# Development of Macroinvertebrate Biocriteria for Streams of Minnesota's Lake Superior Watershed

Technical Report # NRRI/TR-2000/19

Kevin Stroom
Carl Richards

**Natural Resources Research Institute**
University of Minnesota - Duluth

July 2000

# Table of Contents

## List of Figures



## List of Tables



## Appendices

## Abstract

Genus-level macroinvertebrate data from reference and disturbed streams were used to develop and test biocriteria for 1st - 3rd order streams in Minnesota's Lake Superior Watershed (LSW). Fifteen metrics, most used elsewhere, were investigated for utility. Five metrics failed because of high correlation or inability to differentiate disturbed streams. Ten metrics were combined into a multimetric index. Metric values were scored relative to metric biocriteria according to EPA protocols and summed for each stream. The minimum reference stream score defined the index biocriterion. Eleven urban or agricultural/rural stream scores were compared to the index biocriterion to test its ability to reveal impairment in test streams. Reference and disturbed stream scores were statistically different ($p < 0.01$). Urban streams were better separated from the reference condition ($p < 0.001$) than Ag/Rural streams ($p < 0.01$). The index biocriterion detected impairment in eight of the disturbed streams, while two streams scored within the safety buffer where a judgement of impairment was uncertain but possible. One stream scored slightly above the biocriterion. Several metrics which were useful elsewhere were also effective here, while others were not. The locally-tailored multimetric index and associated biocriteria developed here were effective in assessing stream ecosystem health in appropriate areas of the LSW. GIS analysis of subwatershed land use/cover showed that even moderate percentages of developed + hay/pasture/grass land covers (12-15%) and developed + hay/pasture/grass + roads (15-17%) resulted in some streams scoring as impaired. Therefore, LSW streams appear to be relatively fragile environments requiring careful watershed management.

Methodologies involving chironomid inclusion, sample processing, sample size, and reach location were investigated. Chironomid abundance varied among Reference streams by more than an order of magnitude, though % chironomids varied less. Thus, inclusion of family-level chironomid data may increase variability in metrics involving abundance. A high percentage of taxa were "rare" at levels of ≤2 and ≤4 individuals per replicate; 34.0 and 48.0 % respectively for Reference streams and 44.2 and 61.2 for Disturbed streams. Thus, subsampling may strongly influence richness metric values, suggesting whole LSW samples should be processed for maximum assessment effectiveness. A richness/area curve showed continual increase through 5 replicates, thus, collecting fewer samples may reduce biocriteria effectiveness in the LSW if samples are composited. Metric scores were often significantly different in the three Kimball Cr. sample reaches, thus reach location may be an important consideration in developing biocriteria in the LSW.

## Introduction

Macroinvertebrates are increasingly being used to supplement chemical testing to more adequately assess water resource quality. The motivation for more complete assessment comes from the goal of the Clean Water Act and an amendment known as the Water Quality Act of 1987. "Section 101a states that the Act's primary objective is to 'restore and maintain the chemical, physical, and biological integrity of the nation's waters'" (Gibson 1996). Sampling stream organisms gives a more biologically relevant analysis of the effects of pollutants or degradation of water quality (Gibson 1996). Sampling macroinvertebrate, fish, or periphyton communities from a group of least-impaired (reference) streams within a relatively homogeneous landscape area (ecoregion) enables the development of biocriteria for that region. A biocriterion is a numerical standard for a biological community metric. A single, cumulative biocriterion is often developed by combining individual metrics into a multimetric index and summing metric scores (Barbour *et al.* 1995). Examples of such indices are the Index of Biotic Integrity (IBI; Karr 1981), the Benthic Index of Biotic Integrity (B-IBI; Kerans and Karr 1994), and the Invertebrate Community Index (ICI; DeShon, 1995). Individual streams in the ecoregion can then be compared to these reference streams. Those not meeting the biocriterion are suspected of being impaired by one or several anthropogenic activities in the watershed.

The Lake Superior Watershed (LSW) is highly prized by the residents of Minnesota and very important to the state for its tourism value. The highway running the length of the watershed along the coastline is one of only three federally designated scenic highways in the U.S. Increasing development and recreation pressures on the coastline and inland may have negative consequences for the area's ecosystems. To date, only a few quantitative studies of macroinvertebrate communities have been done for the LSW (Surber 1922, Surber 1924, Smith and Moyle 1944, Waters 1981, Richards and Host 1994). These studies either covered only parts of the LSW or sampled relatively few streams. The Minnesota Pollution Control Agency and USEPA are currently doing extensive macroinvertebrate studies in a more broad area which includes the LSW. This study had several goals: 1) to define baseline macroinvertebrate conditions in this still relatively pristine watershed for use in future water resource management, 2) to establish biocriteria from this reference condition, 3) to provide guidance for future assessments in the LSW and similar watersheds in other geographical areas, 4) to contribute in the endeavor to find effective, broadly applicable metrics by testing common, promising metrics in a new geographical area, and 5) to add to the existing evidence for the effectiveness of bioassessment for determining water resource health.

## Description of the study area

The LSW (Fig. 1) is one of 4 watersheds that make up Minnesota's portion of the Lake Superior Basin. The St. Louis R., Nemadji R., and Cloquet R. watersheds were not included in this study. The LSW is a narrow band of land running along the shoreline of Lake Superior for 302 km (189 mi.) and comprises an area of 5,688 km$^2$, or 2,184 mi$^2$ (MPCA 1975). The LSW is at the USGS Accounting Unit scale and is designated as Northwestern Lake Superior, # 040101. It is comprised of Hydrologic Units 04010101 and 04010102.

Except where noted, much of the following description of the LSW is derived from Smith and Moyle (1944). The LSW is part of the Canadian Shield geological formation. Relatively thin, glacially-derived soil lies over metamorphic and igneous bedrock. Inland topography is relatively flat, but elevation drops quickly nearing the Lake Superior coastline. This topographical pattern occurs along the length of the LSW, to varying degree, leading to similar longitudinal gradient patterns of LSW streams. In general, the streams flow perpendicularly toward Lake Superior. The combination of the area's geology with the narrow watershed width results in strong gradient changes over short distances, giving sections of LSW streams characteristics similar to those in more mountainous regions. Headwater reaches are typically low gradient, slow water stretches with silt and soft, organic substrates. Middle reaches have moderate gradients with gravel/cobble substrates. Lower reaches have high gradients with large cobble/boulder substrates, cascades, and waterfalls. Surface runoff is the main factor controlling stream flow (MPCA 1975) and stream flows are highly variable. For example, maximum/minimum flows of 1273/2.00, 583/2.00 and 2120/2.00 cfs are reported for Gooseberry R., Cascade R., and Beaver R. respectively. This variability is moderated to some extent in streams that have lakes or wetlands in their headwaters which provide a more stable source of water (MPCA 1975).

The majority of the LSW is heavily forested with second-growth mixed hardwoods and conifers. Extensive forest harvest and fires occurred from the late 1890's to 1925. Much of the upper half of the LSW lies in the Superior National Forest. LSW land uses include forestry, urban, rural low density residential, and a small amount of low-intensity agriculture, which occurs almost exclusively in the southern ⅓ of the LSW. Human population is concentrated in the southern ⅓ of the watershed in Duluth and its adjacent areas. Other urban developments are small, primarily coastal towns.

LSW streams have relatively soft water compared to other Minnesota regions. Streams in the southern portion of the LSW are considered medium hard while those in the northern portion are considered soft (MPCA 1975). Turbidity is low, but coloration is strong due to naturally high levels of dissolved organic compounds (MPCA 1975). Smith and Moyle (1944) classified LSW streams as "moderately fertile" based on hardness levels. Aquatic macrophytes are not abundant and generally absent from middle and lower reaches. Production of benthic fauna is low.

It is important to note that we attempted to combine a "watershed approach" and "ecoregion approach" to management. A basic assumption in developing biocriteria is that they hold for streams within a defined, physically homogeneous land area (ecoregion) with similar climate, geology/soils, vegetation, and physiography. The region's streams can then also be considered physically similar. The watershed approach has become recognized as an important spatial framework for managing water resources, but it still must be done within a framework of natural ecoregions (Omernik 1995). The LSW is completely contained within the Northern Lakes and Forests USEPA Ecoregion, # 50 (Omernik 1987). However, Omernik's original ecoregions have often not been sufficiently detailed for biological assessment (Omernik 1995). A finer-scale ecoregion, landscape ecosystem subsection X.9, named the North Shore (Lake Superior) Highlands (Albert 1995), is fully contained in USEPA ecoregion 50 and has similar dimensions to the LSW, except that the LSW boundary extends somewhat further inland than the subsection in the northern portion of the study area.. This same area is the North Shore subsection scale unit of the Minnesota Department of Natural Resources Ecological Classification System. Thus, the majority of the LSW can be considered quite homogeneous based on two geographical classifications of different scale. Our sample sites generally were contained within the intersection of the LSW and North Shore subsection. Four streams had headwater portions in an adjacent subsection, though sample reaches were in the North Shore subsection, while 2 streams had sample reaches and corresponding upstream watersheds outside of, though very close to, the North Shore subsection.

## Methods

*Stream selection and classification*

Reference streams

Proper selection of reference streams is critical in developing biocriteria (Yoder and Rankin 1995). Reference streams were not selected at random but were subjectively evaluated for high quality (Hughes 1995, Karr and Chu 1999). Streams chosen for establishing the reference condition were minimally influenced by human activity. Even in this relatively pristine area, it was difficult to find completely undisturbed streams. Possible disturbance factors include logging, roads and ditching, residential development, agriculture, introduction of brook trout (*Salvelinus fontinalis*), and stream habitat modification for trout.

The LSW was divided into three equal segments and a similar number of streams were chosen from each segment to evenly represent the LSW. Segments were numbered 1 - 3 from south to north. A total of 20 streams were included in the Reference set (Table 1). These included six streams sampled during 1989 - 1992, mostly from segment 1, and fourteen streams sampled in 1997 and 1998, mostly from segments 2 and 3. For 1997-98 sampling, logistical constraints necessitated that streams have easy road access.

Streams should be classified regarding their physical characteristics when forming biocriteria (Barbour *et al.* 1997, Karr and Chu 1999). One important classification is stream size. First, second, and third order streams were sampled, while the larger LSW streams were excluded. Stream order (Strahler 1957) was determined from 1:24,000 USGS topographical maps using tributaries marked in solid blue. In addition, a GIS was used to calculate watershed area from the sample reach. Reference streams ranged from 713 ha to 20596 ha with a mean = 4646.0, SD = 2624.9 (Table 7).

Originally, reference and disturbed streams were classified by several means, some of which were subjective and qualitative. Land use/cover data was available for some subwatersheds in segment 1 (Johnston *et al.* 1991, Richards and Host 1994). These data revealed a range of human influence in these streams. Where agricultural or urban land use occurred above trivial amounts in the watershed, streams were excluded from the Reference set.

Logging is a common land use in LSW headwater areas. Areas sampled generally had no noticeable nearby logging or disturbed riparian areas. Logging had occurred near the riparian zone and adjacent to the study reach on one side of Irish Creek, however, the stream was included in the Reference set because a significant buffer had been left along the river and the cut did not continue upstream of the reach.

Minnesota DOT county maps were used to select sites that had relatively low road densities (including mapped logging roads) in the subwatersheds. Streams chosen as references generally did not have roads paralleling them. Some streams had rural road crossings upstream of the sample site, especially in segment 1. Road runoff may have influenced these sites, though these crossings were generally at least 1.5 km upstream of the sample reach. Several of these crossings were gravel roads which do not receive winter salt applications. Many streams in sections 2 and 3 have only one mapped crossing other than the coastal highway.

Development influences were observed on some streams. Rural houses sometimes occurred near road crossings along streams. When this situation was encountered, sampling was done above the home site to eliminate possible impacts from the development (e.g. leaking septic systems).

Some biological disturbance may have occurred in LSW streams relative to pre-settlement times due to the introduction of brook trout (*Salvelinus fontinalis*), which are now wild. Native Lake Superior brook trout were likely found only in the very lowest reaches of tributaries due to barrier falls near stream mouths which prevented trout from colonizing upstream reaches. Trout conceivably could affect macroinvertebrate abundance or community composition by predation. However, brook trout in a similar stream type did not significantly affect invertebrate taxa diversity (Allan 1982). Moderate trout habitat improvement occurred on a few study streams including within two reference reaches (Irish Cr. and Poplar R.). Improvements at Junco Cr. were a few hundred meters above the sample reach. In these streams, sampling was done away from the relatively few habitat structures to minimize possible influences. Since virtually all permanent LSW streams have trout populations, the presence of trout was disregarded. Due to sample site proximity to roads and the open fishing season, sample substrates may have had short term and probably minimal disturbance by wading anglers, two of whom were encountered at one stream.

Disturbed streams

Data were available or collected for eleven streams with disturbance levels beyond the background historical whole-watershed disturbance (Table 1). Disturbed streams were those that did not meet criteria for reference streams above. Significantly disturbed streams were somewhat difficult to find. Six streams were shown to be impaired to some degree in previous studies (Richards *et al.* 1992, Richards and Host 1994). Seven of the streams are within the city limits of Duluth or adjacent towns. Five of these are especially influenced by urbanization. Six streams had some agriculture and low density residential development within their watersheds (Johnston *et al.* 1991, Richards and Host 1994). Within the Disturbed set, streams were further classified, somewhat subjectively, as either agriculture with low density residential development (Ag/Rural) or Urban. Mean watershed sizes (ha) were: Disturbed = 3075.5, Ag/Rural = 4337.7, and Urban = 1561.0. Three Disturbed streams lie just outside of the LSW, a few km south of the LSW border, though still within Subsection X.9. Seven streams are in LSW segment 1 and one stream, the Flute Reed R., lies in segment 3. It appeared to be organically enriched because of the presence of large amounts of filamentous algae and elevated turbidity. MnDNR personnel also report it to have compromised water quality (S. Persons, pers. comm.). Detailed water quality data to confirm impairment were not available for most LSW streams.

GIS confirmation of classification

After data analyses were conducted using the original watershed classifications, a GIS was used to quantify land cover in both the Reference and Disturbed watersheds to confirm the classifications. The coverage used was Minnesota Land Use and Cover: 1990's Census of the Land, produced in 1999 by MnDNR from several data sources. It has 8 classes of land cover. Road percentages were quantified with TIGER data. The MnDNR 1995 watershed coverage was modified to delineate the actual watershed area for each particular sample site, using a 30 m DEM to digitize boundaries.

*1997-99 Sample Collection*

Sampling methods generally followed the protocol of Karr and Chu (1999). Single habitat sampling was conducted, with riffle habitat chosen because of the high taxa richness typically found in riffles. Middle reaches were sampled because of the abundance of riffle habitat found there and because existing data were collected from these areas. An effort was made to sample reaches not immediately receiving lake water inputs. Sample reaches were generally a minimum distance of 3 - 4 km from significant lake water inputs. The reaches were also generally at least 50 m upstream from the bridge crossing. In two instances, good riffle habitat was absent upstream and downstream reaches were sampled. These reaches were sampled about 75m from the crossing in an attempt to minimize road influences.

Physical characteristics of riffles varied among streams. Some streams had highly defined riffle/pool sequences while others contained pocket water with small patches of riffles between larger substrates. Therefore, samples were not taken at a consistent location in the riffle (i.e. always at the head, mid, or tail areas). In attempting to reduce variability, priority was given to finding riffles having similar substrate size, depth, and flow characteristics. Generally, riffles having substrate of roughly 5 - 10 cm diameter were sampled. A few streams had little riffle area of this type and somewhat larger or smaller sized substrates were sampled. No riffles having substrates smaller than about 2.5 - 3 cm were sampled. Current speed varied by site, but sample riffles had some amount of surface turbulence. Riffles with very fast current and shallow depth (< 2-3 cm) were avoided as this riffle type was not found at all sites. Water depths of 5 - 20 cm were sought, with 8 - 13 cm being the most common depth. These depths are at the low end of the depths suggested by Karr and Chu (1999).

A D-frame kick net was used to collect the macroinvertebrates. The net measured 12 in. (base), by 6.75 in. (height), by 6.5 in. (depth) with 500 μm mesh netting. Five replicates were collected working upstream. An approximately 0.09 $m^2$ (1 $ft^2$) area immediately upstream of the net was disturbed by hand, first scrubbing larger rocks individually with the aid of cotton gloves in the mouth of the net. Scrubbed rocks were quickly inspected to find remaining invertebrates. The remaining smaller rocks and sand were agitated with fingers to a depth of approximately 3-6 cm. The contents of the net were washed into a plastic tub and strained through a 500 μm mesh soil sieve to remove the water. This material was washed from the sieve into a polypropylene bag with ethanol from a squirt bottle.

