Effects of different sampling effort and taxonomic resolution on assessment metrics

by

Kyung Seo Park

A thesis submitted
In partial fulfillment of the requirements
For the degree of
Master of Science
Natural Resources and Environment
at The University of Michigan
August 2007

Thesis Committee:

Professor Michael J. Wiley, Chair

Professor James S. Diana

ACKNOWLEDGMENTS

There are so many who have contributed to my research and graduate life over past three years. Starting with those at School of Natural Resources and Environment (SNRE), I would like to first thank Professor Michael Wiley, who has great enthusiasm and knowledge and guided me with kindness and patience. I give special thanks to Professor James Diana and Mrs. Barbara Diana whose guidance kept me on track during the graduate life. Both Professors Wiley and Diana helped me realize my purpose for being here and where I want to go after graduate study and beyond. I will never forget their constructive advice on how to be a good scientist and not just a person with a professional degree.

I thank Dr. Catherine Riseng for her good humor and being like an older sister to me. She not only helped with data collection and analysis but also shared her experience in research as well as life in general. I extend special thanks to Dr.'s Shin Suk Choi, Young Pyo Hong, Gwang Guk An, Jin Won Seo and Mr. Jong Hun Shim who since undergraduate years have supported and encouraged my interest in stream ecology. I thank members of Wiley Lab, especially Ethan Bright, Beth Sparks-Jackson, Shaw Lacy, Lori Ivan, Matt Ladewig, Suting Chen, Yu-Chun Kao who helped me with field sampling, lab work, data analysis, and English corrections. I am grateful to Kevin Bolon, Huiling Niu, Hee Sun Chung, and many Korean friends in Ann Arbor for being such good friends and listening. I would like to extend personal thanks to the SNRE

professors and staff for providing me an opportunity to pursue my professional career and MS degree.

Finally, this work would not have been possible without moral support from Chan Seo and Jong Seo, my brothers, Keun Young Kim, and especially my parents Seung Il Park and Keum Ja Lee who always believe in me.

TABLE OF CONTENTS

ABSTRACT

I investigated the effect of different degrees of sampling effort and taxonomic resolution on macroinvertebrate assessment indicator metrics. Specifically, I examined the relationship between a typical rapid assessment approach involving a low-effort sampling (LES) and a more thorough and intensive sampling method, or a high-effort sampling (HES) across a set of watersheds with varying degrees of agricultural impact. Seven macroinvertebrate indicator metrics were significantly different ($\alpha = 0.05$) between low- and high-effort samplings, but response patterns of the indicator metrics were strongly correlated between effort levels. Likewise, metrics based on different levels of taxonomic resolution were significantly correlated each other; however, based on t-tests, metric scores between family- and genus-levels were inconsistent. Normalizing regression models using landscape features were employed to assign standardized impairment status to each site. The models for indicator metrics using HES or genus level identification had higher R^2 and F-ratios than did LES or family level identification. However, there was no significant difference ($\alpha= 0.05$) in normalized indicator scores between levels of sampling effort or taxonomic resolution. Normalized impairment classifications were also significantly ($\alpha= 0.05$) correlated among all combinations of sampling effort and taxonomic resolution. Assessments produced using very different degrees of effort yielded almost identical results when metric results were normalized. LES and family level identification is more cost-

effective since they require less effort and time. However, the higher statistical significance and precision of HES or genus level resolution suggests that where precision is a determining factor, a higher level of sampling effort should be considered.

I. INTRODUCTION

The need for monitoring and assessment programs to monitor, manage, and protect current ecosystems and natural resources is increasing in proportion to industrial development and population size. Natural systems face a wide scope of anthropogenic impacts including hydraulic modifications of channel characteristics for flood control, destruction of habitat by point and non-point source pollution, and conflicting land use interests (Seelbach and Wiley 1996, Hughes and Hunsaker 2002, MDNR 2002, HRWC 2003, Brenden et al. 2006, Riseng et al. 2006). Human activities not only directly influence the biological diversity and population balance of natural systems (Merritt and Cummins 1996, Allan et al. 1997, Lammert and Allan 1997, Wang et al. 2001, Wang et al. 2003, Riseng et al. 2004), but also affect human quality of life (Bradley and Altizer 2007, Esbah 2007). Decision makers, resource managers, and community planners need to initially evaluate current environmental conditions in order to develop appropriate strategies to protect resources and properties (Rabeni and Sowa 1996, Higgins et al. 1999, Seelbach et al. 2002, Riseng et al. 2006).

To quantify environmental change, monitoring and assessment methods have developed in many ways including fundamental surveys, integrated assessments, and development of models to predict environmental impacts (Hughes et al. 1986, Cairns and Pratt 1993, Merritt and Cummins 1996, Seelbach et al. 2002, Wiley et al. 2002, Baker et al. 2005). State agencies, federal agencies, and environmental groups have

used various ecological indicators; such as soil contaminants, air quality, water chemistry, and the presence/absence or abundance of biological organisms to assess environmental conditions (Merritt and Cummins 1996, Olsen et al. 1999, Fore and Yoder 2003). Macroinvetebrate assemblages, in stream and river studies, have often been a preferred indicator of environment conditions due to their sensitivity to pollution, relative immobility, ease of collection, and quantity of taxa and individuals (Hellawell 1986, Rosenberg and Resh 1993, Merritt and Cummins 1996). As a result, the macroinvertebrate assemblages have played an important role in the development of many analytical methods and indicator metrics to assess biological changes resulting from anthropogenic impacts (Hilsenhoff 1987, Johnson et al. 1993, Resh and Jackson 1993, Wiley et al. 2002, Fore and Yoder 2003).

Different levels of effort in field sample collection are one of the most pivotal factors affecting indicator metrics. The question of how much to sample affects costs in time and money (Yoon et al. 1998, Cao et al 2002), accuracy in ecological analysis (Morin 1997), and quality of resulting management and protection plans (Fore and Yoder 2003). Sampling effort includes sampling devices, work effort, taxonomic resolution, specific habitat selection, sample size, and study site characteristics (Merritt and Cummins 1996, Wiley et al. 2002, Fore and Yoder 2003). The sampling effort necessary to provide appropriate data has been addressed by several studies (Metcalfe-Smith and Maio 2000, Larsen et al. 2001, Cao et al. 2002, Fore and Yoder 2003). Several studies suggest that emphasis should be placed on level of effort that enhances

accuracy and precision for predicting current condition and possible changes (Merritt and Cummins 1996, Larsen 1997, Wiley et al. 2002, Fore and Yoder 2003). However, others note that excessive sampling effort can result in wasted time and money (Brewer and McCann 1982, Yoon et al. 1998) especially if it does not enhance the interpretation of data.

An explicit comparison of different degrees of sampling effort can clarify both efficiency of sampling and data accuracy. Sampling for macroinvertebrate assemblages has been often classified as either qualitative or quantitative, which are well explained in many documents (U. S. EPA 1998, Merritt and Cummins 1996, MDEQ 1997, Fore and Yoder 2003). Previous studies have compared effects of sample size (Brewer and McCann 1982, Yoon et al. 1998), sampling area, subsampling procedures (Vinson and Hawkins. 1996), and sampling devices (Kroger 1972, Mason 1976) on aquatic insect communities and taxa richness. While descriptive comparisons are often done for different methodologies (Merritt and Cummins 1996, Olsen et al. 1999, Fore and Yoder 2003), quantitative comparisons of different degrees of sampling effort to test for practical difference in biological data and to evaluate monitoring data sets are rare (Wiley et al. 2002, Riseng et al. 2006).

A variety of assessment modeling methods has been developed to predict reference conditions thereby helping to evaluate current status of streams and rivers due to anthropogenic impacts (Turak et al. 1999, Olden and Jackson 2001, Wiley et al. 2002, Riseng et al. 2006). Many assessments are challenged by scarcity of historical

data for estimating reference conditions, are implemented with limited data, or include wide scales of reference conditions (Seelbach et al. 2002, Wiley et al. 2002, Riseng et al. 2006). Recently, in the Great Lakes region (USA), regional ecological normalization has been useed to help clarify assessment assumptions and integrate multiple types of indicator metrics (Wiley et al. 2002). This normalization uses a linear model to predict reference condition from biological assemblages and a suite of landscape variables expected to influence ecological condition. The effects of different degrees of sample collection effort and taxonomic resolution on the normalization approach have not yet been examined (Wiley et al. 2002, Riseng et al. 2006).

