United States
Environmental Protection
Agency

Report on the Regional Environmental Monitoring and Assessment Program Study of Wadeable Streams in the Driftless Area Ecoregion in Western Wisconsin

RESEARCH AND DEVELOPMENT

EPA 600/R-06/165
November 2006
www.epa.gov

Report on the Regional Environmental Monitoring and Assessment Program Study of Wadeable Streams in the Driftless Area Ecoregion in Western Wisconsin

Michael A. Miller[1], Alison C.C. Colby[1], and Paul Kanehl[2]

[1]Wisconsin Department of Natural Resources
Bureau of Fisheries Management and Habitat Protection
101 South Webster St.
Madison, Wisconsin 53707

[2]Wisconsin Department of Natural Resources
Bureau of Integrated Science Services
2801 Progress Rd.
Madison, Wisconsin 53716

U.S. Environmental Protection Agency
Office of Research and Development
Washington, DC 20460

Notice

The United States Environmental Protection Agency (EPA) through its Office of Research and Development funded and managed the research described here via a grant (#R-829537-01-0). It has been reviewed by the EPA and approved for publication.

Mention of trade names or commercial products does not constitute endorsement or recommendation by EPA for use.

Foreword

The mission of the Ecological Exposure Research Division (EERD) within the National Exposure Research Laboratory (NERL) of the U.S. Environmental Protection Agency (EPA) is to improve the scientific basis for understanding, measuring, and protecting biological integrity so that EPA and other resource agencies can make sound, defensible environmental decisions. Our research is primarily focused on the development, evaluation, and implementation of new methods to assess ecosystem condition, to evaluate biotic responses to environmental stressors, and to predict future vulnerability of natural populations, communities, and ecosystems. The scale of our research ranges from molecular to ecosystem levels of biological organization and addresses immediate as well as emerging environmental threats.

EPA's Environmental Monitoring and Assessment Program (EMAP) is a research program to develop the tools necessary to monitor and assess the status and trends of national ecological resources. EMAP's goal is to develop the scientific understanding for translating environmental monitoring data from multiple spatial and temporal scales into assessments of current ecological condition and forecasts of future risks to our natural resources. EMAP focuses on surveys based on probability (i.e., random) sampling designs to estimate condition with a known level of uncertainty. Regional EMAP (REMAP) was initiated to test the applicability of the EMAP approach to answer questions about ecological conditions at regional and local scales. REMAP proposals are submitted through EPA's Regional Offices to the Office of Research and Development (ORD). ORD carries out scientific review of proposals, and qualified proposals are funded as cooperative agreements.

This report describes the results of a REMAP agreement awarded to the Wisconsin Department of Natural Resources (WDNR) to compare random and modified-random sampling designs. The EPA Project Officer was Bhagya Subramanian, and Karen Blocksom, a statistician in EERD, provided assistance with data analysis. The report was prepared by WDNR and has undergone EPA review. EERD is publishing this report to make these findings more widely available, given their potential significance for state or tribal agencies that, like WDNR, are considering wider adoption of random-sampling approaches in their water quality monitoring programs.

Florence Fulk
Acting Director
Ecological Exposure Research Division

Table of Contents

Notice 2

Foreword 3

Tables 6

Figures 8

1. Project Summary 10

- **1.1 Background and Rationale** 10
- **1.2 Project Objectives and Focus of This Report** 10
- **1.3 Methods** 11
- **1.4 Results** 11
- **1.5 Importance to the Science of Environmental Monitoring** 12
- **1.6 Other Analyses Planned for Wisconsin REMAP Data** 13

2. Study Area 14

- **2.1 Watershed Delineation, Land Use Quantification, and Water Quality Stressor Gradient Development** 14
- **2.2 Stream Resources in Study Area** 14

3. Stream Population Sampling Design 17

- **3.1 Random, Modified-Random, Reference, and Targeted Stream Sites Selection** 17
- **3.2 Identification of Sampling Sites** 18
- **3.3 Temporal Sampling Frame** 18

4. Stream Site Sampling Protocols 19

- **4.1 Physical Habitat** 19
- **4.2 Water Quality and Water Chemistry** 19
- **4.3 Periphyton** 20
- **4.4 Benthic Macroinvertebrates** 20
- **4.5 Fish** 21

5. Data Analytical Methods 22

- **5.1 Random, Bridge, and Reference Site Comparisons** 22
- **5.2 Probability Estimates and Evaluation of Population Distributions** 22
- **5.3 Land Use Evaluation** 23

6. Results 24

- **6.1 Sample Size Summary** 24
- **6.2 Watershed Land Use Identification and Quantification - Relationships Among Agricultural Land Use and Physical, Chemical, and Biological Measures** 25
- **6.3 Random, Modified-Random, and Reference Site Comparisons** 30

6.4 Comparison of Reference Sites with Random and Modified-Random Assessment Sites 38
6.5 Probability Estimates and Evaluation of Population Distributions 46
6.6 Targeted vs. Random Stream Assessment 58
6.7 Effects of Geographical Distances Between Random and Modified-Random Sites 59
7. Discussion 61
8. Acknowledgements 66
9. References 67
Appendix A. Comparison of Spring and Fall Water Chemistry Parameters 70
Appendix B. Relationships of Stream Order with Physical Habitat, Water Quality and Water Chemistry, and Biological Measures 72
Appendix C. Fish Assemblage Data Analyses 76

Tables

Table 1. Pearson correlations between watershed land use and in-stream physical habitat measures (N = 78). 25

Table 2. Pearson correlations between watershed land use and water quality and water chemistry measures. 26

Table 3. Pearson correlations between in-stream physical habitat, watershed land use, and biological measures. 28

Table 4. Pearson correlations between water quality and water chemistry measures, and biological measures. 29

Table 5. Weighted paired t-test comparisons of stream physical habitat measures from X and B stream assessment sites. 31

Table 6. Weighted paired t-test comparisons of water quality and water chemistry measures from the X and associated B assessment sites. 33

Table 7. Weighted paired t-test comparisons of macroinvertebrate metrics and indices collected from X and associated B assessment sites. 35

Table 8. Weighted paired t-test comparisons of fish metrics collected from X and B assessment sites. 37

Table 9. Two sample t-test comparisons for physical habitat measures collected at the X and reference sites. 39

Table 10. Two-sample t-test comparisons of water quality and water chemistry values between the X and reference stream sites 41

Table 11. Two sample t-test comparisons of the X and reference site macroinvertebrate metrics. 43

Table 12. Two-sample t-test comparisons of the X and reference site fish metrics. 45

Table 13. Physical habitat reference condition thresholds. 46

Table 14. Water chemistry and water quality reference condition thresholds. 49

Table 15. Macroinvertebrate metric reference condition thresholds. 53

Table 16. Fish metric reference condition thresholds. 55

Table A 1. Paired t-test comparisons between spring and fall water quality and water chemistry measures collected at the X stream sites. 70

Table B 1. ANOVA results of stream order and physical habitat measures collected at the X sites. 72

Table B 2. Tukey HSD pairwise comparison probabilities for physical habitat measures that produced significant ANOVA results. 72

Table B 3. ANOVA of stream order and water quality and water chemistry measures from the X stream sites. 73

Table B 4. Tukey HSD pairwise comparison for the percent oxygen saturation. 73

Table B 5. ANOVA of stream order and macroinvertebrate measures from the X stream sites. 74

Table B 6. ANOVA of stream order and fish assemblage measures from the X stream sites (N=42). 75

Table B 7. Tukey HSD pairwise comparison probabilities for fish metrics that produced significant ANOVA results. 75

Table C 1. Fish species captured at X stream sites. 77

Table C 2. Fish species captured at reference stream sites. 78

Figures

Figure 1. Random and least disturbed reference stream sites in the REMAP study area. 16

Figure 2. Bivariate scatterplots comparing physical habitat measurements collected at the X and associated B sites. 30

Figure 3. Bivariate scatterplots comparing water quality and water chemistry measurements collected from the X and associated B sites. 32

Figure 4. Bivariate scatterplots comparing macroinvertebrate metrics collected at the X and associated B sites. 34

Figure 5. Bivariate scatterplots comparing fish metrics collected at the X and associated B sites. 36

Figure 6. Bar charts of weighted means and standard deviations (SD) for X and B sites and simple random sample mean and SD for reference sites for physical habitat measures. 38

Figure 7. Bar charts of means and standard deviations for water quality and water chemistry measures collected at the X, B, and reference stream sites. 40

Figure 8. Bar charts of mean and standard deviation of macroinvertebrate metrics collected at the X, B, and reference stream sites. 42

Figure 9. Bar charts of mean and standard deviation of fish metrics collected at the X, B, and reference stream sites. 44

Figure 10. Cumulative distribution function curves of physical habitat measures collected at the X and B sites. 47

Figure 11. The percentages of stream miles not meeting the reference condition threshold values for physical habitat measures. 48

Figure 12. Cumulative distribution function curves of water quality measures collected *in situ* at the X and B sites. 50

Figure 13. Cumulative distribution function curves of laboratory analyzed water chemistry measures collected at the X and B sites. 51

Figure 14. The percentages of stream miles not meeting the reference condition threshold values for water quality and water chemistry measures. 52

Figure 15. Cumulative distribution function curves of macroinvertebrate metrics collected at the X and B sites. 54

Figure 16. The percentages of stream miles not meeting the reference condition threshold values for macroinvertebrate metrics. 55

Figure 17. Cumulative distribution function curves for fish metrics collected at the X and B sites. 56

Figure 18. The percentages of stream miles not meeting the reference condition threshold values for fish metrics. 57

Figure 19. Cumulative distribution function curves of fish IBI scores from biologists' targeted sites and REMAP X sites, compared to the reference condition threshold value. Box plots and distribution of the biologist's targeted and REMAP X sites data is shown in figure on right. 58

Figure 20. Scatterplots showing relationships between the distance between the X and B sites, and the absolute value of the differences in physical, chemical, and biological measures collected at these sites. 59

Figure A 1. Box and whisker plots of spring and fall water quality and water chemistry measures collected at the X-sites. 71

Figure C 1. Percent frequency of occurrence of fish species at the X, B, and reference sites. 79

1. Project Summary

1.1 Background and Rationale

Hundreds of millions of dollars have been spent in Wisconsin on watershed management and restoration activities, yet certain farming, and suburban and urban development practices continue to significantly impact a large portion of the state's waters, and in some areas of the state severe stream degradation is readily apparent. Improvements in the assessment of Wisconsin stream resources are needed to provide a comprehensive evaluation of the state's waters. Greater understanding of land use factors affecting water resources will help improve land and water resource management, and broader dissemination of this information is needed to improve the political decision-making processes for these management activities.

To date, stream monitoring conducted by the WDNR has primarily been targeted sampling to provide information for stream-specific management issues. Physical, chemical, and biological data are often collected from either highly degraded streams affected by polluted run-off, or from high quality streams where game fish management or stream habitat enhancement efforts are being evaluated. This resulting data set can be strongly biased if used for making inferences about broad-scale conditions of stream resources. Spatial clustering of the WDNR's current sampling effort on a relatively small proportion of the State's streams and a focus on larger streams that support adult game fish also limit the ability to make meaningful statements about Wisconsin's entire stream population.

Beginning in 2003, the Wisconsin Department of Natural Resources (WDNR), with support from the EPA's Regional Environmental Monitoring and Assessment Program (REMAP), conducted an assessment of the physical, chemical, and biological conditions of wadeable streams in the Driftless Area ecoregion in western Wisconsin using a probabilistic sampling design. The Driftless Area ecoregion encompasses 20 percent of Wisconsin's total land area and contains 21 percent of the State's perennial stream miles.

1.2 Project Objectives and Focus of This Report

The Wisconsin REMAP study was conducted to address the following objectives:

Objective 1:
Use a statistically valid probabilistic sampling design to assess the condition of the entire wadeable stream population in the Driftless Area ecoregion, and evaluate whether a modified-random sample survey design (sampling randomly selected stream reaches near road-accessible access points) characterizes individual or populations of streams similarly to a truly randomized survey sampling design.

Objective 2:
Evaluate whether targeted stream sampling routinely conducted by the WDNR characterizes stream fish populations similarly to a probabilistic sampling design.

Objective 3:
Investigate whether the effects of agricultural land use can be detected in the quality of in-stream physical habitat or water chemistry measures, and whether changes in stream physical or chemical characteristics influence the biological assemblages in streams.

Objective 4:
Evaluate whether one biotic assemblage is more discriminating of specific stressors such as percent agricultural land within a watershed, in-stream or riparian physical habitat degradation, or chemical pollutants, than others, and whether there are significant differences in the estimated miles of degraded streams depending upon the biotic assemblage used in the estimation.

Objective 5:
Compare macroinvertebrate taxonomic data collected using two different field sampling protocols, differing laboratory sub-sample sizes, and differing levels of laboratory taxonomic resolution, to evaluate whether certain field or lab protocols are more effective at detecting and measuring various types of environmental stressors. Additionally, develop a multi-metric macroinvertebrate index for the Driftless Area ecoregion and subsequently the entire state, using the measured physical and chemical explanatory data, and the more rigorous protocols identified.

This report presents the results of data analyses addressing Objectives 1-4 for macroinvertebrate and fish assemblages. Algae were collected and processing of these samples is underway. Analyses to address Objective 5 are in progress.

1.3 Methods

Watershed land use, riparian and in-stream habitat, field and laboratory-analyzed water chemistry, periphyton, macroinvertebrate, and fish data were collected from randomly selected stream sites (n = 60), and an associated “modified-random" sampling site on each of these streams, that was accessed via a road crossing nearest the randomly selected sampling site. These data allowed us to evaluate whether sampling randomly selected stream sites at the nearest road crossings would significantly bias the assessment results compared to sampling truly random sites. Least disturbed reference sites were sampled to develop stream-quality expectations based on objective reference condition criteria (n = 22). Also, WDNR biologists provided fish assemblage data from 60 targeted stream sites they had sampled, and which were thought to be representative of the range and modal condition of stream resources in the study area. We compared fish data from these targeted sites to that from the REMAP random sites to evaluate whether targeted sampling characterized the Driftless Area ecoregion stream population similarly to the stratified random survey sampling design.

1.4 Results

The mean distance between the random (probabilistic) and the associated modified-random assessment site on each stream was 701 meters. Study results show no significant differences between the random and modified-random assessment sites for all 9 physical habitat measures, 11 water chemistry measures, 7 macroinvertebrate metrics, or 8 fish metrics.

We provide evidence that targeted sampling routinely done by WDNR characterized the biological condition of stream resources in the Driftless Area ecoregion differently than probabilistic sampling. Cumulative distribution plots of fish Index of Biotic Integrity (IBI) scores from targeted stream sites sampled by WDNR regional biologists indicate that 65% of the stream sites in the study area were not meeting the reference condition threshold value of 60, versus the estimate of 80% of the random sample sites not meeting this threshold based on the probabilistic sampling design results.

Row cropping is the dominant land use in the study area followed by forest cover. Study results indicate there were no significant relationships between the percentage of agricultural land in each watershed and the degradation of various in-stream physical habitat features, but water column nitrate-nitrite concentrations and conductivity were positively correlated with percent watershed agricultural land. Fish assemblages appeared to be insensitive indicators of water quality, but sensitive macroinvertebrate taxa declined with increased Kjeldahl nitrogen, ammonia, total phosphorus and total dissolved phosphorus concentrations, and were positively correlated with dissolved oxygen concentration, percent dissolved oxygen saturation and water transparency.

Physical, chemical, and biological measures from least disturbed reference stream sites were used to develop reference conditions. There were significant differences between the random sample population (and by inference the entire stream population in the Driftless Area ecoregion of Wisconsin) and the reference condition for a number of stream physical, chemical, and biological measures. As examples, we documented that 75% of the random sample population streams were degraded based on measures of stream bank erosion; 70% of the random sample population had higher total dissolved phosphorus concentrations than the reference condition; approximately 75% of the streams were degraded based on the percentage of sensitive macroinvertebrate taxa present; and 74% of the streams showed impairment based on fish IBI scores.

1.5 Importance to the Science of Environmental Monitoring

Randomization is an important aspect of sample survey design as it allows objectivity in the sample selection process, and in the evaluation of sources of error or sampling variability (Larsen 1995, Peterson et al. 1999). Probabilistic sampling designs can reduce sampling bias and increase the representativeness of sample data used to make inferences about the target population. Previous experience from employing probabilistic sampling designs in Wisconsin have shown disadvantages when using this type of sampling design including: 1) Reduced accessibility and increased travel times to more remote sites can significantly increase sampling effort and cost; 2) Reduced site accessibility can result in having to use more transportable and often less effective sampling gear; and 3) Streams highly degraded by poor land management are often associated with landowners unwilling to allow access to streams on their property due to distrust of government agencies or fear of legal action. With the REMAP project, we evaluated whether random, modified-random, and non-random (targeted) sampling designs provided similar results. These findings are being used to evaluate fundamental aspects of stream sampling designs, which should help improve the monitoring and assessment of stream resources in Wisconsin and elsewhere.

Knowledge of what the physical, chemical, or biological conditions of a stream should be in the absence of human disturbances, and understanding of how streams respond to anthropogenic stressors, are important to guide assessment and management efforts (Karr and Chu 1999, Davies et al. 2000). The Wisconsin REMAP project initiated the collection of physical, chemical, and biological data from least disturbed stream sites throughout western Wisconsin to help define ecoregional expectations (reference conditions). This methodology will subsequently be applied to the entire state. Land use types within the watersheds upstream from each study site were quantified to evaluate how land use influence stream habitat and water quality characteristics, which in turn can influence the biotic assemblages.

1.6 Other Analyses Planned for Wisconsin REMAP Data

Numeric biological criteria can provide objective quantifiable measures of ecosystem health. Biological indices have been developed for Wisconsin streams using macroinvertebrate and fish assemblage data (Hilsenhoff 1987, Lyons 1992a, Lyons et al. 1996). Hilsenhoff's Biotic Index (HBI) uses macroinvertebrate taxa data from samples collected from riffles to assess organic pollution that manifests itself as reduced dissolved oxygen concentrations (Hilsenhoff 1987). Alternative or refined macroinvertebrate field or lab protocols may result in the collection of data that are more discriminating of specific types of stream impairment other than organic pollution. Refined macroinvertebrate indices would also add additional measures of stream condition to corroborate physical, chemical, and biotic measures such as fish or diatom indices, and can be applied to streams where these other biological indices are insensitive. Two commonly used macroinvertebrate field sample collection protocols (single (riffle) habitat and proportional multi-habitat sampling), differing laboratory sub-sampling levels (100, 300, 500 fixed-count), and differing levels of taxonomic identification (family and genus levels, versus lowest practical level) are also being evaluated as part of the Wisconsin REMAP project. These results will be used to characterize the method performance (rigor) of various aspects of the WDNR's macroinvertebrate field and lab protocols. These data will also be used to develop a multi-metric macroinvertebrate index for western Wisconsin streams, and this process will subsequently be applied to develop multi-metric macroinvertebrate indices for the entire state.

2. Study Area

The geographic extent of the REMAP study area is 7,418,000 acres (20.6% of Wisconsin's total land area), and encompasses the Western Coulee and Ridges and the Southwest Savanna ecoregions in western Wisconsin (Figure 1), which are based on the United States Department of Agriculture – Forest Service's National Hierarchical Framework of Ecological Units (Keys et al. 1995). These ecological units closely approximate Omernik's Level III Driftless Area Ecoregion (Omernik 1987) and are collectively referred to as the Driftless Area in this report. The Driftless Area extends into the states of Minnesota, Iowa, and Illinois, and remained unglaciated during the most recent glacial period, whereas the rest of Wisconsin and most of the upper Midwest was covered with ice. As a result, the Driftless Area has a "mature" drainage system characterized by deeply incised valleys and few natural lakes or wetlands relative to the rest of Wisconsin. The relatively steep topography creates a significant water flow gradient and valley bottoms are close to the water table, resulting in streams with high baseflow that are dominated by coldwater fish assemblages. The soils of the Driftless Area are well-drained silty loess over dolomite, limestone, or sandstone. Land use patterns closely follow spatial differences in slope, with field corn, soybean, alfalfa, and pastureland situated on the ridge tops and in the valley bottoms, and hardwood forests dominate the steeply sloped valley sides. The Driftless Area was selected for the REMAP study area because the amount of topographic relief in this area provides clearly delineated watersheds, and land use (and presumably stream quality) varies significantly among watersheds, relative to other areas of the state (Keys et al. 1995).

2.1 Watershed Delineation, Land Use Quantification, and Water Quality Stressor Gradient Development

We used Geographic Information System (GIS) technology and a digital elevation model to delineate the watershed area upstream of each random and least disturbed reference stream site. Wisconsin Initiative for Statewide Cooperation on Landscape Analysis and Data (WISCLAND) GIS land cover data were used to quantify land use within the study area watersheds. The WISCLAND data were derived from LANDSAT Thematic Mapper satellite imagery acquired from fly-overs beginning in August 1991 and ending in May 1993. The on-the-ground resolution of each WISCLAND digital pixel is 30 square meters. We quantified land use data to develop a land use disturbance gradient based on the total proportion of agricultural (row crop) acreage found within each watershed upstream of each stream assessment site. Higher proportions of row crop acreage within each watershed were equated with greater potential in-stream impacts. Agricultural land use is the major land "disturbance" and cause of degradation to Wisconsin streams (Robertson et al. 2006). In particular, row crops such as field corn and soybeans are the major contributors of sediment, and chemical fertilizers and livestock manure are the major sources of nitrogen and phosphorus for the study area streams. It is recognized that riparian land use, buffer width, plant species composition and linear extent, cropland slope, crop rotations, and numerous geological and meteorological factors strongly influence pollutant delivery to streams, but the level of effort needed to adequately characterize and quantify these factors were beyond the scope of this study.

2.2 Stream Resources in Study Area

The REMAP study area has an estimated 3,560 individual (19.5 % of statewide total number) and 8,840 miles (21.1% of Wisconsin's total stream mileage) of perennial streams. Perennial streams in the study area were identified and numbers and miles quantified from a 1:24,000-scale GIS statewide hydrography layer developed by the WDNR from digitized 7.5 minute U. S. Geological Survey (USGS) quadrangle maps. The numbers of perennial stream miles by Strahler (1957) stream order for the REMAP stream population are as follows: order 1 (1,031 mi.), order 2 (2,448 mi.), order 3 (2,282 mi.), and order 4 (1,548 mi.). The REMAP target population of first through fourth order streams comprise 3,540 individual (19.3% of Wisconsin's total number) and 7,310 miles (17.4% of Wisconsin's total mileage of perennial streams) of streams. In western Wisconsin, fifth order streams typically transition from wadeable to non-wadeable; therefore, only first through fourth order streams were included in the target and sample populations.