Sampling was done in mid- to late-September. Water levels were quite low, as is typical for the season. In addition, 1997 and 1998 summers were abnormally dry. A fifth replicate was collected in 1998 for two streams having only four replicates collected in 1997.

*Sample processing and identification*

In the lab, the ethanol was decanted from the sample bags into a 500 μm or 250 μm soil sieve such that most material remained in the bag. Material from the sieve and bag were washed with water into a glass tray. Macroinvertebrates were picked from the debris with the aid of a 2x magnifier against a white and black background to facilitate finding the invertebrates. All invertebrates were picked from each replicate, as suggested by Karr and Chu (1999). All samples were rechecked and the remaining invertebrates (generally early instars or chironomids) were collected.

Insects were the only macroinvertebrates identified and used in the study, as they are the primary invertebrates in LSW streams. Non-insect taxa comprise less than 5 % of the total taxa (Richards and Host 1994). Insects were identified to genus using keys in Merritt and Cummins (1996). Chiromomidae were identified only to family. The number of individuals in each taxon was recorded for each replicate. Most Gomphidae were early instars and difficult to identify to genus. *Ophiogomphus* appeared to be prevalent and all Gomphidae were classified as *Ophiogomphus*. Classification of Ceratopogonidae was somewhat difficult, partly because those collected often keyed to an endpoint with two genera, *Bezzia* and *Palpomyia*. Others seemed to key to *Probezzia*, though differences were very minor and not clear-cut. Based on habitat preference for the three taxa in Merritt and Cummins (1996), all Ceratopogonidae were classified as *Palpomyia*. Two early instar Trichoptera taxa could not be positively identified to genus. One appeared to be and was considered Limnephilidae while the other was not conclusive and was named "unknown 1".

*Historical data*

Historical macroinvertebrate data from LSW streams collected in 1989-92 were utilized to increase the number of reference and disturbed streams. These data were quite compatible with the later data with regard to sampling date, habitat, collecting gear, processing, and identification level. A few streams were sampled with 500 μm mesh Surber or Hess samplers. Influence of different sampling gear was believed to be non-problematic here since metrics were not usually dependent on absolute numbers, mesh size was the same, and whole samples were picked. Sample size was three, or more commonly four, instead of five. Ceratopogonidae, Empididae, Hydroptilidae, and Limnephilidae were only identified to family. For first three families, one genus was highly predominant in the 1997-99 samples so earlier identifications were converted to that genus. Metric values from two years of data were averaged for the Lester River.

*Metric and biocriteria development*

<u>Choice and development of metrics</u>

A large number of metrics are available for use in bioassessments (Barbour *et al.* 1997). Fifteen candidate metrics were chosen to test for utility in the LSW (Table 2). The candidate metrics were from four broader metric classifications: richness, tolerance, functional feeding group/habit, or population attributes. Most of the candidate metrics have been determined to be useful in detecting human-induced impairment of streams in several studies (Fore *et al.* 1996, Karr and Chu 1999), or are sometimes used in multimetric indexes. ABND, a moderately common metric (Barbour *et al.* 1997), has shown mixed results regarding usefulness in bioassessment, sometimes being effective in detecting impairment (Kerans and Karr 1994) while other times not (Resh and Jackson 1993, Fore *et al.* 1996, Rothrock *et al.* 1998). In addition to the often useful CLG, %CLG was assessed. %CHIR has been investigated commonly (Karr

Sample processing

The percentage of "rare" taxa per replicate was calculated for two levels of rarity; ≤2 and ≤4 individuals. These numbers were chosen because perfect sample homogenization, an unlikely scenario, would be required in order to consistently find these taxa in ½ and ¼ subsamples.

Chironomidae inclusion

Detailed chironomid identification is time intensive and thus was beyond the scope of the project. This presents an opportunity to assess the effectiveness of bioassessment without such identifications. The necessity or effectiveness of inclusion of chironomids was assessed in several ways. As mentioned, Chironomidae were only included in two candidate metrics, TAXA and % CHIR, and at family level. Some other metrics would not otherwise include chironomids (e.g. EPHEM), or could not because identification was not sufficient to properly class them (e.g. BI, % TOL). In addition to the among stream variation of % CHIR, mean abundance/replicate was calculated for each Reference stream and comparisons were made. Among replicate variation of abundance within streams was assessed with the CV. Lastly, an assessment was made regarding the effectiveness of the biocriteria and index when chironomids were included in only a minor way.

Reach location

Kimball Cr. provided an opportunity to examine possible longitudinal community differences in LSW streams. It has a largely undisturbed watershed down to the mouth and road access at 1) the intersection of the slow headwaters and faster middle section, 2) very close to the mouth at Lake Superior, and 3) a location about halfway between these points. Straight-line distances between upstream and middle reaches, and middle and lower, were 1.2 and 2 km respectively.

The three reaches had visibly different geomorphology. The upper site was relatively low gradient pocket water with slow water upstream, while the middle site was within a continuous stretch that contained well defined riffle/pool sequences. The near-mouth reach had higher gradient, with nearly continuous riffle/run habitat. Kimball Creek has been divided up into sections of similar geomorphology by the USFS (unpublished data). Each of the three study reaches was contained in a different USGS section. Two of these sections differ by Rosgen (1994) classification (upstream = C, middle = B, downstream = B; USFS data). We calculated gradient of the reach and immediately adjacent channel from USGS topographical maps by measuring stream length from the 2nd elevation line above the sample site to the 2nd elevation line below and dividing the difference by stream distance between points. These gradients were 31.0, 29.2, and 56.5 m/km for the upper, middle, and lower reaches respectively. Therefore, the upstream and middle reaches differed by channel type (related to gradient), and the middle and lower reaches differed by measured gradient.

Riffle habitat was sampled in each reach. Variances within metrics among upstream, middle, and near-mouth sites were much different, suggesting different sampling distributions (Elliott 1977). Therefore, data from the three reaches were compared using the non-parametric Kruskal-Wallis ANOVA, which is little affected by unequal variance (Zar 1999).

## Results

*General LSW aquatic insect community attributes*

A total of 84 insect taxa (81 genera) were identified from the Reference streams and 91 taxa (88 genera) for all streams. The unknown Trichoptera taxa is not included in these figures. Mean composite richness for Reference streams was 36.6 (STD 5.63) with a range of 29 - 51. The five streams with the highest composite taxa richness were Gooseberry R. (51), Kimball Cr. (43), Sucker R. (43), Cascade R. (42), and Junco Cr. (41). Composite taxa richness for 5-replicate Reference streams is given in Table 3. Abundances of the various insect taxa varied greatly. The percentage of taxa in the combined orders Ephemeroptera, Plecoptera, and Trichoptera was high. The Reference stream mean was 73.9 % with a range of 68.0 - 81.4 %. This mean would be lower had chironomids been included below family level. Chironomids comprised approximately 33.9 % percent of the total abundance in Reference streams.

Our findings regarding community composition are similar to those of previous studies. The order percentages observed were similar to those reported by Smith and Moyle (1944). Waters (1986) also found high abundance of the combined orders Ephemeroptera, Plecoptera, and Trichoptera, and of chironomids in three Cook Co. (LSW) streams.

*Analysis of metrics*

Pearson's correlation coefficients are summarized in Table 4. The EPT, INTOL, and CLG metrics were strongly correlated ($\geq 0.93$) with TAXA. The EPT metric was also strongly correlated with CLG, TRIC, INTOL and %TOL. Correlations of TAXA with EPHM and PLEC were weak, though the correlations with TRIC and %TOL were just above the strong correlation threshold (0.81 and 0.82 respectively). Most other correlations were weak.

Among stream CVs (Fig. 2) were generally low for richness metrics, between 0.133 - 0.261. CVs were higher for composition and functional metrics, from 0.291 - 0.789, with exceptions of 0.172 for %DOM and 0.169 for %CLG. The %CHIR metric had a moderate CV of 0.448. CVs for all candidate metrics were <1.00, the exclusion boundary.

Results of the Mann-Whitney and Kruskal-Wallis tests comparing individual metrics values from the Reference streams to the Disturbed streams and Ag/Rural or Urban streams respectively were generally significantly different at $\alpha = 0.05$ with most significantly different at $\alpha = 0.005$ or lower (Table 5). The %TOL metric was not significantly different for either test but the Kruskal-Wallis $p = 0.052$, suggesting that Urban group, which systematically had higher values, was nearly significantly different from the Reference streams while the Ag/Rural group was not. The %CLG and %CHIR metrics were not significantly different in either test.

Metrics were further examined by comparing plots of metric values grouped by stream disturbance types. In general, Urban streams were better separated from the Reference streams than the Ag/Rural streams (Fig. 3). There was almost always at least a small amount of overlap between the Reference and Disturbed values. TAXA, TRIC, PLEC, ABND, INTOL, BI, and CLG showed good separation. The EPT metric had the best separation, with no overlap between the Reference and Disturbed stream values. The HYD/TRI values overlapped most of the Reference range, however many of the values were at the upper end or outside of the Reference range. The %TOL metric had a similar overlap pattern, yet six of the Disturbed streams scored outside of the 75th percentile. The %PRED metric separated 6 of 11 Disturbed sites below the $25^{th}$ Reference percentile. Though the %CLG metric, investigated as an alternative to CLG, was weakly correlated with TAXA and had a very low CV, it did not behave as expected. Only 3 of the 11 Disturbed streams had lower %CLG values than the 25th Reference percentile and 5 of the 11 values exceeded the highest Reference value. With the exception of the EPHEM and BI metrics, the Ag/Rural streams were generally clustered at the low end and outside of the Reference range for the accepted metrics.

From the above analyses, five metrics were excluded from the final index. EPT, INTOL, and CLG were dropped because of high correlation with TAXA. TAXA was retained over the others because of its widespread use and effectiveness. Lack of statistically significant differences between Reference and Disturbed sites, and poor separation of Reference and Disturbed streams for both %CHIR and %CLG led us to exclude these metrics from the index.

Ten metrics passed our acceptance criteria and comprised the final index. Though TRIC and %TOL correlations with TAXA fell in the strongly correlated range, they were included in the index because they were just above the strong threshold and were not strongly correlated with any other candidate metric. In addition, these metrics were not strongly correlated with TAXA for the Ag/Rural or Urban stream sets ( $r =$ 0.40 and 0.55 respectively for %TOL, and 0.74 and 0.79 for TRIC). Since the correlations were not strong for the Disturbed streams, there was not a duplication or magnification of scoring effect by a redundant measurement. An allowance was also made for the %TOL metric regarding the statistical difference criteria since the Kruskal-Wallis p-value was so close to 0.05. The final index retains metrics from the four metric classifications: richness, tolerance, functional feeding group and habit, and population (Table 2). Biocriteria and scoring divisions for accepted metrics are presented in Table 6.

*Index Biocriterion and comparison to Disturbed streams*

Reference streams scored between 38 and 48 (Fig. 4) with a mean score of 44.9. Final scores of the Urban streams were much lower than Reference streams, ranging from 12 to 24 with a mean of 19.2. Thus, some of these streams scored near the minimum possible score. Streams suspected of being impaired due to agriculture and low density residential development (Ag/Rural) scored either slightly lower or at the very low end of the Reference condition, ranging from 28 to 40 with a mean of 34.0. The range of the Reference final scores was quite narrow, especially the interquartile range, suggesting that the Reference streams are quite similar. The range for the Disturbed streams was much wider, as might be expected due to the fact that there were varied levels of disturbance in their watersheds.

Box plots show a definite difference in final score ranges (Fig. 4). Reference and Disturbed streams showed highly significant difference (Mann-Whitney $p < 0.001$). A Kruskal-Wallis test with the Disturbed streams divided as Ag/Rural and Urban also had strong statistical difference ($p < 0.001$). A Dunn test (Zar 1999) for significant differences among pairs of stream sets showed that Reference vs. Ag/Rural was significantly different ($p < 0.01$), Reference vs. Urban was highly significantly different ($p = 0.001$), while Urban vs. Ag/Rural was not significantly different at ($p > 0.50$).

*GIS land cover analysis of study watersheds*

The GIS land cover analysis confirmed our stream classifications as disturbed land was consistently much higher in Disturbed classified subwatersheds (Table 7). This is partly to be expected since GIS data was originally used to classify some streams. Urban classified streams were within the city limits of Duluth and had the greatest developed land, ≥ 7.6 %, and developed + hay/pasture/grass, ≥ 26.2 %. Ag/Rural streams had a combined developed + hay/pasture/grass of > 4.9 %. All Reference streams had developed + hay/pasture/grass values ≤ 3.2 %, with most ~ 1.0 %. The Encampment R., a Reference stream, could be placed into the Disturbed category since it had about 3 times more disturbed land than the typical Reference stream, though it had less disturbed land than any other Disturbed stream. Percent roads was correlated with the amount of combined developed land, much higher in Urban than Reference, with Ag/Rural intermediate, though closer to Reference. The exception to these results was the Flute Reed River, which though classified as disturbed, had low percentages of disturbed lands in the range of most Reference streams.

*Methodology investigations*

Sample size

Cumulative taxa richness continued to increase for 1 - 5 composited replicates (Fig. 5) and did not level off within the cumulative area sampled (max. area = 0.45 $m^2$). The taxa richness increase for a composited sample of increasing replicate number showed a systematic decrease, with the change in richness down to 1.6 when moving from a 4 to 5 replicate composited sample.

Sample processing

A relatively high number of taxa were found in low abundance in the replicates. The mean percentage of taxa represented by ≤ 2 and ≤ 4 individuals in a replicate for Reference streams was 34.0 and 48.0 % respectively. For Disturbed streams, these percentages were higher, 44.2 and 61.2 %.

Chironomid inclusion

Results of the %CHIR metric assessment are presented above. Chironomid abundance varied among Reference streams by more than an order of magnitude, ranging from a mean of 32.6 chironomids/replicate (E. Beaver R.) to 663.4 (Cascade R.). The mean chironomids/replicate among streams was 165.4, with the CV = 1.030. Chironomids/replicate variation was also fairly high within streams, with a mean CV of 0.559 among Reference streams. Examples of the more severe cases of variation are East Split Rock R., which ranged from 28 - 177 chironomids/replicate and Kimball Cr., which ranged from 27 - 678.

Reach location

By simple visual inspection of invertebrate storage vials, it was obvious that biomass was much less in the lower site than the uppermost site. Seven of the 10 final index metrics varied systematically along the stream's longitudinal profile. In general, there was a progression of increasing TAXA, EPHEM, TRICH, ABND, HYD/TRI, and %PRED moving upstream. The PLEC and %DOM metrics showed no definite trend, though the low reach had the highest %DOM and lowest PLEC values. Though %DOM was much higher in the low reach, though the dominant taxa there were primarily very intolerant ones. The %TOL and BI metrics showed the reverse trend with the upper reach having higher values (Table 8). This was probably due to the increased abundance of Odonata moving upstream. The near-mouth, high gradient reach consistently had the lowest values for richness and abundance metrics. Kruskal-Wallis tests of the 10 final index metrics showed significant differences among the three reaches for 8 of the metrics at $\alpha = 0.10$. (Table 6). All eight metrics were significantly different between at least 2 locations at $\alpha = 0.05$. Significant differences typically involved the lower reach (8 metrics), but differences did occur between upper and middle reaches for 3 metrics (Table 8).

Calculation of final scores for the three reaches showed that the middle reach scored highest (48), with the upper reach similar but a bit lower (46), and the near-mouth reach scoring a fair amount lower than either (40).

## Discussion

*Analysis of metrics*

Several studies/reviews have recently sought to find metrics which are effective at detecting impairment (Barbour *et al.* 1992, Fore *et al.* 1996, Thorne and Williams 1997, Rothrock *et al.* 1998, Karr and Chu 1999, Kennen 1999, Maxted *et al.* 2000). Though comparison of metrics among studies is somewhat confounded by the different levels of macroinvertebrate identification, we found very similar patterns of effectiveness or lack thereof for metrics that have recently been tested in other geographic areas. For example, the B-IBI includes ten metrics, based on their systematic variation across a watershed disturbance gradient in several studies (Karr and Chu 1999). We tested nine of these metrics, all of which were effective at detecting impairment.