In this study, I investigated the effects of different degrees of field collection effort and taxonomic resolution on assessment metrics. Specifically the relationship between a typical rapid assessment approach, the Michigan Department of Environmental Quality (MDEQ) Procedure 51 protocol, and the more effort intensive collection methods employed by the Michigan River Inventory (Seelbach and Wiley 1996) and the Huron River Watershed Council (Wiley et al. 2002, HRWC 2003) were examined. Thus, my first objective was to compare samples for taxonomic composition and assessment results of all indicator metrics produced by these different levels of sampling effort and taxonomic resolution. For this objective, I hypothesized that additional effort or higher resolution of taxonomic analyses would produce more precise assessment data of all indicator metrics. My second objective was to provide advice for resource managers and researchers with expected results that different

degrees of effort in sample collection and taxonomic resolution affect efficiency and accuracy of assessments.

II. METHODS

1. Study Area

A total of 52 sites were sampled for macroinvertebrate assemblage data at five representative basins in Lower Peninsula of Michigan and northern Ohio (Figure 1) to compare the effects of different degrees of sampling effort (low-effort sampling, LES and high-effort sampling, HES) and taxonomic resolution (family- and genus-level) on sample taxonomic composition and assessment indicator metrics. The study sites were located in five watersheds in the Midwestern U.S. characterized by different proportion of agricultural land use, water temperature, and stream flow (Table 1, Figure 1 and 2): Crane Creek (OH), Mill Creek (MI), and three tributaries of Muskegon River (MI), Bigelow Creek, Brooks Creek, and Cedar Creek. In particular, these five watersheds were distinctly characterized by different proportion of agricultural land use, water temperature, and stream flow.

Crane Creek (13 sites) is a small agricultural tributary located in the western basin of Lake Erie that flows northeasterly. The stream length and drainage area were approximately 32.2 km long and 143.5 km^2, respectively (Wells 2001). Topology of Crane Creek is relatively flat from the head water to the estuary (Kasat 2006). Mean percentages of agricultural and urban land uses were 84.1% (SD= 4.18, n=13) and 6.28% (SD= 1.77, n=13) (Table 1) and tributaries of the upper watershed are primarily converted to agricultural ditches (Ohio DNR 1996, Kasat 2006).

Mill Creek (11 sites), located in Michigan's southeastern Lower Peninsula, is the largest tributary (approximately 374.7 km^2) of the Huron River. For these sites, agricultural land use averaged 40.67% (SD= 15.63, n=11) while urban land use averaged 5.11% (SD= 1.96, n=11) (Table 1). However, the Mill Creek watershed has increasing pressure from local urbanization and population growth (Seelbach and Wiley 1996, HRWC 2003). The surficial geology of Mill Creek is nearly 50 percent glacial till and approximately 25 percent each glacial outwash and end moraine (Hay-Chmielewski et al. 1995). Mill Creek hydrology is a mix of runoff and groundwater contributions due to a combination of rolling till plain topography that generates runoff and prevents extensive infiltration and higher basin slope that helps to provide groundwater to stabilize baseflow (Seelbach and Wiley 1996).

The Muskegon River (28 sites), located in western mid-Michigan, is the second longest river and the third largest watershed in Michigan (MRWA 2005). Agricultural and urban land use dominated the watershed (33.4 % and 9.6 %, respectively) and urban land use was relatively minor (9.6 %) (MCD 2004, Riseng et al. 2006). Study sites were located in three of the primary tributaries of the lower portion of the river; Bigelow Creek, Brooks Creek, and Cedar Creek (Table 1, Figure 1 and 2). Sites on three studied watersheds had an average of 35.0 % (SD= 19.75, n=28) agricultural and 3.5 % (SD= 2.22, n=28) urban land uses, similar to the Muskegon River watershed as a whole.

2. Levels of Sampling Effort and Taxonomic Resolution

To investigate the effects of different degrees of sampling effort on taxonomic composition and assessment metrics of macroinvertebrates, I examined the relationship between a typical rapid assessment approach involving a low-effort sampling (LES) and the more effort intensive collection method (high-effort sampling, HES) used by the Michigan Rivers Inventory (MRI) (Seelbach and Wiley 1996, Wiley et al. 2002, HRWC 2003). Field sampling with LES and HES was conducted at the same time and location for each site. The reach length was 12 times the average stream width and ranged from 50m to 250m. Detailed descriptions of macroinvertebrate sampling methods are provided in elsewhere (LES in MDEQ 1996, Merrit and Cummins 1996, MDEQ 1997; HES in Seelbach and Wiley 1996, Riseng et al. 2006).

The low-effort sampling (LES) is described in MDEQ Procedure 51 (MDEQ 1996, and MDEQ 1997). The LES is used for characterizing the structure of invertebrate communities in terms of relative abundances of each taxon rather than absolute density (Moulton et al. 2002). Survey for this study was conducted in accordance with the Great Lakes Environmental Assessment Section (GLEAS) Procedure 51 (MDEQ 1997). D-frame dip nets (250 μm mesh) were used to sample macroinvertebrate assemblages for 30 minutes at each site by one person. Kicking, dipping, and sweeping were used for general sampling with the dip net, and hand-picking was used for areas with boulders, debris, and logs. Samples from all habitats

were combined in a basket and then 100 organisms were randomly selected from the composite sample for further analysis (Merritt and Cummins 1996, MDEQ 1997, Riseng et al. 2006). The 100 selected organisms were preserved in 70 % ethanol and returned to the laboratory for identification and enumeration (Merritt and Cummins 1996).

The objective of high-effort sampling (HES) is to obtain as many different macroinvertebrate taxa as are present in a sampling reach (Fore and Yoder 2003). The HES was conducted over 2 person-hours per reach, while larger streams were received proportionally more effort. All habitats in the reach were sampled. In practice one person-hour each was targeted toward erosional or depositional habitat, but sometimes center and edge if only one type of habitat occurred. A variety of sampling methods were used including D-frame dip nets with 250 μm mesh, kick screens, and hand picking. Collectors recorded taxon name and relative abundance from erosional and depositional habitat types on site field sheets and representative specimens were placed into vials by habitat type for lab validation of taxonomy (Riseng et al. 2006).

In order to investigate the effects of different level of taxonomic resolution on sample taxonomic composition and assessment indicator metrics, macroinvertebrates from low- and high-effort samplings were identified to family- and genus-level in the laboratory. Generally aquatic insects were identified to family- and genus-level while all other groups were identified to family- or order-level (Annelida, Crustacea, and Mollusca). Then, numbers of taxa for each indicator metric were counted to compare

the effects of family- and genus-level taxonomic identification on assessment metrics.

All identified macroinvertebrate taxa were counted and classified into seven indicator metrics: number of total taxa (*I-totaxa*), number of EPT (Ephemeroptera, Plecoptera, and Tricoptera) taxa (*I-EPT*), average Macroinvertebrate Biotic Index (MBI) taken from Hilsenhoff or EPA established biotic index values (*I-MBI*; Hilsenhoff 1987, USEPA 2006), number of sensitive taxa (*I-sens*), number of metabolic conformer taxa (*I-metc*), number of surface dependent taxa (*I-surf*), and number of surface dependent taxa divided by number of total taxa (*I-suta*). A tolerance value for each taxon ranged from 0 to 11 (Hilsenhoff 1987, USEPA 2006) and the average MBI score of each site was calculated by averaging sum of a published tolerance value for each taxon collected (Riseng 2006). These seven macroinvetebrate indicator metrics were used as the basis for my comparisons of sample taxonomic overlap, linear regression models, and normalized assessment scores and classifications for different degrees of sampling effort and taxonomic resolution (Wiley et al. 2002, Riseng et al. 2006).

3. Assessment Process and Data Analysis

My regional models included variables at both site and catchment scale to develop the best linear models for all macroinvertebrate indicator metrics. Site-based variables were measured in the field and included stream width, reach slope, water temperature (2005), and stream flow. For mean July water temperatures, I used field collected temperatures for each site in Crane Creek, OH (Kasat 2006) and used

predicted July water temperatures from landscape-based Krieging models (Brenden et al. 2006) for all Michigan stream sites.