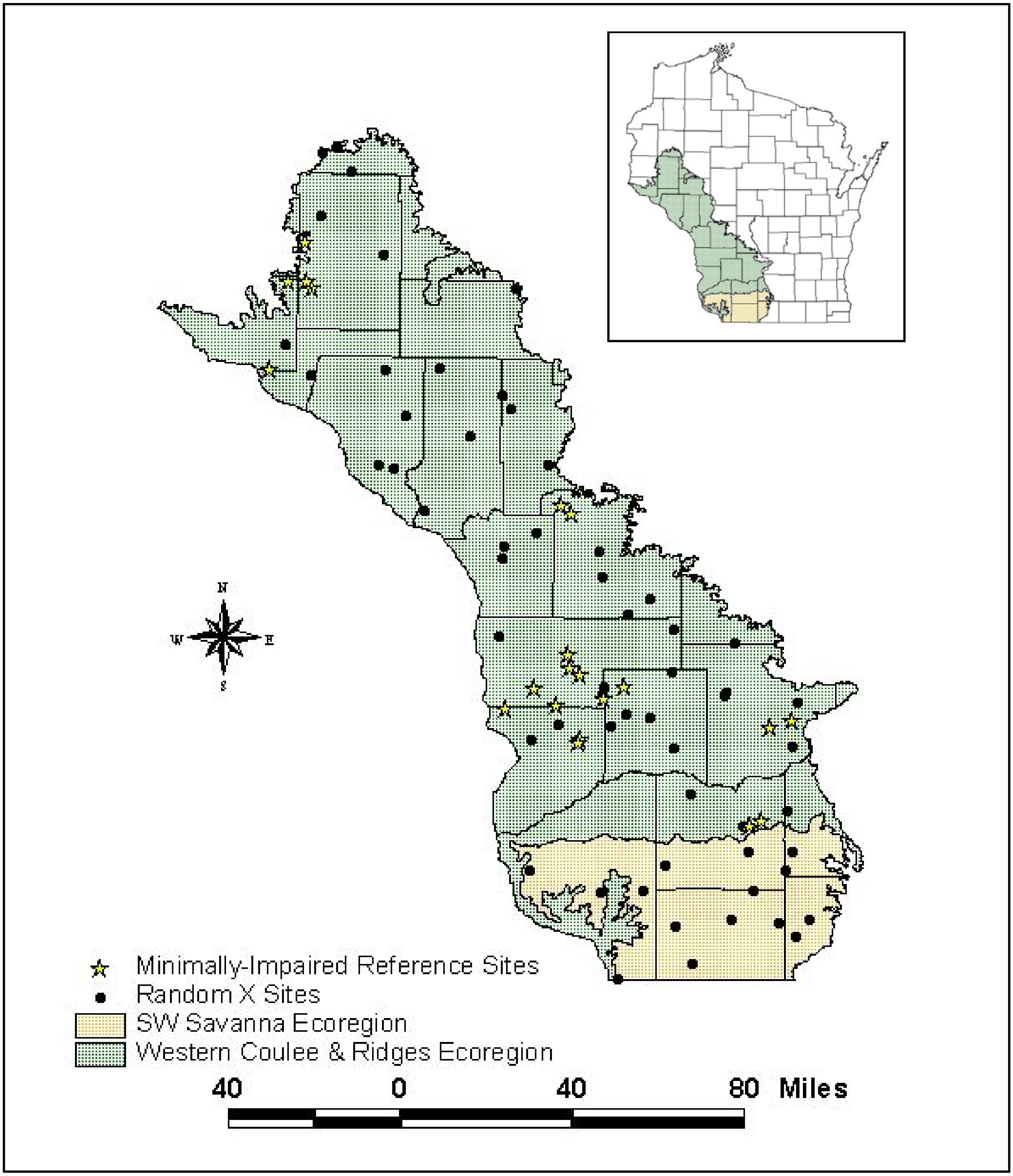

Figure 1. Random and least disturbed reference stream sites in the REMAP study area.

3. Stream Population Sampling Design

3.1 Random, Modified-Random, Reference, and Targeted Stream Sites Selection

The EPA's Environmental Monitoring and Assessment Program (EMAP) probabilistic sampling design was used to randomly select stream sites from the target population (Stevens and Olsen 1999, 2004). The sampling design was weighted so that equal numbers of first through fourth order streams were included in the sample population. The EPA Office of Research and Development (ORD) staff in Corvallis, Oregon used an unequal probability random tessellation stratified (RTS) sampling design described by Stevens and Olsen (1999) and Stevens (1997), and the WDNR 1:24000 - scale hydrography layer to identify 100 potential sample sites. From these sample sites, 15 perennial stream sites in each of stream orders one through four (n = 60) were randomly selected throughout the Driftless Area ecoregion. EPA - ORD also identified an additional 100 over-sample sites to use as replacements if original sites were rejected. The sample population for each stream order was weighted to account for the number of stream miles these populations represented in the target population. These initial weights were calculated assuming that only the first 100 sites would be evaluated. Because many of the original sites were rejected and replaced with over-sample sites, weights were adjusted by multiplying the total stream length calculated for the study area by the original weight, and dividing this number by the sum of the weights of the final sample population. The WDNR hydrography layer was also used to generate summary statistics on numbers and miles of streams statewide and in the study area.

Reference Site Selection

Selection of reference stream sites was based on evaluation of watershed land ownership (i.e., county, state, or federal lands tended to have less agricultural or urban land) and land use, using the WISCLAND database. This *a priori* method used to select candidate watersheds and reference stream sites was based on guidance developed by EPA and others (Hughes et al. 1986, Gibson et al. 1996). Reconnaissance of each candidate reference site was conducted to verify that there were no apparent watershed or riparian land use factors significantly degrading the site, and to do a cursory evaluation of in-stream physical habitat conditions prior to designating it as a reference site. To develop reference conditions we collected physical, chemical, and biological data from 22 least disturbed reference stream sites located throughout the study area (Figure. 1). We sampled a total of four first-order, four second-order, ten third-order, and four fourth-order least disturbed reference sites during the 2003 field season for the same parameters as the random and modified-random sites.

Regional Targeted Sampling Sites

Regional WDNR biologists that assess and manage stream resources in the REMAP study area were asked to select a total of 60 stream sites that were recently sampled and thought to be representative of the range and modal condition of stream resources in the study area. Fish data were collected by the biologists from these non-random sites using the same sampling protocols used in the REMAP study, and these data were then compared to the fish data collected at the REMAP random sites.

3.2 Identification of Sampling Sites

EPA-ORD provided the latitude and longitude coordinates to identify the mid-point of each randomly selected stream sampling site. Each corresponding modified-random sampling site was located either upstream or downstream of the randomly selected site, at the nearest "easy" access point (typically accessed from a roadway or driveway bridge that crossed each stream). We dropped candidate streams from the sample population if there were intervening tributaries between the random and modified-random sites that resulted in a Strahler stream order difference of more than one order between the random and modified-random sites. In this report, random sites are subsequently referred to as "X" sites and the modified-random sites are referred to as "B" (Bridge) sites. The B sites were located sufficient distances (typically 10 x the mean stream wetted width (MSW)) away from road crossings or driveways to avoid hydraulic influences of bridge abutments, culverts, or other manmade structures on stream physical habitat characteristics, or that created artificial fish habitat.

3.3 Temporal Sampling Frame

We followed the sampling index periods of the WDNR's standard operating procedures (SOPs) when collecting stream physical habitat and fish assemblage data, and macroinvertebrate and water chemistry samples. We collected stream physical habitat data and water chemistry samples during stream baseflow conditions following spring snowmelt and rain event high-flow conditions. Fish sampling was done from mid-June through mid-September. The fish sampling index period avoids spring migratory movements of spawning catostomids, and high numbers of young-of-the-year salmonids and centrarchids that exceed typical stream carrying capacities, as well as fall upstream spawning migrations of salmonids and downstream overwintering migrations of ictalurids and centrarchids. We collected macroinvertebrate samples in the fall of 2003. The fall index period for benthic macroinvertebrate sampling allows for a greater portion of aquatic invertebrates to be identified since most taxa are of sufficient size and development for identification. In-situ water chemistry (instantaneous readings from an electronic water quality meter) and laboratory-processed grab samples were collected in the spring (June) and again in the fall (August – September) during baseflow conditions at each X, B, and reference stream site during the 2003 field season (Appendix A).

4. Stream Site Sampling Protocols

4.1 Physical Habitat

In-stream and riparian physical habitat characteristics were measured or visually estimated at 12 transects within each X, B, and reference stream assessment reach. We used WDNR SOPs for assessing stream physical habitat at all of the study sites. Stream assessment reach lengths for physical habitat and fish assemblage sampling were based on 35 x MSW. An assessment length of 35 X MSW generally encompasses three run-riffle-pool sequences, and typically most fish species found within stream reaches in Wisconsin will be encountered within this length of stream surveyed (Lyons 1992b). For streams with a MSW less than 2.9m, the minimum assessment reach was 100m long. At some sites, the reach lengths were slightly lengthened if it resulted in starting and/or ending the assessment reach at a riffle, which enabled more efficient fish collection since blocknets were not used. At each habitat assessment transect, measures of water and sediment depth, bankfull water depth, overhead canopy; and visual estimates of percent substrate composition were collected at four points equally spaced along the transect line. Bank erosion was measured and riparian land use and land cover characteristics were visually estimated along each of the transect lines extending laterally 10 meters into the upland riparian zones. The distance between bends was measured and divided by the mean stream width. Throughout the report this ratio is referred to as ‘distance between bends’ to abbreviate.

4.2 Water Quality and Water Chemistry

We collected 6 *in situ* (metered) water quality measures at all X, B, and reference stream sites in the spring and fall during baseflow conditions. A Yellow Springs Instruments Company Model 85 electronic water quality meter was used to collect instantaneous measures of water temperature, dissolved oxygen concentration, dissolved oxygen percent saturation, and conductivity. An Orion Quikchek™ pen was used to measure pH. Both meters were calibrated following the manufacturer’s instructions prior to the start of each field day using air calibration for dissolved oxygen, and chemical standards for conductivity and pH. To measure water column transparency we used a transparency tube. This device consists of a 125 cm long x 4.5 cm diameter Plexiglas cylinder sealed at the bottom with a 4.5 cm diameter Secchi disk. The transparency tube was filled with stream water while at each assessment site, and the water was drained from the tube (reducing the water column height) until the Secchi disk became visible when looking down into the tube through the water column. The height of the water column was measured by reading a height scale on the side of the tube. For some assessment sites (particularly reference sites), the transparency of the water was often high enough that the Secchi disk could be seen when the tube was completely filled with water (122 cm), so that an exact transparency value could not be measured, since the true value was “off-scale”. In these instances, a value of 122 cm was assigned.

Five water chemistry parameters were measured. We collected one laboratory-analyzed water chemistry grab sample during baseflow in the spring and fall at each X, B, and reference site. The parameters analyzed included: total phosphorus, total dissolved phosphorus, total Kjeldahl nitrogen, ammonia, and nitrate - nitrite nitrogen. Samples were preserved (acidified) and placed on ice in the field, refrigerated upon return from the field, and submitted to the Wisconsin State Laboratory of Hygiene within specified holding times for analysis, following the lab SOPs.

4.3 Periphyton

To assess diatom assemblages we collected periphyton samples at all X, B, and reference sites. Samples were collected in August and September of 2003 from cobble or smaller rocks, or in the absence of mineral substrates, soft sediment samples were collected. At sites with rock substrate, nine rocks (5cm – 25cm in diameter) were taken from riffle areas. An area 4 cm square was measured on the upper surface of each of the rocks with a ruler and delineated with a scalpel. The scalpel was then used to scrape the delineated area on each rock, and the material adhering to the scalpel and loosened on the rock was washed with de-ionized water into a clean pan. The same delineated areas were then scrubbed with a clean nylon brush and both the brush head and brushed area of the rocks were again rinsed into the sample collection pan. The composite sample was poured from the pan into a 1L bottle and diluted to the smallest volume possible, either 250mL, 500mL, 750mL, or 1000mL, and the dilution volume was recorded. The capped 1L bottle was vigorously shaken to suspend and homogenize the sample. A clean turkey baster was then used to remove a 40mL aliquot from the 1L bottle and transfer it into a 60mL bottle. The 40mL aliquot was preserved with 1.6mL of a 50% glutaraldehyde solution. The preserved samples were placed on ice in the field and refrigerated upon arrival to the lab. At sites lacking coarse substrate, a glass Petri dish (49mm inside diameter) was inverted and placed on streambed fine sediment (sand, clay, silt), and a stainless steel spatula without openings in the blade was used to trap the fine sediment within the Petri dish. The fine sediment was then deposited into a clean sample pan. A total of nine fine-sediment samples were collected and composited in the sample pan. The fine sediment periphyton samples were homogenized, sub-sampled, and preserved using the sample methods applied to the coarse substrate periphyton samples.

Although the collection and analysis of periphyton samples from the X, B, and reference sites were not part of the original study proposal, these samples were collected and are currently being processed. Upon completion of the analytical results, a manuscript will be submitted to a scientific journal.

4.4 Benthic Macroinvertebrates

We collected macroinvertebrate samples at all X, B, and reference stream sites in the fall of 2003 using a D-framed net with 500 micron mesh, following WDNR SOPs. One kick sample was collected from a single riffle (course gravel or cobble substrate) located within the habitat and fish sampling reach where water velocities created erosional habitat. The net frame was placed at arm's-length downstream of the sample collector, who using the toe or heel of their boot disturbed the substrate to a depth of approximately 5 cm. Sampling continued for approximately 3 minutes, typically at which time there was a fist-sized wad of debris in the net and it was evident that over 100 organisms had been collected. In the absence of coarse substrate, overhanging riparian vegetation, or grass and leaf "snags" were sampled. The D-frame net was used to sweep and jab the overhanging vegetation for approximately 5 minutes, or the net was placed downstream of snags, and the tree branches holding the snags were disturbed until the organic debris was dislodged and washed into the net. In addition, a 20-jab proportional-habitat sample was collected along a 100m-long reach within the habitat and fish assessment reach at each X, B, and reference stream site following protocols described by Barbour et al. (1999). A rangefinder was used to estimate each 100m-long sampling reach. For both the single riffle and proportional habitat samples, rocks, twigs, and other coarse debris were removed from the net while making certain to rinse or remove by hand organisms clinging to the debris and placing them back into the sample. Samples were placed in containers identified with internal and external labels, preserved with an 80% ethanol solution, and transported to the lab for processing.

We used a 125+ organism fixed-count random sub-sampling method in the lab to process all of the macroinvertebrate samples, following WDNR SOPs. In the lab, each field sample was evenly distributed in a gridded pan, and a random number was used to select a grid square. All organisms within the chosen grid-square were removed, enumerated, and stored for later taxonomic identification. If fewer than 125 organisms were removed from the initial grid square, subsequent randomly selected grid squares were picked in their entirety until the target number of 125+ organisms was reached. Guidelines provided by Hilsenhoff (1987) were followed for determining which taxonomic groups were identified and the level of taxonomic resolution that was applied to the taxa in the sub-samples. Lifestages and groups of organisms not included in the enumeration and identification included: adult insects, empty or sealed Trichoptera cases, Hemiptera, Coleoptera (non-dryopids), Collembola, Mollusca, Annelida, Decapoda, Nematoda, Nematomorpha, Hydracarina, and Turbellaria.

The taxonomic analyses are still being conducted on the 20-jab samples and only the results of the riffle samples are included in this report. The single-riffle and 20-jab sample data will be part of a detailed Performance-Based System study (Miller et al.: in prep.) that follows guidelines developed by Diamond et al. (1996). The study will evaluate the rigor of macroinvertebrate field and lab methods to help refine WDNR biological assessment protocols and macroinvertebrate indices for streams.

4.5 Fish

We used WDNR SOPs to assess fish assemblages by electrofishing each of the X, B, and reference sites. The majority of the stream sites were sampled by a three-person crew that used two handheld electrodes powered by a DC generator mounted in a small tow barge, pulled upstream by one crew member. Each crewmember used a hand-held net to capture fish. In streams too narrow or shallow to be negotiated by the tow barge, a two-person crew, each with a hand-held net, used one battery-powered backpack electrofishing unit with one anode to stun fish. All stream habitats were thoroughly sampled, and an effort was made to capture all fish greater than 25mm total length. The fish were placed in the tow barge live well or in buckets, and subsequently identified to species, measured, enumerated, and released. If field identification of fish specimens were uncertain, specimens were preserved for laboratory identification.

5. Data Analytical Methods

5.1 Random, Bridge, and Reference Site Comparisons

A total of 9 physical habitat, 11 water quality and water chemistry, 7 macroinvertebrate, and 8 fish metrics were analyzed and are presented in this report. Prior to analyses, variables were "weighted" using an assigned weight that accounted for the total number of stream miles represented by that site in the target population. Weighted means and Horvitz-Thompson estimates of standard deviation (Diaz-Ramos et al. 1996) were calculated for X and B sites using the PSURVEY.ANALYSIS package (v. 2.7) developed by EPA's EMAP program (available at http://www.epa.gov/nheerl/arm/analysispages/software.htm) for the free software R (R: A Language and Environment for Statistical Computing, Vienna, Austria; http://www.r-project.org/). An alpha level of 0.05 was applied to determine the significance of all tests reported.

Bivariate scatterplots were created for each metric to visually compare the X and B site data, and Pearson correlation coefficients were calculated to identify any significant correlations. Paired t-tests incorporating the weighted means and standard deviations were used with a Bonferroni adjustment to determine if significant differences existed for any of the variables collected, between the X and B sites.

The weighted means and standard deviations were plotted alongside the simple random sample mean and variance estimates calculated from reference sites for each variable. Two-sample t-tests assuming unequal variances were performed using the weighted mean and variance values for the X sites and the simple random sample mean and variance for reference sites. A Bonferroni adjustment was applied to each group of variables (i.e., habitat, chemistry, macroinvertebrates, fish) to determine significance of differences at an overall Type I error rate of 0.05.

5.2 Probability Estimates and Evaluation of Population Distributions

Empirical cumulative distribution functions (CDFs) of X and B site data were generated to estimate the percentage of the X and B stream populations that met reference condition thresholds for various physical habitat, chemical, and biological metrics. The 25th percentile (for those metrics where lower values indicated lower environmental quality) or the 75th percentile (for those metrics where higher values indicated lower environmental quality) of physical habitat, chemical, and biological measures calculated from the reference sites were used to develop the reference condition thresholds. The CDFs and percentage estimates were calculated using the PSURVEY.ANALYSIS package in R. A cumulative distribution function summarizes the overall distribution of some variable (x-axis) measured at many random sites. The probability that any variable within the population will be less than or greater than some specified value can be estimated from the CDF curve (Sokal and Rohlf, 1981). If X and B CDFs provide equivalent data, then their population estimates should be similar and their confidence intervals should overlap. Kolmogorov-Smirnov two-sample tests were used to determine whether the X and B probability distributions differed statistically (Sokal and Rohlf, 1981). CDF plots were also used to compare the WDNR regional biologist's targeted fish sample data with the REMAP X sites fish sample data.

5.3 Land Use Evaluation

A digital elevation model was used to delineate watershed areas upstream of each X, B, and reference site. WISCLAND land cover data were used to estimate percent land use types within each of these watersheds. Pearson correlation coefficients were calculated to evaluate relationships between stream reach physical habitat, water chemistry, and biological assemblage attributes, and the percent row crop agriculture, grassland, and forest cover within each watershed. Stepdown Bonferroni adjustments were applied to determine p-value significance (Legendre and Legendre, 1998).

6. Results

6.1 Sample Size Summary

Of the 100 random sites and 100 additional over-sample sites provided by EPA-ORD, we rejected a total of 71 candidate sites during their initial reconnaissance. Reasons for dropping sites included: stream channel was dry at assessment site (38 sites), with the majority of these in Strahler first order (68%) and second order (24%) streams; Strahler stream order changes between the random and modified-random assessment sites were greater than 1 stream order (10 sites); assessment site was in a wetland without a well-defined stream channel (6 sites); greater than 50% of the assessment site fell within an artificial impoundment or area dammed by beaver (<u>Castor</u> <u>canadensis</u>) making these sites too deep to wade (5 sites); assessment site was actually part of Mississippi River backwater channels (5 sites); access to private property was denied (4 sites); no stream channel was located at the random site coordinates (2 sites); a 3 meter-tall barrier fence prevented access to the assessment site (1 site). Two of the 60 streams accepted after reconnaissance were subsequently dropped because they dried-up prior to collecting all of the fish, habitat, or macroinvertebrate data. The final study population consisted of 58 stream sites. In all analyses, the term “percentage of stream miles” refers specifically to stream miles that were accessible, had water present, and were wadeable.

Physical habitat data were collected at 57 of the 58 X-B paired sites. One of the 58 sites could not be sampled at the X site due to the presence of dangerous livestock. Four streams had their X and B sites combined for habitat sampling, because upon delineating the assessment site it was found that the X and B sampling sites overlapped.

Water chemistry data were collected at the X and B sites from 56 of the 58 streams. Two of the 58 streams had their X and B sites combined when collecting water chemistry data. Some sampling sites did not have water quality or chemistry data collected from either the spring or fall sampling period. Reasons included: suspect meter readings (significantly outside the range of values routinely encountered, or a meter was not meeting the calibration standard while in the field); grab samples were not collected because recent rainfall resulted in non-baseflow conditions at the site; or as with conductivity, no spring data were collected due to a meter malfunction. Only spring water quality and chemistry data were used for analyses, with the exception of conductivity measures, since the spring sample size was greater.

Macroinvertebrates were sampled from 57 of the 58 random sites, and fish were surveyed at 55 of the 58 sites. Fish were not sampled at two sites with dangerous livestock, and inclement weather prevented sampling at another site. Four streams had their X and B sites combined for fish assessments, because upon delineating the assessment sites it was found that the X and B assessment reaches overlapped. Only those stream sites where a minimum of 25 individual fish were captured were used in computing fish IBI scores.

6.2 Watershed Land Use Identification and Quantification - Relationships Among Agricultural Land Use and Physical, Chemical, and Biological Measures

The primary land use in the study area, determined from the 1991 – 1993 WISCLAND (LANDSAT) data, was agriculture (40%) followed by forest cover (37%), grassland (13%), wetland (5%), open water (2%), urban land (1%), and miscellaneous other (2%). Watershed area upstream from the X sites ranged from 102 acres to 81,881 acres, with an average size of 7,268 acres. For the reference sites, watershed sizes ranged from 123 acres to 34,450 acres, with an average watershed size of 5,532 acres.

The total percentage of land area identified as row cropland within the X site watersheds ranged from zero to 82%, with an average of 19%. The total land area identified as row cropland in the reference site watersheds ranged from zero to 41%, with an average of 14.5%. Forest cover in the X site watersheds ranged from zero to 86% with an average of 43%, and in the reference site watersheds forest cover ranged from 15% to 99%, with an average of 41%. Pearson correlations were calculated to evaluate whether percent agricultural land, percent grassland, or percent forest cover within watersheds had detectable relationships with various riparian and in-stream physical habitat measures. Data collected at both the random X sites and the minimally disturbed reference sites were used in these analyses. No significant correlations were observed among the in-stream physical habitat measures and the surrounding watershed land use (Table 1). Stepdown Bonferroni adjusted p-values equal 1.00 for all correlations in Table 1.

