Lakes. *BioScience* 54: 919-929, doi:10.1641/0006-3568(2004)054[0919:BTWBIT]2.0.CO;2
- Hulme PE (2006) Beyond control: wider implications for the management of biological invasions. *Journal of Applied Ecology* 43:835-847, doi:10.1111/j.1365-2664.2006.01227.x
- Jackson DA, Harvey HH (1997) Qualitative and quantitative sampling of lake fish communities. *Canadian Journal of Fisheries and Aquatic Sciences* 54: 2807-2813, doi:10.1139/cjfas-54-12-2807
- Kolar CS, Lodge DM (2002) Ecological predictions and risk assessment for alien fishes in North America. *Science* 298:1233-1236, doi:10.1126/science.1075753
- Levine JM, D'Antonio CM (2003) Forecasting biological invasions with increasing international trade. *Conservation Biology* 17: 322-326, doi:10.1046/j.1523-1739.2003.02038.x
- Lindgren JP, Ongstad PW, Spurrier JR (1997) The St. Louis Bay fishery 1986-1994. Minnesota Department of Natural Resources, Division of Fish and Wildlife, Fisheries report F-29-R(P)-16, St. Paul, Minnesota, USA
- Ludwig HR, Leitch JA (1996) Interbasin transfer of aquatic biota via anglers' bait buckets. *Fisheries* 21: 14-18, doi:10.1577/1548-8446(1996)021<0014:ITOABV>2.0.CO;2
- Mackenzie DI, Royle JA (2005) Designing occupancy studies: general advice and allocating survey effort. *Journal of Applied Ecology* 42: 1105-1114, doi:10.1111/j.1365-2664.2005.01098.x
- Magnuson JJ, Benson BJ, McLain AS (1994) Insights on species richness and turnover from long-term ecological research: fishes in north-temperate lakes. *American Zoologist* 34: 437-451
- Mao CX, Colwell RK, Chang J (2005) Estimating the species accumulation curve using mixtures. *Biometrics* 61: 433-441, doi:10.1111/j.1541-0420.2005.00316.x
- Mehta SV, Haight RG, Homans FR, Polasky S, Venette RC (2007) Optimal detection and control strategies for invasive species management. *Ecological Economics* 61: 237-245, doi:10.1016/j.ecolecon.2006.10.024
- Meyers JH, Simberloff D, Kuris AM, Carey JR (2000) Eradication revisited: dealing with exotic species. *Trends in Ecology and Evolution* 15: 316-320, doi:10.1016/S0169-5347(00)01914-5
- Mills EL, Leach JH, Carlton JT, Secor CL (1993) Exotic species in the Great Lakes: a history of biotic crises and anthropogenic introductions. *Journal of Great Lakes Research* 19: 1-54
- Palmer MA, Poff NL (1997) Heterogeneity in streams: the influence of environmental heterogeneity on patterns and processes in streams. *Journal of the North American Benthological Society* 16: 169-173, doi:10.2307/1468249
- Peterson JT, Rabeni CF (1995) Optimizing sampling effort for sampling warmwater stream fish communities. *North American Journal of Fisheries Management* 15: 528-541, doi:10.1577/1548-8675(1995)015<0528:OSEFSW>2.3.CO;2
- Pimentel D, Zuniga R, Morrison D (2005) Update on the environmental and economic costs associated with alien-invasive species in the United States. *Ecological Economics* 52: 273-288
- Rew LJ, Maxwell BD, Dougher FL, Aspinall R (2006) Searching for a needle in a haystack: evaluating survey methods for non-indigenous plant species. *Biological Invasions* 8: 523-539, doi:10.1007/s10530-005-6420-2
- Ricciardi A (2006) Patterns of invasion in the Laurentian Great Lakes in relation to changes in vector activity. *Diversity and Distribution* 12: 425-433, doi:10.1111/j.1366-9516.2006.00262.x
- Rixon CAM, Duggan IC, Bergeron NMN, Ricciardi A, MacIsaac HJ (2005) Invasion risks posed by the aquarium trade and live fish markets on the Laurentian Great Lakes. *Biodiversity Conservation* 14: 1365-1381, doi:10.1007/s10531-004-9663-9
- Rosenzweig ML (1995) Species diversity in time and space. Cambridge University Press, Cambridge, UK
- Smith KL, Jones ML (2008) Allocation of sampling effort to optimize efficiency of watershed-level ichthyofaunal inventories. *Transactions of the American Fisheries Society* 137: 1500-1506, doi:10.1577/T07-024.1
- Smokorowski KE, Pratt TC (2007) Effect of a change in physical structure and cover on fish and fish habitat in freshwater ecosystems - a review and meta-analysis. *Environmental Reviews* 15: 15-41, doi:10.1139/A06-007
- Stevens DL, Olson AR (2004) Spatially-balanced sampling of natural resources in the presence of frame imperfections. *Journal of the American Statistical Association* 99: 262-278, doi:10.1198/016214504000000250
- Stortz KR, Sydor M (1980) Transports in the Duluth - Superior harbor. *Journal of Great Lakes Research* 6: 223-231
- Swanson TA (1999) Spatial and temporal variations of macrobenthos in the St. Louis river estuary: food available for benthophagous fishes. M.Sc. thesis, University of Minnesota, Duluth, Minnesota, USA
- Turner AM, Trexler JC (1997) Sampling aquatic invertebrates from marshes: evaluating the options. *Journal of the North American Benthological Society* 16: 694-709, doi:10.2307/1468154
- Vanderploeg HA, Nalepa TF, Jude DJ, Mills EL, Holeck KT, Liebig JR, Grigorovich IA, Ojaveer H (2002) Dispersal and emerging ecological impacts of Ponto-Caspian species in the Laurentian Great Lakes. *Canadian Journal of Fisheries and Aquatic Sciences* 59: 1209-1228, doi:10.1139/f02-087
- Voigts DK (1976) Aquatic invertebrate abundance in relation to changing marsh vegetation. *American Midland Naturalist* 95: 313-322, doi:10.2307/2424396
- Willis TV, Magnuson JJ (2000) Patterns in fish species composition across the interface between streams and lakes. *Canadian Journal of Fisheries and Aquatic Sciences* 57: 1042-1052, doi:10.1139/cjfas-57-5-1042