Catchment scale data were used as primary control factors in my regional modeling (Roth 1994, Allan et al. 1997, Wiley et al. 2002, Seelbach et al. 2002, Wang et al. 2003, Riseng et al. 2006) and I summarized the catchment scale landscape variables (drainage area, land use, and geology) using Geographic Information System (ESRI 2005). The GIS maps were delineated by the Michigan and Ohio Departments of Natural Resources (MDNR and ODNR) from U.S. Geological Survey 1:24,000 scale topographic maps and modified for each site using 1:250,000 scale resolution (Wiley et al. 2002, Baker et al. 2005, Riseng et al. 2006). Catchment and riparian buffer (100m) areas were applied to summarize proportion of land use categories (urban, agriculture, forest, range, wetland, forested wetland, non-forested wetland, and water) and surficial geology categories (coarse-till, outwash, and ice-contact) by using land-cover and geology maps from 1998 MRI System and from Division of Geological Survey, ODNR (Brenden et al. 2006, Riseng et al. 2006). Also, I calculated the proportions of urban land cover above 10% (Brabec et al. 2002) and agricultural land cover above 25% to capture only strong stressor effects of highly developed land covers (Riseng et al. 2006). Forest land use was not used for models due to the strong inverse correlation with agricultural data.

The above environmental data were placed into two categories: non-stressor (natural) and stressor variables. Natural landscape variables which could influence

stream invertebrate assemblages were drainage area, mean July water temperature, stream slope and width, stream discharge, low to high flow ratio, and geology. Anthropogenic stressor variables that potentially affected stream invertebrate assemblages were urban and agricultural land uses (Brabec et al. 2002, Riseng et al. 2006).

I used multiple linear regression in a regional ecological normalization process using a multiple linear model to produce assessment scores for degrees of sampling effort and taxonomic resolution (Wiley et al. 2002, Baker et al. 2005, Riseng et al. 2006). I constructed MLR models for each macroinvertebrate metric and effort-resolution combination using a 2x2 factorial combination of sampling effort and taxonomic resolution (LES & family, LES & genus, HES & family, and HES & genus). Each MLR was constructed from independent site-, catchment- and buffer-scale (100m) variables (Wiley et al. 2002) having ecological, biological, and statistical significance for the specific macroinvertebrate indicator variable. Independent variables included nonstressors such as drainage area, mean July water temperature, stream discharge, ratio of low to high flow, and stream width and stressor variables such as agricultural and urban land use in the catchment and buffer. MLR models were selected that maximized R^2 and significance of regression coefficients and used for calculation of expected ecological condition. All indicator variables used for MLR models were significant at $p < 0.05$ in the model. Before using the independent variables in the models, the integer 1 was added to the variable and then transformed to

natural log form to meet assumptions of normality for all variables.

I calculated normalized assessment scores for each indicator metric using the MLR models for the expected condition of each site, then calculated the difference between expected and actual condition. Reference condition scores were estimated by setting stressor variables (agricultural and urban land use) in MLR models to zero. I then calculated deviation values for each indicator metric by subtracting the expected value from the observed value. For *I-MBI* and *I-surf*, the deviation values were calculated by subtracting the observed value from the expected value for each site because an increase in those taxa indicates a decline in ecological condition. Finally, the deviation values were scaled by dividing the deviation by the standard deviation of the modeled reference expectation to produce a normalized score scaled by standard deviation units.

A composite normalized score was calculated to represent final assessment condition for each site by averaging the normalized scores for some significant and stable indicator metrics. These metrics included number of total taxa, number of EPT taxa, number of metabolic conformer taxa, and MBI. A normalized score close to zero indicated no impact by anthropogenic stressors because the difference between the observed and expected values was near zero. Positive or negative normalized scores indicated that a site was better or worse than expected based on the predictive model. I established a general assessment classification based on normal distributions and standardized scores (Z-scores). Normalized scores above 0.5 were assigned

"exceptional," scores between -0.5 and 0.5 were assigned "good," scores below -0.5 and above -1.0 were assigned "threatened," scores below -1.0 and above -2.0 were assigned "poor," and scores below -2 were assigned "very poor."

Independent samples t-test and Pearson correlation were used to compare the raw macroinvertebrate data sets and normalized scores between and among all indicator metrics for different degrees of sampling effort and taxonomic resolution using SPSS 12.0 (SPSS, Inc. 2003). Also, Chi-Square test was used for comparison of impairment classification among a 2x2 combination of sampling effort and taxonomic resolution. Multiple linear regression models, box plots, and statistical summaries (mean, median, standard deviation, minimum, and maximum) were performed in Datadesk (Velleman and Velleman 1988).

III. RESULTS

Different degrees of sampling effort influenced taxa numbers of each indicator metric. Numbers of collected taxa using HES had greater numbers of taxa compared to the numbers using LES for all indicator metrics from 52 studied sites (Table 2, Figure 3). Independent samples t-tests showed that macroinvertebrate data of each indicator metric between low- and high-effort samplings were significantly different ($\alpha = 0.05$, df = 102) at both family- and genus-level identification (Table 3). This indicated that different degrees of sampling effort affected number of taxa for each indicator metric and HES included more taxonomic data than the LES, while number of families and genera increased with sampling effort.

Different levels of taxonomic resolution were inconsistent in describing the effect of family- and genus-level identification on taxonomic composition. Different levels of taxonomic resolution showed that mean taxa numbers or values were significantly higher at genus levels than at family level except for *I-MBI* (Table 2, Figure 3). Different levels of taxonomic resolution also influenced number of taxa for each indicator metric in HES (Table 3). However, taxa numbers in LES had inconsistent independent samples t-test results for each indicator metric between family- and genus-level identification. These results indicated that family-level identification was more efficient than genus-level identification in low effort sampling, but genus-level identification provided more detailed taxonomic information than

family-level identification in HES.

Different degrees of sampling effort and taxonomic resolution did not affect *I-MBI*. Means and medians of the MBI scores were not significantly different ($\alpha = 0.05$, df = 102) between different degrees of sampling effort and between taxonomic resolution (Table 3). *I-MBI* produced the same results regardless of degrees of sampling effort and taxonomic resolution because a tolerance value of taxon at family level was calculated by a composite averaging tolerance value of collected genera in the family.

I-surf and *I-suta* did not show consistent correlations to other indicator metrics (Table 4). Compared to the significant difference for number of taxa between different degrees of sampling effort and taxonomic resolution, all indicator metrics except for *I-surf* and *I-suta* were significantly correlated to each other reflecting similar relationships to stream conditions ($\alpha = 0.05$, n = 52). *I-surf* and *I-suta* showed positive correlations to *I-MBI*, and these indicators had negative correlations to other indicator metrics (Table 5). MLR models for metrics of *I-surf* and *I-suta* were also not developed because ecologically reasonable models with significance of $p < 0.10$ could not be constructed with available data.

HES and genus-level identification produced a more detailed taxonomic composition than LES and family-level identification. Total number of taxa at each site using LES and family-level identification overlapped highly with the taxa list using HES or genus-level identification, respectively (Table 5). At family level, 81.22% of

mean total taxa collected using LES were also collected using HES, while only 55.31% of those collected using HES were also collected using LES. At genus level, 75.45% of mean total taxa collected using LES were also collected using HES while only 48.28% of those collected using HES were also sampled using LES. Therefore, LES and family-level identification did not provide as much detail for biological data of macroinvertebrate assemblages compared to HES and genus-level identification.

MLR models of HES had better fits and used landscape variables that have more ecological meaning than those of LES. Model statistics of constructed regression models for each indicator metrics showed that models of HES had higher R^2 values and F-ratios ($\alpha = 0.05$, $n = 52$) than those of LES at both taxonomic levels (Table 6 and 7). In the MLR models, *I-EPT*, *I-MBI*, and *I-metc* included nearly similar landscape variables at both LES and HES (Table 6). However, *I-totaxa* and *I-sens* at HES were supported by more landscape variables than at LES. MLR models for the HES explained over 65.4% of the variance (Table 7), indicating that the HES models produced better fits to the data, represented by higher R^2s and F-ratios, than LES models. This also indicated that models of HES were explained better with landscape variables and more detailed data than those of LES.