Table 1. Pearson correlations between watershed land use and in-stream physical habitat measures (N = 78).

Habitat Measures		% Agriculture	% Grassland	% Forest
% Sand, silt, and clay sediments	r_s = p-value =	-0.163 0.153	0.138 0.230	0.125 0.276
Bank erosion	r_s = p-value =	-0.069 0.550	0.114 0.318	0.041 0.721
Depth of fine substrate	r_s = p-value =	-0.001 0.993	0.087 0.446	-0.068 0.552
Sinuosity	r_s = p-value =	0.049 0.667	0.093 0.420	-0.052 0.649
Distance between bends	r_s = p-value =	0.132 0.250	0.063 0.581	-0.178 0.120
Riffle length	r_s = p-value =	-0.042 0.716	0.058 0.615	0.049 0.670
Pool length	r_s = p-value =	0.020 0.866	-0.022 0.851	-0.001 0.993
Riparian buffer	r_s = p-value =	-0.185 0.106	0.062 0.590	0.178 0.120
Width:Depth ratio	r_s = p-value =	-0.139 0.224	-0.080 0.486	0.173 0.129

Pearson correlations and associated p-values were calculated to evaluate relationships between watershed land use and stream water quality and water chemistry measures, and these original p-values are presented in Table 2. In addition, Stepdown Bonferroni adjustments were applied and the resulting adjusted p-values are provided in parentheses. Nitrate-nitrite and conductivity were found to be significantly correlated with percent agriculture and forest cover.

Table 2. Pearson correlations between watershed land use and water quality and water chemistry measures. P values in bold are significant when a Stepdown Bonferroni adjustment is applied.

Water Quality and Chemistry		% Agriculture	% Grassland	% Forest
N-Kjeldahl	r_s =	0.109	-0.058	-0.152
	p-value =	0.332 (1.00)	0.610 (1.00)	0.177 (1.00)
	N =	80	80	80
NH_3	r_s =	0.118	-0.015	-0.171
	p-value =	0.298 (1.00)	0.892 (1.00)	0.130 (1.00)
	N =	80	80	80
NO_3NO_2	r_s =	0.606	-0.149	-0.586
	p-value =	**<.001 (0.003)**	0.188 (1.00)	**<.001 (0.003)**
	N =	80	80	80
Total Dissolved Phosphorus	r_s =	0.113	-0.143	-0.169
	p-value =	0.317 (1.00)	0.207 (1.00)	0.135 (1.00)
	N =	80	80	80
Total Phosphorus	r_s =	-0.017	-0.0417	-0.037
	p-value =	0.878 (1.00)	0.714 (1.00)	0.742 (1.00)
	N =	80	80	80
Dissolved O_2	r_s =	-0.064	0.228	0.047
	p-value =	0.572 (1.00)	0.042 (1.000)	0.678 (1.00)
	N =	80	80	80
% O_2 Saturation	r_s =	0.099	0.113	-0.101
	p-value =	0.380 (1.00)	0.319 (1.00)	0.373 (1.00)
	N =	80	80	80
Temperature	r_s =	0.235	-0.170	-0.207
	p-value =	0.036 (0.999)	0.132 (1.00)	0.066 (1.00)
	N =	80	80	80
Transparency	r_s =	-0.069	0.194	0.031
	p-value =	0.544 (1.00)	0.085 (1.00)	0.788 (1.00)
	N =	80	80	80
pH	r_s =	0.272	-0.051	-0.217
	p-value =	0.018 (0.513)	0.661 (1.00)	0.060 (1.00)
	N =	76	76	76
Conductivity	r_s =	0.544	-0.093	-0.543
	p-value =	**<.001 (0.003)**	0.429 (1.00)	**<.001 (0.003)**
	N =	75	75	75

Stepdown Bonferroni adjusted p-values are in parentheses.

Pearson correlations were also calculated to evaluate the sensitivity of the macroinvertebrate and fish assemblage metrics to specific environmental stress factors such as high percentages of cropland in the watersheds, in-stream or riparian physical habitat degradation, or chemical pollutants. The percentage of EPT macroinvertebrate genera present, HBI score, percentage of 'tolerant' fish individuals present, and fish IBI score were compared with 9 in-stream habitat measures, 3 watershed-scale land use measures, and 11 water quality and water chemistry measures (Tables 3, 4). Macroinvertebrate measures showed stronger correlations with physical habitat data than the fish measures, based on the Pearson correlations presented in Table 3, and statistically significant correlations were observed between macroinvertebrate measures and 7 of the 11 water quality and water chemistry measures when Stepdown Bonferroni adjustments were applied (Table 4).

Table 3. Pearson correlations between in-stream physical habitat, watershed land use, and biological measures.

Habitat Parameters		% EPT Genera		HBI Score		% Tolerant Fish Individuals		IBI Score	
% Sand, silt,	r_s =		-0.247		0.177		0.055		-0.062
and clay	p-value* =	0.030	(1.00)	0.127	(1.00)	0.676	(1.00)	0.640	(1.00)
sediments	N =		77		76		61		60
Bank	r_s =		-0.360		0.261		0.372		-0.198
erosion	p-value =	0.001	(0.062)	0.023	(0.866)	0.003	(0.150)	0.130	(1.00)
	N =		77		76		61		60
Depth of	r_s =		-0.126		0.281		0.036		-0.104
fine	p-value =	0.275	(1.00)	0.014	(0.592)	0.768	(1.00)	0.430	(1.00)
substrate	N =		77		76		61		60
Sinuosity	r_s =		-0.227		0.112		0.281		-0.256
	p-value =	0.048	(1.00)	0.336	(1.00)	0.028	(1.00)	0.049	(1.00)
	N =		77		76		61		60
Distance	r_s =		-0.284		0.144		0.130		-0.119
between	p-value =	0.012	(0.533)	0.214	(1.00)	0.319	(1.00)	0.365	(1.00)
bends	N =		77		76		61		60
Riffle	r_s =		0.232		-0.332		-0.181		0.204
length	p-value =	0.043	(1.00)	0.003	(0.150)	0.163	(1.00)	0.117	(1.00)
	N =		76		76		61		60
Pool	r_s =		0.052		0.012		0.007		0.016
length	p-value =	0.656	(1.00)	0.915	(1.00)	0.958	(1.00)	0.902	(1.00)
	N =		77		76		61		60
Riparian	r_s =		0.271		-0.168		-0.273		0.328
buffer	p-value =	0.017	(0.688)	0.146	(1.00)	0.033	(1.00)	0.011	(0.477)
	N =		77		76		61		60
Width:Depth	r_s =		0.191		-0.249		-0.241		0.120
ratio	p-value =	0.096	(1.00)	0.030	(1.00)	0.061	(1.00)	0.362	(1.00)
	N =		77		76		61		60
%	r_s =		-0.259		0.283		0.071		-0.265
Agriculture	p-value =	0.020	(0.796)	0.012	(0.533)	0.589	(1.00)	0.041	(1.00)
	N =		80		79		61		60
% Grassland	r_s =		0.209		-0.274		0.012		0.198
	p-value =	0.063	(1.00)	0.015	(0.599)	0.928	(1.00)	0.129	(1.00)
	N =		80		79		61		60
% Forest	r_s =		0.232		-0.254		-0.062		0.205
	p-value =	0.038	(1.00)	0.024	(0.881)	0.635	(1.00)	0.116	(1.00)
	N =		80		79		61		60

*Stepdown Bonferroni adjusted p-values are in parentheses.

Table 4. Pearson correlations between water quality and water chemistry measures, and biological measures. P-values in bold are significant when a Stepdown Bonferroni adjustment is applied.

Water Quality and Chemistry		% EPT Genera		HBI Score		% Tolerant Fish Individuals		IBI Score	
N-Kjeldahl	r_s =		-0.491		0.564		0.263		-0.374
	p-value* =	**<.001**	**(0.004)**	**<.001**	**(0.004)**	0.041	(0.851)	0.003	(0.099)
	N =		79		78		61		60
NH_3	r_s =		-0.429		0.537		0.232		-0.348
	p-value =	**<.001**	**(0.004)**	**<.001**	**(0.004)**	0.073	(1.00)	0.007	(0.176)
	N =		79		78		61		60
NO_3NO_2	r_s =		-0.011		0.101		-0.259		-0.001
	p-value =	0.921	(1.00)	0.377	(1.00)	0.044	(0.874)	0.996	(1.00)
	N =		79		78		61		60
Total Dissolved P	r_s =		-0.346		0.418		-0.059		-0.241
	p-value =	0.002	(0.058)	**<.001**	**(0.004)**	0.651	(1.00)	0.064	(1.00)
	N =		79		78		61		60
Total P	r_s =		-0.361		0.388		0.055		-0.318
	p-value =	**<.001**	**(0.004)**	0.005	(0.143)	0.674	(1.00)	0.013	(0.335)
	N =		78		78		61		60
Dissolved O_2	r_s =		0.312		-0.487		-0.357		0.332
	p-value =	0.005	(0.143)	**<.001**	**(0.004)**	0.005	(0.143)	0.010	(0.250)
	N =		79		78		61		60
% O_2 Saturation	r_s =		0.150		-0.397		-0.216		0.172
	p-value =	0.187	(1.00)	0.003	(0.099)	0.094	(1.00)	0.190	(1.00)
	N =		79		78		61		60
Temperature	r_s =		-0.350		0.215		0.280		-0.316
	p-value =	**0.002**	**(0.007)**	0.058	(1.00)	0.030	(0.669)	0.014	(0.335)
	N =		79		78		61		60
Transparency	r_s =		0.435		-0.413		-0.267		0.373
	p-value =	**<.001**	**(0.004)**	**0.002**	**(0.007)**	0.038	(0.834)	0.003	(0.099)
	N =		79		78		61		60
pH	r_s =		-0.218		-0.036		0.091		-1.071
	p-value =	0.060	(1.00)	0.763	(1.00)	0.491	(1.00)	0.594	(1.00)
	N =		75		74		60		59
Conductivity	r_s =		-0.138		-0.018		0.157		-0.159
	p-value =	0.243	(1.00)	0.882	(1.00)	0.235	(1.00)	0.232	(1.00)
	N =		74		73		59		58

*Stepdown Bonferroni adjusted p-values are in parentheses.

6.3 Random, Modified-Random, and Reference Site Comparisons

Stream Physical Habitat Measures

Each of the 9 physical habitat measures collected at the X sites was compared with its associated B site value using Pearson correlations and bivariate scatterplots (Figure 2). Large correlation coefficients (r) and significant p-values were observed for 8 of the 9 habitat variables. The only physical habitat variable that did not show a significant correlation between the X and B sites was mean total sum of pool habitat length.

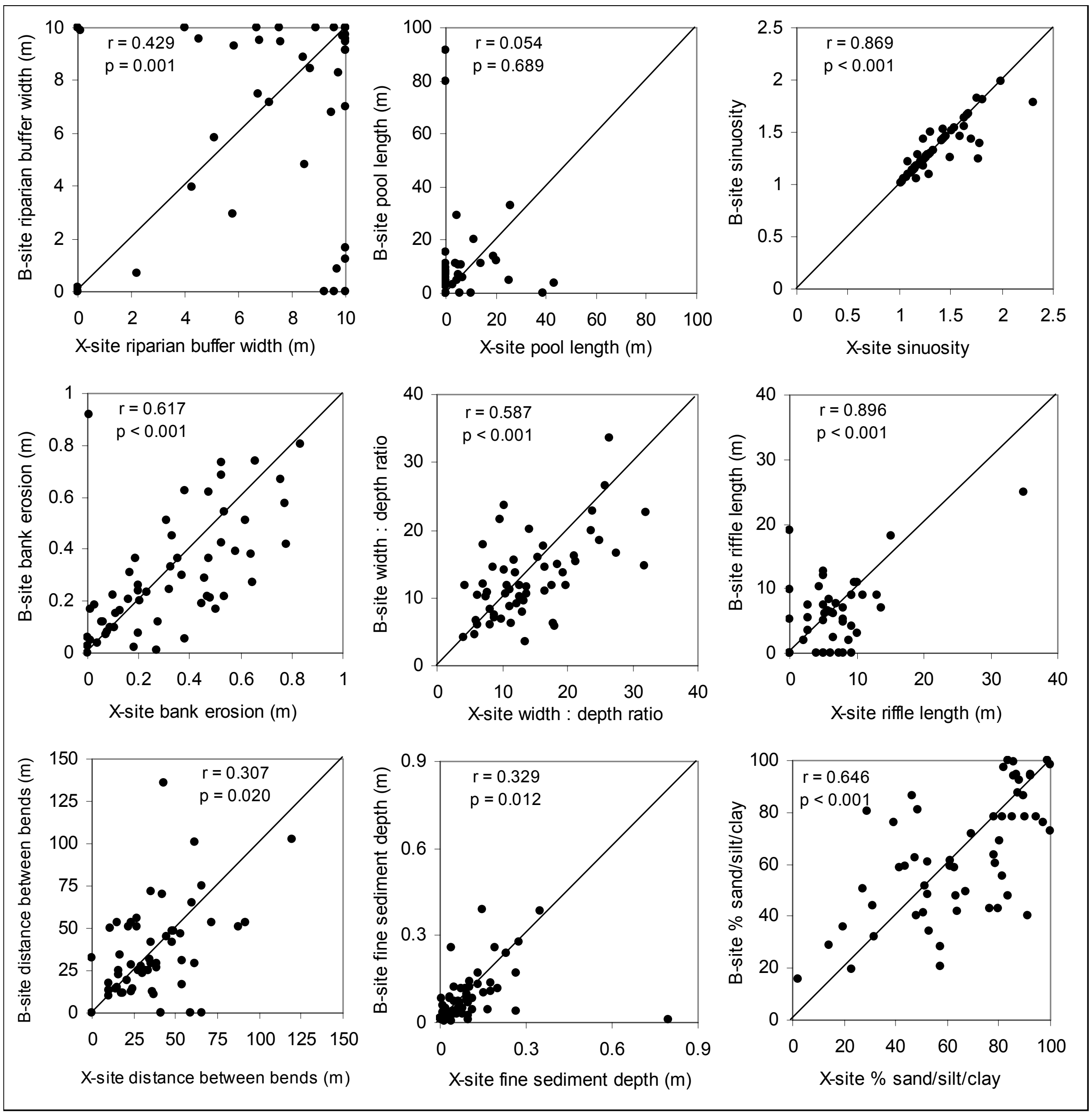

Figure 2. Bivariate scatterplots comparing physical habitat measurements collected at the X and associated B sites.

Weighted paired t-tests were used to determine if significant differences existed in physical habitat data between the X and B sites. The weighted paired t-tests did not produce statistically significant p-values for any of the 9 physical habitat measures, indicating that there were no significant differences between the X and B sites for these variables (Table 5).

Table 5. Weighted paired t-test comparisons of stream physical habitat measures from X and B stream assessment sites.

X vs B Site Physical Habitat Measures (n = 57)	t-value*	p-value
Mean Riparian Buffer Width (m)	1.68	0.0981
Mean Length of Pools (m)	-1.61	0.1125
Sinuosity	1.84	0.0712
Mean Width of Bank Erosion (m)	0.36	0.7218
Width : Depth Ratio	-1.64	0.1060
Mean Length of Riffles (m)	1.74	0.0878
Mean Distance Between Bends (m)	1.85	0.0697
Mean Depth of Fines (m)	-0.52	0.6081
Percent Sand, Silt, and/or Clay Sediments	-0.05	0.9609

*Degrees of freedom = 56.

Water Chemistry and Water Quality Measures

Each of the 11 water quality and water chemistry measures collected at the X sites was compared with its associated B site value using Pearson correlations and bivariate scatterplots (Figure 3).

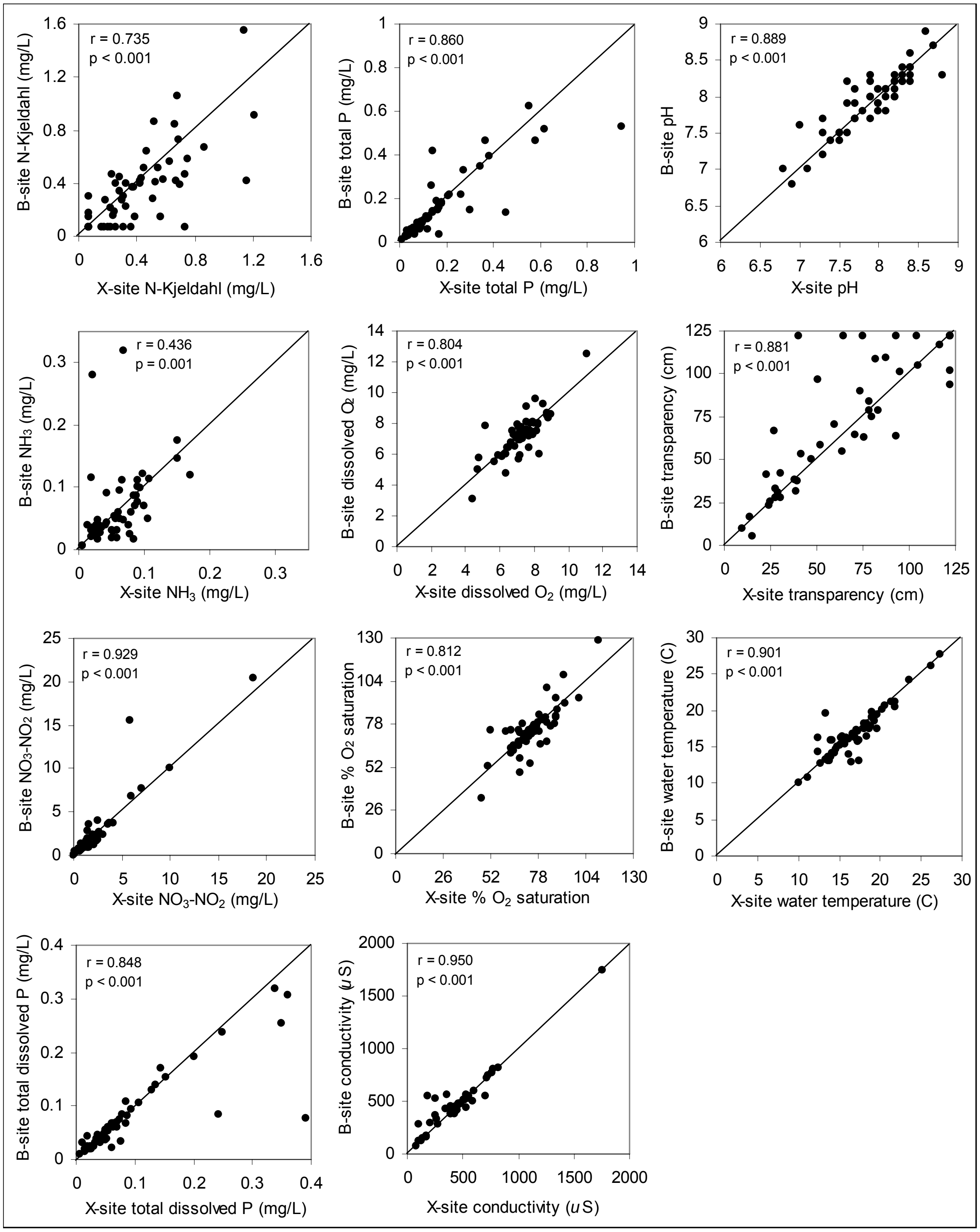

Figure 3. Bivariate scatterplots comparing water quality and water chemistry measurements collected from the X and associated B sites.

Large correlation coefficients (r) with significant p-values were observed for all 11 water quality and chemistry variables (Figure 3), showing a strong linear relationship between X site and B site data. The results of weighted paired t-tests evaluating water quality and water chemistry measures show no significant differences between the X and B site data for all 11 variables (Table 6).

Table 6. Weighted paired t-test comparisons of water quality and water chemistry measures from the X and associated B assessment sites.

X vs B Site Water Chemistry Measures (n = 57)	**t-value***	**p-value**
Total Phosphorus (mg/L)	1.69	0.0965
Total Dissolved Phosphorus (mg/L)	2.88	0.0056 (0.0617)†
NH_3 (mg/L)	-0.08	0.9344
NO_3-NO_2 (mg/L)	-2.28	0.0264 (0.2904)
N-Kjeldahl (mg/L)	1.09	0.2809
Dissolved Oxygen (mg/L)	1.04	0.3010
Percent Dissolved Oxygen Saturation	1.06	0.2946
pH (df=55)	-1.89	0.0644
Conductivity (uS/cm) (df=52)	-1.26	0.2140
Water Temperature (C)	0.21	0.8377
Transparency (cm)	-1.53	0.1319

*Degrees of freedom = 56, unless otherwise noted.
†Value in parentheses is the Bonferroni-adjusted p-value.

Macroinvertebrate Metrics

Figure 4 shows the relationships between the 7 macroinvertebrate metrics collected at each X and associated B site. Pearson correlation coefficients and bivariate scatterplots show significant p-values and correlations (r) between X and B site data for all 7 macroinvertebrate metrics, indicating strong linear relationships between the X and B sites.