Richness metrics were particularly effective. However, some of these metrics were highly correlated and therefore lacked utility in a multimetric index for the LSW. Kerans and Karr (1994) found significant correlations between richness metrics but not between richness and composition metrics. The results of our correlation analyses were similar. In our study, EPT taxa comprised a high percentage of the total number of taxa in all streams and our correlation coefficients were very high between TAXA and three other richness metrics, EPT, INTOL, and CLG. Karr (pers. comm.) contends that high correlation is not automatically bad as long as the correlated metrics are providing different biological information. It was felt that these highly correlated richness metrics were not providing different information here since clingers and intolerant taxa generally are in the orders Ephemeroptera, Plecoptera, and Trichoptera. Barbour *et al.* (1992) also found that TAXA and EPT had a high correlation (0.83) in their minimally impaired study streams and suggested that this would occur in streams having a high diversity of these three orders. These correlations in our study may have been lower somewhat lower if chironomid genera had been included, if species-level identifications were made, or if non-insect taxa were included. In these cases, EPT taxa might comprise a smaller proportion of the taxa. Therefore, these correlations may not occur here if other identification protocols are followed. Taxa Richness has been effective in studies using identification protocols similar to ours (Resh and Jackson 1993, Kleindl 1995, Patterson 1996).

The EPT richness metric, though not included in the index, separated the Reference and Disturbed streams better than any other candidate metric. It was the only metric that had no overlap in the ranges of Reference and Disturbed values. Use of an index is felt to be a more robust assessment of stream condition (Resh and Jackson 1993, Barbour *et al.* 1995), but it appears that using the EPT metric could be a good quick-screening procedure for assessing stream health in the LSW. The EPT metric has been useful in Ohio's ICI (DeShon 1995) New York's biocriteria (Bode and Novak 1995), and in North Carolina's water quality assessment program (Lenat 1993). It is recommended as a good metric 'in terms of effort relative to information' in a rapid assessment (Resh 1995).

Tolerance based metrics were also generally effective. Barbour *et al.* (1992) found the Hydropsychidae/Trichoptera metric useful in detecting impairment and suggested using it in place of the variable EPT/Chironomidae metric. A metric of *Hydropsyche+Cheumatopsyche*/Trichoptera worked well in a study by Maher (pers. comm.). Our modified version of this metric to Hydropsychinae/Trichoptera performed quite well in our study. Some of the overlap with Reference values could be due to the fact that hydropsychid species vary in pollution tolerance, even within Hydropsychinae (Hilsenhoff 1987). It appears that Hydropsychinae species in the LSW streams are generally those which have lower tolerance values. The moderately tolerant *Ceratopsyche slossonae* (HBI value = 4) was a common species based on limited species level ID for some streams and *Ceratopsyche* dominated over the generally more tolerant *Cheumatopsyche* and *Hydropsyche*. Also, since hydropsychids are filter feeders, perhaps much of the sample variation is due to specific characteristics of riffle habitat, such as current velocity, which were not selected for in detail in our study. Similarly, the effectiveness of the %TOL metric may be somewhat reduced for detecting impairment because the water quality of the broader area has historically been high, and therefore there may be relatively few tolerant taxa indigenous to the area and available to colonize degraded streams. Though redundant with TAXA, %TOL was retained using Karr's rationale above, since it was unclear why these should be positively correlated. The ineffectiveness of the %CHIR metric is discussed below.

The functional group/habit metrics showed mixed usefulness. While the CLG metric was effective here (though redundant) and elsewhere (Karr and Chu 1999), the alternative %CLG was not. Clinger taxa should be impacted by increased sedimentation filling in substrate crevices used by clingers and thus %CLG should decrease with impairment. This result was not observed. For example, Amity Cr. had the highest substrate embeddedness of streams measured by Richards and Host (1994) yet it had nearly the highest %CLG value of all study streams. Our results with this metric differ from those of Maxted *et al.*(2000), where it was effective. %PRED was useful as has been found in several other studies (Karr and Chu 1999).

The two population metrics were useful. ABND performed well for LSW streams, though typically it performs poorly (Resh and Jackson 1993, J. Karr, pers. comm. 2000). Its effectiveness here may be due to the very high proportion of Ephemeroptera, Plecoptera, and Trichoptera in our samples. Kennen (1999) found a metric of EPT abundance to be effective in bioassessment. %DOM (3 taxa), effective about half of the time elsewhere (Karr and Chu 1999), performed well here. Other studies have found versions of this metric to be both useful (Rothrock *et al.* 1998, Kennen 1999, Karr and Chu 1999) and not (Thorne and Williams 1997, Maxted *et al.* 2000).

Barbour *et al.* (1992) suggest assessing metric effectiveness with the CV. We observed the same pattern of higher CVs for proportional metrics versus richness metrics as they found. There is also strong similarity of CVs for our four metrics-in-common. Our CVs were also similar to those reported by Rothrock *et al.* (1998) from streams with varied impairment levels for our three metrics-in-common (Taxa Richness, Abundance, and HBI), though ours were all were somewhat higher. Rothrock *et al.* (1998) also found that the CV was higher for their lone proportional metric.

Our study supports other evidence that metrics found promising elsewhere are indeed effective across different regions and in streams of different types. Maxted *et al.* (2000) completed a study which evaluated many metrics, eleven of which were similar to ours. Two significant study differences were 1) substantially different stream habitat and thus habitat sampled (unconsolidated substrate v. rocky), and 2) inclusion of genus-level chironomids in certain metrics. Despite these differences, most of these metrics showed similar results as to diagnostic usefulness. The 7 metrics sharing effectiveness in detecting impairment were: INTOL, %TOL, EPT, EPHEM, TRICH, HBI, TAXA. Additionally, both studies found %CHIR at least sometimes not useful. Three metrics behaved oppositely: %CLG, %DOM, and HYD/TRI (though subtly different in composition between studies).Though neither study used all of these metrics for reasons described by the respective authors, it is of interest that the set performed so similarly across different stream types and contributes to the assessment of their worthiness to be included in a pool of metrics with general utility across wide geographic areas. Also of interest is that relevant metrics behaved similarly despite our not having included Chironomidae, suggesting that it may be valid to exclude them and thus attain time and cost savings in bioassessments.

Despite the finding that many metrics perform well across wide geographic areas, the importance of locally tailoring an index (Resh and Jackson 1993, DeShon 1995, Karr and Chu 1999) rather than blindly applying one developed elsewhere is also confirmed by comparing our results with other studies and developed indices, since not all metrics behaved the same between geographic areas. Even when there is good or complete agreement of metric effectiveness, as with the CPMI (Maxted *et al.* 2000) and B-IBI respectively, our correlation analysis showed that several of these metrics were redundant in the LSW. Stream habitat differences or assemblage differences among regions are possible factors in metric usefulness discrepancies. For example, Maxted *et al.* (2000) note that Plecoptera are rare and add minor contribution to their metrics, while they are abundant and important in the LSW.

*Biocriteria*

Some studies define narrative classes of impairment (i.e. non-, moderate, or severe) based on the final scores (Barbour *et al.* 1997), but no such specific classification was made in our study. We derived a numerical biocriterion, the lowest final Reference stream score, which results in test streams either being classed non-impaired, impaired, or impairment uncertain. Kerans and Karr (1994) suggest that further classification is subjective and that a feel for the level of impairment can be gained by simply looking at the relative scores.

*Comparison of Disturbed streams to the Reference condition*

The final multimetric index functioned quite well, as most Disturbed streams, with some level of suspected impairment, had lower final scores than the Reference range. The Urban streams were especially well separated from the Reference streams. Kennen (1999) found that the probability of detecting impairment correlated positively with area of urban land upstream of the sample site. The index also detected impairment in most of the Ag/Rural streams, which is fairly impressive since agriculture is non-intensive in the LSW and typically forms only a small percentage of these watersheds. Rothrock *et al.* (1998) found that aquatic bio-integrity was compromised in agricultural watersheds relative to "pristine" and silvicultural watersheds. One Disturbed (Ag/Rural) stream, the Flute Reed R., scored within the Reference range, though at the very low end. Though observations of the stream and watershed reveal it to be somewhat impaired, it was not detected with our biocriteria. Barbour *et al.* (1995, 1997) and DeShon (1995) state that the individual metric scores should be evaluated in addition to the final score. Individual metric analysis was inconclusive for this stream. Several of the metrics which received scores of 3 were quite close to the 5 range, so the stream came close to scoring even higher. Barbour *et al.* (1997) suggest that some Reference streams will periodically score relatively low. It is possible that the Reference stream which scored below the Flute Reed R. scored lower than they typically would. Unfortunately, we lacked independent data for testing additional disturbed streams to further examine the effectiveness of our index. This was in part due to the limited number of disturbed streams in the study area. Most available disturbed (all the Urban) streams of large enough size were used in the development of our biocriteria.

Buffers are often used when classifying sites with regard to impairment severity. The ICI of Ohio (Yoder and Rankin 1995) allows a 4 point buffer below the Reference condition, such that streams within this borderline region are not considered impaired. A buffer of similar percentage relative to the total score for our study is 3 points. Two other Disturbed streams scored just 2 points below the Reference condition. Therefore, as judged by the final score biocriterion with this buffer, 3 of the Disturbed streams were not detected by the biocriterion. Without ancillary data to determine that water quality and habitat are degraded in these streams, it is difficult to tell whether the biocriteria actually failed or whether these streams were in fact not impaired.

All Reference streams scored in Hilsenhoff's (1987) "Excellent" HBI water quality category. Interestingly, the Disturbed streams all scored as "Excellent" or "Very Good". It appears that the HBI analysis alone may not be sufficient for detecting impairment at disturbance levels or for disturbance types found in the LSW study area. Because the HBI measures organic pollution, this result suggests that organic pollution may not be the primary factor responsible for impairment seen in the Disturbed streams.

Many invertebrate metrics were shown to correlate with watershed size at the sample site in Ohio's ICI (DeShon 1995). How well this holds for the LSW and our sampling methods is unknown. It is possible that such an effect influences our results with regard to the conclusion of impairment in some streams since these streams were generally on the small end of the range of streams included in the Reference set. However, first order Reference streams scored considerably higher in general than the small disturbed streams. Also, the metrics do not change a great deal over the range of watershed sizes we studied (DeShon 1995). Maxted *et al.* (2000) did not find catchment-size influences in their metrics when the larger streams were omitted (as we also omitted). Lastly, the same habitat was sampled in all streams which should reduce the influence of habitat variability related to stream size. Since several of the Disturbed streams scored well below the Reference range, they likely were actually impaired.

The GIS analysis of land cover showed that LSW streams begin to show impairment at quite low % watershed disturbance levels. Impairment determinations were able to be made for streams with as low as 11.9 % developed + hay/pasture/grass lands, and 14.9 % developed + hay/pasture/grass + roads. Roads were not included in the "developed" classification. Since road percentages followed a similar pattern as other disturbed classes, impairment may be related to roads as well. Both of these classifications probably contribute to the observed impairment. The discrepancy in land cover and classification (based on water quality observations) for the Flute Reed R. may be due to proximity of disturbance to the stream. A road closely follows the streambed and has badly eroding ditches, perhaps therefore having stronger influence than road percentage would suggest. This stream did not score as impaired, but did score low in the Reference range. These results suggest that LSW streams are quite fragile environments which can be easily degraded and which may require close management attention to prevent their decline.

*Methodology investigations*

Sample size

Stream investigators frequently face the problem of deciding how many replicates to collect. If a small sample area can effectively represent the community of a reach, bioassessment costs can be reduced. Sampling a larger area may increase the ability of the metrics to detect more subtle impairment by maximizing richness differences, especially if a composited sample is analyzed. An argument has been made that sampling a small area of habitat does not give a good estimate of richness (Larsen and Herlihy 1998). Barbour *et al.* (1997) recommend collecting an area of 2 $m^2$, which is significantly larger than that collected here (0.45 $m^2$). Larsen and Herlihy (1998) report that with genus level insect identification and inclusion of other macroinvertebrates at family or order, "1-2 new taxa were found, per Surber sample, even after 40-50 samples had been collected." In contrast, Karr and Chu (1999) concluded that three Surber samples (0.27 $m^2$) was adequate to represent a reach and that a sample size of five did not add much precision. Since five replicates were collected for this study, an analysis was made to see if this conclusion is true with generic level identifications for the LSW. Our results did show increasing taxa richness through the maximum of 5 replicates collected, so there may be some additional benefit to collecting more replicates. However, it appears that this benefit may be small for LSW streams since genus-level taxa richness increase was reduced to near 1.5 taxa with the addition of the $5^{th}$ replicate. It appears there may be one or fewer new taxa on average with a sixth replicate. Apparently, LSW streams are somewhat taxa-poor in comparison those studied by Larsen and Herlihy (1998). Collecting a sixth replicate here is likely not cost-effective. Therefore, a 2 $m^2$ sample area is likely not necessary for LSW streams.

Sample processing

Subsampling saves time and resources and some consider it to be an appropriate approach provided that certain methods are followed. Karr and Chu (1999) suggest that picking the whole sample is better than subsampling, in part because missing rare or intolerant taxa "would exclude some of the strongest biological signals about the condition of places." Our study seems to support their opinion, since we found rather high percentages of rare taxa in our replicates. It appears that several of those taxa collected in a sample replicate would often be missed in processing, resulting in lower richness values. Without perfect sample homogenization, there is at least a 1 in 4 chance of missing taxa with ≤4 individuals if a ¼ subsample is picked. This increases to a 1 in 2 chance for taxa with ≤2 individuals if a ½ subsample is picked. If a replicate has 6 taxa with ≤ 2 individuals (not uncommon here), at least 3 taxa would often be missed. Because several metrics rely on richness differences to reveal impairment, reducing the differential would weaken the ability of the biocriteria to detect less obvious impairment or to detect impairment early. Indeed, Cao *et al.* (1998) found that taxa richness differences between reference and impaired sites were lessened when rare taxa were deleted from the analysis. Larsen and Herlihy (1998) have discussed the influence of whole vs fixed-count enumeration on taxa richness indicators. They have shown that when taxa evenness is low, taxa richness values will be low relative to the actual richness of the sample when smaller counts are performed. Because of the high natural % Dominance values found in replicates of the present study, a fixed-count subsampling procedure could be quite biased toward high abundance taxa and therefore many other taxa would be missed, weakening the ability of richness metrics to detect impairment. The high % Dominance and rare taxa percentages in the Reference replicates found here suggest that it is important to pick whole sample units for LSW streams.

Chironomid inclusion

Processing and identifying Chironomidae can be very time consuming and therefore, costly. Investigators may try using chironomids at the family level in order to save funds. However, metrics using chironomids at the family level (e.g. % Chironomidae and EPT/Chironomidae) have often not been able to distinguish between unimpaired and impaired sites (Barbour *et al.* 1992, Resh and Jackson 1993, Thorne and Williams 1997, Karr and Chu 1999, Kennen 1999). These metrics were found to be highly variable among sites in the above studies/reviews. CVs among replicates for EPT/Chironomidae typically were higher relative to several other metrics in a study by Rothrock *et al.* (1998). Also, year to year variability in proportion of chironomids was found to be very high by Kerans and Karr (1994). The %CHIR metric was not significantly different between Reference and Disturbed streams in the present study. This finding goes against the theoretical expectation of an increase in the percentage of chironomids in disturbed streams (Fore *et al.* 1994, Karr and Chu 1999). In addition, chironomid abundance did not respond as expected, since many of the Disturbed streams had fewer rather than more chironomids than Reference streams. Variability of chironomid abundance was fairly high within streams and quite high among streams. Therefore, the use of metrics that incorporate chironomids at the family level, such as %CHIR, is probably not beneficial in the LSW either. Resh and Jackson (1993) concluded that the % chironomids metric is "not consistently accurate enough to be used effectively in rapid assessment approaches." Because of the high variability, including chironomids in other metrics such as ABND may increase variability, thus obscuring impairment trends.

It appears that when chironomids are not identified below family, excluding them from the analyses would reduce expenses while still allowing a solid assessment of stream health. Detection of impairment was still possible in the present study with only minimal use of chironomids. Even when identified more precisely, it may be beneficial to exclude them from abundance metrics because of the variability in their abundance. Kerans and Karr (1994) stated that Chironomidae "need to be investigated further before their utility in indexes can be determined fully."

Field notes of study riffle characteristics suggest that detailed characteristics such as velocity and periphyton standing stock may affect chironomid abundance. For example, Beaver R. had relatively fast riffles and few chironomids. Fast riffles on Cascade R. had few chironomids while slower riffles with more periphyton had high abundances of chironomids.