MLR models for each indicator metric showed significant effects ($\alpha = 0.05$, $n = 52$) with different levels of taxonomic resolution. Indicator metrics for genus-level identification had higher R^2 values and F-ratios than those for family-level identification at both different levels of sampling effort (Table 7). The model statistics

explained that genus-level identification influenced MLR models and explained more of the variance than family-level identification. Therefore, accuracy of MLR models with landscape variables was improved when genus-level identification was used rather than family-level identification.

Normalized scores and composite scores were not influenced by different degrees of sampling effort. Mean normalized scores at HES were lower than those at LES for both levels of taxonomic resolution except for *I-MBI* at genus-level identification (Table 8, Figure 4). However, statistical significance ($\alpha = 0.05$, $df = 102$) indicated that normalized scores of all indicators had no differences between LES and HES except for *I-sens* (Table 9). Normalized scores for all indicator metrics showed highly significant correlations ($\alpha = 0.05$, $df = 102$) regardless of degrees of sampling effort at both family and genus levels (Table 10). Chi-square tests showed that 5 categories of impairment classification were not significantly different ($\alpha = 0.05$, $df = 102$) among 2x2 factorial combinations of sampling effort and taxonomic resolution. Also, about 52 % and 60 % of impairment classification between degrees of sampling effort gave the same results at family- and genus-level identification, and about 79 % and 85 % between levels of taxonomic resolution gave the same results at LES and HES, respectively (Table 11). The normalization process could show consistent significant correlation among indicator metrics despite data transformation.

Normalized assessment scores for each indicator metric were not significantly different ($\alpha = 0.05$, $df = 102$) between family- and genus-level identification at both

degrees of sampling effort (Table 9). Normalized assessment scores for each indicator metric showed strong correlations ($\alpha = 0.01$) between levels of taxonomic resolution at both degrees of sampling effort (Table 10). Also, about 52 % and 60 % of impairment classification between levels of taxonomic resolution were matched at LES and HES, respectively (Table 11). Therefore, family-level identification was more efficient to produce normalized assessment scores while genus-level identification produced better model statistics.

Macroinvertebrate data and MLR models were significantly affected by different degrees of sampling effort and taxonomic resolution, indicating that HES and genus-level identification included more taxa information. Therefore, HES and genus-level identification explained more accurate data for biological and ecological interpretation than LES and family-level identification. However, normalized assessment scores and impairment classifications for each indicator metric were not significantly different ($\alpha = 0.05$, $df = 102$) between different degrees of sampling effort and taxonomic resolution. This indicated that LES and family-level identification was more efficient to produce normalized assessment scores with less sampling and identification intensity.

IV. DISCUSSION

High-effort sampling and genus-level identification affected taxonomic composition and ecological models, but did not affect assessment scores and impairment classification by the ecological models. Regional ecological normalization using linear models can adjust statistically distinct biological data sets from different degrees of sampling effort and taxonomic resolution and then produce an almost similar final assessment. Regional normalizing model has been recently used as an alternative assessment (Wiley et al. 2002, Baker 2005 et al., Riseng et al. 2006), because previous monitoring and assessment approaches have had limitations in predicting a reference condition and comparing various assessment results (Gallant et al. 1989, Claessen et al. 1994. Seelbach et al. 2002, Wiley et al. 2002, Riseng et al. 2006). Several recent studies also have shown that the normalizing model that can integrate multiple types of data and indicator metrics, determine the effect of different spatial scales, and communicate relative risks of impairment (Wiley et al. 2002, Riseng et al. 2006). Comparison of different levels of sampling effort (LES and HES) and taxonomic resolution (family- and genus-level identification) demonstrate that the normalizing models can integrate multiple types of data from different methods (Wiley et al. 2002).

In addition, the MLR models suggest that multiple types of methods can be modeled and interpreted with landscape features (Wiley et al. 2002, Riseng et al. 2006).

Each indicator metric used almost similar landscape-scale variables to construct an MLR model with four combinations of sampling effort and taxonomic resolution. Inclusion of these similar independent variables using different levels of both sampling effort and taxonomic resolutions indicated that landscape-scale features strongly influenced macroinvertebrate assemblage regardless of sampling design. Colleted biological data directly reflected the condition of a catchment area. The theory of using catchment and landscape features has been proposed by various studies (Hynes 1975, Frissell et al. 1986, Allan et al. 1997, Poff 1997, Wang et al. 2003) and has been used recently for various assessment, recruitment, and management studies (Wiley et al. 2002, Baker 2005, Riseng et al. 2006).

In this study, normalization was also used to rescale differences in biological data from both different degrees of sampling effort and taxonomic resolution (Wiley et al. 2002, Baker et al. 2005, Riseng et al. 2006). Normalization is generally used for mathematical transformation and to produce unitless scores in order to facilitate interpretation. Independent samples t-test results showed that macroinvertebrate indicator metrics were significantly different ($\alpha = 0.05$) due to different levels of sampling effort. However, normalized scores did not show any significant difference ($\alpha = 0.05$) between different degrees of sampling effort. This comparison of raw and normalized data sets demonstrates that normalization using the MLR models did well in moderating differences of sampling methodologies (Wiley et al. 2002, Baker et al. 2005). Normalized scores for each indicator metric showed highly significant

correlations between different levels of sampling effort and taxonomic resolution , as did raw data indicator metrics.

Different degrees of sampling effort affect observed taxon richness. The higher compositional overlap observed with HES can describe more detailed and accurate sample taxonomic composition than LES, suggesting that different degrees of sampling effort affected the sample taxonomic composition differently. Other studies have shown that LES can be used to easily detect general impairment status (Merritt and Cummins 1996), while HES can help to represent a more extensive impact on local environment (Moulton et al. 2002, Fore and Yoder 2003). However, normalized scores of each indicator metric using regression models did not show any significant difference between different levels of sampling effort. Also, impairment classification based on composite scores showed no significant differences between different degrees of sampling effort at both family and genus levels. Since regression models of indicator metrics of HES had higher R^2 values and significant F-ratios, using HES can increase stability of sampled data and minimize sampling errors among studied sites. The resulting regression models can be used to explain more exact effects of landscape variables on biological data.

Impairment classifications showed final impairment status for each site and how assessment results were different using four methodological combinations of sampling effort and taxonomic resolution. The impairment classification between different degrees of sampling effort (52 % at family level and 60 % at genus level) and

taxonomic resolutions (79 % at LES and 85 % at HES) was well matched between sampling efforts, while one category difference of impairment classification was mostly observed in comparison of impairment classification. The classification difference might have several causes: anthropogenic pressures, analysis error, or sampling error. The first two may be related to macroinvertebrate abundance, which affects statistical accuracy and is influenced by natural variability (Gerrodette 1987, Buckland et al. 2000, Fore and Yoder 2003). Most sites in this study were from categories of threatened to very poor composition, due to high anthropogenic development. As a result, populations were quite low. The number of indicator taxa collected could be significantly different using higher sampling effort and taxonomic resolution. Second, difference of impairment classification was related to specific dominant taxa and relative abundance of some specific indicator taxa group. Also, indicator metrics of both different degrees of sampling effort were designed to consider total number of indicator taxa so that assessment result would only be dependent on number of indicator taxa: a site with equal number of indicator taxa, but with different relative abundance of individuals would be hard to explain. Finally, types of substrates or sampling devices might also affect sample collection (Resh 1979, Merritt and Cummins 1996). Low-effort sampling (LES) could be restricted to subsample macroinvertebrate samples from a basket in which all samples were mixed with various materials of stream substrates, such as thick algae, debris, mud, and leaves. Conversely, in high-effort sampling (HES), all invertebrate taxa could be collected whenever

sampled. However, normalized scores from both samplings were not significantly different each other despite the above concerns.

Almost all metrics showed similar results through all analyses except indicator metrics related to surface dependent taxa. I used two indicator metrics for the surface dependent taxa: number of surface dependent taxa (*I-surf*) and relative abundance of the surface dependent taxa (*I-suta*). I could not develop MLR models related to surface dependent taxa, due to poor fit of the data ($\alpha = 0.10$). In addition, compared to other indicator metrics, the statistical analyses could not show clear correlation and independent samples t-test results for comparisons of different levels of sampling effort and taxonomic resolution. The potential for sampling bias for the surface dependent is obvious because most taxa in this group are free to move and to escape from the collector unlike other indicator taxa and collectors' methodology can influence collection of the surface dependent taxa. While this indicator is currently generally used in monitoring and assessment programs in Michigan, these results suggest that the surface dependent taxa should be rejected as a reliable indicator.