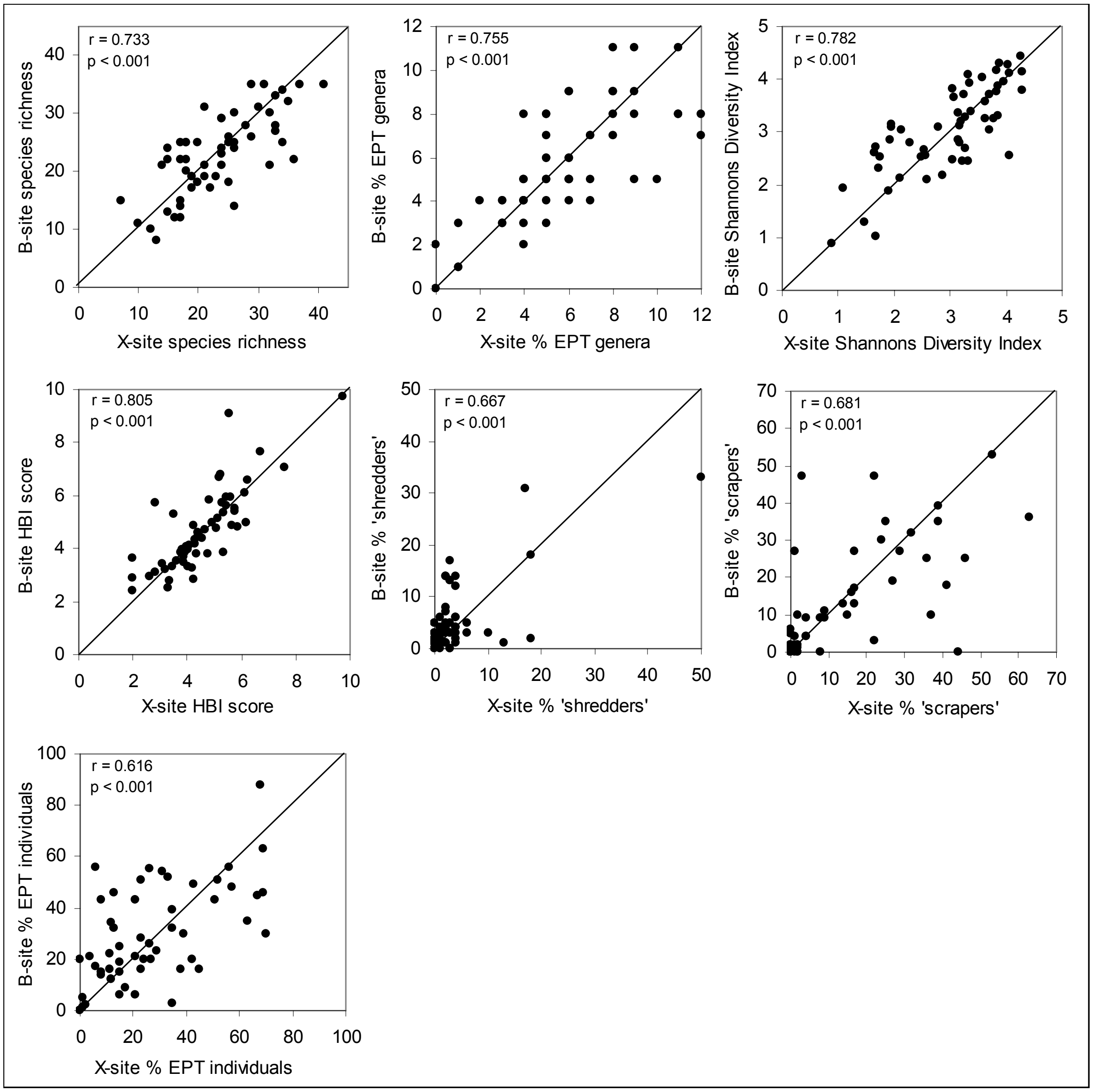

Figure 4. Bivariate scatterplots comparing macroinvertebrate metrics collected at the X and associated B sites.

Weighted paired t-tests were used to compare macroinvertebrate samples collected at the X and associated B sites. The results show no significant differences between these paired sites for any of the 7 metrics (Table 7).

Table 7. Weighted paired t-test comparisons of macroinvertebrate metrics and indices collected from X and associated B assessment sites.

X vs B Site Macroinvertebrate Metrics (n = 56)	t-value*	p-value
Species richness	0.42	0.6739
Percent EPT genera present	0.02	0.9865
Shannon's Diversity Index score	-0.46	0.6473
Hilsenhoff's Biotic Index score	-0.69	0.4962
Percent Shredders	-1.23	0.2244
Percent Scrapers	0.33	0.7406
Percent EPT individuals present	0.12	0.9056

*Degrees of freedom = 55.

Fish Assemblage Metrics

Bivariate scatterplots and Pearson correlation coefficients show that significant correlations exist between X and B site data for all 8 fish metrics analyzed (Figure 5).

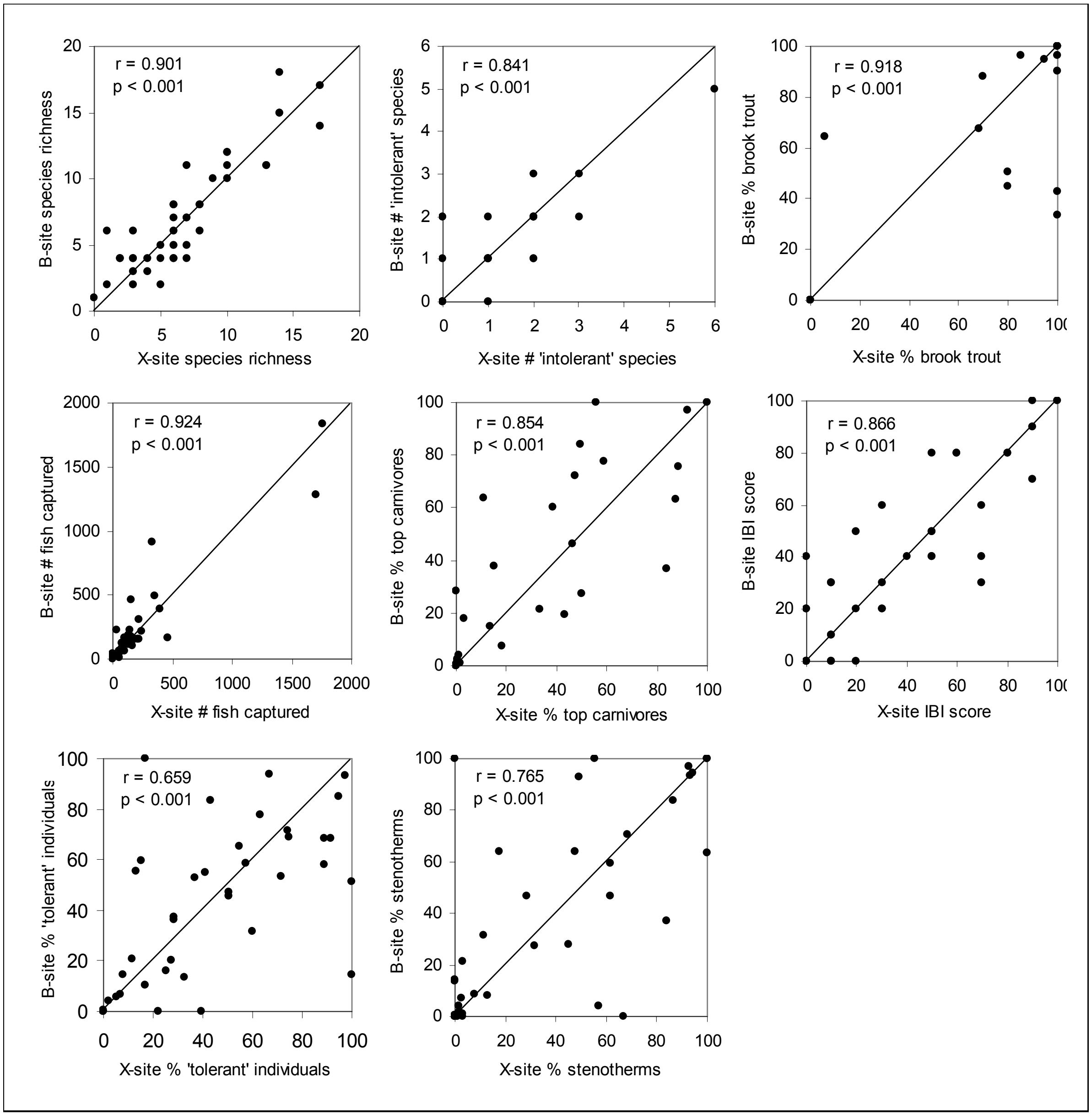

Figure 5. Bivariate scatterplots comparing fish metrics collected at the X and associated B sites.

Weighted paired t-tests were used to determine if significant differences existed between the X and B sites. Of the 58 streams sampled, 42 streams had sufficient numbers of fish at both the X and associated B sites with which to make paired-site t-test comparisons. The results of the weighted paired t-tests show no significant differences between X and B sites (Table 8).

Table 8. Weighted paired t-test comparisons of fish metrics collected from X and B assessment sites.

X vs B Site Fish Community Metrics (n = 42)	**t-value***	**p-value**
Species Richness	-2.17	0.0358 (0.2509)†
Number of Intolerant Species	0.06	0.9553
Percent of Salmonid Individuals that are Brook Trout	1.08	0.2883
Percent of Individual that are Top Carnivore Species	-1.58	0.1219
Coldwater IBI Score	-0.77	0.4471
Percent of Individuals that are 'Tolerant' Species	-0.34	0.7368
Percent of Individuals that are Stenothermal Cool / Coldwater Species	0.20	0.8406

*Degrees of freedom = 41.

†Value in parentheses is the Bonferroni-adjusted p-value.

6.4 Comparison of Reference Sites with Random and Modified-Random Assessment Sites

Stream Physical Habitat Measures

Bar charts were used to visually compare data for 9 physical habitat measures collected at the X, B, and reference sites (Figure 6).

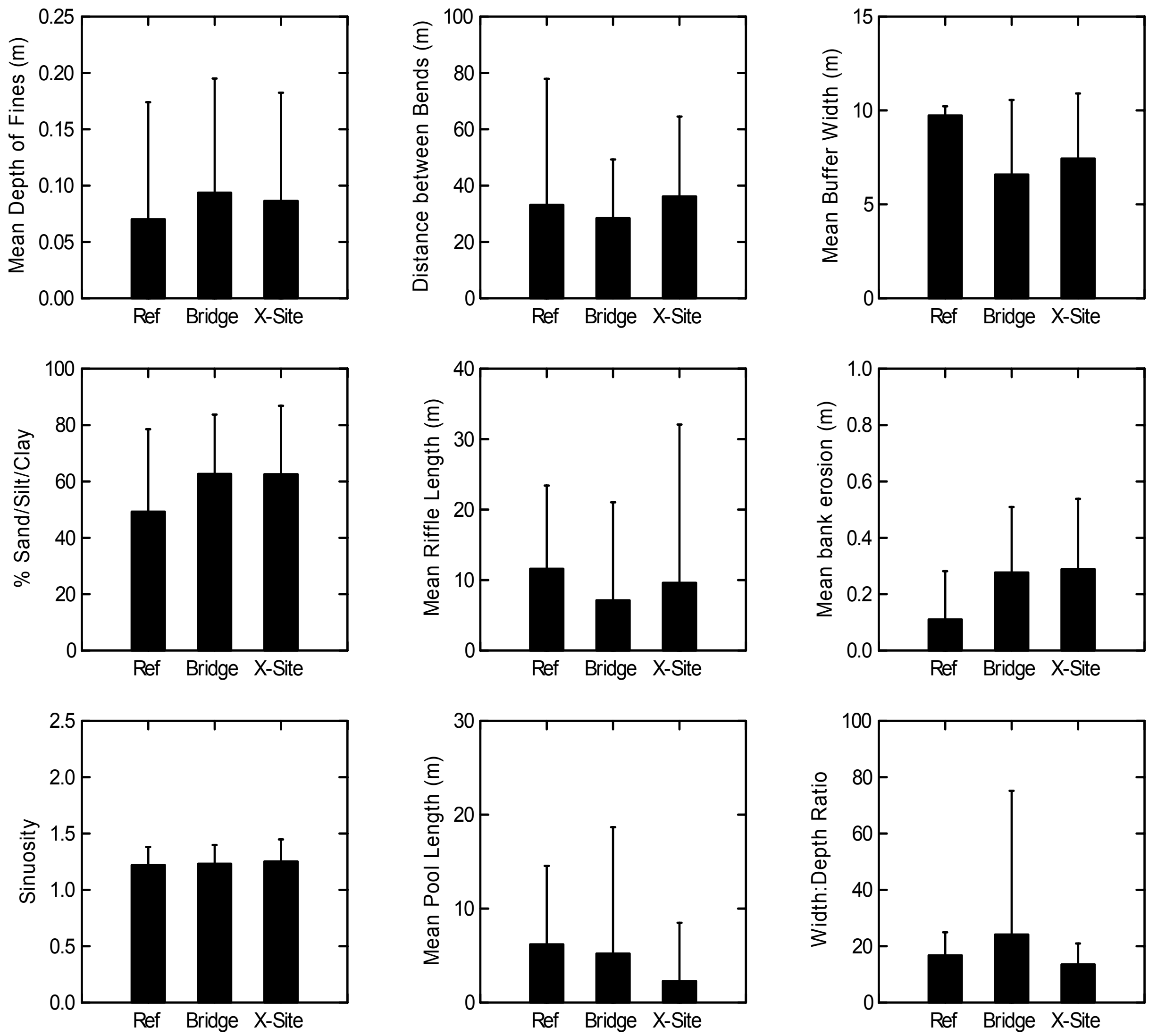

Figure 6. Bar charts of weighted means and standard deviations (SD) for X and B sites and simple random sample mean and SD for reference sites for physical habitat measures.

Two-sample t-tests were used to compare the 9 physical habitat measures collected at the X sites with data collected at the least disturbed reference sites. The mean width of stream bank erosion and the mean width of riparian buffer were the only measures that showed a significant difference between the X sites and the reference sites (Table 9).

Table 9. Two sample t-test comparisons for physical habitat measures collected at the X and reference sites. Tests are based on the difference (X site mean - reference mean) and use weighted values for X sites.

X vs Reference Site Habitat Measures	t-value	p-value*
Mean Depth of Fines (m)	0.631	0.5300
Mean Distance Between Bends (m)	0.287	0.7746
Mean Riparian Buffer Width (m)	-4.869	**<0.0001**
Percent Sand, Silt, and/or Clay Sediments	1.855	0.0675
Mean Length of Riffles (m) (df=75)	-0.504	0.6159
Mean Width of Bank Erosion (m)	3.570	**0.0006**
Sinuosity	0.742	0.4605
Mean Length of Pools (m) (df=75)	-1.913	0.0596
Width:Depth Ratio	-1.575	0.1195

*P-values in bold are significant ($p \leq 0.006$) when a Bonferroni correction is applied for an overall Type I error rate of 0.05. Degrees of freedom = 76, unless otherwise noted.

Water Quality and Water Chemistry Measures

Bar charts were used to visually compare data for 11 water chemistry and water quality measures among the X, B, and reference sites (Figure 7).

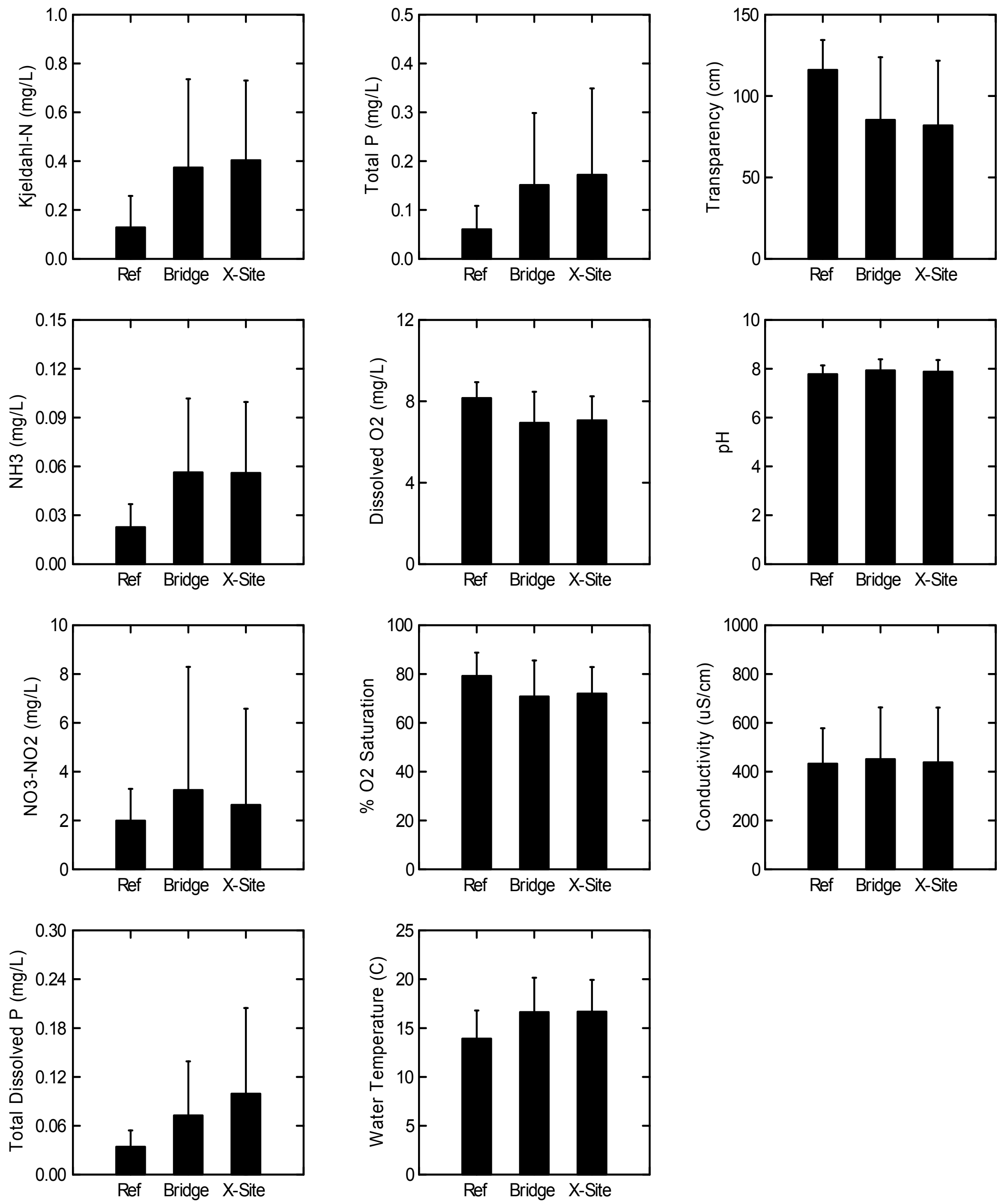

Figure 7. Bar charts of means and standard deviations for water quality and water chemistry measures collected at the X, B, and reference stream sites. Values for X and B sites are weighted.

Two-sample t-tests were used to compare the 6 in-situ water quality and 5 lab-analyzed grab-sample water chemistry measures collected at the X and reference sites. Eight of these 11 measures were found to be significantly different at the X sites versus the reference sites (Table 10).

Table 10. Two-sample t-test comparisons of water quality and water chemistry values between the X and reference stream sites. Tests are based on the difference (X site mean - reference mean) and use weighted values for X sites.

X vs Reference Site Water Chemistry Measures	t-value	p-value*
Kjeldahl-N (mg/L)	5.410	**<0.0001**
Total Phosphorus (mg/L)	4.376	**<0.0001**
Transparency (cm)	-5.243	**<0.0001**
NH_3 (mg/L)	5.137	**<0.0001**
Dissolved Oxygen (mg/L)	-4.850	**<0.0001**
pH (df = 74)	0.953	0.3436
NO_3-NO_2 (mg/L)	1.100	0.2748
Percent Dissolved Oxygen Saturation	-2.910	**0.0047**
Conductivity (uS/cm) (df = 73)	0.122	0.9035
Total Dissolved Phosphorus (mg/L)	4.470	**<0.0001**
Water Temperature (C)	3.745	**0.0003**

*P-values in bold are significant ($p \leq 0.005$) when a Bonferroni correction is applied for an overall Type I error rate of 0.05. Degrees of freedom = 78, unless otherwise noted.

Macroinvertebrate Metrics

Bar charts were used to visually compare data among X, B, and reference sites for macroinvertebrate metrics (Figure 8).

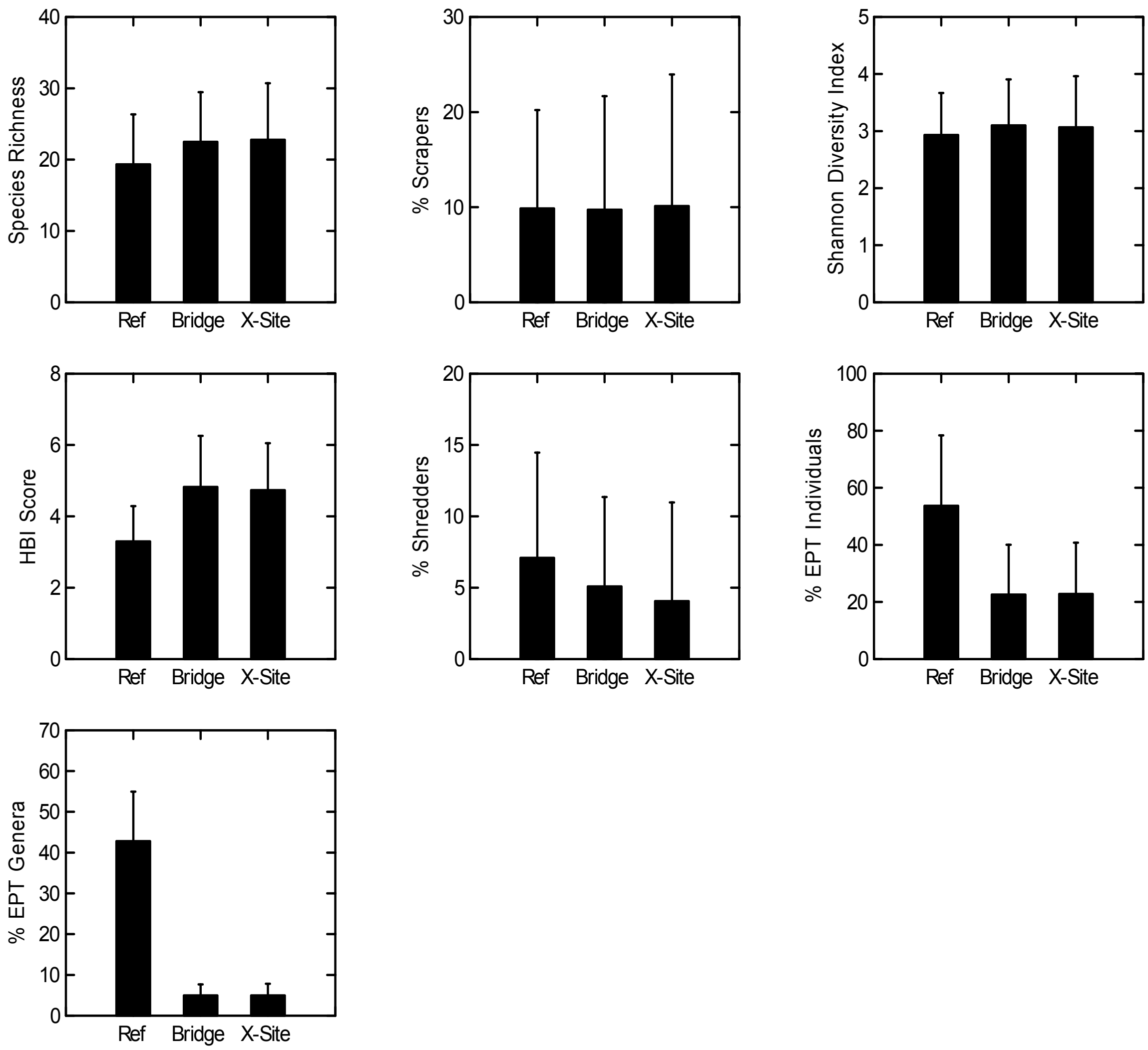

Figure 8. Bar charts of mean and standard deviation of macroinvertebrate metrics collected at the X, B, and reference stream sites. Values for X and B sites are weighted.