Reach location

Because LSW stream geomorphology changes quickly and significantly from headwater to mouth, it may be useful or necessary to stratify stream reaches even within this rather fine-scale landscape ecoregion. We found that eight of the 10 index metrics differed between upstream or middle and near-mouth reaches on Kimball Cr., while few metrics differed between upstream and middle reaches (upstream was actually in the "middle" section of the river as defined earlier). These results suggest that our biocriteria as developed may not be applicable to all LSW riffle sites, but rather to middle reach riffles specifically. Replication at additional LSW streams is needed before a conclusion can be made, however in a sense we replicated for this effect by testing many metrics for reach differences. In addition, Waters (1986) found that invertebrate biomass was consistently greater in upstream sites than the near-mouth sites on three LSW streams; Kimball Cr. and 2 adjacent streams. These findings together suggest that one or more physical factors are influencing the invertebrate community. Marchant *et al.* (1999) found that longitudinal gradient affected the macroinvertebrate community over a large spatial scale and cited several other studies with similar results. They conclude that "it seems longitudinal [macroinvertebrate] gradients are present no matter what spatial scale has been used in the study". Their large spatial scale analysis did not reveal a trend of taxa richness with river slope, but did with substratum character, though that character varied much more than ours did within Kimball Cr. Marchant *et al.* (1999) report in an earlier study of a smaller spatial scale more similar to ours, that macroinvertebrates were correlated with variables such as benthic organic material and mean particle size of substratum. The latter seem to fit with our site as the upstream and near mouth sites appeared to differ regarding the amount of organic material and small substrate particles (pers. obs.).

Such longitudinal influence on metrics could be very influential in bioassessment because it would then be helpful to carefully select similar sites in order to reduce biocriteria variability and enhance their discriminatory power. Care would also need to be taken to compare only similar sites of test streams to avoid erroneous conclusions. We found that the lower site, though not likely influenced by watershed disturbance, had a final score lower than the middle and upper sites. The upper reaches had among the highest Reference scores, while the near-mouth site was at the very low end of the Reference range, near the "gray area". Karr and Chu (1999) have stated the importance of classifying streams by several factors including gradient. Community differences in LSW streams may be due to the hydrologic regime differences related to the high longitudinal variability of topography and gradient (i.e. flow velocities and the associated substrate types and sizes, and CPOM) of these streams.

Karr and Chu (1999) discussed the need to strike a balance in classification between too coarse and too fine due to excess variability and high cost respectively. Because of the particularly high value of LSW streams, it may be useful to stratify streams at the reach level rather than at the ecoregion level to get greater resolution of the reference condition and thus potentially better ability to detect impairment or discover impairment earlier. Marchant *et al.* (1999) suggest that "unless there is a close correspondence between the patterns of distribution of the biotic communities and ecoregions, such a strategy [ecoregion-based sampling] is likely to confuse rather than simplify the monitoring of river environments." It may be that the heterogeneous topography of the LSW study area could lead to some such confusion. Of added interest, development is occurring especially along the near-shore areas of the LSW, so having biocriteria developed for near-mouth reaches may be important. Therefore, watershed or cultural characteristics unique to local regions may warrant stratification of sites beyond even a fine-scale ecoregion strata, especially in watersheds of high value, where the added cost may be justified.

*Future refinements of biocriteria*

This is a first attempt at defining biocriteria for the LSW. Further refinement of these biocriteria may be beneficial. According to casual observation, stream order does not seem to describe stream size well in this watershed. Others have noted that stream order is not a very good measurement for determining stream size (Hughes and Omernik 1981, DeShon 1995). As discussed, stream size (as determined by watershed area) can influence metric values (DeShon 1995). Therefore, future studies in the LSW should classify streams according to watershed size and construct separate biocriteria for each classification if differences are found or if the largest LSW streams are included. The number of streams sampled in the present study is probably not sufficient for determining this. Also, a broad range of stream sizes is necessary to determine trends. Since the smallest and largest streams were excluded from sampling in the present study, trends may not be revealed were this done with the present data set. Classification of reaches according to gradient may also increase the precision of the metrics.

Perhaps more stringent criteria should be used for reference stream classification. Based on the land cover analysis here, most of our Reference streams had disturbed lands of less than 1.0 % of the watershed area. Since four of our Reference streams had 1 - 5 % disturbed land, our biocriteria may be more liberal than they should be, which may reduce the ability of our index to detect impairment.

The current study was based on data from a single year per stream. Data from multiple years should be obtained and averaged in order to account for any natural year to year variability in metric values and to minimize the effects of this temporal variation. Our data did contain some variation in sampling years.

Scoring divisions may need to be redefined for some metrics. Scoring for individual metrics can be a complicated process because metrics do not all respond in a linear fashion over a range of impairment (Karr and Chu 1999). The equal scoring divisions used in the present study assume a linear response, which is a possible source of error. However, discovery of non-linear relationships of metrics to impairment severity requires values from a collection of many sites having a gradient of impairment. In the LSW, such a range may be difficult to find in since much of the area is relatively undisturbed. The current literature does not contain much generalized information on the detailed relationships of metrics along disturbance gradients.

Additional metrics should perhaps be analyzed to see if they also are able to detect impairment in LSW streams. Though we tested many of the promising metrics, other metrics have been found to be useful in detecting impairment (Barbour *et al.* 1992, Karr and Chu 1999). The addition of these metrics may increase the separation of the Disturbed stream final scores from the Reference scores.

Finally, as physical and chemical data become available, they should be compared to final scores of LSW streams to assess how well the determination of "impaired" corresponds with similar determinations using these more conventional means. This would allow further examination of the effectiveness and importance of using biocriteria to monitor water quality. Biocriteria may detect impairment that chemical assessments miss (Yoder 1991).

*Conclusions*

Our study was an initial effort to characterize the reference condition of LSW streams and develop a multimetric index useful for detecting stream impairment. Even so, several important findings resulted from it. 1) Metrics previously found to be effective at detecting stream impairment in other geographic locations were also effective in the LSW, though some were highly correlated. 2) Our study confirms that "rapid bioassessment", as advocated by Barbour *et al.* (1997) and Karr and Chu (1999) is a useful water resource monitoring tool. Efforts were streamlined by sampling one habitat, sampling a relatively small area of stream bottom, genus-level insect identification, and particularly, family-level identification and minimal use of Chironomidae. Our results further substantiate the use of the sampling and processing protocol recommended by Karr and Chu (1999). This scale of effort was effective as the final multimetric index detected strongly impaired (Urban) streams and in general also detected even moderately impaired (Ag/Rural) ones. The effectiveness of the index without conducting detailed chironomid identification has significance for cost-reductions of bioassessments. 3) We discovered special characteristics of the area's macroinvertebrate community which have implications for sampling stratification, the lab processing of samples, and metric selection which reveal the importance of regionalizing the procedures involved in bioassessment. 4) It appears that LSW streams are fragile environments which therefore require careful management of land use to prevent degradation. 5) Biocriteria, such as were developed in this study, appear to be an effective way to monitor the biological integrity and water quality of water resources in the LSW.

## References


Albert, D. A. 1995. *Regional Landscape Ecosystems of Michigan, Minnesota, and Wisconsin: A Working Map and Classification.* Gen. Tech. Rept. NC-178. US Dept. of Agriculture, Forest Service, North Central Forest Experiment Station, St. Paul, MN.

Allan, J. D. 1982. The Effects of Reduction in Trout Density on the Invertebrate Community of a Mountain Stream.. *Ecology* 63(5):1444-1455.

Barbour, M. T., J. L. Plafkin, B. P. Bradley, C. G. Graves, R. W. Wisseman. 1992. Evaluation of EPA's Rapid Bioassessment Benthic Metrics: Metric Redundancy and Variability Among Reference Stream Sites. *Environmental Toxicology and Chemistry* 11(4):437-449.

Barbour, M. T., J. B. Stribling, and J. R. Karr. 1995. Multimetric Approach for Establishing Biocriteria and Measuring Biological Condition. Pages 63-78 in W. S. Davis and T. P. Simon (editors). *Biological Assessment and Criteria: Tools for Water Resource Planning and Decision Making.* Lewis Publishers, Boca Raton, FL.

Barbour, M. T., J. Gerritsen, B. D. Snyder, and J. B. Stribling. 1997. *Revision to Rapid Bioassessment Protocols For Use in Streams and Rivers: Periphyton, Benthic Macroinvertebrates, and Fish.* EPA 841-D-97-002. USEPA Office of Water, Washington, D.C.

Bode, R.W. and M.A. Novak. 1995. Development and Application of Biological Impairment Criteria for Rivers and Streams in New York State. Pages 97-107 in W. S. Davis and T. P. Simon (editors). *Biological Assessment and Criteria: Tools for Water Resource Planning and Decision Making.* Lewis Publishers, Boca Raton, FL.

Cao, Y., D. D. Williams, and N. E. Williams. 1998. How Important are Rare Species in Aquatic Community Ecology and Bioassessment? *Limnology and Oceanography* 43(7):1403-1409.

DeShon, J. E. 1995. Development and Application of the Invertebrate Community Index (ICI). Pages 217-243 in W. S. Davis and T. P. Simon (editors). *Biological Assessment and Criteria: Tools for Water Resource Planning and Decision Making.* Lewis Publishers, Boca Raton, FL.

Elliott, J. M. 1997. *Some Methods for the Statistical Analysis of Samples of Benthic Invertebrates.* Freshwater Biological Association, Scientific Publication Number 25.

Fore, L. S., J. R. Karr, and R. W. Wisseman. 1996. Assessing Invertebrate Responses to Human Activities: Evaluating Alternative Approaches. *Journal of the North American Benthological Society* 15(2):212-231.

Gibson, G. R. Jr. (editor). 1996. *Biological Criteria: Technical Guidance for Streams and Small Rivers.* EPA 822B-96-001. U. S. Environmental Protection Agency, Office of Water Regulations and Standards, Washington, D.C.

Hilsenhoff, W. L. 1987. An Improved Biotic Index of Organic Stream Pollution. *The Great Lakes Entomologist* 20(1):31-39.

Houghton, D. 1999. PhD. candidate, Dept. of Entomology, Univ. of Minnesota-St. Paul.

Hughes, R. M. 1995. Defining Acceptable Biological Status by Comparing with Reference Conditions. Pages 31 - 48 in W. S. Davis and T. P. Simon (editors). *Biological Assessment and Criteria: Tools for Water Resource Planning and Decision Making.* Lewis Publishers, Boca Raton, FL.

Hughes, R. M., and J. M. Omernik. 1981. Use and Misuse of the Terms Watershed and Stream Order. Pages 320-326 in L. A. Krumholz (editor). *Warmwater Streams Symposium.* Southern Division, American Fisheries Society, Lawrence, KS.

Johnston, C. A., B. Allen, J. Bonde, J. Sales, and Paul Meysembourg. 1991. *Land Use and Water Resources in the Minnesota North Shore Drainage Basin.* NRRI Technical Report NRRI/TR-91/07. Natural Resources GIS Laboratory, University of Minnesota - Duluth.

J-NABS. 1996. Bridges: Integrating Basic and Applied Benthic Science. *Journal of the North American Benthological Society* 15(3):381-399.

Karr, J. R. 1981. Assessment of Biotic Integrity Using Fish Communities. *Fisheries* 6:21-27.

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

Kennen, J. G. 1999. Relation of macroinvertebrate community impairment to catchment characteristics in New Jersey streams. *Journal of the American Water Resources Association* 35(4):939-955.

Kerans, B. L. and J. R. Karr. 1994. A Benthic Index of Biotic Integrity (B-IBI) For Rivers of the Tennessee Valley. *Ecological Applications* 4(4):768-785.

Kleindl, W. J. 1995. A benthic index of biotic integrity for Puget Sound lowland streams, Washington, USA. MS thesis, University of Washington, Seattle.

Larsen, D. P. and A. T. Herlihy. 1998. The Dilemma of Sampling Streams for Macroinvertebrate Richness. *Journal of the North American Benthological Society* 17(3):359-366.

Lenat, D. R. 1993. A Biotic Index for the Southeastern United States: Derivation and List of Tolerance Values, with Criteria for Assigning Water-Quality Ratings. *Journal of the North American Benthological Society* 12(3):279-290.

Maher, A. 2000. Fairfax County, VA Dept. of Public Works and Environmental Services.

Marchant, R., A. Hirst, R Norris, and L. Metzeling. 1999. Classification of macroinvertebrate communities across drainage basins in Victoria, Australia: consequences of sampling on a broad spatial scale for predictive modeling. *Freshwater Biology* 41(2):253-268.

Maxted, J. R., M. T. Barbour, J. Gerritsen, V. Poretti, N. Primrose, A. Silva, D. Penrose, and R. Renfrow. 2000. Assessment framework for mid-Atlantic costal plain streams using benthic macroinvertebrates. Journal of the North American Benthological Society 19(1):128-144.

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

MPCA (Minnesota Pollution Control Agency). 1975. *Study of Minnesota's Tributaries to Lake Superior.* St. Paul, MN.

Omernik, J. M. 1987. Ecoregions of the Conterminous United States. *Annals of the Association of American Geographers* 77(1):118-125.

Omernik, J. M. 1995. Ecoregions: A spatial framework for environmental management. Pages 49-62 in W. S. Davis and T. P. Simon (editors). *Biological Assessment and Criteria: Tools for Water Resource Planning and Decision Making.* Lewis Publishers, Boca Raton, FL.

Patterson, A. J. 1996. The effect of recreation on biotic integrity of small streams in Grand Teton National Park. MS thesis, University of Washington, Seattle.

Persons, S. 1999. Minnesota Dept. of Natural Resources.

Resh, V.H. 1995. Freshwater Benthic Macroinvertebrates and Rapid Assessment Procedures for Water Quality Monitoring in Developing and Newly Industrialized Countries. Pages 167-177 in W. S. Davis and T. P. Simon (editors). *Biological Assessment and Criteria: Tools for Water Resource Planning and Decision Making*. Lewis Publishers, Boca Raton, FL.

Resh, V. H. and J. K. Jackson. 1993. Rapid assessment approaches to biomonitoring using benthic macroinvertebrates. Pages 195-233 in D.M. Rosenberg and V.H. Resh (eds). *Freshwater Biomonitoring and Benthic Macroinvertebrates*. Chapman and Hall, New York.

Rosgen, D. L. 1994. A Classification of Natural Rivers. *Catena* 22(3):169-199.

Richards, C. and G. Host. 1994. Examining Land Use Influences on Stream Habitats and Macroinvertebrates: A GIS Approach. *Water Resources Bulletin* 30(4):729-738.

Richards, C., P. Tucker, and F. Kutka. 1992. *A Comparison of Macroinvertebrate Communities, Habitat, and Water Chemistry Along the Length of Miller Creek.* NRRI Technical Report, Number: NRRI/TR-92/13, University of Minnesota-Duluth.

Rothrock, J. A., P. K. Barten, and G. L. Ingman. 1998. Land Use and Aquatic Biointegrity in the Blackfoot River Watershed, Montana. *Journal of the American Water Resources Association* 34(3):565-581.

Smith, L., and J. Moyle. 1944. *A Biological Survey and Fishery Management Plan for the Streams of the Lake Superior North Shore Watershed.* Minnesota Department of Game and Fish., St. Paul, MN.

Strahler, A. N. Quantitative Analysis of Watershed Geomorphology. Transactions of the American Geophysical Union 38:913-920.

Surber, T. 1922. *A Biological Reconnaissance of Streams Tributary to Lake Superior, Baptism River to Devil Track River.* Manuscript Report, MN Dept. of Conservation, Bureau of Fish. Res., St. Paul, MN.

Surber, E. W. 1924. *Streams Tributary to Lake Superior: Knife River to Baptism River.* Manuscript Report, MN Dept. of Conservation, Bureau of Fish. Res., St. Paul, MN.

Thorne, R. S. J. and W. P. Williams. 1997. The response of benthic macroinvertebrates to pollution in developing countries: a multimetric system of bioassessment. Freshwater Biology 37:671-686.

USFS (United States Forest Service). Unpublished data. Superior National Forest - Tofte, MN Station.

Waters. T. F. 1986. *The Benthos of Three Minnesota Streams Tributary to Lake Superior And Its Relationship to Acid Precipitation.* Miscellaneous Publication 40, Agricultural Experiment Station, University of Minnesota, St. Paul, MN.