The indicator metric for sensitive taxa should also be reconsidered for normalized assessment scores. Normalized scores of sensitive taxa showed significant difference ($\alpha = 0.01$) between low-effort (LES) and high-effort (HES) sampling at both family- and genus-level identification, whereas other indicators had no significant difference in normalized assessment scores. Sensitive taxa were limited to some specific families that included many sensitive genera. Determination of sensitive taxa is

relatively subjective, but these can be crucial both to reflect instream conditions and related landscape variables in regression models. Each taxon was identified as "sensitive" based on published tolerance values (USEPA 2006) at genus level. In this study, only a few sensitive taxa were ever collected at each site and counts of the sensitive taxa was strongly affected by sampling effort. Thus, this might have influenced the regression models which produced biased scores like the surface dependent taxa.

One of most important questions is whether differences in accuracy and efficiency of different degrees of sampling effort and taxonomic resolution can lead to different assessment results. For degrees of sampling effort, LES can be more cost-efficient because this sampling effort spends less time on sampling, fewer invertebrates are collected for identification, and less human effort is needed. LES produced the same normalized assessment scores and impairment classification as HES, but LES had less information of sample taxonomic composition. However, assessment results from HES are supported by better sample taxonomic composition, stronger statistical significance, and more landscape variables than from LES. Thus, the HES may have less risk for temporal and seasonal environmental changes, so that HES with higher accuracy can show more reliable assessment estimation of current condition (Li et al. 2001, Fore and Yoder 2003). For levels of taxonomic resolution, genus level also showed relatively higher sample taxonomic composition and accuracy for MLR models with statistical evidence than family level. However, the family level was more

efficient in producing normalized assessment scores due to less time effort for identification. The genus-level identification would be recommended for assessment modeling based on landscape features because some indicators are based on each species rather than a family, which frequently might result in analytical errors to explain ecological landscape effects on biological assemblages.

V. References

Allan, J. D., D. L. Erickson, and J. Fay. 1997. The influence of catchment land use on stream integrity across multiple spatial scales. Freshwater biology 37: 149-162.

Baker, E. A., K. E. Wehrly, P. W. Seelbach, L. Wang, M. J. Wiley, and T. Simon. 2005. A multimetric assessment of stream condition in the Northern Lakes and Forests ecoregion using spatially explicit statistical modeling and regional normalization. Transactions of the American Fisheries Society 134(3): 697-710.

Brabec, E., S. Schulte and P. Richards. 2002. Impervious surfaces and water quality: a review of current literature and its implications for watershed planning. Journal of Planning Literature 16: 499-514.

Bradley, C. A., and S. Altizer. 2007. Urbanization and the ecology of wildlife diseases. Trends in Ecology and Evolution 22 (2): 95-102.

Brenden, T. O., R. D. Clark,A. R. Cooper, P. W. Seelbach, L. Wang, S. S. Aichele, E G. Bissell, and J. S. Stewart. 2006. A GIS framework for collecting, managing, and analyzing multi-scale landscape variables across large regions for river conservation and management. *In* R. Hughes, L. Wang, and P.W. Seelbach, (eds.), Landscape Influences on Stream Habitats and Biological Communities. American Fisheries Society Symposium 48, Bethesda, Maryland.

Brewer, R., and M. T. McCann. 1982. Laboratory and field manual of ecology. Saunders College Publishing. New York. 269 p.

Buckland S. T., I. B. J. Goudie, and D. L. Borchers. 2000. Wildlife population assessment: past developments and future directions. Biometrics 56: 1-12.

Cairns, J., Jr., and J. R. Pratt. 1993. A history of biological monitoring using benthic macroinvertebrates. pp. 10-27. *In* D. M. Resenberg and V. H. Resh (eds.). Freshwater biomonitoring and benthic macroinvertebrates. Champman and Hall, New York.

Cao, Y., D. P. Larsen, R. M. Hughes, P. L. Angermeier, and T. M. Patton. 2002. Sampling effort affects multivariate comparisons of stream assemblages. Journal of the North American Benthological Society 21 (4): 701-714.

Claessen, F. A. M., F. Klijn, J. P. M. Witte, and J. G. Nienhuis. 1994. Ecosystem classification and hydro-ecological modeling for national water management. pp 199-222. *In* F. Klijn, editor. Ecosystem classification for environmental management. Kluwer, Dordrecht, The Netherlands.

Esbah, H. 2007. Land use trends during rapid urbanization of the city of Aydin, Turkey. Environmental management 39:443-459.

ESRI. 2005. ArcGIS. Version 9.1. (Build 722), Redlands, CA.

Fore, L., and C. O. Yoder. 2003. The design of a biological community trend monitoring program for Michigan wadeable streams. MI/DEQ/WD-03/086. Michigan Department of Environmental Quality (MDEQ). Surface Water Quality Division. Lansing, Michigan.

Frissell, C. A., W. J. Liss, C. E. Warren, and M. D. Hurley. 1986. A Hierarchical Framework for Stream Habitat Classification - Viewing Streams in a Watershed Context. Environmental Management 10: 199-214.

Gallant, A. L., T. R. Whittier, D. P. Larsen, J. M. Omernik, and R. M. Hughes. 1989. Regionalization as a tool for managing environmental resources. EPA/600/3-89/060. U. S. Environmental Protection Agency, Corvallis, Oregon.

Gerrodette, T. 1987. A power analysis for detecting trends. Ecology 68: 1364-1372.

Hay-Chmielewski, E.M., P. Seelbach, G. Whelen, and D. Jester. 1995. Huron River Watershed Assessment. Michigan Department of Natural Resources, Fisheries Division. Lansing, Michigan.

Hellawell, J. M. 1986. Biological indicators of freshwater pollution and environmental management. Elsevier, London. 546 pp.

Higgins, J., M. Lammert, and M. Bryer. 1999. Including aquatic targets in ecoregional portfolios: guidance for ecoregional planning teams. Designing a geography of hope, Update 6, The Nature Conservancy, Arlington, Virginia.

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

Hughes, R. M., D. P. Larsen, and J. Omernik. 1986. Regional reference sites: a method for assessing stream potentials. Environmental Management 10 (5): 6291-635.

Hughes, R.M., and C.T. Hunsaker. 2002. Effects of landscape change on aquatic biodiversity and biointegrity. pp 309-329. *In* Gutzwiller, K. J., editor. Applying landscape ecology in biological conservation. Springer-Verlag, New York.

Huron River Watershed Council (HRWC). 2003. Mill Creek Subwatershed Management Plan, Huron River Watershed (HUC 04090005), Ann Arbor, Michigan.

Hynes, H. B. N. 1975. The stream and its valley. Internationale Vereinigung für theoretische und angewandte Limnologie Verhandlungen 19: 1-15.

Johnson, R. K., T. Wiederholm, and D. M. Rosenberg. 1993. Freshwater biomonitoring using individual organisms, populations, and species assemblages of benthic macroinvertebrates. Chapman and Hall, New York.

Kasat R. J. 2006. Nutrient dynamics in a small agricultural Lake Erie tributary. M.S. thesis, University of Michigan, Ann Arbor, MI.

Kroger, R. L. 1972. Underestimation of standing crop by the Surber sampler. Limnology and Oceanography 17 (3): 475-478.

Lammert, M., and J. D. Allan. 1997. Assessing biotic integrity of streams: effects of scale in measuring influence of land use/cover and habitat structure on fish and macroinvertebrates. Environmental Management 23: 257-270.

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

Larsen, D. P., T. M. Kincaid, S. E. Jacobs, and N. S. Urquhart. 2001. Designs for evaluating local and regional scale trends. BioScience 12: 1069-1078.

Li, J., A. Herlihy, W. Gerth, P. Kaufmann, S. Gregory, S. Urquhart, and D. P. Larson. 2001. Variability in stream macroinvertebrates at multiple spatial scales. Freshwater Biology 46: 887-897.

Mason, J. C. 1976. Evaluating a substrate tray for sampling the invertebrate fauna of small streams with comments on general sampling problems. Archiv für Hydrobiologie 78: 51-70.

Merritt, R. W., and K. W. Cummins. 1996. An introduction to the aquatic insects of North America. Kendall Hunt Publishing, Dubuque, Iowa.