Two sample t-tests comparing 7 macroinvertebrate metrics for data collected at the X and reference sites show significant differences in HBI scores, and the percentage of genera and individuals in the sample that are Ephemeroptera, Plecoptera and Trichoptera (EPT) taxa. The lower HBI scores and the higher percentages of EPT genera or EPT individuals found at the reference sites relative to the X sites indicate that the reference streams have greater biological quality and a greater percentage of environmentally sensitive taxa than the X stream sites. No significant differences were detected between the X and reference stream populations for species richness, Shannon's Diversity Index, the percent 'shredders', and the percent 'scrapers' present (Table 11).

Table 11. Two sample t-test comparisons of the X and reference site macroinvertebrate metrics. Tests are based on the difference (X site mean - reference mean) and use weighted values for X sites.

X vs Reference Site Macroinvertebrate Metrics	t-value	p-value*
Species Richness	1.900	0.0611
Percent Scrapers	0.091	0.9279
Shannon's Diversity Index Score	0.691	0.4915
Hilsenhoff's Biotic Index Score (HBI)	5.307	**<0.0001**
Percent Shredders	-1.689	0.0953
Percent EPT Individuals	-5.439	**<0.0001**
Percent EPT Genera	-14.772	**<0.0001**

*P-values in bold are significant ($p \leq 0.007$) when a Bonferroni correction is applied for an overall Type I error rate of 0.05. Degrees of freedom = 77.

Fish Metrics

Bar charts were used to visually compare distributions of fish metrics among X, B, and reference sites (Figure 9).

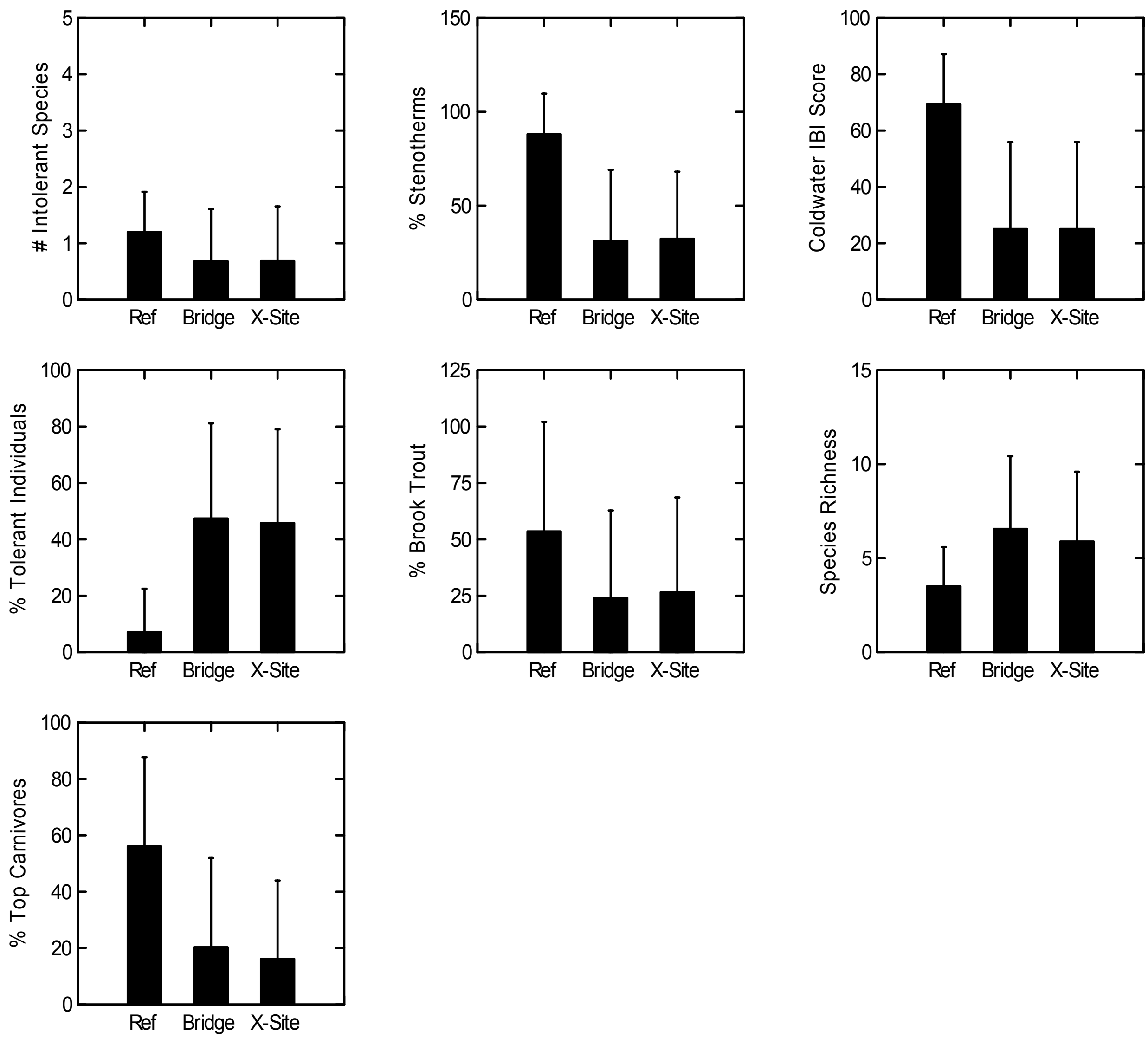

Figure 9. Bar charts of mean and standard deviation of fish metrics collected at the X, B, and reference stream sites. Values for X and B sites are weighted.

Fish IBI scores were significantly different between the X and reference sites, when compared using two-sample t-tests. The mean IBI score for the X sites population was 34 (narrative rating of "Fair"), versus 70 (narrative rating of "Good") for the reference condition. In addition, there were significant differences found in species richness, percentage of individuals captured that are considered "tolerant", percentage of individuals captured that are top carnivores, and the percentage of individuals that are cool/coldwater stenotherms (as defined by Lyons 1996). There were no significant differences observed between the X stream population and the reference condition for the number of intolerant species present or the percentage of salmonids captured that were brook trout. (Table 12)

Table 12. Two-sample t-test comparisons of the X and reference site fish metrics. Test is based on the difference (X site mean - reference mean).

X vs Reference Site Fish Metrics	t-value	p-value*
Number of Intolerant Species (df=59)	-2.319	0.0239
Percent of Individuals that are Stenothermal Cool or Coldwater Species	-7.427	**<0.0001**
Coldwater IBI score	-7.021	**<0.0001**
Percent of Individuals that are Tolerant Species	6.165	**<0.0001**
Percent of Salmonid Individuals that are Brook Trout (df=59)	-2.087	0.0412
Species Richness (df=59)	3.183	**0.0023**
Percent of Individuals that are Top Carnivore Species	-4.632	**<0.0001**

*P-values in bold are significant ($p \leq 0.007$) when a Bonferroni correction is applied for an overall Type I error rate of 0.05. Degrees of freedom = 58, unless otherwise noted.

6.5 Probability Estimates and Evaluation of Population Distributions

Physical Habitat Cumulative Distributions

Empirical CDFs were plotted to further compare differences between the X and B site data, and to estimate the percentage of the X and B sample populations that met reference condition threshold values. For each physical habitat measure, either the 25th or 75th percentile calculated from the least disturbed reference sites population was used as a reference condition threshold (Table 13). The upper 75th percentiles calculated from least disturbed reference sites were used as the thresholds for the percentage of sand, silt, and clay substrates, the mean depth of fine sediments, the width:depth ratio, the mean distance between bends, and the mean erosion width. The 25th percentile was used as the threshold for mean riparian buffer width.

The X and B site cumulative distribution functions and their confidence intervals overlap for all of the physical habitat measures analyzed (Figure 10), indicating the population estimates of these measures are not significantly different between the random and modified-random sites. The results of Kolmogorov-Smirnov tests (D_{max} values) also found no significant differences (all D_{max} p-values > 0.05).

Table 13. Physical habitat reference condition thresholds.

Habitat Measures	Reference Condition Threshold	Impairment Criteria
% sand, silt, and clay sediments	66.8%	> 66.8% sand, silt and clay
mean depth of fine sediment (m)	0.122	> 0.122m mean depth of fines
width : depth ratio	18.4	> 18.4 width : depth ratio
mean distance between bends (m)	42	> 42m distance between bends
mean riparian buffer width (m)	9.5	< 9.5m mean buffer width
mean bank erosion width (m)	0.104	> 0.104m mean bank erosion

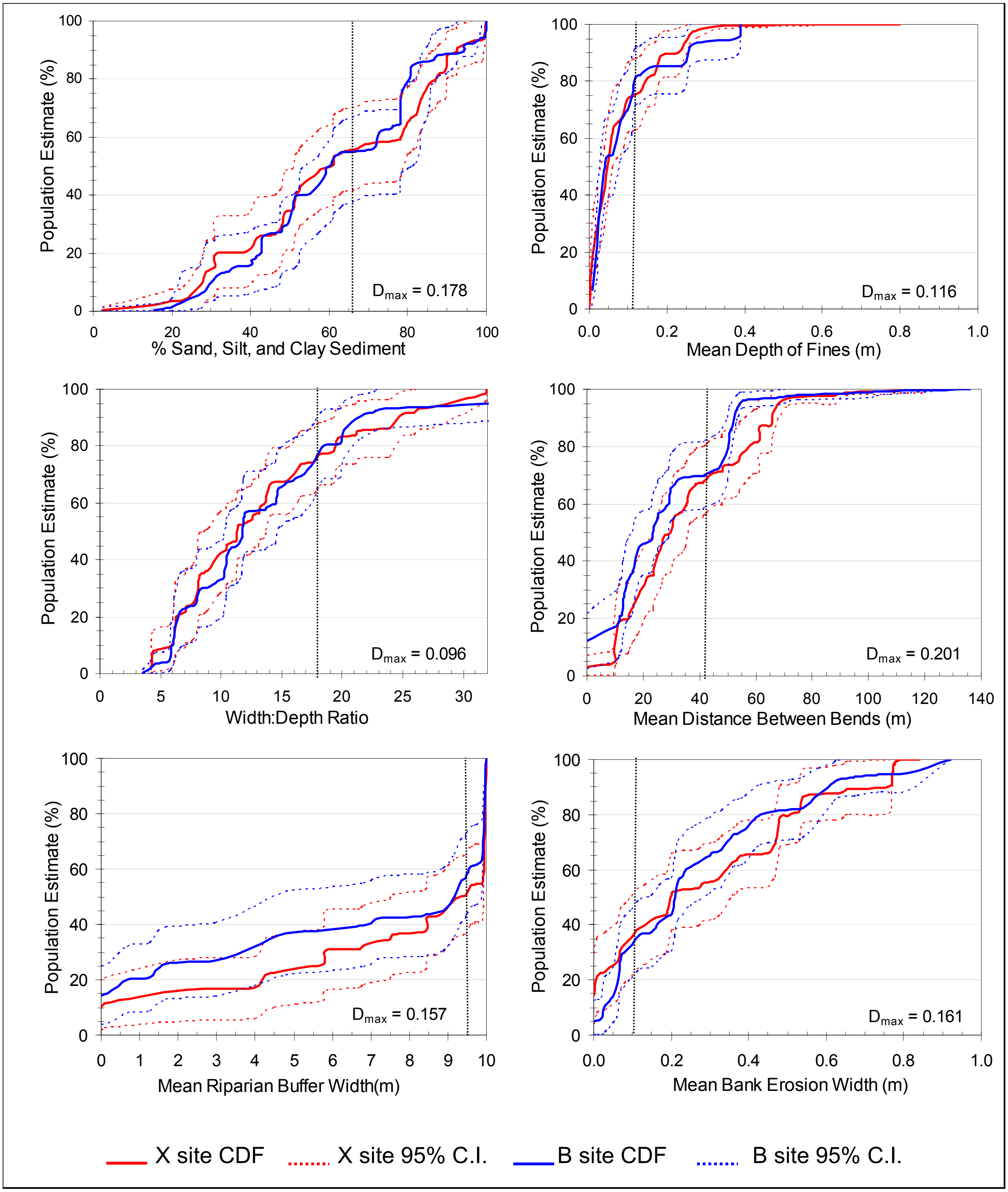

Figure 10. Cumulative distribution function curves of physical habitat measures collected at the X and B sites. The stippled lines represent 95% confidence intervals around the distribution plots. The vertical lines represent the reference condition threshold values.

Reference condition threshold values were applied to the CDF data to approximate the percent of the stream population (miles) that did not meet these criteria and therefore received a ‘poor’ rating (Figure 11). Overlapping error bars show no significant differences between the X site population and the B site population for all physical habitat measures presented in Figure 11.

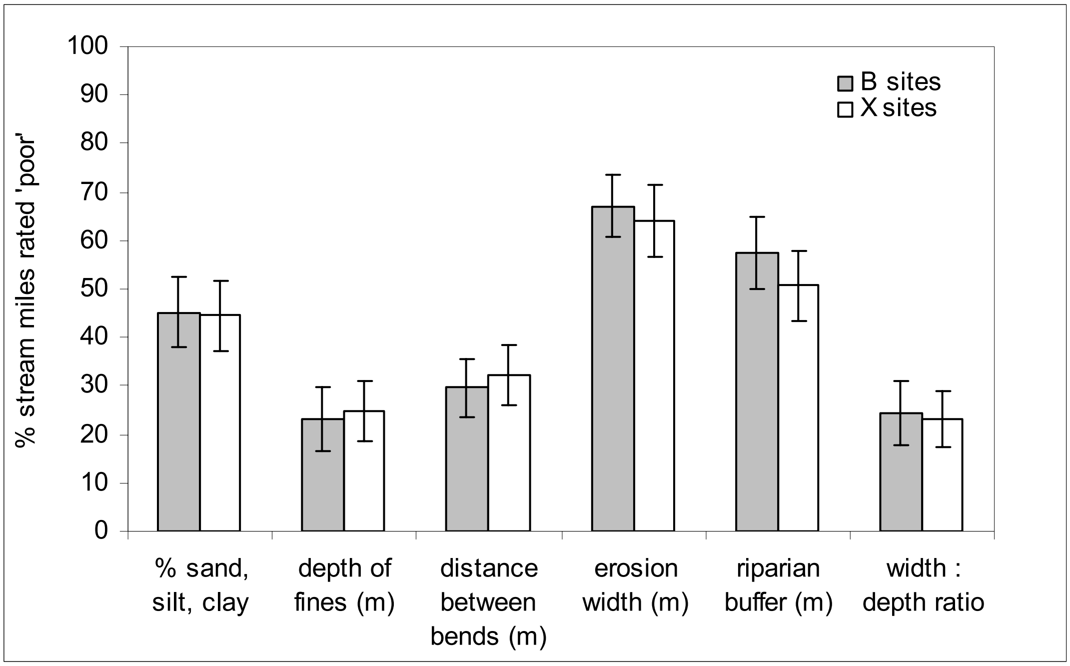

Figure 11. The percentages of stream miles not meeting the reference condition threshold values for physical habitat measures. The error bars equal 1 standard error.

Water Quality and Water Chemistry Cumulative Distributions

Cumulative distribution functions were plotted for 10 water quality and water chemistry measures. Either the 25th or 75th percentile of each water quality and water chemistry measure calculated from the least disturbed reference sites were used to set reference condition thresholds (Table 14). The 75th percentile calculated from the reference sites was used as the threshold for Kjeldahl-N, NH_3, NO_3-NO_2, total dissolved P, and total P. The 25th percentile was used as the threshold for the percentage dissolved oxygen saturation, dissolved oxygen concentration, and water clarity.

Table 14. Water chemistry and water quality reference condition thresholds.

Water Chemistry & Quality Measures	Reference Condition Threshold	Impairment Criteria
% Oxygen Saturation	73.3%	< 73.3% dissolved oxygen saturation
Transparency (cm)	122	< 122cm water transparency
Dissolved Oxygen (mg/L)	7.6	< 7.6 mg/L dissolved oxygen conc.
Kjeldahl-N (mg/L)	0.15	> 0.15 mg/L concentration
NH_3 (mg/L)	0.028	> 0.028 mg/L concentration
NO_3-NO_2 (mg/L)	2.63	> 2.63 mg/L concentration
Total Dissolved P (mg/L)	0.04	> 0.04 mg/L concentration
Total P (mg/L)	0.07	> 0.07 mg/L concentration

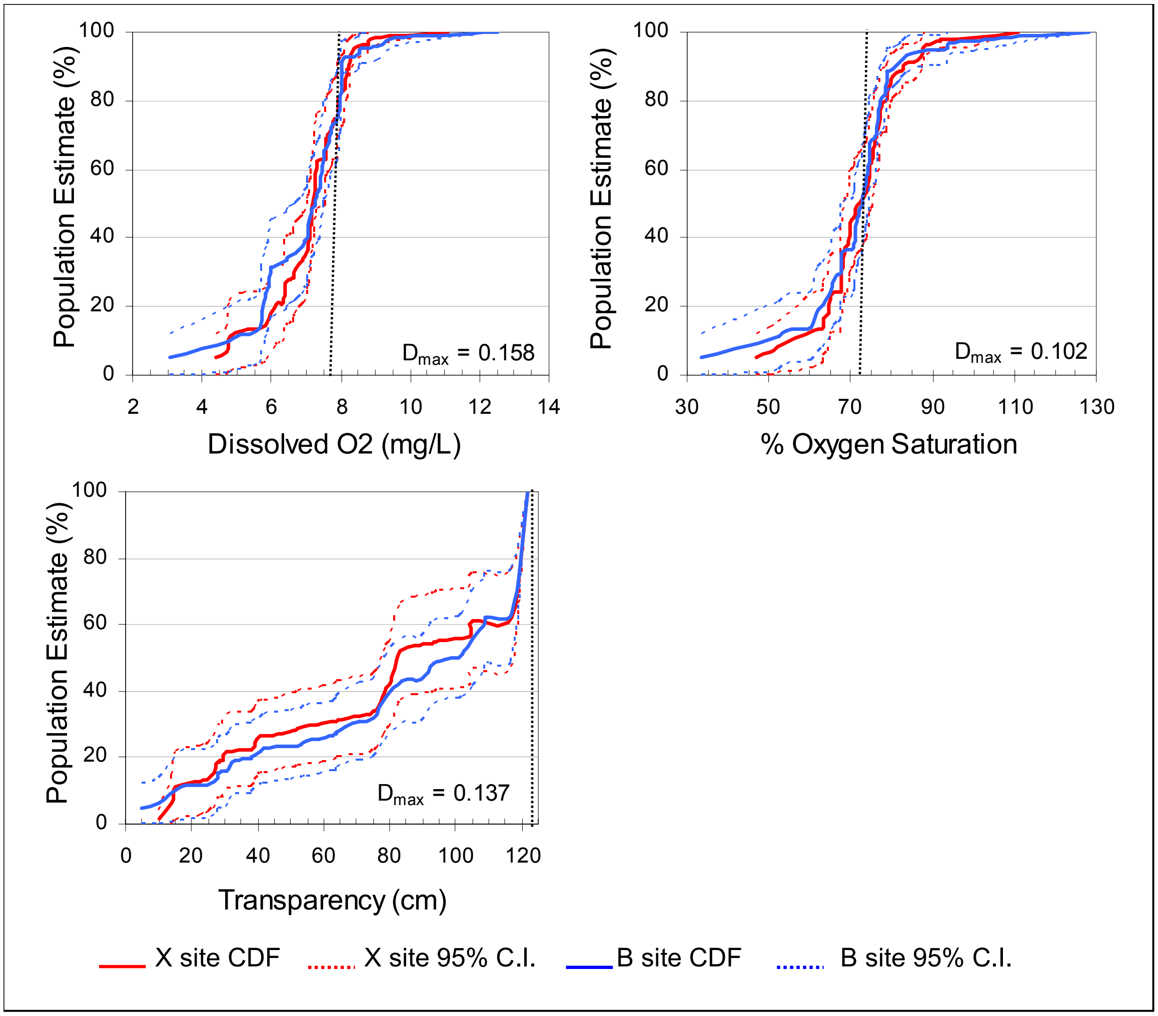

Figure 12. Cumulative distribution function curves of water quality measures collected *in situ* at the X and B sites. The stippled lines represent 95% confidence intervals around the distribution functions. The vertical lines represent the reference condition threshold values.

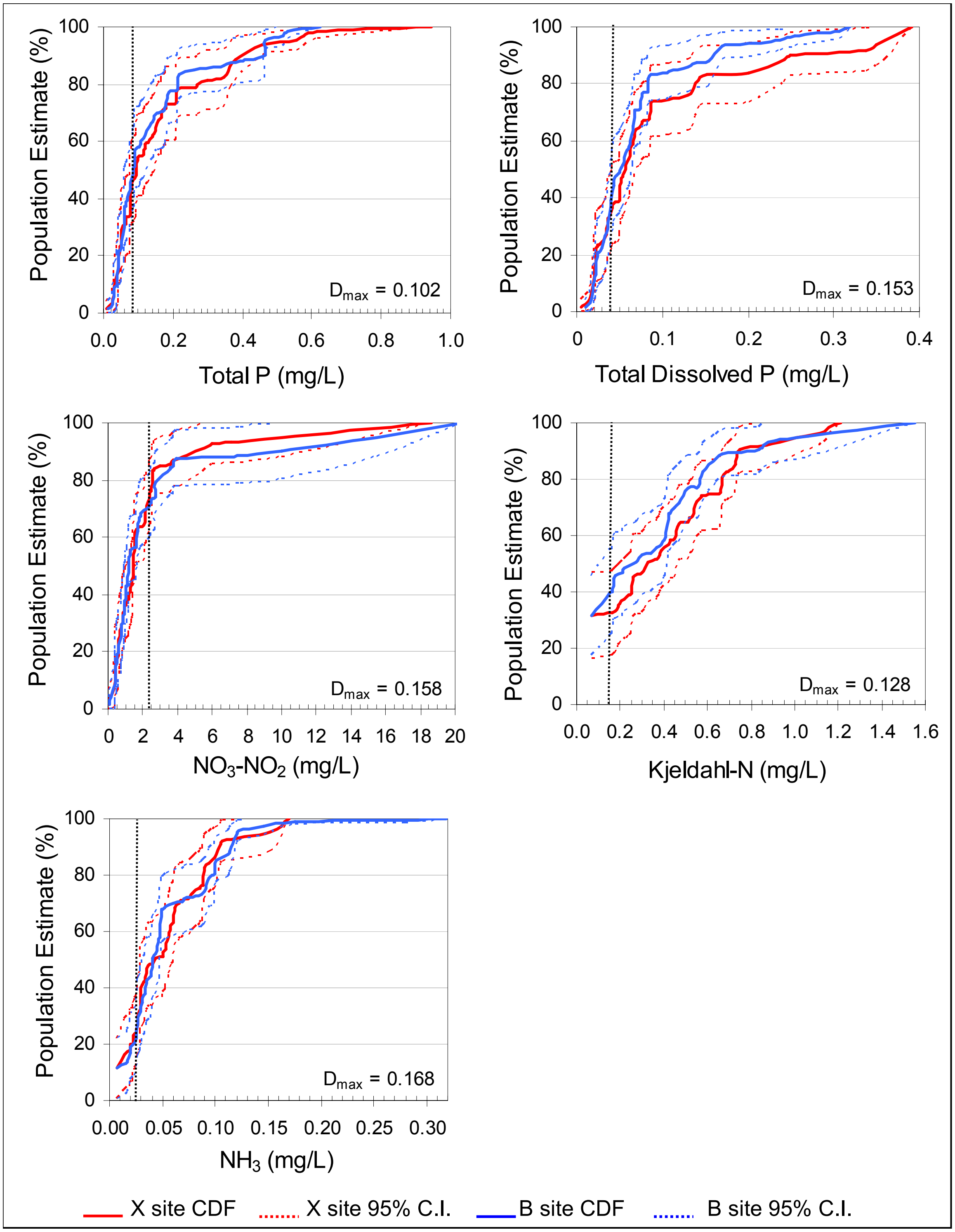

Figure 13. Cumulative distribution function curves of laboratory analyzed water chemistry measures collected at the X and B sites. The stippled lines represent 95% confidence intervals around the distribution plots. The vertical lines represent the reference condition threshold values.