Yoder, C. O. 1991. Answering some concerns about biological criteria based on experiences in Ohio. Pages 95-104 in G. H. Flock (ed.). *Water Quality Standards for the 21st Century, Proceedings of a National Conference.* U.S. Environmental Protection Agency, Office of Water, Washington D.C.

Yoder, C. O. and E. T. Rankin. 1995. Biological Criteria Program Development and Implementation in Ohio. Pages 109-144 in W. S. Davis and T. P. Simon (editors). *Biological Assessment and Criteria: Tools for Water Resource Planning and Decision Making.* Lewis Publishers, Boca Raton, FL.

Zar, J. H. 1999. *Biostatistical Analysis.* 4th edition. Prentice Hall, Upper Saddle River, NJ.

## Acknowledgments

This study was a M.S. project for KS at the University of Minnesota. KS thanks Paul and Jan Stroom for encouragement and Paul for help with fieldwork. Thanks to Carl Richards for conceiving the study and for guidance along the way. Roger Moon and Ray Newman made helpful comments on earlier manuscript versions. Ralph Holzenthal, Valerie Brady, and David Houghton provided various help and guidance regarding insect identifications and/or distributions. Paul Tucker and Nancy Kirsch identified macroinvertebrates from the early sampling period. Gerry Sjerven assisted with the LSW map production and Mike Edlund constructed the permutation computer program. Fisheries personnel at the Minnesota DNR offices in Grand Marais and Finland provided helpful stream information.

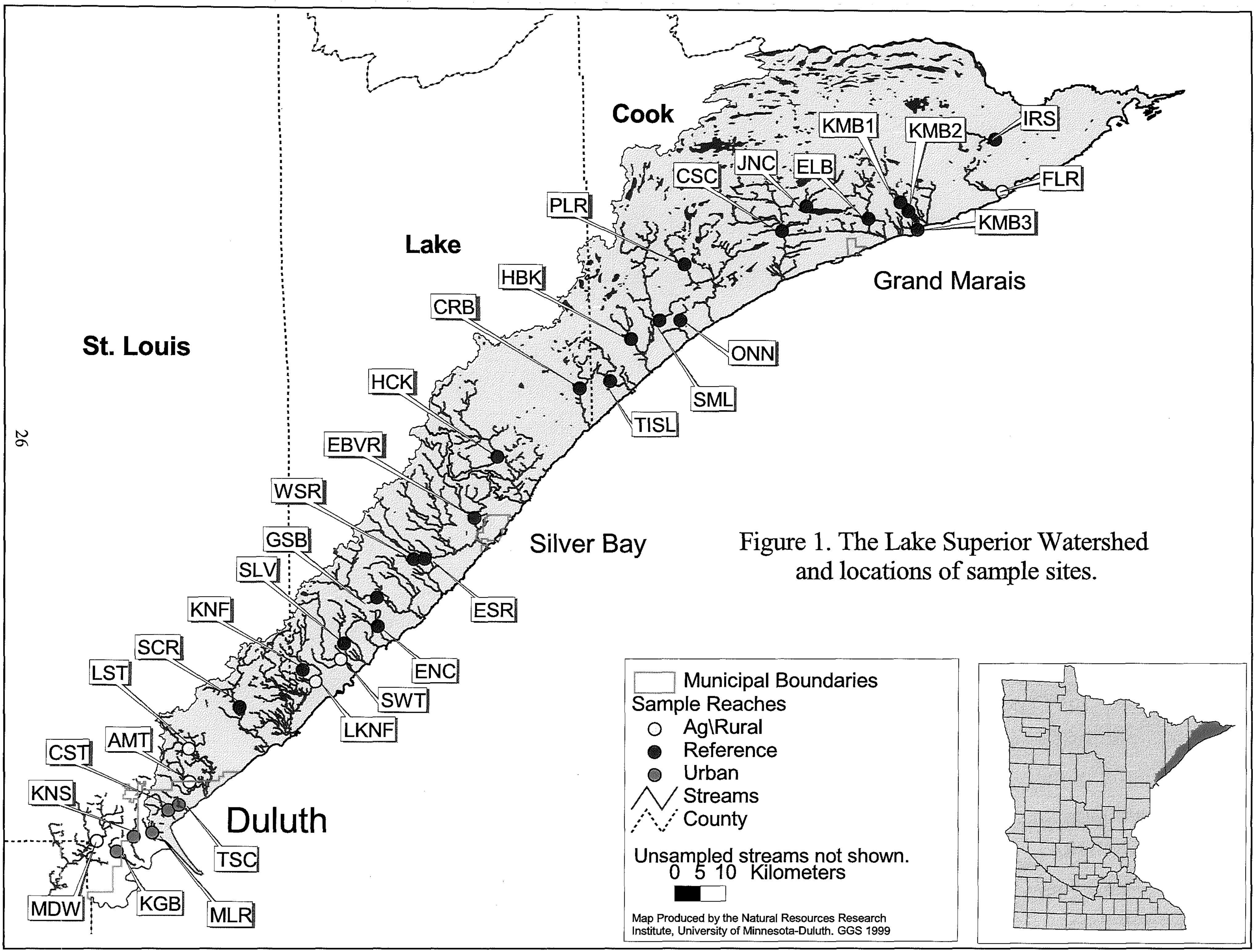



Figure 1. The Lake Superior Watershed and locations of sample sites.

Fig. 2. Coefficients of variation for metrics among Reference streams.

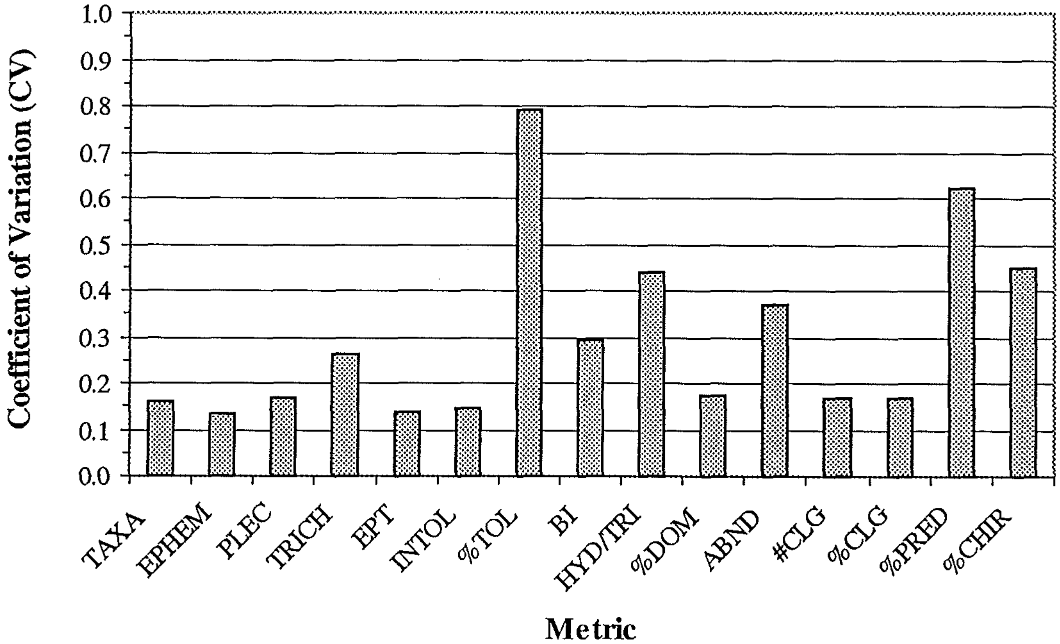


Figure 3. Spread of data points for individual streams. ◆ = Reference, ○ = Ag/Rural, △ = Urban. Some Reference points overlap.

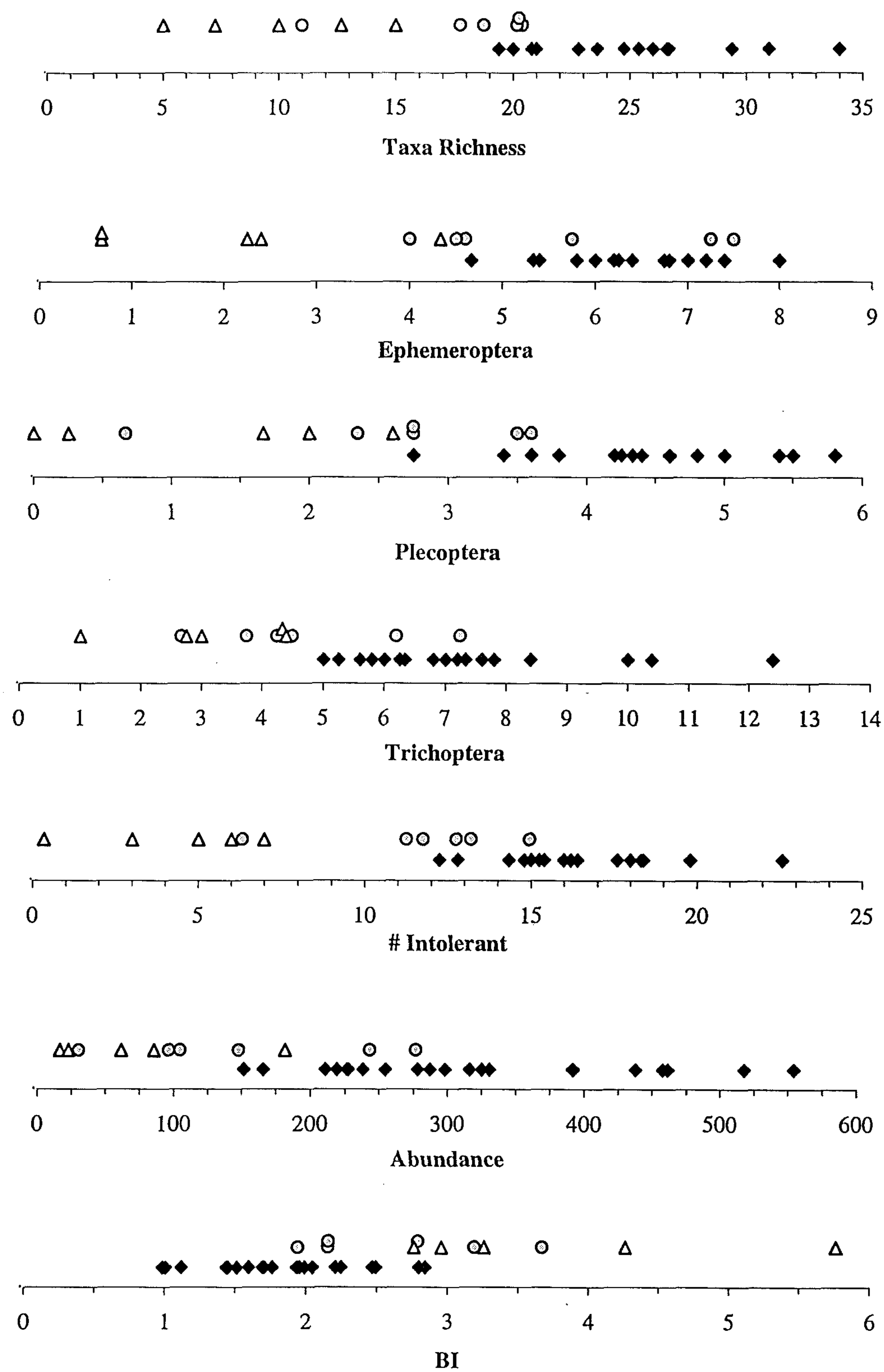


Figure 3. Continued.

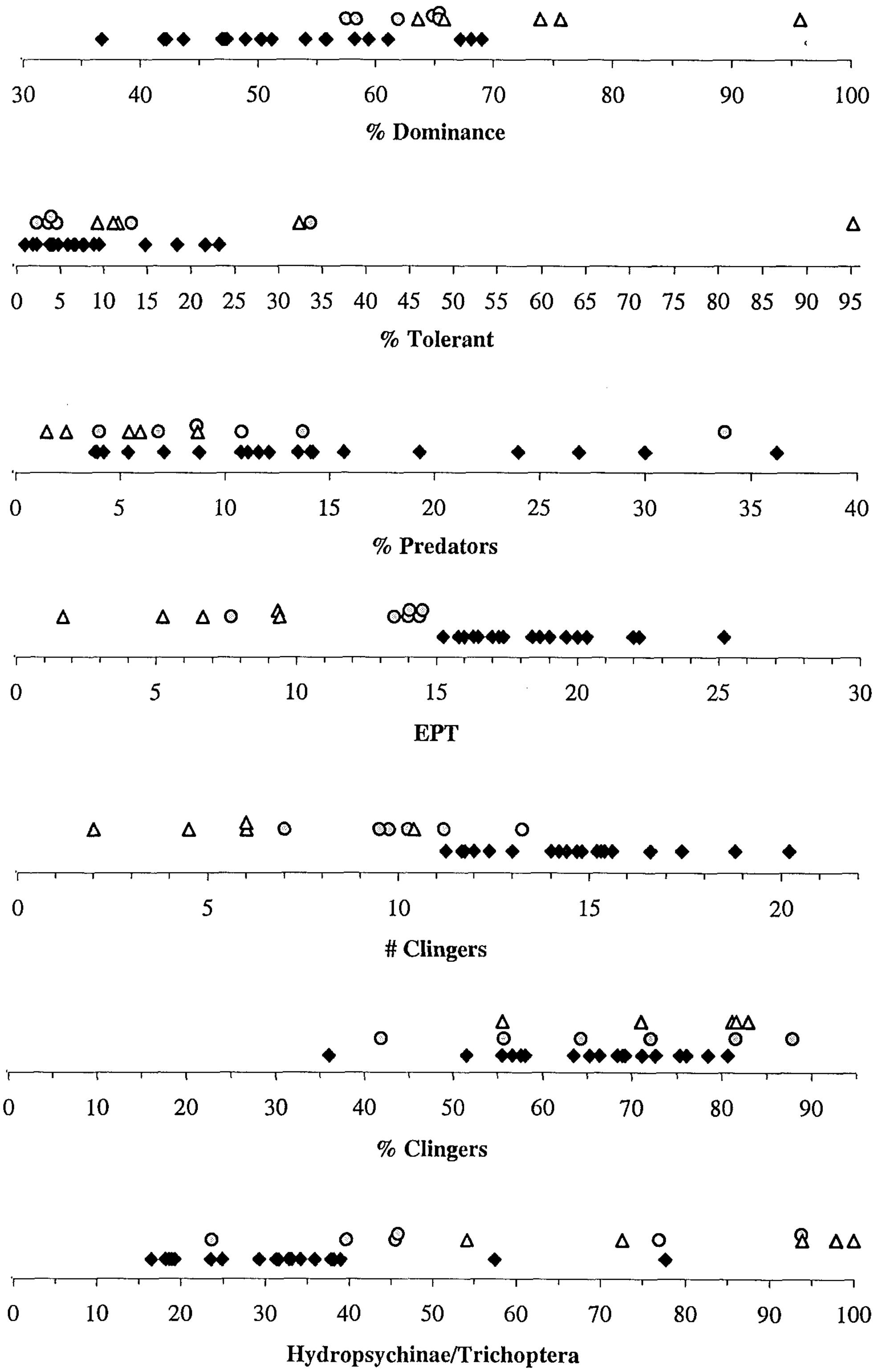


Figure 3. Continued.

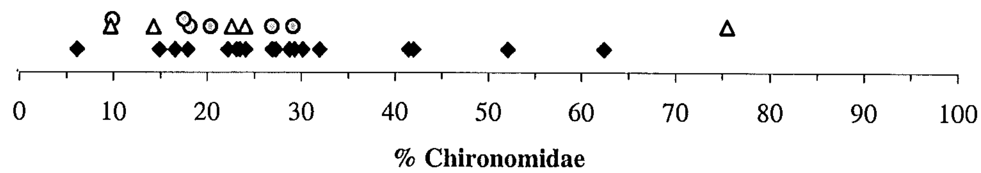


Figure 4. Ranges of final scores for stream groups. Box plots represent the $10^{th}$, $25^{th}$, median, $75^{th}$, and $90^{th}$ percentiles. A number 2 indicates two overlapping stream scores.