Metcalfe-Smith, J. L., and J. D. Maio. 2000. Effect of sampling effort on the efficiency of the timed search method for sampling freshwater mussel communities. Journal of the North American Benthological Society 19 (4): 725-732.

Michigan Department of Environmental Quality (DEQ). 1996. Update of GLEAS Procedure 51 Metric Scoring and Interpretation. Staff report. Lansing, Michigan.

Michigan Department of Environmental Quality (DEQ). 1997. Qualtitative biological and habitat survey protocols for wadeable streams and rivers. Michigan Department Environmental Quality (MDEQ), Surface Water Quality Division, Great Lakes and Environmental Assessment Section, GLEAS Procedure 51 (Revised 2002). Lansing, Michigan.

Michigan Department of Natural Resources (MDNR). 2002. Huron River Plan (Livingston, Washtenaw, Oakland, Monroe and Wayne Counties). Michigan Department of Natural Resources, Fisheries Division. Lansing, Michigan.

Morin, A. 1997. Empirical models predicting population abundance productivity in lotic systems. Journal of North American Benthological Society 16: 319-337.

Moulton, S. R., J. G. Kennen, R. M. Goldstein and J. A. Hambrook. 2002. Revised protocols for sampling algal, invertebrate and fish communities as part of the National Water-Quality Assessment Program. USGS Open-File Report 02-150. Reston, Virginia.

Muskegon Conservation District (MCD). 2004. Muskegon River and AOC Watershed. (Source: http://www.deq.state.mi.us/documents/deq-ess-nps-muskegon-river.pdf). Ann Arbor, Michigan.

Muskegon River Watershed Assembly (MRWA). 2005. Preserving, protecting and restoring the Muskegon River Watershed. Annual Report. Big Rapids, MI

Ohio Department of Natural Resources. 1996. Introduction to Soil Region of Ohio. (http://www.dnr.state.oh.us/tabid/9068/default.aspx). Division of Soil and Water Conservation. Columbus, Ohio.

Olden, J. D., and D. A. Jackson. 2001. Fish-habitat relationships in lakes: gaining predictive and explanatory insight by using artificial neural networks. Transactions of the American Fisheries Society 130: 878-897.

Olsen, A., R. J. Sedransk, D. Edwards, C. A. Gotway, W. Liggett, S. Rathbun, K. H. Reckhow, and L. J. Young. 1999. Statistical issues for monitoring ecological and natural resources in the United States. Environmental Monitoring and Assessment 54: 1-45.

Poff, N. L. 1997. Landscape filters and species traits: Towards mechanistic understanding and prediction in stream ecology. Journal of the North American Benthological Society 16: 391-409.

Rabeni, C. F., and S. P. Sowa. 1996. Integrating biological realism into habitat restoration and conservation strategies for small streams. Canadian Journal of Fisheries and Aquatic Sciences 53 (Supplement 1): 252-259.

Resh, V. H. 1979. Sampling variability and life history features: basic considerations in the design of aquatic insect studies. Journal of the Fisheries Research Board Canada 36: 290-311.

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

Riseng, C. M., M. J. Wiley, and R. J. Stevenson. 2004. Hydrologic disturbance and nutrient effects on benthic community structure in Midwestern U. S. streams: a covariance structure analysis. Journal of the North American Benthological Society 23: 309-326.

Riseng, C. M., M. J. Wiley, R. J. Stevenson, T. Zorn, and P. W. Seelbach. 2006. Comparison of coarse versus fine scale sampling on statistical modeling of landscape effects and assessment of fish assemblages of the Muskegon River, Michigan. pp 555-575. *In* R. Hughes, L. Wang, and P. W. Seelbach, editors, Landscape influences on stream habitats and biological communities. American Fisheries Society Symposium 48, Bethesda, Maryland.

Rosenberg, D. M., and V. H. Resh (eds.). 1993. Freshwater biomonitoring and benthic macroinvertebrates. Chapman and Hall, New York.

Roth, N. N. 1994. Land use, riparian vegetation, and stream ecosystem integrity in an agricultural watershed. Masters Thesis. University of Michigan, Ann Arbor.

Seelbach, P. W. and M. J. Wiley. 1996. An assessment of the potential for ecological rehabilitation and restoration in Mill Creek. Huron River Watershed Council (HRWC). Ann Arbor, Michigan.

Seelbach, P. W., M. J. Wiley, P. A. Soranno, and M. T. Bremigan. 2002. Aquatic conservation planning: using landscape maps to predict ecological reference conditions for specific waters. pp 454-478. *In* Gutzwiller, K. J., editor. Applying landscape ecology in biological conservation. Springer-Verlag, New York.

SPSS, Inc. 2003. SPSS for Windows v12.0. SPSS, Inc, Chicago, Illinois.

Turak, E., L. K. Flack, R. H. Norris, J. Simpson, and N. Waddell. 1999. Assessment of river condition at a large spatial scale using predictive models. Freshwater Biology 41:283-298.

U. S. Environmental Protection Agency (EPA). 1998. Environmental monitoring and assessment program. EMAP – surface waters: field operations and methods for measuring the ecological condition of wadeable streams. EPA/620/R-94/004F. Office of Research and Development, National Exposure Research Laboratory, Research Triangle Park, North Carolina.

U. S. Environmental Protection Agency (EPA). 2006. The evaluation of methods for creating defensible, repeatable, objective and accurate tolerance values for aquatic taxa. EPA/600/R-06/045. Office of Research and Development, National Exposure Research laboratory, Cincinnati, OH.

Velleman, P. F., and A. Y. Velleman. 1988. Datadesk: statistics and reference guide. Odesta Corporation, Northbrook, Illinois.

Vinson, M. R., and C. P. Hawkins. 1996. Effects of sampling area and subsampling procedure on comparisons of taxa richness among streams. Journal of North American Benthological Society 15: 392-399.

Wang , L., J. Lyons, P. Kanehl, and R. Bannerman. 2001. Impacts of urbanization on stream habitat and fish across multiple spatial scales. Environmental Management 28: 255-266.

Wang, L., J. Lyons, P. Rasmussen, P. Seelbach, T. Simon, M. Wiley, P. Kanehl, E. Baker, S. Niemela, and P. Stewart. 2003. Watershed, reach, and riparian influences on stream fish assemblages in the Northern Lakes and Forest ecoregion, USA. Canadian Journal of Fisheries and Aquatic Sciences 60: 491-505.

Wells, S. 2001. Crane Creek Progress Report 1999. United States Fish and Wildlife Service. Oak Harbor, Ohio.

Wiley, M. J., P. W. Seelbach, K. Wehrly, and J. S. Martin. 2002. Regional ecological normalization using linear models: a meta-method for scaling stream assessment indicators. pp 201-223. *In* T.P. Simon, editor. Biological response signatures: indicator patterns using aquatic communities. CRC Press, Boca Raton, Florida.

Yoon, Il Byong, Y. J. Bae, T. H. Ro, S. J. Lee, and J. H. Park. 1998. Determination on the optimal sample size in the aquatic insect community analysis – Pangtae Creek model. The Korean Journal of Ecology 21 (5-1): 409-418.

Table 1. Summary of landscape and stream characteristic variables used for normalizing linear regression models.