All 10 water quality and water chemistry measurements have overlapping X and B site cumulative distribution function curves (Figures 12, 13), showing no significant differences between the X and B sampling sites. The results of Kolmogorov-Smirnov tests (D_{max} values) also found no significant differences (all D_{max} p-values > 0.05). Reference condition threshold values were applied to the CDF data and estimates of the percent of stream miles not meeting these criteria were determined (Figure 14). The overlapping error bars show similarities between the X site population and the B site population for the 8 water quality and water chemistry measures analyzed in Figure 14.

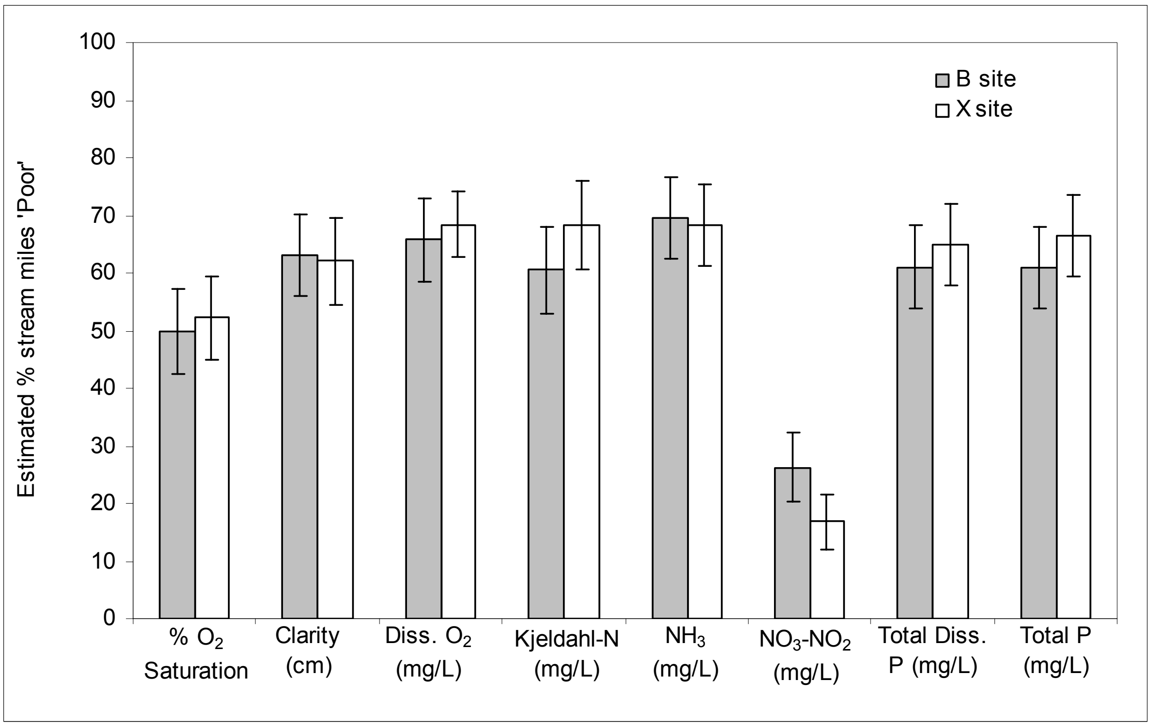

Figure 14. The percentages of stream miles not meeting the reference condition threshold values for water quality and water chemistry measures. The error bars equal 1 standard error.

Macroinvertebrate CDFs

Cumulative distribution function curves were created for 4 macroinvertebrate measures. The 75th percentile of reference site measurements was used as the threshold for the HBI scores. The 25th percentile was used as the threshold for species richness, the percentage of EPT genera and EPT individuals present. (Table 15)

Table 15. Macroinvertebrate metric reference condition thresholds.

Macroinvertebrate Metrics	Reference Condition Threshold	Impairment Criteria
HBI	3.92	HBI score > 3.92
Species Richness	16	< 16 species
% EPT Genera	35%	< 35% of genera are EPT taxa
% EPT Individuals	31%	< 31% of individuals are EPT taxa

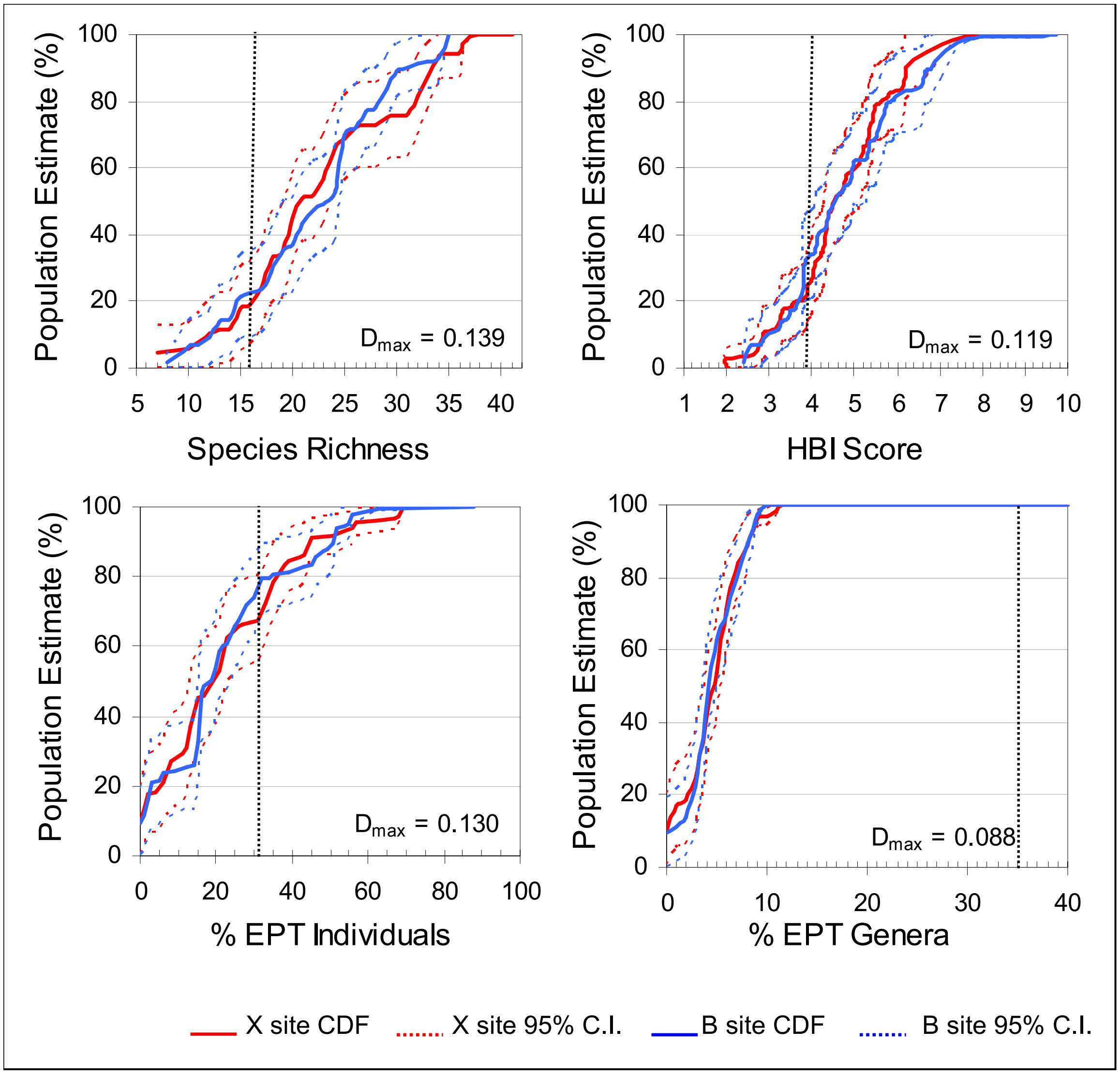

Figure 15. Cumulative distribution function curves of macroinvertebrate metrics collected at the X and B sites. The stippled lines represent the 95% confidence intervals around the cumulative distribution estimates. The vertical lines represent the reference condition threshold values.

The X site and B site data produced overlapping CDFs and confidence intervals for species richness, HBI score, the percentage of EPT genera and the percentage of EPT individuals (Figure 15). The results of Kolmogorov-Smirnov tests (D_{max} values) also found no significant differences (all D_{max} p-values > 0.05). Reference condition threshold values were applied to the CDF data and estimates of the percent of stream miles not meeting these criteria were determined (Figure 16). The overlapping error bars show no statistical differences between the X site population and the B site population for all macroinvertebrate metrics in Figure 16. The reference condition threshold value for the percentage of EPT genera present in a sample was 35 percent, and no sites in the sample populations met this criteria.

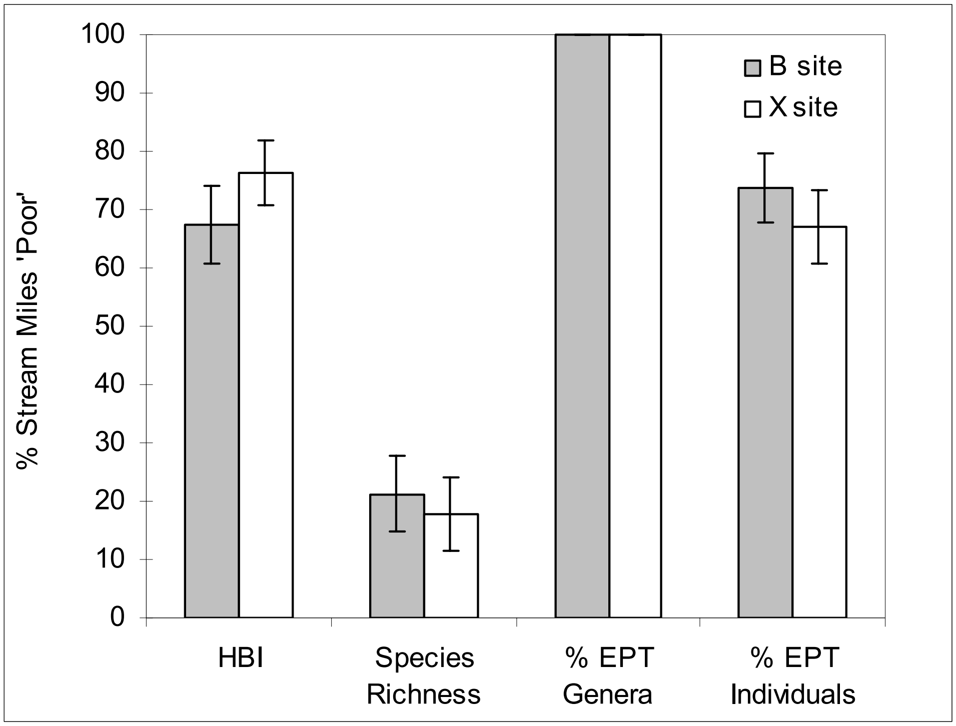

Figure 16. The percentages of stream miles not meeting the reference condition threshold values for macroinvertebrate metrics. The error bars equal 1 standard error.

Fish Cumulative Distributions

Cumulative distribution functions were plotted for the 4 fish metrics (Figure. 17). The 75th percentile calculated from the reference site measurements was used as the threshold for the percentage of 'tolerant' individuals. The 25th percentile was used as the threshold for the percentage of stenothermal individuals present, the percentage of top carnivore individuals present, and the fish IBI score. (Table 16)

Table 16. Fish metric reference condition thresholds.

Fish Metrics	**Reference Condition Threshold**	**Impairment Criteria**
% Top Carnivore Individuals	24.4%	Any site where < 24.4% of individuals were not top predators
% Stenothermal Individuals	83.5%	Any site where < 83.5% of individuals were not stenotherms
% 'Tolerant' Individuals	3.40%	Any site where > 3.40% of individuals are tolerant species
IBI Score	60	Any site where fish IBI score is < 60

The CDF results show significant overlap between the X site and B site data, indicating that X and B site data provide equivalent population estimates for fish metrics. The results of Kolmogorov-Smirnov tests (D_{max} values) also found no significant differences (all D_{max} p-values > 0.05).

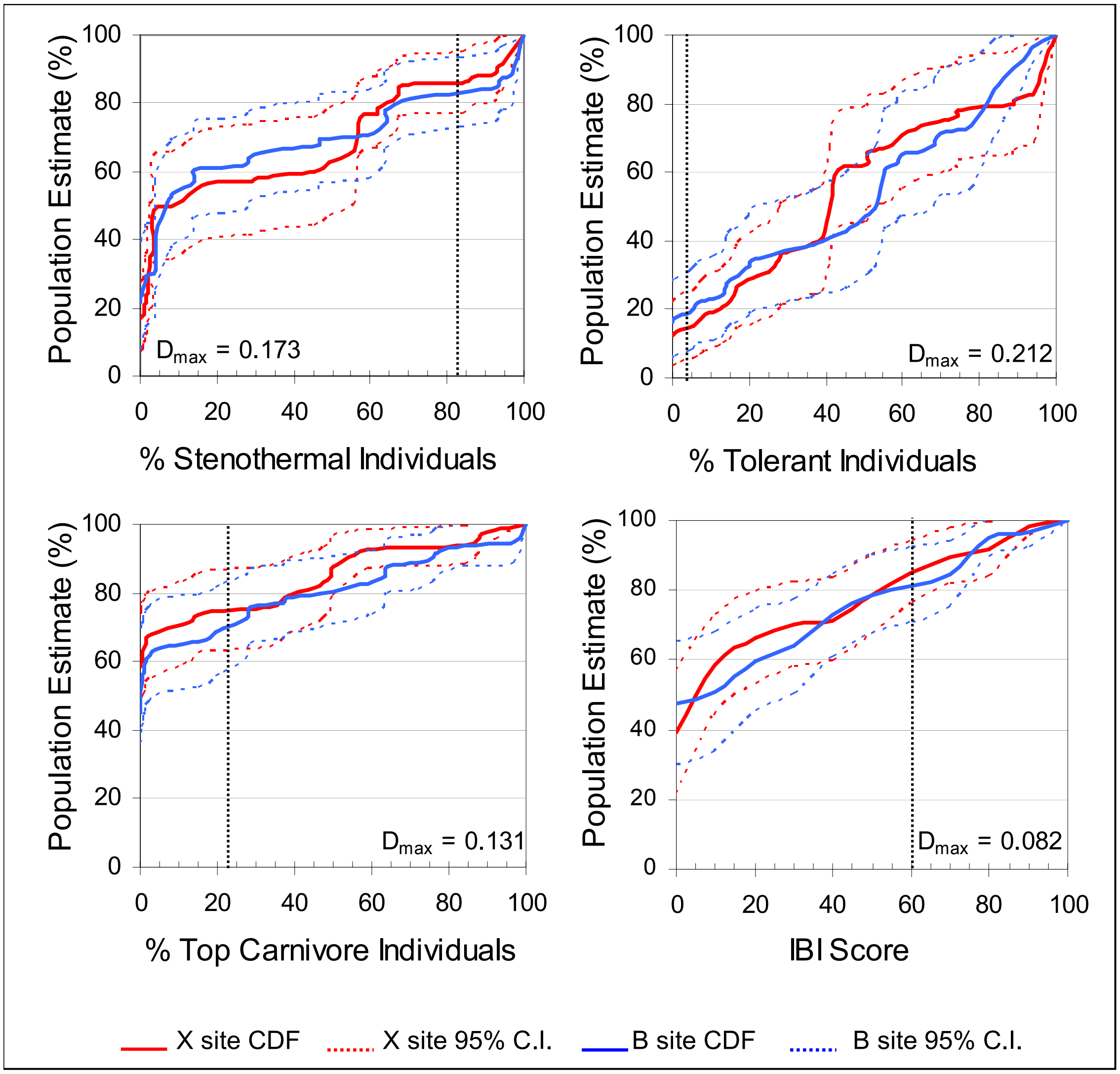

Figure 17. Cumulative distribution function curves for fish metrics collected at the X and B sites. The stippled lines represent 95% confidence intervals around the distribution plots. The vertical lines represent the reference condition values.

Fish metric reference condition threshold values were applied to the CDF data and estimates of the percent of stream miles not meeting these criteria were determined (Figure 18). The overlapping error bars show no statistical differences between the X site population and the B site population for all fish metrics.

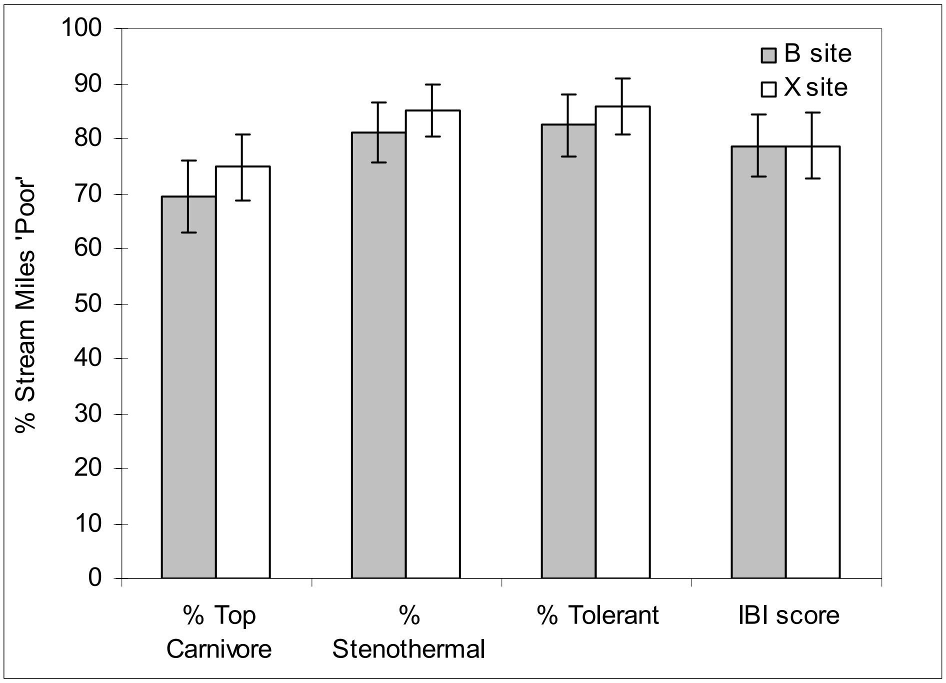

Figure 18. The percentages of stream miles not meeting the reference condition threshold values for fish metrics. The error bars equal 1 standard error.

6.6 Targeted vs. Random Stream Assessment

Fish Index of Biotic Integrity Cumulative Distribution Function Plots of Random and Regional Biologist's Targeted Sampling Sites

Cumulative distribution function curves and box plots show differing distributions of fish IBI scores between the REMAP X and WDNR regional biologist's targeted sampling sites (Figure. 19). The cumulative distribution functions show that 80 percent of the REMAP random sample population (using weighted data) and 65 percent of the WDNR targeted stream population received an IBI score of 60 or less, thereby not meeting the reference condition threshold for fish IBI score.

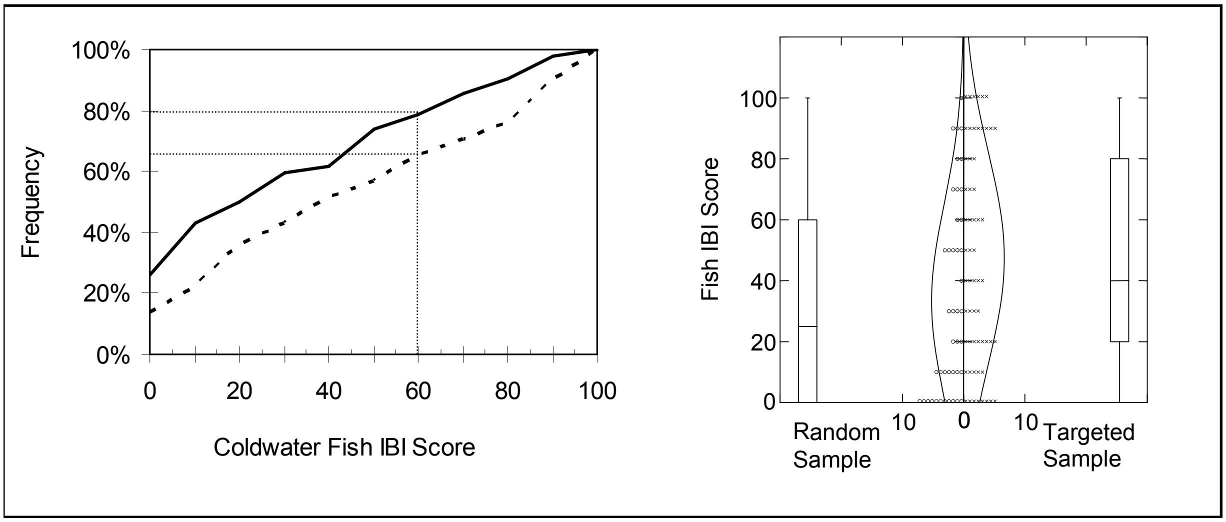

Figure 19. Cumulative distribution function curves of fish IBI scores from biologists' targeted sites (solid bold line) and REMAP X sites (stippled bold line), compared to the reference condition threshold value (stippled vertical line). Box plots and distribution of the biologist's targeted and REMAP X sites data is shown in figure on right.

6.7 Effects of Geographical Distances Between Random and Modified-Random Sites

Excluding 3 sites where the X and B reaches overlapped, the average distance from the X sites to the B sites was 701 meters, with a minimum distance of 106 meters and a maximum distance of 2,283 meters. We investigated the relationships between the distance from X to B sampling sites for the following measures: width:depth ratio, percentage of fine substrate, percentage of EPT genera present, HBI score, number of fish captured, IBI score, fish species richness, dissolved total phosphorus, and NH_3.