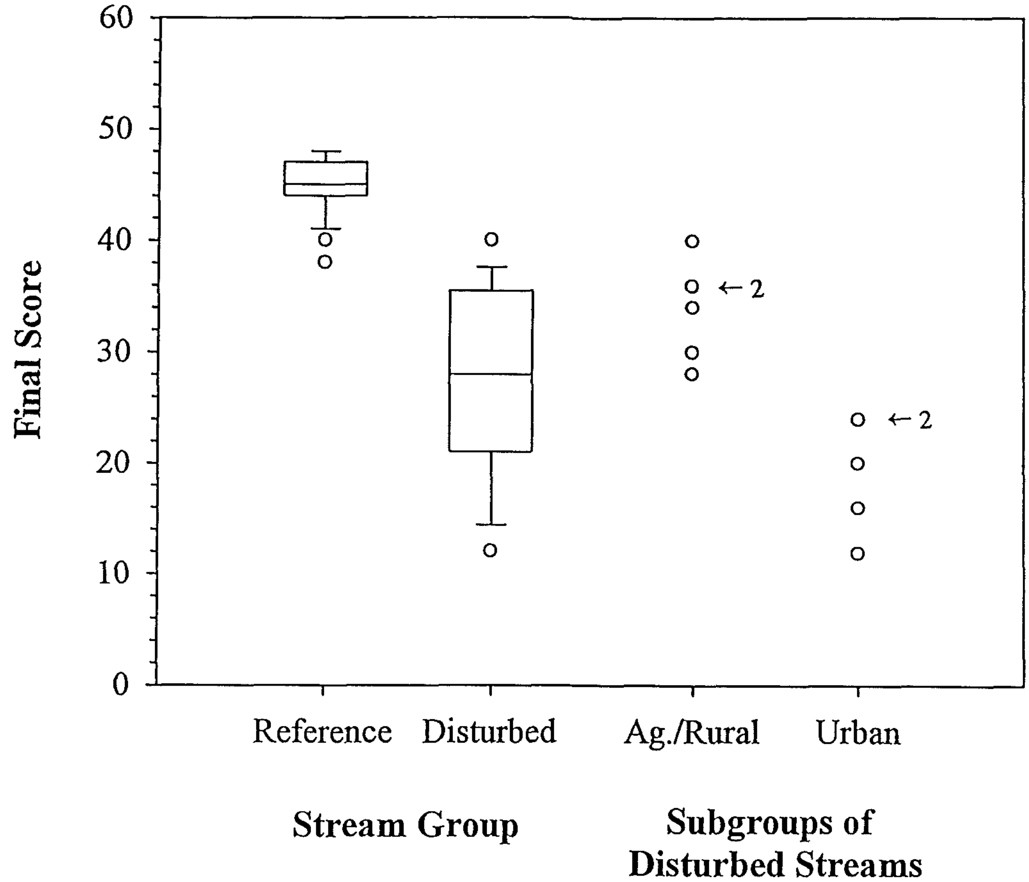


Fig. 5. Taxa Richness/Area curve for Reference streams with 5 replicates. Graphed Richness values are the grand means of 14 Reference streams calculated from individual stream means of all permutations of each composited sample formed of 1 - 5 replicates. Numbers along graph line are the incremental changes in taxa richness between adjacent points.

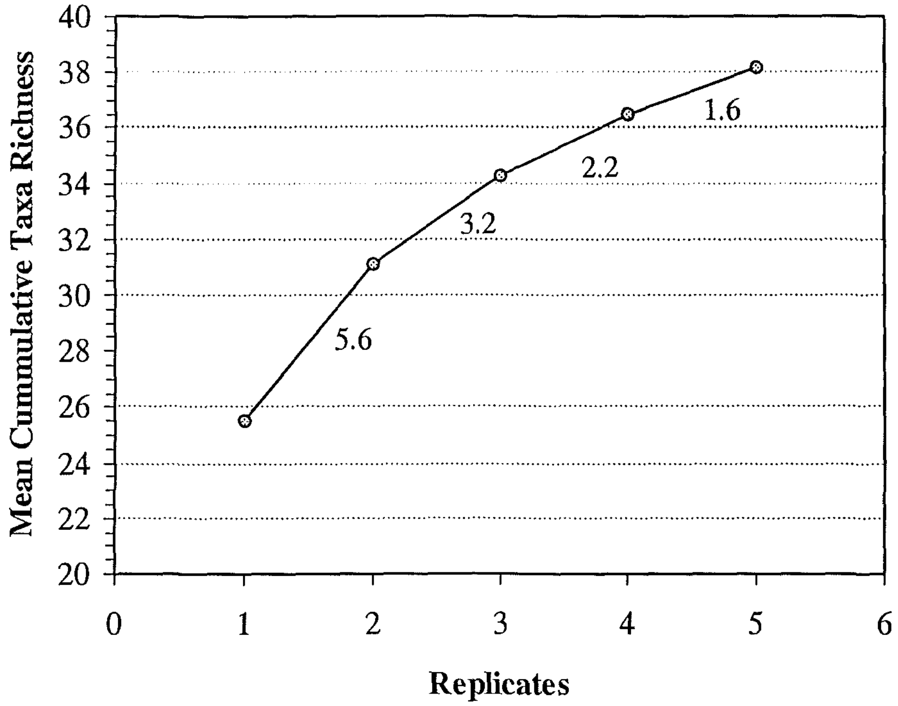


Table 1. Name abbreviations, watershed segment, county and specific sampling locations, disturbance type, stream order, sample year, sample size (n) and sample gear for Reference and Disturbed streams (listed from south to north).

| *Reference Stream* | *Abbr.* | *Seg.* | *Co.* | *Location* | *Disturbance* | *Order*[*] | *Year* | *n* | *Type* |
|---|---|---|---|---|---|---|---|---|---|
| **Sucker R.** | SCR | 1 | SL | St. Louis Co. 40†, upstr. | Minimal | 3** | 98 | 5 | K |
| **Knife R.** | KNF | 1 | L | Lake Co. 302†, upstr. | Minimal | 3 | 89 | 4 | K |
| **Encampment R.** | ENC | 1 | L | Lake Co. 3, upstr. | Minimal | 3 | 89 | 4 | K |
| **Gooseberry R.** | GSB | 1 | L | Lake Co. 3, upstr. | Minimal | 3** | 98 | 5 | K |
| **Silver Cr.** | SLV | 1 | L | Lake Co. 2, upstr. | Minimal | 1 | 89 | 3 | K |
| **W. Split Rock R.** | WSR | 1 | L | Lake Co. 3, upstr. | Minimal | 2** | 89 | 3 | H |
| **E. Split Rock R.** | ESR | 1 | L | Lake Co. 3, upstr. | Minimal | 2 | 89 | 4 | H |
| **E. Beaver R.** | EBVR | 2 | L | Lake Co. 4, upstr. | Minimal | 2** | 97 | 5 | K |
| **Hockamin Cr.** | HCK | 2 | L | SFR 701, upstr. | Minimal | 2** | 97 | 5 | K |
| **Caribou R.** | CRB | 2 | L | Lake Co. 8, downstr. | Minimal | 3** | 97 | 5 | K |
| **Two Island R.** | TISL | 2 | C | Cook Co. 1, upstr. | Minimal | 2** | 97-98 | 5 | K |
| **Heartbreak Cr.** | HBK | 2 | C | NFD 166, upstr. | Minimal | 2** | 97 | 5 | K |
| **Sixmile Cr.** | SML | 2 | C | Cook Co. 2, upstr. | Minimal | 1 | 97 | 5 | K |
| **Onion Cr.** | ONN | 2 | C | North border, Section 35 | Minimal | 1 | 92 | 3 | K |
| **Poplar R.** | PLR | 3 | C | NFD 164, downstr. | Minimal | 2** | 97 | 5 | K |
| **Cascade R.** | CSC | 3 | C | NFD 157, upstr. | Minimal | 3** | 97 | 5 | K |
| **Junco Cr.** | JNC | 3 | C | Cook Co. 57, upstr. | Minimal | 2** | 97 | 5 | K |
| **Elbow Cr.** | ELB | 3 | C | Cook Co. 12, upstr. | Minimal | 3** | 97 | 5 | K |
| **Kimball Cr.** | KMB | 3 | C | NFD 304, upstr. | Minimal | 3** | 97-98 | 5 | K |
| **Irish Cr.** | IRS | 3 | C | ~ 1.2 mi. W. of 16, upstr. | Minimal | 2 | 97-98 | 5 | K |
| ***Disturbed Streams*** | | | | | | | | | |
| **Midway R.** | MDW | Out | SL | St. Louis Co. 19, upstr. | A | 3 | 90 | 3 | K |
| **Kingsbury Cr.** | KGB | Out | SL | Pionk Drive, downstr. | U | 1 | 91 | 3 | S |
| **Keenes Cr.** | KNS | Out | SL | Okerstrom Rd., downstr. | U | 1 | 90 | 3 | K |
| **Miller Cr.** | MLR2 | 1 | SL | U Trinity Rd., upstr. | U | 2 | 91 | 3 | S |
| **Chester Cr.** | CST | 1 | SL | Chester Park above ski jump | U | 2 | 90 | 4 | K |
| **Tischer Cr.** | TCH | 1 | SL | Columbus Ave., upstr. | U | 2** | 99 | 5 | K |
| **E. Br. Amity Cr.** | AMT | 1 | SL | St. Louis Co. 37, upstr. | A | 2 | 90 | 4 | K |
| **Lester R.** | LST | 1 | SL | St. Louis Co. 37, upstr. | A | 3** | 89, 90 | 3,4 | K |
| **L. E. Br. Knife R.** | LKNF | 1 | L | Lake Co. 12♯, upstr. | A | 2 | 89 | 4 | K |
| **Stewart R.** | SWT | 1 | L | Lake Co. 2, upstr. | A | 3** | 89 | 4 | K |
| **Flute Reed R.** | FLR | 3 | C | Cook Co. 69‡, downst. | A | 3** | 98 | 5 | K |

*Order determined from 1:24.000 USGS topographical maps.
**Stream originates from a lake or contains a lake along its course above sample site.
†Northernmost crossing, ♯Southernmost crossing, ‡First crossing from Lake Superior on # 69
SFR = State Forest Road
NFD = National Forest Development road

Disturbance
U = Urban
A = Agriculture and low-density residential development

County
C = Cook
L = Lake
SL = St. Louis

Type
K = Kick
S = Surber
H = Hess

Table 2. Candidate metrics for multimetric index.

| *Metric* | *Abbrev.* | *Class*[*] | *Description of Metric* |
|---|---|---|---|
| Taxa Richness[†] | TAXA | R | Number of insect taxa. |
| # Ephemeroptera taxa[†] | EPHM | R | Number of Ephemeroptera taxa. |
| # Plecoptera taxa[†] | PLEC | R | Number of Plecoptera taxa. |
| # Trichoptera taxa[†] | TRIC | R | Number of Trichoptera taxa. |
| # EPT taxa | EPT | R | Number of Ephemeroptera, Plecoptera, and Trichoptera taxa. |
| # Intolerant taxa | INTOL | T | Number of intolerant taxa. |
| % Tolerant taxa[†] | %TOL | T | Percent of total abundance which are tolerant. |
| Biotic Index[†] | BI | T | Weighted average of tolerance values of each taxa. |
| Hydropsychinae/Trichopt.[†] | HYD/TRI | T | Abundance ratio of Hydropsychinae to Trichoptera. |
| % Chironomids | %CHIR | T | Percent of total abundance that are chironomids. |
| % Dominance[†] | %DOM | P | Percentage of the 3 most abundant taxa. |
| Total Abundance[†] | ABND | P | Total number of insect individuals. |
| # Clingers | CLG | F/H | Number of taxa classified as clingers. |
| % Clingers | %CLG | F/H | Percent of total abundance which are clingers. |
| % Predators[†] | %PRED | F/H | Percent of total abundance which are predators. |

[*]R = Richness and Composition; T = Tolerant/Intolerant; F/H = Functional Feeding Group, Habit; P = Population Attribute
[†]Included in the final index

Table 3. Composite Taxa Richness for Reference streams (n = 5 Surber samples).

| *Stream name* | *Composite Taxa* |
|---|---|
| Caribou R. | 37 |
| Cascade R. | 42 |
| E. Beaver R. | 38 |
| Elbow Cr. | 34 |
| Gooseberry R. | 51 |
| Heartbreak Cr. | 29 |
| Hockamin Cr. | 40 |
| Irish Cr. | 29 |
| Junco Cr. | 41 |
| Kimball Cr. | 43 |
| Poplar R. | 37 |
| Sixmile Cr. | 37 |
| Sucker R. | 43 |
| Two Island R. | 33 |
| *Mean* | 38.1 |

Table 4. Pearson's correlation coefficients (r) for all pairs of metrics using Reference stream data. Coefficients with strong correlations (≥ 0.80) are in bold.

| | EPHM | PLEC | TRIC | ABND | INTOL | BI | %DOM | %TOL | %PRED | EPT | CLG | %CLG | HYD/TRI | %CHIR |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| TAXA | 0.47 | 0.63 | **0.81** | 0.67 | **0.93** | 0.66 | -0.64 | **0.82** | 0.11 | **0.96** | **0.94** | -0.03 | 0.44 | 0.30 |
| EPHM | | 0.19 | 0.07 | 0.25 | 0.45 | -0.10 | -0.43 | 0.37 | 0.00 | 0.45 | 0.51 | -0.04 | 0.27 | 0.20 |
| PLEC | | | 0.35 | 0.60 | 0.70 | 0.24 | -0.23 | 0.33 | -0.14 | 0.62 | 0.56 | 0.07 | 0.69 | 0.09 |
| TRIC | | | | 0.62 | 0.72 | 0.65 | -0.48 | 0.78 | 0.03 | **0.88** | 0.79 | 0.19 | 0.29 | 0.22 |
| ABND | | | | | 0.64 | 0.38 | -0.18 | 0.48 | -0.18 | 0.73 | 0.70 | 0.17 | 0.61 | 0.07 |
| INTOL | | | | | | 0.50 | -0.59 | 0.71 | -0.06 | **0.90** | 0.91 | 0.08 | 0.39 | 0.33 |
| BI | | | | | | | -0.61 | 0.69 | 0.53 | 0.65 | 0.57 | -0.28 | 0.14 | 0.40 |
| %DOM | | | | | | | | -0.60 | -0.27 | -0.57 | -0.58 | 0.10 | 0.14 | -0.32 |
| %TOL | | | | | | | | | 0.22 | **0.81** | **0.83** | 0.06 | 0.17 | 0.24 |
| %PRED | | | | | | | | | | -0.02 | -0.06 | -0.82 | -0.31 | 0.07 |
| EPT | | | | | | | | | | | **0.94** | 0.15 | 0.51 | 0.26 |
| CLG | | | | | | | | | | | | 0.12 | 0.42 | 0.20 |
| %CLG | | | | | | | | | | | | | 0.18 | 0.16 |
| HYD/TRI | | | | | | | | | | | | | | -0.16 |

Table 5. Probability results from the Mann-Whitney and Kruskal-Wallis difference tests for all metrics.

| | TAXA | EPHM | PLEC | TRIC | EPT | INTOL | %TOL | BI | HYD/TRI | DOM | ABND | CLNG | %CLG | %PRED | %CHIR |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2 way* | <.001 | .004 | <.001 | <.001 | <.001 | <.001 | .215 | <.001 | <.001 | .001 | <.001 | <.001 | .173 | .043 | .076 |
| 3 way** | <.001 | .002 | <.001 | <.001 | <.001 | <.001 | .052 | .004 | .001 | .002 | <.001 | <.001 | .272 | .033 | .201 |

*Reference v. Disturbed (=Ag./Rural + Urban)
**Reference v. Ag/Rural v. Urban

Table 6. Scoring divisions for final index metrics. The top row represents the individual metric biocriteria.

| *Score* | TAXA | EPHM | PLEC | TRIC | %TOL | BI | HYD/TRI | %DOM | ABND | %PRED |
|---|---|---|---|---|---|---|---|---|---|---|
| 5 | > 21.0 | > 5.8 | > 4.2 | > 6.0 | < 9.5 | < 2.24 | < 38.0 | < 59.4 | > 226.8 | > 7.1 |
| 3 | 10.5 - 21.0 | 2.9 - 5.8 | 2.1 - 4.2 | 3.0 - 6.0 | 9.5 - 29.8 | 2.24 - 4.12 | 38.0 - 61.9 | 59.4 - 79.7 | 113.4 - 226.8 | 3.6 - 7.1 |
| 1 | < 10.5 | < 2.9 | < 2.1 | < 3.0 | 29.8 - 50 | 4.12 - 6.00 | > 61.9 | > 79.7 | < 113.4 | < 3.6 |