	Total sites (n= 52)		Bigelow Creek (n= 5)		Brooks Creek (n= 14)		Cedar Creek (n= 9)		Crane Creek (n= 13)		Mill Creek (n= 11)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Drainage area (km^2)	70.33	87.80	37.41	21.41	48.17	50.44	58.07	41.81	40.94	36.98	158.25	147.62
Stream Slope	0.0016	0.0017	0.0026	0.0027	0.0019	0.0014	0.0010	0.0009	0.0008	0.0009	0.0021	0.0022
Width (m)	5.65	4.25	4.97	2.96	4.22	2.83	7.56	4.74	5.21	5.04	6.73	4.65
Estimated Temperature (2005)	13.19	4.22	10.52	1.52	12.36	2.52	14.58	2.14	17.67	4.10	9.01	2.50
Predicted July mean temperature (ºC)	19.09	3.07	17.48	0.57	16.52	0.72	16.18	0.93	22.80	2.16	21.11	0.59
Percentage agricultural land >0.25 within watershed	44.71	0.31	0.00	0.00	50.88	5.19	15.16	23.4	84.1	4.18	34.85	24.31
Percentage agricultural land within watershed	48.46	26.34	8.13	3.71	50.88	5.19	25.21	16.3	84.1	4.18	40.67	15.63
Percentage agricultural land within 100m buffer	39.07	29.04	2.71	0.57	30.70	6.62	19.37	13.3	84.0	6.85	29.32	14.53
Percentage urban land within watershed	4.55	2.34	1.27	0.13	4.52	2.57	3.26	0.97	6.28	1.77	5.11	1.96
Percentage urban land within 100m buffer	4.12	2.89	0.59	0.04	4.13	2.98	2.40	1.06	7.10	2.55	3.61	1.32
Percentage non-forest wetland within 100m buffer	4.58	4.33	2.68	0.26	4.42	2.21	8.06	6.46	0.55	0.17	7.57	3.51
Flow (m^3/s)	0.37	0.49	0.40	0.35	0.29	0.36	0.53	0.47	0.07	0.06	0.68	0.74
Q90Y/Q10Y (Low and high flow yield ratio)	0.18	0.14	0.44	0.09	0.11	0.03	0.35	0.14	0.08	0.00	0.10	0.03
Q75 (m^3/s)	0.35	0.37	0.33	0.21	0.21	0.23	0.48	0.41	0.24	0.21	0.59	0.53
Q90 (m^3/s , low flow)	0.24	0.25	0.28	0.19	0.13	0.15	0.39	0.35	0.16	0.15	0.35	0.32

Table 2. Means and medians of macroinvertebrate taxa for each indicator metric with each level of sampling effort and taxonomic resolution (n=52). LES indicates low-effort sampling and HES indicates high-effort sampling.

Dependent variable	Sampling effort	Taxonomic resolution	Mean	Median	SD	Min	Max
# Taxa	LES	Family	14.13	14.50	3.70	5.00	21.00
		Genus	15.54	16.00	4.41	5.00	24.00
	HES	Family	21.73	22.00	5.47	8.00	32.00
		Genus	25.35	26.00	7.22	9.00	39.00
# EPT	LES	Family	4.50	4.00	2.89	0.00	10.00
		Genus	5.40	4.50	3.67	0.00	13.00
	HES	Family	6.56	6.50	3.76	0.00	14.00
		Genus	8.58	9.00	5.01	0.00	18.00
MBI	LES	Family	5.91	6.00	1.19	3.28	8.54
		Genus	5.83	5.90	1.24	3.39	8.42
	HES	Family	5.82	5.67	1.09	3.64	8.64
		Genus	5.70	5.64	1.18	3.27	8.93
# Sensitive	LES	Family	2.40	2.00	2.30	0.00	9.00
		Genus	3.62	3.00	3.16	0.00	10.00
	HES	Family	3.96	4.00	3.24	0.00	13.00
		Genus	6.15	5.00	5.21	0.00	19.00
# Surface dependent	LES	Family	2.12	2.00	1.32	0.00	5.00
		Genus	4.54	4.00	3.12	0.00	14.00
	HES	Family	4.54	4.00	2.10	1.00	11.00
		Genus	9.62	9.00	4.76	1.00	21.00
# Surface Dependent /# Taxa	LES	Family	0.15	0.13	0.10	0.00	0.44
		Genus	0.29	0.28	0.19	0.00	0.82
	HES	Family	0.21	0.20	0.08	0.05	0.39
		Genus	0.38	0.37	0.15	0.08	0.70
# Metabolic conformers	LES	Family	3.63	3.00	2.77	0.00	9.00
		Genus	4.17	3.00	3.22	0.00	11.00
	HES	Family	5.58	5.00	3.66	0.00	15.00
		Genus	6.85	7.00	4.50	0.00	16.00

Table 3. Independent T-tests of macroinvertebrate indicator metrics between LES and HES at each taxonomic resolution and between levels of taxonomic resolution at each sampling effort.

	t-test for Equality of Means				
	t	df	Sig (2-tailed)	Mean Differece	Std. Error difference
LES vs. HES in family level					
# Taxa	-8.294	102	**0.000**	-7.596	0.916
# EPT	-3.128	102	**0.002**	-2.058	0.658
Macroinvertebrate Biotic Index (MBI)	0.374	102	0.709	0.084	0.223
# Sensitive	-2.828	102	**0.006**	-1.558	0.551
# Surface dependent	-7.039	102	**0.000**	-2.423	0.344
# Surface dependent / # Taxa	-3.134	102	**0.002**	-0.056	0.018
# Metabolic conformers	-3.048	102	**0.003**	-1.942	0.637
LES vs. HES in genus level					
# Taxa	-8.357	102	**0.000**	-9.808	1.174
# EPT	-3.686	102	**0.000**	-3.173	0.861
Macroinvertebrate Biotic Index (MBI)	0.568	102	0.571	0.134	0.236
# Sensitive	-3.001	102	**0.003**	-2.538	0.846
# Surface dependent	-6.430	102	**0.000**	-5.077	0.790
# Surface Dependent /# Taxa	-2.437	102	**0.017**	-0.082	0.033
# Metabolic conformers	-3.485	102	**0.001**	-2.673	0.767
Family vs. genus level in LES					
# Taxa	-1.759	102	0.082	-1.404	0.798
# EPT	-1.395	102	0.166	-0.904	0.648
Macroinvertebrate Biotic Index (MBI)	0.331	102	0.742	0.079	0.238
# Sensitive	-2.233	102	**0.028**	-1.212	0.543
# Surface dependent	-5.154	102	**0.000**	-2.423	0.470
# Surface dependent / # Taxa	-4.665	102	**0.000**	-0.141	0.030
# Metabolic conformers	-0.914	102	0.363	-0.538	0.589
Family vs. genus level in HES					
# Taxa	-2.877	102	**0.005**	-3.615	1.257
# EPT	-2.325	102	**0.022**	-2.019	0.868
Macroinvertebrate Biotic Index (MBI)	0.583	102	0.561	0.129	0.222
# Sensitive	-2.576	102	**0.011**	-2.192	0.851
# Surface dependent	-7.035	102	**0.000**	-5.077	0.722
# Surface Dependent /# Taxa	-7.180	102	**0.000**	-0.166	0.023
# Metabolic conformers	-1.577	102	0.118	-1.269	0.805

Table 4. Pearson correlations among macroinvertebrate indicator metrics between degrees of sampling effort at each taxonomic resolution and between levels of taxonomic resolution at each sampling effort. One star indicates significance at $p \leq 0.05$, and two stars indicate indicates significance at $p \leq 0.01$.