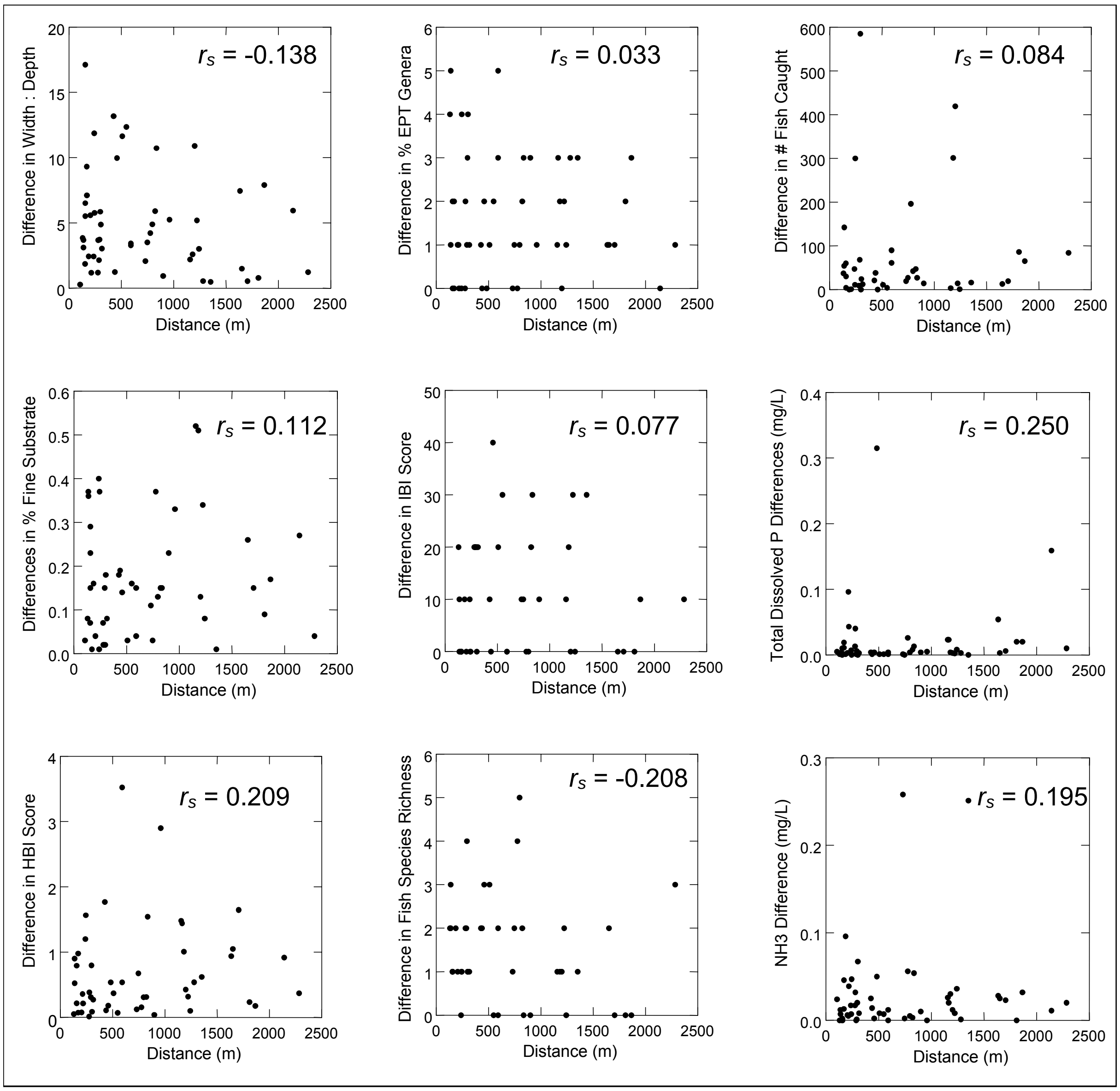

Figure 20. Scatterplots showing relationships between the distance between the X and B sites, and the absolute value of the differences in physical, chemical, and biological measures collected at these sites.

Spearman correlation coefficients (r_s) indicate no significant relationships between the distance between the X and associated B assessment sites and any of the physical, chemical, or biological measures reported (Figure 20). From these coefficients, we determined that the distance between a random sampling site and the closest bridge sampling site did not appear to influence the strength of the correlations between the X and B sites data.

7. Discussion

Characterization of Stream Resources in the Driftless Area

A primary objective of this study was to characterize the physical, chemical, and biological conditions of stream resources in the Driftless Area ecoregion using the EMAP probabilistic sampling design. Previous sampling efforts in the Driftless Area and entire state have primarily been targeted sampling to gather data to address stream specific data needs. These targeted sampling efforts may have induced intentional or unintentional biases, and have produced data for which confidence values cannot be estimated. This REMAP study is the first broad-scale assessment of stream resources by the WDNR that has produced data of known quality and applied objective numeric criteria to judge whether individual or populations of streams are physically, chemically, or biologically degraded. The 25th or 75th percentile values for various physical, chemical, and biological measures from the reference stream sites were used to set reference condition values (management expectations). Study results indicate that the percentages of the sample population (and by inference the entire Driftless Area stream population) not meeting their potential ranged from 77 to 100 percent depending upon the physical, chemical, or biological criteria used. While setting management goals at the 25th or 75th percentiles is a scientifically defensible approach, setting resource management goals should perhaps be viewed as a societal decision. Stream assessment findings and the methods used to set reference conditions will stimulate further discussion within WDNR on which individual or combinations of measures should be used to assess stream quality, and what numeric criteria thresholds for these various parameters should be applied to judge stream quality.

Differences Between Random and Modified-Random Sample Sites

Another primary objective of this study was to evaluate whether sampling randomly selected stream segments at sites accessed from road crossings would provide results comparable to a truly randomized sampling design. None of the 35 physical, chemical, or biological, parameters evaluated were significantly different between the random (X) and modified-random (B) sampling sites.

Given the relative homogeneity of water chemistry parameters due to mixing in lotic systems, intuitively, few differences would be expected to exist for these measures between the X and B sites. However, it is possible that tributaries intervening between the X and B sampling sites can change the concentration of chemical parameters, water temperature, or dissolved oxygen concentrations, if the tributaries have differing water chemistry concentrations, temperature, or dissolved oxygen concentrations than the receiving stream. Intervening groundwater or point-source inputs could also have similar affects.

None of the seven macroinvertebrate metrics analyzed were significantly different between and X and B sites. Most macroinvertebrate taxa are relatively sessile organisms and have been shown to be more strongly influenced by local habitat or reach-scale environmental factors than fish, which show a stronger response to watershed-scale influences (Barbour et al. 1999, Lammert and Allan 1999). Given that macroinvertebrates are thought to respond more strongly to site-specific or reach-scale influences, the lack of X and B site differences in the macroinvertebrate measures may be a more sensitive test of bias being induced by the modified-random sampling design than the fish assemblage data.

Probabilistic Sampling Design Issues

Fundamental principles of probabilistic sampling are that every population element in the target population has a known (and non-zero) probability of being sampled, and it is critically important to rigorously define both the target population and the elements it's comprised of (Cochran 1977). A major objective of the Wisconsin REMAP study was to estimate the number of stream miles in the study area that were meeting physical, chemical, and biological, reference condition criteria. Therefore, a continuous sampling design was applied, where assessment measurements were taken at or in the vicinity of randomly selected points (Larsen 1997). For population elements that are spatially-static such as stream physical habitat (and perhaps less influenced by dynamic elements such as flowing water), sampling at the modified-random sites violates key principles of probabilistic sampling (in essence, the target population becomes the lengths of all streams that are within some distance of road crossings that field crews are willing to travel to reach an assessment site). For spatially-dynamic population elements (water chemistry parameters, fish, or macroinvertebrates that are strongly influenced by upstream land use and the ambient conditions of flowing water) is it less clear (at least to the authors) what the spatial boundaries of these population elements are. It is hoped that this REMAP report will generate further research and discussion on the validity of using road accessible stream sampling sites to characterize stream target populations.

The mean distance between the random and modified-random assessment sites in this study was approximately 701 meters (0.44 mi.). For other ecoregions in Wisconsin or areas outside the state with lower road density, greater land cover or land use heterogeneity, higher potential for intervening point sources of pollution, or greater topographic relief than the REMAP study area, there may be a greater potential for differences between random and modified-random assessment reaches. For example, in a significant proportion of northern Wisconsin road density is about 30 percent less than that of the REMAP study area. This increases the potential distance between random and road-accessible sites, and may result in greater observed differences between these sites. However, this portion of northern Wisconsin is also characterized by more homogeneous land cover and land use, and human population densities are less than in the REMAP study area, which may result in fewer observed differences between sites.
It will be interesting to evaluate these types of interactions in Wisconsin's other ecoregions.

Sample Population Site Selection

The finding that 71 of the randomly selected stream sites were rejected during the field reconnaissance effort to reach a sample population of 60 streams (nearly 120% of the original sample population) is of significance for stream assessment studies and programs that use probabilistic sampling designs. Sample designs must include sufficient over-sample populations to maintain target sample sizes, given that significant numbers of random sites are likely to be rejected. Also, project planning must incorporate a sufficient amount of time for map work to identify site locations, and for field time for reconnaissance of assessment sites and subsequent replacement of rejected sites. The fact that over half of sites were dropped due to dry stream channels indicates a significant amount of error in the WDNR's perennial stream hydrography database for the Driftless Area Ecoregion. Only 5% of the random sites were dropped due to landowner access denial, which is significantly less than what the investigators had expected prior to the start of the study.

Selection of Reference Sites and Application of Reference Condition Data

While macroinvertebrate and fish assemblage indices are increasingly used to objectively assess stream resources in Wisconsin, development of reference conditions for stream physical habitat and water chemistry measures provides additional objective numeric criteria with which to evaluate the condition of individual and populations of streams. In addition, the macroinvertebrate and fish assemblage data collected at reference sites allowed calibration of statewide biological indices specifically for the study area streams, thereby increasing the accuracy of the stream assessments.

The selection of least disturbed reference sites for this study was the first time WDNR has applied an *a priori* site selection of reference sites based on GIS data of percentages of watershed land use types and stream physical habitat characteristics (Hughes et al. 1986). Streams in watersheds with high proportions of county, state, or federal land ownership, and low proportions of agricultural land were evaluated as potential least disturbed reference sites. Study results show that on average the size of the least disturbed watersheds were smaller than the mean size for the random sample population watersheds. As watershed area increases there is a greater likelihood that agricultural and other land uses that negatively affect stream quality will occur within a watershed, increasing the potential cumulative negative impacts in higher order streams. Since stream size (Strahler order, or flow volume) is strongly influenced by watershed size, there presumably are fewer least disturbed large streams relative to small streams in Wisconsin. The influence of stream order on physical, chemical, and biological parameters is further evaluated in Appendix B. Stream order may have implications for setting management goals for larger streams. If reference conditions are developed based on data from lower order (smaller) streams, applying these criteria to higher order streams may result in unrealistic management goals for larger wadeable streams.

A total of 60% (n = 21) of the 35 parameter comparisons between the random and reference stream populations were significantly different. The remaining 14 parameters that were not significantly different may vary little among streams in the study population, and may be indiscriminate measures of stream condition in the study area. Thus, the lack of differences between the random and associated modified-random sample pairs for these same parameters may be more a function of insensitive ecological measures than a lack of differences between random and modified-random sampling sites.

Response of Stream Physical Habitat and Water Chemistry to Watershed Land Use

No significant relationships were detected between watershed percent cropland, percent grassland, or percent forest cover, and riparian or in-stream physical habitat characteristics (e.g., sediment depth). With increasing agricultural land, particularly row cropping, the potential for sediment delivery to streams increases. While the proportion of cropland in individual watersheds can influence sediment delivery, proximity of cropland to streams, cropland slope, riparian land use, width and linear extent, and plant species composition of riparian buffers, and a number of other factors can influence nutrient and sediment delivery to streams. An earlier study conducted by the USGS in Wisconsin showed that streams in the Driftless Region had twice the sediment and nutrient loading rate compared to the statewide average (Corsi et al. 1997). Based on Natural Resource Conservation Service cropland erosion rate data and the USGS study findings, it is evident that significant amounts of sediment are being delivered to the study area streams, but given the relatively steep topography and resulting stream gradient, the majority of the sediment flows through these stream systems and is not deposited on the streambeds. In addition, the WISCLAND land use data used to quantify percent watershed agricultural land was based on 10-12 year old LANDSAT data. So while the general land use patterns have changed little in these watersheds over recent time, the dated land use data could have weakened the relationships between land use and measurable in-stream physical habitat degradation. The fact that some of the reference stream watersheds had significant amounts of agricultural land suggests that with sufficient riparian protection and proper agricultural land management, agricultural productivity and high quality streams can co-exist.

Response of Biotic Indices to Percent Watershed Agriculture, Stream Sediment and Water Chemistry Variables

The results of the Pearson correlation analyses indicate stronger responses of macroinvertebrates than fish to increasing watershed disturbance. Based on adjusted p-values, neither the macroinvertebrate nor the fish metrics were strongly related to habitat or land use variables. However, the two macroinvertebrate metrics examined were significantly correlated to several water chemistry and water quality measures. The fish IBI score was only weakly correlated with Kjeldahl nitrogen and water clarity. These results are surprising, as various studies have provided evidence that the structure and function of stream fish communities are altered with increasing proportions of agricultural land within watersheds (Fausch et al., 1990, Meador and Goldstein, 2003). The EPT metric indicated reductions in the quality of the benthic community in response to increasing total dissolved phosphorus concentration and, more weakly, streambed sedimentation. Various studies have shown that as nutrient concentrations increase in streams there is an increase in filamentous algae growth which displaces the periphytic (diatom) community that is an important food source for many macroinvertebrate taxa (Lillie et al. 2003). Similarly, with increasing sedimentation there is a corresponding loss of available coarse substrate (rubble, cobble, and gravel) and the associated interstitial spaces that are critical habitat for many macroinvertebrate taxa (Lillie et al. 2003). The lack of strong relationships between metrics and measures of habitat and land use may be a function of the metrics selected, as some metrics may show responses only to specific types of disturbance. Other factors related to the land cover data also could affect the relationships observed, including the age of the LANDSAT data, the distribution of agricultural land within individual watersheds, and the quality of the riparian corridor along individual streams.

Comparison of Random Sample Population with WDNR Biologist's Targeted Sample Population

The WDNR has primarily relied on targeted sampling to provide data on stream conditions for site-specific, regional and statewide resource assessments. While it is critically important to continue to collect data for stream-specific assessments and management decisions, the REMAP study results indicate that using targeted data to make broad spatial inferences about the quality of stream populations can lead to erroneous estimations of stream quality. In general, the emphasis of stream monitoring in western Wisconsin has been to target high quality streams to assess the condition of trout fisheries or to determine the potential for these streams to support brown trout or the more environmentally-sensitive native brook trout. A key finding of this study was that the bias of targeted sampling resulted in an overestimation of stream quality when extrapolated to the population-level, findings similar to that of Hughes et al. (2000). Bias induced from targeted sampling can of course also underestimate the quality of regional stream populations if the focus of the targeted sampling is directed toward streams impaired by industrial or municipal waste discharges, or polluted urban or agricultural run-off, as is typically the focus of Total Maximum Daily Load watershed assessments.

Implications for WDNR's Wadeable Stream Monitoring Program

The REMAP study results provide valuable insights for improving the rigor of the WDNR's wadeable stream monitoring program. This study is the first time the WDNR has conducted a probabilistic survey on a broad geographic scale. Results also indicate macroinvertebrate and fish assemblages differ in their power to detect physical habitat degradation and chemical pollution. In addition, this study provides evidence that the targeted sampling design routinely used overestimates the quality of stream resources in western Wisconsin. The preliminary findings that a modified-random sampling design appears to induce little bias in the assessment of stream quality may allow a more efficient and cost-effective stream sampling effort in Wisconsin, but additional study is needed to more rigorously evaluate the utility of applying a road-accessible sampling design. Finally, the process of using numeric criteria to objectively determine whether individual or populations of streams are meeting their potential will be applied statewide which will help reduce WDNR resource managers' reliance on subjective, qualitative, resource evaluations.

8. Acknowledgements

A number of WDNR staff contributed to this study. Among those, we thank Sue Acre, Bob Busch, Bill Ceelen, Walt Jaeger, and Joanne Tooley. In addition, a number of EPA staff were key contributors to this project and the authors are very thankful for their support. Dr. Florence Fulk – NERL, and Sarah Lehmann and Edward Hammer of Region 5 helped secure funding and initiate the project. Dr. Tony Olsen, NHEERL-WED provided the sample population - draw and statistical guidance; Drs. Robert Hughes, Tony Olsen, Steve Paulsen, and Paul Ringold of NHEERL provided review comments; and Karen Blocksom of NERL provided statistical guidance, and technical reviews of earlier report drafts.

9. References

Barbour, M.T., J. Gerritsen, B.D. Snyder, and J.B. Stribling. 1999. *Rapid Bioassessment Protocols for Use in Streams and Wadeable Rivers: Periphyton, Benthic Macroinvertebrates, and Fish, Second Edition.* EPA 841-B-99-002. U. S. Environmental Protection Agency, Office of Water; Washington, D. C.

Cochran, W.G. 1977. *Sampling Techniques, 3rd edition.* J. Wiley and Sons, New York, New York.

Corsi, S.D., J. Graczyk, D.W. Owens, and R.T. Bannerman. 1997. *Unit-area Loads of Suspended Sediment, Suspended Solids, and Total Phosphorus from Small Watersheds in Wisconsin.* U.S. Geological Fact Sheet No. USGS-FS-195-97.

Davies, N.M., R.H. Norris, and C.T. Martin. 2000. Prediction and assessment of local stream habitat features using large-scale catchment characteristics. *Freshwater Biology* 45:343-369.

Diamond, J.M., M.T. Barbour, and J.B. Stribling. 1996. Characterizing and comparing bioassessment methods and their results: a perspective. *Journal of the North American Benthological Society* 15:713-727.

Diaz-Ramos, S., D.L. Stevens, Jr., and A.R. Olsen. 1996. EMAP Statistical Methods Manual. EPA/620/R-96/002. U.S. Environmental Protection Agency, Office of Research and Development, National Health Effects and Environmental Research Laboratory, Western Ecology Division, Corvallis, Oregon.

Fausch, K.D., J. Lyons, J.R. Karr, and P.L. Angermeier. 1990. Fish communities as indicators of environmental degradation. *American Fisheries Society Symposium* 8:123-144.

Gibson, G.R., Jr., M.T. Barbour, J.B. Stribling, J. Gerritsen, and J.R. Karr. 1996. *Biological Criteria: Technical Guidance for Streams and Small Rivers*. EPA/822-B-96-001. U.S. Environmental Protection Agency, Office of Science and Technology, Washington, DC.

Hilsenhoff, W.L. 1987. An improved biotic index for organic stream pollution. *Great Lakes Entomologist* 20:31-39.

Hughes, R.M., D.P. Larsen, and J.M. Omernik. 1986. Regional reference sites: a method for assessing stream potentials. *Environmental Management* 10:629-635.

Hughes, R.M., S.G. Paulsen, and J.L. Stoddard. 2000. EMAP-Surface Waters: a multiassemblage probability survey of ecological integrity in the U.S.A. *Hydrobiologia* 422:429-443.

Karr, J.R. and E.W. Chu. 1999. *Restoring Life in Running Waters: Better Biological Monitoring*. Island Press, Washington, D.C.

Keys, J.E., Jr., C.A. Carpenter, S.L. Hooks, F.G. Koeneg, W.H. McNab, W.E. Russell, and M.L. Smith. 1995. *Ecological Units of the Eastern United States--First Approximation.* Technical Publication R8-TP 21. Map (scale 1:3,500,000). U.S. Department of Agriculture, Forest Service, Atlanta, Georgia.

Lammert, M., and J.D. Allan. 1999. Assessing Biotic Integrity of streams: effects of scale in measuring the influence of land use/cover and habitat structure on fish and macroinvertebrates. *Environmental Management* 23:257-270.

Larsen, D.P. 1995. The role of ecological sample surveys in the implementation of biocriteria. In: Davis, W.S. and T.P. Simon (Editors), *Biological Assessment and Criteria: Tools for Water Resource Planning*. CRC Press Inc. Boca Raton, Florida.

Larsen, D.P. 1997. Sample survey design issues for bioassessment of inland aquatic ecosystems. *Human and Ecological Risk Assessment* 3:979-991.

Legendre, P. and L. Legendre. 1998. *Numerical Ecology*. Elsevier Science, Amsterdam, The Netherlands.

Lillie, R.A., S.W. Szczytko, and M.A. Miller. 2003. *Macroinvertebrate Data Interpretation Guidance Manual.* Wisconsin Dept. of Natural Resources, Madison, Wisconsin. Pub. SS-965-2003.

Lyons, J. 1992a. *Using the Index of Biotic Integrity (IBI) to Measure Environmental Quality in Warmwater Streams of Wisconsin*. General Technical Report NC-49. North Central Forest Experiment Station, U.S. Department of Agriculture, St. Paul, Minnesota.

Lyons, J. 1992b. The length of stream to sample with a towed electrofishing unit when fish species richness is estimated. *North American Journal of Fisheries Management* 12:198–203.

Lyons, J., L. Wang, and T.D. Simonson. 1996. Development and validation of an index of biotic integrity for coldwater streams in Wisconsin. *North American Journal of Fisheries Management* 16:241-256.

Meador, M.R., and R.M. Goldstein. 2003. Assessing water quality at large geographic scales: relationships among land use, water, physiochemistry, riparian conditions, and fish community structure. *Environmental Management* 31:505-517.

Miller, M.A., J.D. Diamond, J.B. Stribling, and K.L. Schmude. (In prep.) Development and application of a performance-based comparability framework using a comparison of macroinvertebrate field and laboratory protocols.

Omernik, J.M. 1987. Ecoregions of the conterminous United States. *Annals of the Association of American Geographers* 77:118-125.

Petersen, S.A., N.S. Urquhart, and E.B. Welch. 1999. Sample representativeness: a must for reliable regional lake condition estimates. *Environmental Science and Technology* 33:1559-1565.

Robertson, D.M., D.J. Graczyck, P.J. Garrison, L. Wang, G. LaLiberte, and R. Bannerman. 2006. Nutrient concentrations and their relationships to the biological integrity of streams in Wisconsin. U.S. Geological Survey Prof. Paper 1722.

Sokal, R.R., and F.J. Rohlf. 1981. *Biometry 2nd Edition*. W.H. Freeman and Co. New York, NY.

Stevens, D.L., Jr. 1997. Variable density grid-based sampling designs for spatial populations. *Environmetrics* 8:167-95.

Stevens, D.L., Jr., and A.R. Olsen. 1999. Spatially restricted surveys over time for aquatic resources. *Journal of Agricultural, Biological, and Environmental Statistics* 4: 415-428.