Table 7. Watershed size and land cover of individual study streams.

| *Reference Stream* | *Abbr.* | *Watershed size (ha)* | *% Hay/Pasture/ Grassland* | *% Developed* | *% Devloped + Hay/Past./Grass.* | *% Roads* | *% All Disturbed* |
|---|---|---|---|---|---|---|---|
| **Sucker R.** | SCR | 7354.6 | 0.10 | 0.00 | 0.10 | 0.85 | 0.95 |
| **Knife R.** | KNF | 3766.4 | 0.81 | 0.25 | 1.07 | 1.02 | 2.09 |
| **Encampment R.** | ENC | 1988.7 | 2.66 | 0.48 | 3.14 | 1.64 | 4.78 |
| **Gooseberry R.** | GSB | 7948.4 | 0.06 | 0.27 | 0.33 | 0.99 | 1.32 |
| **Silver Cr.** | SLV | 2498.4 | 0.28 | 0.96 | 1.24 | 1.41 | 2.65 |
| **W. Split Rock R.** | WSR | 3035.3 | 0.06 | 0.07 | 0.13 | 0.31 | 0.44 |
| **E. Split Rock R.** | ESR | 3688.7 | 0.06 | 0.05 | 0.11 | 0.16 | 0.27 |
| **E. Beaver R.** | EBVR | 2717.6 | 0.00 | 0.00 | 0.00 | 0.39 | 0.39 |
| **Hockamin Cr.** | HCK | 3952.8 | 0.00 | 0.02 | 0.02 | 0.76 | 0.78 |
| **Caribou R.** | CRB | 2813.0 | 0.00 | 0.54 | 0.54 | 0.17 | 0.71 |
| **Two Island R.** | TISL | 3029.0 | 0.00 | 0.01 | 0.01 | 0.39 | 0.40 |
| **Heartbreak Cr.** | HBK | 3546.6 | 0.00 | 0.00 | 0.00 | 0.42 | 0.42 |
| **Sixmile Cr.** | SML | 2498.4 | 0.00 | 0.00 | 0.00 | 0.25 | 0.25 |
| **Onion Cr.** | ONN | 712.9 | 0.00 | 0.00 | 0.00 | 0.10 | 0.10 |
| **Poplar R.** | PLR | 7674.1 | 0.12 | 0.03 | 0.15 | 0.59 | 0.74 |
| **Cascade R.** | CSC | 20596.8 | 0.10 | 0.05 | 0.14 | 0.60 | 0.74 |
| **Junco Cr.** | JNC | 5282.7 | 0.27 | 0.04 | 0.30 | 0.56 | 0.86 |
| **Elbow Cr.** | ELB | 5265.4 | 0.01 | 0.04 | 0.05 | 0.61 | 0.66 |
| **Kimball Cr.** | KMB | 2323.5 | 0.17 | 0.05 | 0.22 | 0.61 | 0.83 |
| **Irish Cr.** | IRS | 2296.1 | 0.00 | 0.00 | 0.00 | 0.90 | 0.90 |
| *Averages* | | **4645.9** | **0.24** | **0.14** | **0.38** | **0.64** | **1.01** |
| ***Ag/Rural*** | | | | | | | |
| **Midway R.** | MDW | 5962.0 | 24.17 | 2.83 | 27.00 | 5.43 | 32.43 |
| **E. Br. Amity Cr.** | AMT | 1922.2 | 21.00 | 2.31 | 23.31 | 5.12 | 28.43 |
| **Lester R.** | LST | 4872.2 | 11.45 | 0.49 | 11.93 | 2.95 | 14.88 |
| **L. E. Br. Knife R.** | LKNF | 1166.1 | 13.24 | 1.42 | 14.66 | 2.13 | 16.79 |
| **Stewart R.** | SWT | 6007.9 | 3.97 | 0.97 | 4.94 | 1.44 | 6.38 |
| **Flute Reed R.** | FLR | 4195.9 | 0.54 | 0.04 | 0.58 | 0.46 | 1.04 |
| *Averages* | | **4021.1** | **12.39** | **1.34** | **13.74** | **2.92** | **16.65** |
| ***Urban*** | | | | | | | |
| **Kingsbury Cr.** | KGB | 1885.7 | 19.56 | 15.17 | 34.73 | 8.69 | 43.42 |
| **Keenes Cr.** | KNS | 859.8 | 23.55 | 7.59 | 31.14 | 7.08 | 38.22 |
| **Miller Cr.** | MLR | 2181.8 | 9.03 | 33.80 | 42.83 | 16.31 | 59.14 |
| **Chester Cr.** | CST | 1488.7 | 11.02 | 15.23 | 26.24 | 7.01 | 33.25 |
| **Tischer Cr.** | TCH | 1389.9 | 8.33 | 33.03 | 41.37 | 12.42 | 53.79 |
| *Averages* | | **1561.2** | **14.30** | **20.96** | **35.26** | **10.30** | **45.56** |

Table 8. Final Index metric values for upper, middle, and lower sample reaches of Kimball Cr., KruskalWallis p-values, and reach differences.

| | *Means* | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| *Metric* | *Upper* | *Middle* | *Lower* | *Kruskal-Wallis p*‡ | *Corrected p*† | **1** | **2** | **3** |
| Taxa Richness | 31.0 | 26.4 | 16.8 | 0.005 | $p < .001$ | ns | 0.005 | ns |
| Ephemeroptera | 7.2 | 6.8 | 6.0 | 0.466 | $p > .10$ | ns | ns | ns |
| Plecoptera | 4.4 | 5.4 | 3.4 | 0.084 | $0.05 < p < 0.01$ | ns | ns | 0.05 |
| Trichoptera | 10.4 | 6.0 | 4.8 | 0.005 | $p < .001$ | 0.100 | 0.005 | ns |
| % Tolerant | 18.5 | 4.2 | 4.2 | 0.019 | $0.02 < p < 0.01$ | 0.05 | 0.05 | ns |
| BI | 2.49 | 1.40 | 1.22 | 0.009 | $0.002 < p < 0.005$ | 0.05 | 0.020 | ns |
| Hyd/Tri | 31.6 | 31.2 | 20.7 | 0.505 | $p > .10$ | ns | ns | ns |
| % Dominance | 47.2 | 44.2 | 63.3 | 0.046 | $0.02 < p < 0.05$ | ns | 0.01 | 0.001 |
| Abundance | 555.2 | 262.4 | 126.2 | 0.027 | $p = 0.02$ | ns | 0.05 | ns |
| % Predators | 25.2 | 12.2 | 7.9 | 0.008 | $0.001 < p < 0.002$ | ns | 0.02 | ns |

‡Systat version 8.0 (1998)
†Corrected p-values for Kruskal-Wallis test with $n_1 = 5$, $n_2 = 5$, $n_3 = 5$ (Zar 1999)
1 = Upper v. Middle
2 = Upper v. Lower
3 = Middle v. Lower
ns = not significant

Appendix 1. Taxa list, relative abundances, and assigned values and classifications.

Reference Streams

| *Taxon** | *# of Replicates* | *Percent of Replicates* | *Total Count* | *Percent of Total*[†] | *Abundance Rank* | *BI Tol. Value* | *Predator* | *Clinger* |
|---|---|---|---|---|---|---|---|---|
| **Ephemeroptera** (16) | | | | | | | | |
| *Acentrella* | 25 | 27.5 | 59 | 0.203 | 40 | 4 | | |
| *Baetis* | 42 | 46.2 | 175 | 0.601 | 30 | 5 | | |
| *Ameletus* | 1 | 1.1 | 2 | 0.007 | 75 | 0 | | |
| *Ephemera* | 1 | 1.1 | 1 | 0.003 | 79 | 1 | | |
| *Ephemerella* | 89 | 97.8 | 1483 | 5.096 | 8 | 1 | | x |
| *Epeorus* | 82 | 90.1 | 1163 | 3.997 | 9 | 0 | | x |
| *Eurylophella* | 28 | 30.8 | 176 | 0.605 | 29 | 0 | | x |
| *Heptagenia* | 17 | 18.7 | 204 | 0.701 | 26 | 3 | | x |
| *Isonychia* | 2 | 2.2 | 2 | 0.007 | 75 | 2 | | |
| *Leptophlebia* | 1 | 1.1 | 5 | 0.017 | 66 | 1 | | |
| *Leucrocuta* | 64 | 70.3 | 518 | 1.780 | 15 | 1 | | x |
| *Paraleptophlebia* | 90 | 98.9 | 3361 | 11.550 | 3 | 1 | | |
| *Procloeon* | 9 | 9.9 | 19 | 0.065 | 53 | -- | | |
| *Rhithrogena* | 50 | 54.9 | 2126 | 7.306 | 5 | 0 | | x |
| *Serratella* | 20 | 22.0 | 253 | 0.869 | 23 | 2 | | x |
| *Stenonema* | 76 | 83.5 | 746 | 2.564 | 13 | 2 | | x |
| **Odonata** (4) | | | | | | | | |
| *Boyeria* | 32 | 35.2 | 57 | 0.196 | 41 | 5 | x | x |
| *Calopteryx* | 2 | 2.2 | 19 | 0.065 | 53 | 5 | x | |
| *Cordulegaster* | 4 | 4.4 | 4 | 0.014 | 68 | 3 | x | |
| *Ophiogomphus* | 64 | 70.3 | 1074 | 3.691 | 10 | 1 | x | |
| **Plecoptera** (14) | | | | | | | | |
| *Acroneuria* | 76 | 83.5 | 480 | 1.650 | 17 | 0 | x | x |
| *Agnetina* | 17 | 18.7 | 49 | 0.168 | 44 | 2 | x | x |
| *Allocapnia* | 1 | 1.1 | 3 | 0.010 | 73 | 3 | | x |
| *Alloperla* | 1 | 1.1 | 4 | 0.014 | 68 | 0 | | x |
| *Isogenoides* | 10 | 11.0 | 55 | 0.189 | 42 | 0 | x | x |
| *Isoperla* | 67 | 73.6 | 293 | 1.007 | 21 | 2 | x | x |
| *Leuctra* | 42 | 46.2 | 140 | 0.481 | 31 | 0 | | |
| *Nemouridae* | 7 | 7.7 | 19 | 0.065 | 53 | 1 | | |
| *Oemopteryx* | 17 | 18.7 | 31 | 0.107 | 47 | 1 | | |
| *Paracapnia* | 84 | 92.3 | 836 | 2.873 | 12 | 1 | | |
| *Paragnetina* | 21 | 23.1 | 32 | 0.110 | 46 | 1 | x | x |
| *Pteronarcys* | 10 | 11.0 | 14 | 0.048 | 57 | 0 | | x |
| *Taeniopteryx* | 53 | 58.2 | 193 | 0.663 | 27 | 2 | | |
| *Zealeuctra* | 1 | 1.1 | 2 | 0.007 | 75 | 0 | | |
| **Megaloptera** (2) | | | | | | | | |
| *Nigronia* | 28 | 30.8 | 63 | 0.217 | 38 | 0 | x | x |
| *Sialis* | 16 | 17.6 | 23 | 0.079 | 51 | 4 | x | |

Continued.

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Trichoptera** (29) | | | | | | | | |
| *Apatania* | 8 | 8.8 | 20 | 0.069 | 52 | 1 | | |
| *Brachycentrus* | 16 | 17.6 | 55 | 0.189 | 42 | 1 | | x |
| *Ceraclea* | 4 | 4.4 | 4 | 0.014 | 68 | 3 | | |
| *Ceratopsyche* | 90 | 98.9 | 3745 | 12.870 | 2 | 4 | | x |
| *Cheumatopsyche* | 35 | 38.5 | 711 | 2.443 | 14 | 5 | | x |
| *Chimarra* | 35 | 38.5 | 495 | 1.701 | 16 | 2 | | x |
| *Diplectrona* | 8 | 8.8 | 218 | 0.749 | 25 | 0 | | x |
| *Dolophilodes* | 31 | 34.1 | 178 | 0.612 | 28 | 0 | | x |
| *Glossosoma* | 82 | 90.1 | 2820 | 9.691 | 4 | 0 | | x |
| *Goera* | 5 | 5.5 | 10 | 0.034 | 59 | 0 | | x |
| *Helicopsyche* | 8 | 8.8 | 11 | 0.038 | 58 | 3 | | x |
| *Hesperophylax* | 2 | 2.2 | 5 | 0.017 | 66 | 4 | | |
| *Hydropsyche* | 8 | 8.8 | 29 | 0.100 | 49 | 6 | | x |
| *Hydroptila* | 28 | 38.5 | 264 | 0.907 | 22 | 6 | | x |
| *Lepidostoma* | 89 | 97.8 | 1690 | 5.808 | 7 | 1 | | |
| Limnephilidae | 8 | 8.8 | 18 | 0.062 | 56 | -- | -- | |
| *Micrasema* | 37 | 40.7 | 343 | 1.179 | 19 | 2 | | |
| *Molanna* | 2 | 2.2 | 2 | 0.007 | 75 | 6 | | |
| *Mystacides* | 6 | 6.6 | 7 | 0.024 | 61 | 4 | | |
| *Neophylax* | 1 | 1.1 | 8 | 0.003 | 79 | 3 | | x |
| *Neureclipsis* | 22 | 24.2 | 88 | 0.302 | 36 | 7 | | x |
| *Oecetis* | 37 | 40.7 | 133 | 0.457 | 32 | 8 | x | |
| *Oxyethira* | 5 | 5.5 | 7 | 0.024 | 61 | 3 | | x |
| *Paranyctiophylax* | 1 | 1.1 | 1 | 0.003 | 79 | 5 | x | x |
| *Polycentropus* | 40 | 44.0 | 92 | 0.316 | 34 | 6 | | x |
| *Protoptila* | 35 | 38.5 | 314 | 1.079 | 20 | 1 | | x |
| *Psychomyia* | 3 | 3.3 | 9 | 0.031 | 60 | 2 | | x |
| *Rhyacophila* | 30 | 33.0 | 83 | 0.285 | 37 | 0 | x | x |
| Unknown 1 | 19 | 20.9 | 115 | 0.395 | 33 | -- | -- | -- |
| **Coleoptera** (4) | | | | | | | | |
| *Hydroporus* | 1 | 1.1 | 1 | 0.003 | 79 | 0 | x | |
| *Optioservus* | 87 | 95.6 | 2071 | 7.117 | 6 | 4 | | x |
| *Stenelmis* | 14 | 15.4 | 29 | 0.100 | 49 | 5 | | x |
| **Diptera** (15) | | | | | | | | |
| *Antocha* | 32 | 35.2 | 219 | 0.753 | 24 | 3 | | x |
| *Atherix* | 61 | 67.0 | 393 | 1.351 | 18 | 2 | x | |
| *Atrichopogon* | 1 | 1.1 | 1 | 0.003 | 79 | -- | | |
| *Chelifera* | 17 | 18.7 | 30 | 0.103 | 48 | 6 | | |
| Chironomidae | 91 | 100.0 | 14907 | 33.875‡ | 1 | -- | -- | |
| *Chrysops* | 1 | 1.1 | 4 | 0.014 | 68 | 6 | | |
| *Dicranota* | 4 | 4.4 | 6 | 0.021 | 63 | 3 | x | |
| *Hemerodromia* | 69 | 75.8 | 977 | 3.356 | 11 | 6 | x | |

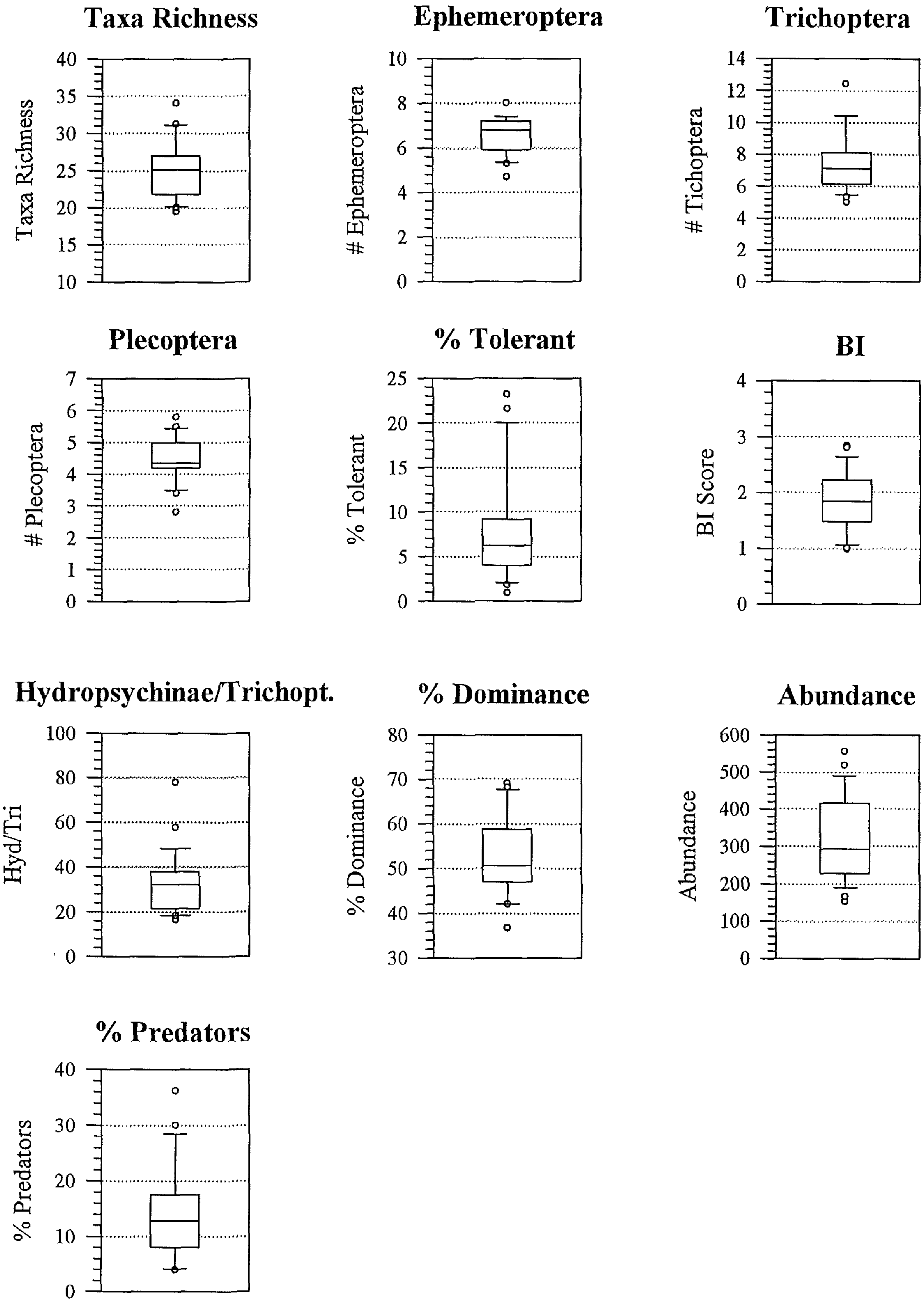
Appendix 2. Metric values for individual streams.

| | All are averages (per 0.09 $m^2$) | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Reference Streams** | **Taxa** | **Ephem** | **Plec** | **Trich** | **EPT** | **Intol** | **Tol** | **BI** | **Hyd/Tri** | **Dom** | **Abund** | **#Clng** | **%Clng** | **Pred** | **% Chiro** |
| Caribou R. | 22.8 | 6.2 | 3.6 | 7.6 | 17.4 | 16.0 | 5.9 | 1.12 | 18.20 | 54.0 | 298.4 | 14.2 | 76.0 | 8.8 | 16.5 |
| Cascade R. | 26.0 | 7.0 | 4.8 | 7.8 | 19.6 | 17.6 | 8.9 | 1.71 | 38.04 | 58.2 | 325.4 | 14.8 | 78.5 | 10.8 | 62.4 |
| E. Split Rock R. | 21.0 | 7.0 | 4.3 | 5.3 | 16.5 | 15.3 | 2.3 | 1.93 | 16.51 | 55.8 | 211.0 | 11.8 | 51.5 | 30.0 | 27.2 |
| East Beaver R. | 25.4 | 7.4 | 5.4 | 6.8 | 19.6 | 16.4 | 6.6 | 1.70 | 77.69 | 67.2 | 518.0 | 15.4 | 69.0 | 5.4 | 6.1 |
| Elbow Cr. | 23.6 | 7.2 | 4.2 | 5.6 | 17.0 | 15.4 | 7.8 | 1.44 | 33.13 | 48.9 | 166.0 | 13.0 | 72.7 | 13.5 | 26.9 |
| Encampment R. | 20.0 | 6.3 | 2.8 | 6.3 | 15.3 | 12.3 | 4.2 | 2.21 | 19.28 | 55.7 | 226.8 | 11.3 | 57.6 | 24.0 | 31.9 |
| Gooseberry R. | 34.0 | 7.4 | 5.4 | 12.4 | 25.2 | 22.6 | 21.6 | 2.84 | 57.50 | 42.0 | 458.2 | 20.2 | 68.4 | 15.7 | 26.9 |
| Heartbreak Cr. | 20.8 | 5.8 | 4.2 | 5.8 | 15.8 | 14.8 | 5.9 | 1.00 | 31.31 | 68.1 | 278.4 | 12.0 | 75.3 | 4.2 | 22.2 |
| Hockamin Cr. | 29.4 | 7.0 | 4.6 | 8.4 | 20.0 | 19.8 | 23.2 | 2.80 | 18.60 | 42.2 | 316.6 | 17.4 | 55.5 | 26.9 | 42.0 |
| Irish Cr. | 19.4 | 6.4 | 3.4 | 6.0 | 15.8 | 12.8 | 4.7 | 1.95 | 18.94 | 47.1 | 151.8 | 12.0 | 71.2 | 12.1 | 23.1 |
| Junco Cr. | 26.6 | 6.0 | 5.8 | 7.2 | 19.0 | 17.6 | 9.5 | 1.76 | 38.19 | 46.9 | 287.4 | 15.2 | 65.3 | 14.1 | 23.4 |
| Kimball Cr. | 31.0 | 7.2 | 4.4 | 10.4 | 22.0 | 18.4 | 18.5 | 2.49 | 31.59 | 47.3 | 554.6 | 18.8 | 63.5 | 19.3 | 30.1 |
| Knife R. | 24.8 | 6.8 | 5.5 | 5.0 | 17.3 | 18.0 | 0.9 | 1.60 | 38.96 | 50.2 | 438.3 | 14.0 | 58.1 | 11.6 | 28.7 |
| Onion R. | 26.7 | 5.3 | 5.0 | 10.0 | 20.3 | 18.3 | 7.6 | 2.47 | 37.85 | 50.3 | 461.7 | 14.7 | 80.7 | 7.1 | 52.1 |
| Poplar R. | 26.0 | 7.2 | 4.2 | 7.0 | 18.4 | 18.4 | 6.8 | 1.45 | 23.54 | 43.7 | 330.8 | 15.6 | 56.6 | 14.1 | 29.3 |
| Silver Cr. | 23.7 | 4.7 | 4.3 | 7.3 | 16.3 | 14.3 | 3.9 | 2.04 | 32.82 | 61.0 | 227.7 | 11.7 | 36.0 | 36.2 | 17.9 |
| Sixmile Cr. | 22.6 | 5.4 | 4.4 | 7.2 | 17.0 | 16.2 | 3.8 | 1.51 | 29.29 | 59.4 | 254.8 | 14.4 | 76.0 | 3.9 | 24.0 |
| Sucker R. | 31.2 | 6.8 | 5.0 | 10.4 | 22.2 | 19.8 | 14.8 | 2.24 | 34.19 | 36.7 | 392.2 | 16.6 | 66.4 | 14.2 | 26.9 |
| Two Island R. | 20.2 | 5.4 | 3.8 | 6.8 | 16.0 | 15.0 | 1.8 | 0.98 | 24.88 | 69.0 | 238.4 | 12.4 | 69.3 | 3.8 | 14.8 |
| W. Split Rock R. | 27.3 | 8.0 | 4.3 | 6.3 | 18.7 | 18.0 | 4.1 | 1.99 | 35.91 | 51.1 | 219.3 | 15.3 | 55.5 | 11.1 | 41.4 |
| **Disturbed Streams** | | | | | | | | | | | | | | | |
| ***Agriculture/Rural*** | | | | | | | | | | | | | | | |
| E. Br. Amity Cr. | 18.8 | 7.5 | 2.8 | 3.8 | 14.0 | 11.8 | 3.6 | 2.15 | 45.56 | 57.5 | 104.8 | 10.25 | 55.7 | 4.0 | 18.1 |
| Flute Reed R. | 20.4 | 4.6 | 3.6 | 6.2 | 14.4 | 13.2 | 4.6 | 2.15 | 39.66 | 58.3 | 147.8 | 11.2 | 64.3 | 8.6 | 29.1 |
| Lester R. (avg.) | 20.2 | 4.5 | 2.4 | 7.3 | 14.0 | 15.0 | 13.2 | 2.79 | 76.94 | 64.9 | 276.9 | 13.25 | 87.9 | 10.8 | 20.3 |
| L. East Br. Knife R. | 20.3 | 7.3 | 2.8 | 4.5 | 14.5 | 11.3 | 33.6 | 3.67 | 93.79 | 65.4 | 243.5 | 9.75 | 81.6 | 6.8 | 17.5 |
| Midway R. | 11.0 | 4.0 | 0.7 | 2.7 | 7.7 | 6.3 | 3.9 | 3.19 | 45.83 | 65.4 | 30.3 | 7.0 | 72.1 | 13.7 | 26.8 |
| Stewart R. | 17.8 | 5.8 | 3.5 | 4.3 | 13.5 | 12.8 | 2.2 | 1.94 | 23.62 | 61.9 | 96.5 | 9.5 | 41.8 | 33.8 | 9.8 |
| ***Urban*** | | | | | | | | | | | | | | | |
| Chester Cr. | 7.25 | 2.3 | 0.3 | 2.8 | 5.3 | 3.0 | 11.6 | 3.26 | 72.62 | 65.8 | 16.3 | 4.5 | 71.0 | 1.5 | 24.0 |
| Keenes Cr. | 12.67 | 4.3 | 2.0 | 3.0 | 9.3 | 7.0 | 9.3 | 2.76 | 93.90 | 63.6 | 85.7 | 6.0 | 55.5 | 6.0 | 22.6 |
| Kingsbury Cr. | 5.00 | 0.7 | 0.0 | 1.0 | 1.7 | 0.3 | 95.3 | 5.76 | 100.00 | 95.7 | 22.7 | 2.0 | 81.1 | 5.4 | 75.5 |
| Miller Cr. 2 | 10.00 | .7 | 1.7 | 4.3 | 6.7 | 5.0 | 11.0 | 2.95 | 54.16 | 75.6 | 61.7 | 6.0 | 81.6 | 8.7 | 9.6 |
| Tischer Cr. | 15.00 | 2.4 | 2.6 | 4.4 | 9.4 | 6.0 | 32.3 | 4.26 | 97.94 | 73.9 | 181.8 | 10.4 | 82.9 | 2.4 | 14.2 |

Appendix 3. Metric ranges for Reference streams.

| | Taxa | Ephem | Plec | Trich | EPT | Intol | Tol | BI | Hyd/Tri | Dom | Abund | # Clng | %Clng | Pred |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Minimum | 17.0 | 4.3 | 2.0 | 4.5 | 10.8 | 12.3 | 0.9 | 1.0 | 3.4 | 37.0 | 136.2 | 11.3 | 36.0 | 4.1 |
| L. quartile | 21.0 | 5.4 | 3.8 | 6.2 | 16.2 | 14.8 | 3.8 | 1.5 | 7.6 | 46.8 | 219.3 | 11.8 | 57.6 | 7.1 |
| Median | 23.5 | 6.2 | 4.3 | 6.8 | 17.1 | 16.1 | 5.6 | 1.9 | 14.5 | 51.5 | 228.3 | 14.1 | 66.4 | 12.5 |
| U. quartile | 26.4 | 7.0 | 4.6 | 7.8 | 18.5 | 18.0 | 11.3 | 2.4 | 18.2 | 60.4 | 420.8 | 15.0 | 74.5 | 15.8 |
| Maximum | 33.4 | 8.0 | 5.5 | 11.8 | 24.6 | 27.0 | 21.8 | 3.5 | 62.5 | 72.5 | 533.0 | 20.2 | 80.7 | 36.2 |

Appendix 4. Box plots of index metrics for Reference streams. Whiskers denote $10^{th}$ and $90^{th}$ percentiles.



Appendix 5. Individual metric and final scores for all streams.

| Reference Streams | Taxa | Ephem | Plec | Trich | Tol | BI | Hyd/Tri | Dom | Abund | Pred | Total Score |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Caribou R. | 5 | 5 | 3 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 48 |
| Cascade R. | 5 | 5 | 5 | 5 | 5 | 5 | 3 | 5 | 5 | 5 | 48 |
| E. Split Rock R. | 3 | 5 | 5 | 3 | 5 | 5 | 5 | 5 | 3 | 5 | 44 |
| East Beaver R. | 5 | 5 | 5 | 5 | 5 | 5 | 3 | 3 | 5 | 3 | 44 |
| Elbow Cr. | 5 | 5 | 3 | 3 | 5 | 5 | 5 | 5 | 3 | 5 | 44 |
| Encampment R. | 3 | 5 | 3 | 5 | 5 | 5 | 5 | 5 | 3 | 5 | 44 |
| Gooseberry R. | 5 | 5 | 5 | 5 | 3 | 3 | 3 | 5 | 5 | 5 | 44 |
| Heartbreak Cr. | 3 | 3 | 3 | 3 | 5 | 5 | 5 | 3 | 5 | 3 | 38 |
| Hockamin Cr. | 5 | 5 | 5 | 5 | 3 | 3 | 5 | 5 | 5 | 5 | 46 |
| Irish Cr. | 3 | 5 | 3 | 3 | 5 | 5 | 5 | 5 | 3 | 5 | 42 |
| Junco Cr. | 5 | 5 | 5 | 5 | 5 | 5 | 3 | 5 | 5 | 5 | 48 |
| Kimball Cr. | 5 | 5 | 5 | 5 | 3 | 3 | 5 | 5 | 5 | 5 | 46 |
| Knife R. | 5 | 5 | 5 | 3 | 5 | 5 | 3 | 5 | 5 | 5 | 46 |
| Onion R. | 5 | 3 | 5 | 5 | 5 | 3 | 5 | 5 | 5 | 3 | 44 |
| Poplar R. | 5 | 5 | 3 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 48 |
| Silver Cr. | 5 | 3 | 5 | 5 | 5 | 5 | 5 | 3 | 5 | 5 | 46 |
| Sixmile Cr. | 5 | 3 | 5 | 5 | 5 | 5 | 5 | 3 | 5 | 3 | 44 |
| Sucker R. | 5 | 5 | 5 | 5 | 3 | 3 | 5 | 5 | 5 | 5 | 46 |
| Two Island R. | 3 | 3 | 3 | 5 | 5 | 5 | 5 | 3 | 5 | 3 | 40 |
| W. Split Rock R. | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 3 | 5 | 48 |
| **Disturbed Streams** | | | | | | | | | | | |
| ***Agriculture/residential*** | | | | | | | | | | | |
| E. Br. Amity Cr. | 3 | 5 | 3 | 3 | 5 | 5 | 3 | 5 | 1 | 3 | 36 |
| Flute Reed R. | 3 | 3 | 3 | 5 | 5 | 5 | 3 | 5 | 3 | 5 | 40 |
| Lester R. (avg.) | 3 | 3 | 3 | 5 | 3 | 3 | 1 | 3 | 5 | 5 | 34 |
| Little E. Br. Knife R. | 3 | 5 | 3 | 3 | 1 | 3 | 1 | 3 | 5 | 3 | 30 |
| Midway R. | 3 | 3 | 1 | 1 | 5 | 3 | 3 | 3 | 1 | 5 | 28 |
| Stewart R. | 3 | 3 | 3 | 3 | 5 | 5 | 5 | 3 | 1 | 5 | 36 |
| ***Urban*** | | | | | | | | | | | |
| Chester Cr. | 1 | 1 | 1 | 1 | 3 | 3 | 1 | 3 | 1 | 1 | 16 |
| Keenes Cr. | 3 | 3 | 1 | 3 | 3 | 3 | 1 | 3 | 1 | 3 | 24 |
| Kingsbury Cr. | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 3 | 12 |
| Miller Cr. | 1 | 1 | 1 | 3 | 3 | 3 | 3 | 3 | 1 | 5 | 24 |
| Tischer Cr. | 3 | 1 | 3 | 3 | 1 | 1 | 1 | 3 | 3 | 1 | 20 |