Stevens, D.L., Jr., and A.R. Olsen. 2004. Spatially-balanced sampling of natural resources. *Journal of the American Statistical Association* 99:262-278.

Strahler, A.N. 1957. Quantitative analysis of watershed geomorphology. *Transactions of the American Geophysical Union* 38:913-920.

Appendix A. Comparison of Spring and Fall Water Chemistry Parameters

Evaluation of Seasonal Effects

Water quality and water chemistry data collected in the spring and fall of 2003 were compared using Bonferroni-adjusted paired t-tests to determine if seasonal differences existed. In the spring and fall of 2003 we measured instantaneous values of dissolved oxygen concentration, dissolved oxygen percent saturation, pH, and water temperature *in situ* with electronic meters, and water transparency with a transparency tube at the X, B and reference stream sites. Total phosphorus, total dissolved phosphorus, NH_3 (ammonia), NO_3-NO_2 (nitrate-nitrite), and total Kjeldahl nitrogen concentrations were also measured using laboratory-analyzed grab samples. Paired t-tests (with Bonferroni adjustments) were used to compare within-site spring and fall differences for these 10 parameters. Water transparency, dissolved oxygen concentration and dissolved oxygen percent saturation showed statistically significant differences between the spring and fall sampling periods (Table A1, Figure A1). Dissolved oxygen concentration, the percent dissolved oxygen saturation, and water column transparency had higher values during the fall sampling period than the spring. Total dissolved phosphorus, NH_3, and Kjeldahl nitrogen also displayed statistically significant concentration differences between the spring and fall samples. Total dissolved phosphorus values were significantly higher in the fall. In contrast, both NH_3 and Kjeldahl nitrogen were significantly higher in the spring samples. Because of these seasonal differences in various water chemistry measures, and because the spring dataset had a greater number of sites sampled, only water chemistry samples collected in spring were used for analyses in this study.

Table A 1. Paired t-test comparisons between spring and fall water quality and water chemistry measures collected at the X stream sites.

Spring Vs Fall Water Chemistry Measures	No. of samples	Degrees of Freedom	p-value
Total Phosphorus (mg/L)*	47	46	0.240
Total Dissolved Phosphorus (mg/L)*	49	48	**0.001**
NH_3 (mg/L)*	49	48	**0.001**
NO_3-NO_2 (mg/L)*	49	48	0.670
N-Kjeldahl (mg/L)*	49	48	**0.043**
Dissolved Oxygen (mg/L)	49	48	**0.000**
Transparency (cm)	43	42	**0.034**
% Oxygen Saturation	43	42	**0.000**
pH	38	37	0.798
Water Temperature (C)	49	48	0.294

*Data were Log_{10} transformed prior to conducting the paired t-tests. P-values in bold are significant.

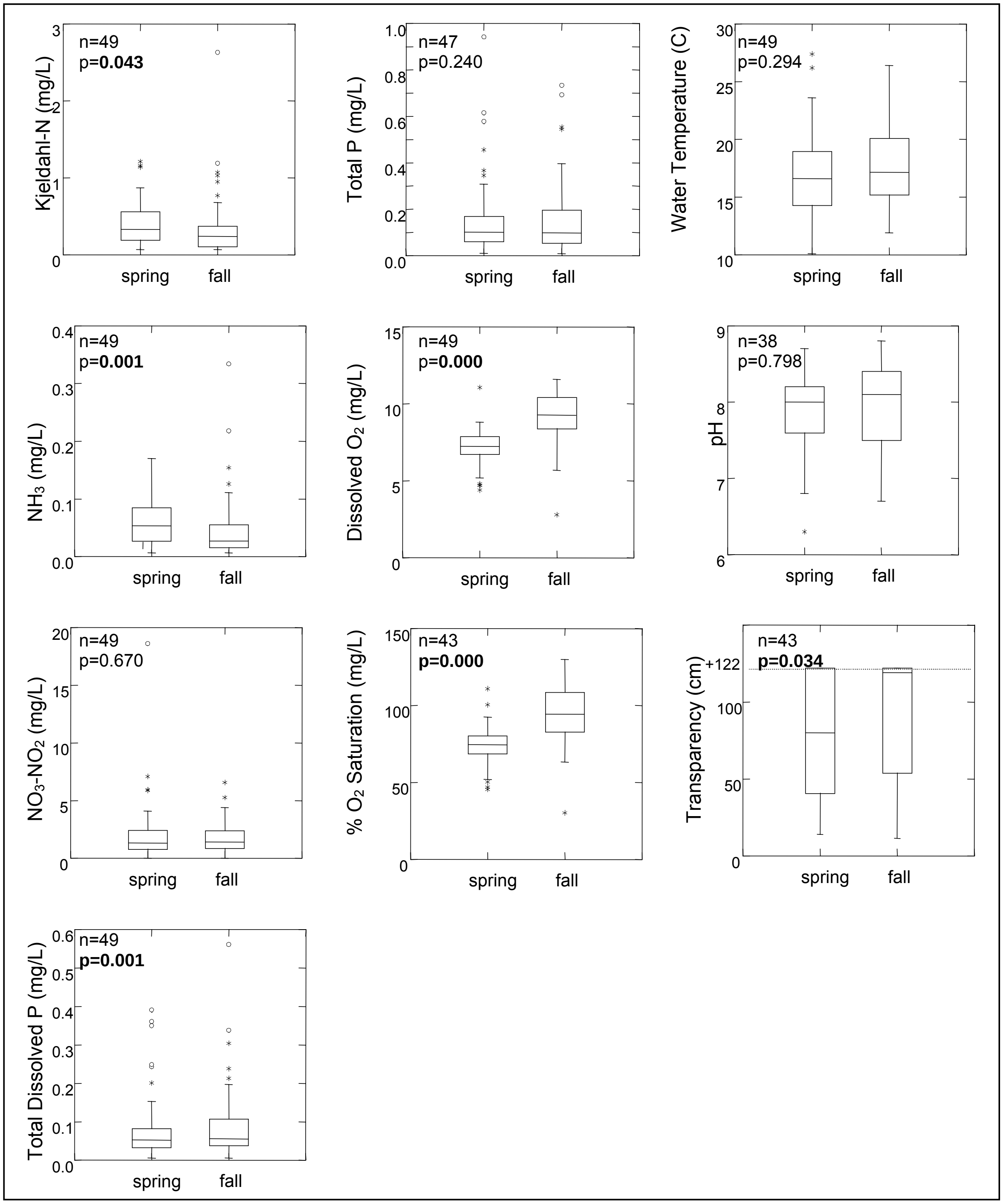

Figure A 1. Box and whisker plots of spring and fall water quality and water chemistry measures collected at the X-sites. P-values represent paired t-test results.

Appendix B. Relationships of Stream Order with Physical Habitat, Water Quality and Water Chemistry, and Biological Measures

Habitat Measures and Stream Order

To evaluate relationships between stream order and physical habitat measures, an ANOVA was applied to each habitat measure collected at the X sites. Significant relationships were detected between stream order and sinuosity, mean distance between bends, and mean length of bank erosion (Tables B1, B2). Sinuosity differed between first and second order sites compared to third and fourth order streams. Width of bank erosion and distance between bends both differed significantly between second and fourth order streams only.

Table B 1. ANOVA results of stream order and physical habitat measures collected at the X sites.

X Site Habitat Measures Vs Stream Order	# of Samples	F-ratio	p-value
Sinuosity*	53	12.551	**<0.001**
Width : Depth Ratio	56	0.869	0.463
Mean Depth of Fines*	53	1.493	0.228
Percent Sand, Silt, and/or Clay Sediments	53	2.144	0.107
Mean Distance Between Bends (m)*	53	3.367	**0.026**
Mean Length of Riffles (m)*	53	0.366	0.778
Mean Length of Pools (m)*	53	1.552	0.213
Mean Width of Bank Erosion (m)*	53	6.795	**0.001**

* Data were $Log_{10}(x+1)$ transformed prior to conducting ANOVA. The percent sand, silt, and clay values were arcsine square root transformed prior to conducting ANOVA. P values in bold are significant.

Table B 2. Tukey HSD pairwise comparison probabilities for physical habitat measures that produced significant ANOVA results. P-values in bold indicate significantly different parameter values between the corresponding stream orders.

Parameter	Stream Order	1	2	3
Sinuosity	2	0.992		
	3	**0.006**	**0.002**	
	4	**0.000**	**0.000**	0.589
Mean Distance Between Bends	2	0.987		
	3	0.895	0.696	
	4	0.069	**0.027**	0.167
Mean Width of Bank Erosion	2	0.100		
	3	0.936	0.211	
	4	0.317	**0.000**	0.061

Water Quality and Water Chemistry Measures and Stream Order

Similar to the ANOVA conducted with the habitat data, we evaluated water quality and water chemistry data collected from the X sites to determine if these measures varied by stream order. Only dissolved oxygen percent saturation showed statistical differences among stream orders (Tables B3, B4).

Table B 3. ANOVA of stream order and water quality and water chemistry measures from the X stream sites.

X Site Water Chemistry Vs Stream Order	# of Samples	F-ratio	p-value
Total Phosphorus (mg/L)*	56	0.506	0.680
Total Dissolved Phosphorus (mg/L)*	56	0.528	0.665
NH3 (mg/L)*	56	0.647	0.588
NO3-NO2 (mg/L)*	56	2.709	0.054
N-Kjeldahl (mg/L)*	56	0.767	0.518
Dissolved Oxygen (mg/L)	57	2.522	0.068
Percent Oxygen Saturation	57	3.662	**0.018**
Transparency (cm)	57	0.536	0.660
pH	57	2.769	0.051
Conductivity (uS)*	47	0.439	0.726
Water Temperature (C)	57	0.398	0.755

*Data were Log_{10} transformed prior to conducting the ANOVA analysis. P values in bold are significant.

Table B 4. Tukey HSD pairwise comparison for the percent oxygen saturation. The p-value in bold indicates significantly oxygen saturation values between second and third order streams.

Stream Order	1	2	3
2	1.000		
3	0.060	**0.049**	
4	0.247	0.224	0.900

Macroinvertebrate Measures and Stream Order

ANOVA results comparing macroinvertebrate data and stream order indicated no significant relationships between stream order and any of the 7 macroinvertebrate metrics or indices. (Table B5)

Table B 5. ANOVA of stream order and macroinvertebrate measures from the X stream sites.

X Site Macroinvertebrate Measures* Vs Stream Order	F-ratio	p-value
Species Richness	0.765	0.520
HBI Score	1.288	0.290
Percent EPT Genera	2.075	0.117
Percent EPT Individuals	0.867	0.465
Shannon's Diversity Index Score	0.713	0.549
Percent Shredders	0.224	0.879
Percent Scrapers	1.303	0.285

*Percent values were arc-sine square root transformed prior to conducting ANOVAS. n = 49

Fish Assemblage Measures and Stream Order

In contrast to the macroinvertebrate data presented above, ANOVA results showed some significant differences among stream orders and fish assemblage measures. Specifically, species richness and the number of "intolerant" species present are significantly related to stream order. (Tables B6, B7)

Table B 6. ANOVA of stream order and fish assemblage measures from the X stream sites (N=42).

X Site Fish Assemblage Measures Vs Stream Order	F-ratio	p-value
Species Richness	4.672	**0.007**
Number of Intolerant Species*	4.601	**0.008**
Percent of Tolerant Individuals	0.558	0.646
Percent Top Carnivores	1.887	0.148
Percent Stenothermal Cool/Coldwater	0.879	0.461
Percent of Salmonids that are Brook Trout	1.228	0.313
IBI Score*	0.993	0.406

* Data were $Log_{10}(X+1)$ transformed prior to conducting ANOVAS. **Data were $Log_{10}(X)$ transformed prior to conducting ANOVAs. Percent values were arc-sine square root transformed prior to conducting ANOVA. P-values in bold indicate significant tests.

Table B 7. Tukey HSD pairwise comparison probabilities for fish metrics that produced significant ANOVA results. P-values in bold indicate significantly different parameter values between the corresponding stream orders.

Parameter	Stream Order	1	2	3
Species Richness	2	0.770		
	3	0.712	1.000	
	4	**0.018**	0.056	**0.043**
Number of Intolerant Species	2	0.200		
	3	0.143	0.999	
	4	**0.005**	0.250	0.247

Appendix C. Fish Assemblage Data Analyses

We captured a total of 54 species and 19,999 individual fish at the X, B, and reference stream sites. Forty different species of fish were captured at the X sites, and eighteen different species of fish were captured at the reference sites (Tables C1, C2). No fish were captured at 11 of the individual X or B stream sites and at 2 of the reference sites. Numbers of individual fish caught per X, B, or reference site ranged from 0 to 1,755. On average, 179 fish were captured at the X and B sites and 101 fish at the reference sites. The number of fish species caught at individual X and B sites ranged from 1 to 18, with an average of 6 species captured per site. The number of fish species captured at each reference site ranged from 1 to 8, with an average of 3 species captured per site.

Brook Trout (*Salvelinus fontinalis*) was the most common species (650 individuals) captured at the reference stream sites, and Central Stoneroller (*Campostoma anomalum*) was the most common species captured at the X sites (1,190 individuals). Two-thirds of the fish species captured at the X sites were not found at the reference sites; conversely, Burbot (*Lota lota*), Rainbow trout (*Oncorhynchus mykiss*), and Slimy Sculpin (*Cottus cognatus*) were captured at the reference sites but at none of the X sites. Species with the greatest dissimilarity in frequency of occurrence between the X and reference sites were Creek Chub (*Semotilus atromaculatus*) occurring in 57% of the X sites and only 5% of the reference sites, White Sucker (*Catostomus commersoni*) occurring in 55% of the X sites and only 14% of the reference sites, and Johnny Darter (*Etheostoma nigrum*) occurring in 45% of the X sites and only 10% of the reference sites. Brook Trout occurred in 57% of the reference sites and only 30% of the X sites, Mottled Sculpin (*Cottus bairdi*) occurred at 38% of the reference sites and only 24% of the X sites, and Brown trout (*Salmo trutta*) occurred at 48% of the reference sites and only 24% of the X sites (Fig. C1).

Table C 1. Fish species captured at X stream sites.

Species	Scientific Name	Total No. of Fish Caught	No. of X-Sites with Species Present	Frequency of species at X-sites
American Brook Lamprey	*Lampetra appendix*	17	5	9.8%
Bigmouth Shiner	*Notropis dorsalis*	49	4	7.8%
Blacknose Dace	*Rhicnichthys atratulus*	284	17	33.3%
Blacknose Shiner	*Notropis heterlepis*	26	4	7.8%
Blackside Darter	*Percina maculata*	4	3	5.9%
Bluegill	*Lepomis macrochirus*	16	3	5.9%
Bluntnose Minnow	*Pimephales notatus*	542	9	17.7%
Brassy Minnow	*Hybognathus hankinsoni*	2	1	2.0%
Brook Stickleback	*Culaea inconstans*	180	20	39.2%
Brook Trout	*Salvelinus fontinalis*	511	15	29.4%
Brown Trout	*Salmo trutta*	208	12	23.5%
Central Mudminnow	*Umbra limi*	4	3	5.9%
Central Stoneroller	*Campostoma anomalum*	1,190	7	13.7%
Channel Catfish	*Ictalurus punctatus*	1	1	2.0%
Common Carp	*Cyprinus carpio*	4	1	2.0%
Common Shiner	*Notropis cornutus*	1,154	9	17.6%
Creek Chub	*Semotilus atromaculatus*	945	29	56.9%
Fantail Darter	*Etheostoma flabellare*	156	14	27.5%
Fathead Minnow	*Pimephales promelas*	106	6	11.8%
Green Sunfish	*Lepomis cyanellus*	22	4	7.8%
Hornyhead Chub	*Nocomis biguttatus*	745	7	13.7%
Johnny Darter	*Etheostoma nigrum*	345	23	45.1%
Longnose Dace	*Rhinichthys cataractae*	95	8	15.7%
Mottled Sculpin	*Cottus bairdi*	260	12	23.5%
Northern Brook Lamprey	*Ichthyomyzon fossor*	3	2	4.0%
Northern Hog Sucker	*Hypentelium nigricans*	2	1	2.0%
Pearl Dace	*Semotilus margarita*	238	3	5.9%
Pumpkinseed	*Lepomis gibbosus*	1	1	2.0%
Quillback	*Carpiodes cyprinus*	4	1	2.0%
Rosyface Shiner	*Notropis rubellus*	34	5	9.8%
Sand Shiner	*Notropis stramineus*	3	1	2.0%
Shorthead Redhorse	*Moxostoma macrolepidotum*	3	3	5.9%
Smallmouth Bass	*Micropterus dolomieu*	25	3	5.9%
Southern Redbelly Dace	*Phoxinus erythrogaster*	227	8	15.7%
Spotfin Shiner	*Notropis spilopterus*	29	3	5.9%
Stonecat	*Noturus flavus*	42	3	5.9%
Suckermouth Minnow	*Phenacobius mirabilis*	32	2	3.9%
Tiger Trout	*Salvelinus fontinalis X Salmo trutta*	1	1	2.0%
Walleye	*Sander vitreus*	2	1	2.0%
White Sucker	*Catostomus commersoni*	762	28	54.9%
Total Number of Fish		**8,275**		

Table C 2. Fish species captured at reference stream sites.

Species	*Scientific Name*	**Total No. of Fish Caught**	**No. of Reference Sites with Species Present**	**Frequency of species at Reference Sites**
American Brook Lamprey	*Lampetra appendix*	55	5	23.8%
Blacknose Dace	*Rhicnichthys atratulus*	35	4	19.1%
Blacknose Shiner	*Notropis heterolepis*	2	1	4.8%
Bluntnose Minnow	*Pimephales notatus*	6	1	4.8%
Brook Stickleback	*Culaea inconstans*	63	6	28.6%
Brook Trout	*Salvelinus fontinalis*	650	12	57.1%
Brown Trout	*Salmo trutta*	625	10	47.6%
Burbot	*Lota lota*	43	2	9.5%
Central Mudminnow	*Umbra limi*	1	1	4.8%
Central Stoneroller	*Camostoma anomalum*	5	1	4.8%
Creek Chub	*Semotilus atromaculatus*	8	1	4.8%
Fantail Darter	*Etheostoma flabellare*	35	2	9.5%
Johnny Darter	*Etheostoma nigrum*	35	2	9.5%
Longnose Dace	*Rhinichthys cataractae*	76	3	14.3%
Mottled Sculpin	*Cottus bairdi*	547	8	38.1%
Rainbow Trout	*Oncorhynchus mykiss*	3	3	14.3%
Slimy Sculpin	*Cottus cognatus*	118	1	4.8%
White Sucker	*Catostomus commersoni*	90	3	14.3%
Total Number of Fish		**2,397**		

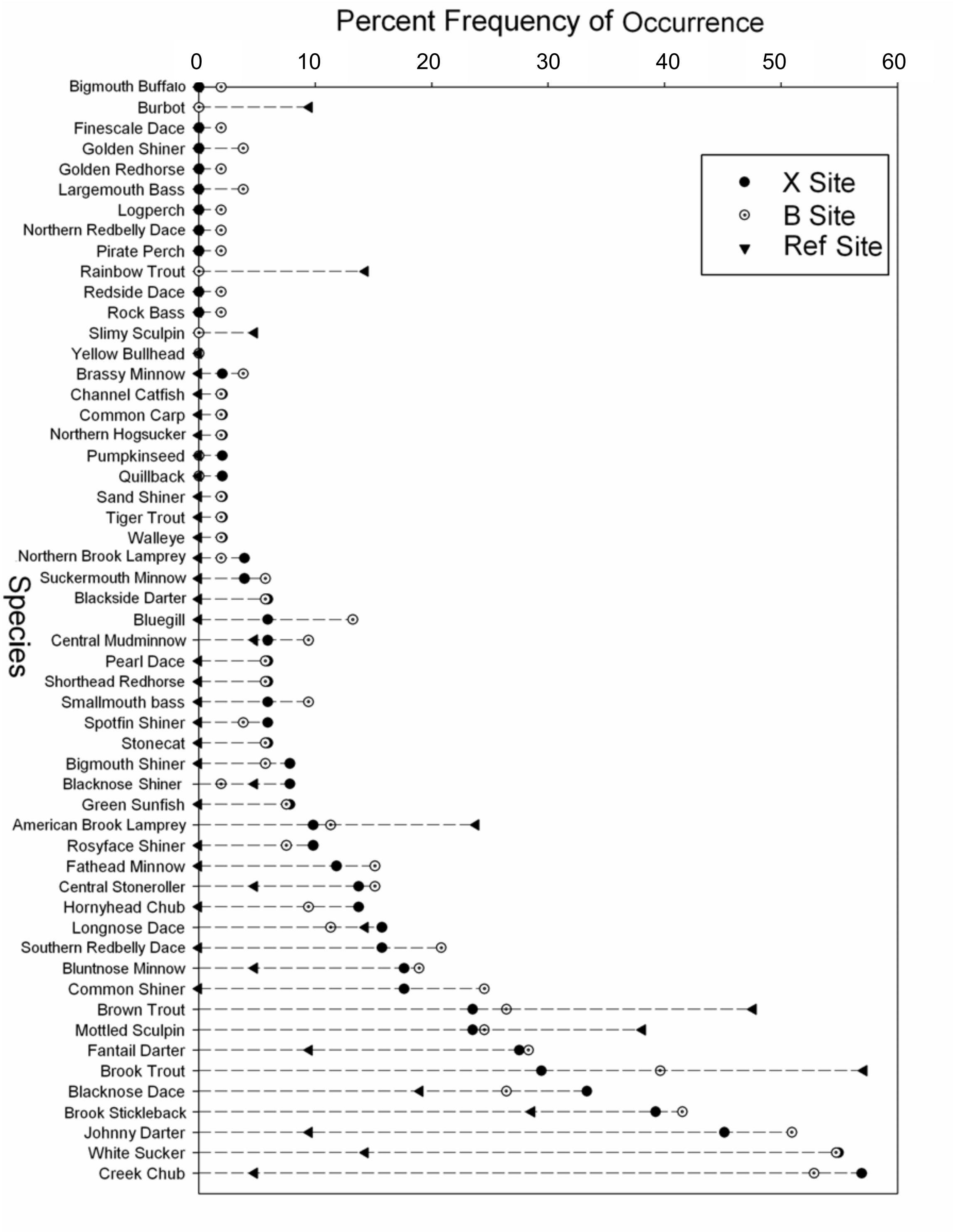

Figure C 1. Percent frequency of occurrence of fish species at the X, B, and reference sites.

United States
Environmental Protection
Agency

Please make all necessary changes on the below label, detach or copy, and return to the address in the upper left-hand corner.

If you do not wish to receive these reports CHECK HERE ☐; detach, or copy this cover, and return to the address in the upper left-hand corner.

PRESORTED STANDARD
POSTAGE & FEES PAID
EPA
PERMIT No. G-35

Office of Research
and Development (8101R)
Washington, DC 20460

Official Business
Penalty for Private Use
$300

EPA 600/R-06/165
November 2006
www.epa.gov