	Family level, HES						
	# Taxa	# EPT	MBI	# Sens	# Sudep	# Su/Ta	# Metc
Family level, LES							
# Taxa	**0.475****	0.481**	-0.511**	0.510**	0.071	-0.166	0.496**
# EPT	0.604**	**0.814****	-0.775**	0.735**	-0.023	-0.404**	0.823**
MBI	-0.475**	-0.805**	**0.857****	-0.767**	0.068	0.400**	-0.826**
# Sensitive (# Sens)	0.409**	0.671**	-0.707**	**0.768****	-0.062	-0.323**	0.729**
# Surface dependent (# Sudep)	-0.020	-0.261	0.190	-0.150	**0.295***	0.390**	-0.273
# Surface dependent / # Taxa (#Su/Ta)	-0.157	-0.406**	0.338*	-0.322*	0.284*	**0.461****	-0.420**
# Metabolic conformers (#Metc)	0.562**	0.804**	-0.782**	0.766**	-0.046	-0.405**	**0.847****
	Genus level, HES						
	# Taxa	# EPT	MBI	# Sens	# Sudep	# Su/Ta	# Metc
Genus level, LES							
# Taxa	**0.573****	0.584**	-0.529**	0.471**	0.187	-0.120	0.559**
# EPT	0.671**	**0.826****	-0.775**	0.728**	0.179	-0.251	0.835**
MBI	-0.550**	-0.813**	**0.875****	-0.805**	-0.170	0.203	-0.839**
# Sensitive (# Sens)	0.503**	0.734**	-0.796**	**0.761****	0.134	-0.207	0.756**
# Surface dependent (# Sudep)	0.133	-0.039	0.080	-0.062	**0.300***	0.258	-0.136
# Surface dependent / # Taxa (#Su/Ta)	-0.072	-0.260	0.249	-0.224	0.247	**0.345***	-0.344*
# Metabolic conformers (#Metc)	0.616**	0.811**	-0.781**	0.735**	0.156	-0.238	**0.854****
	Genus level, LES						
	# Taxa	# EPT	MBI	# Sens	# Sudep	# Su/Ta	# Metc
Family level, LES							
# Taxa	**0.959****	0.636**	-0.305*	0.512**	0.243	-0.113	0.585**
# EPT	0.696**	**0.980****	-0.804**	0.797**	-0.169	-0.428**	0.962**
MBI	-0.473**	-0.831**	**0.964****	-0.844**	0.108	0.269	-0.843**
# Sensitive (# Sens)	0.697**	0.809**	-0.718**	**0.872****	0.065	-0.205	0.819**
# Surface dependent (# Sudep)	0.218	-0.224	0.306*	-0.139	**0.820****	0.806**	-0.235
# Surface dependent / # Taxa (#Su/Ta)	-0.139	-0.431**	0.376**	-0.299*	0.712**	**0.880****	-0.417**
# Metabolic conformers (#Metc)	0.676**	0.961**	-0.808**	0.811**	-0.163	-0.409**	**0.979****
	Genus level, HES						
	# Taxa	# EPT	MBI	# Sens	# Sudep	# Su/Ta	# Metc
Family level, HES							
# Taxa	**0.972****	0.747**	-0.546**	0.486**	0.622**	0.085	0.678**
# EPT	0.801**	**0.978****	-0.844**	0.803**	0.321*	-0.180	0.960**
MBI	-0.658**	-0.864**	**0.971****	-0.847**	-0.290*	0.080	-0.883**
# Sensitive (# Sens)	0.678**	0.835**	-0.864**	**0.915****	0.287*	-0.133	0.867**
# Surface dependent (# Sudep)	0.464**	0.065	-0.006	0.016	**0.839****	0.720**	0.009
# Surface dependent / # Taxa (#Su/Ta)	-0.098	-0.419**	0.352*	-0.300*	0.551**	**0.827****	-0.439**
# Metabolic conformers (#Metc)	0.751**	0.951**	-0.864**	0.842**	0.302*	-0.165	**0.977****

Table 5. Means and medians of compositional overlap (%) of each taxonomic class with LES and HES at family and genus levels (n indicates total number of sites collected).

	LES					HES				
	Mean	Median	SD	Min	Max	Mean	Median	SD	Min	Max
Family level										
Total (n=52)	81.22	82.58	11.94	50.00	100.00	55.31	54.07	13.21	27.27	92.31
Ephemeroptera (n=47)	91.77	100.00	18.79	25.00	100.00	71.38	66.67	24.78	25.00	100.00
Plecoptera (n=29)	93.10	100.00	17.55	50.00	100.00	66.67	50.00	27.82	33.33	100.00
Trichoptera (n=44)	91.78	100.00	16.37	50.00	100.00	63.60	66.67	27.26	16.67	100.00
Odonata (n=47)	82.27	100.00	24.67	33.33	100.00	67.45	50.00	29.10	16.67	100.00
Coleoptera (n=50)	88.67	100.00	20.60	50.00	100.00	62.40	50.00	29.17	20.00	100.00
Heteroptera (n=49)	93.88	100.00	16.56	50.00	100.00	58.43	50.00	29.88	16.67	100.00
Diptera (n=52)	92.72	100.00	16.79	50.00	100.00	55.06	50.00	23.75	20.00	100.00
Others (n=52)	76.84	76.39	21.97	22.22	100.00	60.78	60.00	23.18	16.67	100.00
Genus level										
Total (n=52)	75.45	75.00	13.14	50.00	100.00	48.28	46.95	12.80	21.74	92.31
Ephemeroptera (n=47)	92.16	100.00	18.88	25.00	100.00	67.29	66.67	23.82	20.00	100.00
Plecoptera (n=29)	93.10	100.00	17.55	50.00	100.00	63.39	50.00	26.77	25.00	100.00
Trichoptera (n=44)	82.49	100.00	21.84	25.00	100.00	50.11	50.00	22.83	11.11	100.00
Odonata (n=47)	75.53	100.00	27.11	33.33	100.00	57.08	50.00	28.60	14.29	100.00
Coleoptera (n=50)	77.13	100.00	27.75	33.33	100.00	49.34	40.00	29.10	14.29	100.00
Heteroptera (n=49)	86.73	100.00	22.56	33.33	100.00	50.82	50.00	28.14	16.67	100.00
Diptera (n=52)	86.54	100.00	19.88	50.00	100.00	49.15	50.00	25.27	16.67	100.00
Others (n=52)	76.31	75.00	22.04	22.22	100.00	60.33	60.00	23.11	100.00	100.00

Table 6. Independent variables and coefficients used for best-fit normalizing linear regression models for each indicator metric. One star indicates significance at $p \leq 0.05$, and two stars indicate indicates significance at $p \leq 0.01$.

Dependent variable	Sampling effort	Taxonomic resolution	R^2 (%)	SE	Independent variables						
					Constant	Drainage area	Stream Slope	Width	Field Measured temperature (2005)	Predicted July mean temperature	Proportion of Ag. >0.25 within watershed
# Taxa	LES	Family	30.2	0.2375	2.4170**	0.0761*					
		Genus	42.0	0.2369	2.4078**	0.0980*					
	HES	Family	46.1	0.2077	2.4980**	0.2166**					
		Genus	46.3	0.2371	2.6592**	0.2186**					
# EPT	LES	Family	67.6	0.3964	4.2295**	0.4371**				-1.1422**	-1.7775**
		Genus	72.0	0.4055	4.3617**	0.5076**				-1.1957**	-2.0193**
	HES	Family	77.4	0.3701	6.0241**	0.6405**				-1.8462**	-1.7551**
		Genus	79.4	0.3947	5.1958**	0.5294**				-1.3583**	
MBI	LES	Family	64.6	0.1150	1.1276**	-0.0809*	-43.5615**	-0.1228*		0.3384**	0.5043**
		Genus	64.5	0.1212	1.2282**	-0.1208*	-34.1674**	-0.1632**		0.3559*	
	HES	Family	70.4	0.0919	1.4486**	-0.1216**	-23.7438**			0.2365*	0.2669**
		Genus	72.8	0.1004	1.3182**	-0.0967**	-36.7059**	-0.1334*		0.3077*	0.2693**
# Sensitive	LES	Family	46.0	0.5298	4.0389**				-1.0034**		-1.2455**
		Genus	50.2	0.5959	4.8808**	0.2600**				-1.352*	-1.623**
	HES	Family	65.4	0.4635	7.0065**	0.1616*				-1.5443**	-3.5714**
		Genus	66.4	0.5286	7.9785**	0.2844**				-1.9153**	-3.2185**
# Metabolic conformers	LES	Family	69.0	0.4172	4.2744**	0.3869**	96.3228*			-1.2125**	-1.8240**
		Genus	71.6	0.4230	4.8879**	0.4154**	83.6071*			-1.4018**	-1.9576**
	HES	Family	80.7	0.3553	5.9866**	0.6242**	91.5142**			-1.9128**	-1.7797**
		Genus	78.9	0.4118	6.6205**	0.6732**	80.0852*			-2.1022**	-1.9865**

Table 6. Cont.

Dependent variable	Sampling effort	Taxonomic resolution	Independent variables								
			Proportion of Ag. Within watershed	Proportion of Ag. Within 100m buffer	Proportion of Urban within watershed	Proportion of Urban within 100m buffer	Porportion of Nonforest wetland within 100m buffer	Discharge	Ratio of streamflow yield (90%/10%)	Streamflow 75 %	Streamflow 90 %
# Taxa	LES	Family				-2.7264*	2.0960*				
		Genus				-3.2750*	2.8268**				
	HES	Family				-4.3121**	2.0509*	-0.4876*			
		Genus				-5.4691**	2.1772*	-0.4501*			
# EPT	LES	Family									-1.5204*
		Genus									-1.6926*
	HES	Family									-2.2969**
		Genus		-2.5036**							-1.5383**
MBI	LES	Family					1.3703*				0.6292**
		Genus	0.2934*		2.2597*					0.6786**	
	HES	Family				1.4670*					0.4621**
		Genus				1.8363*					0.6471**
# Sensitive	LES	Family									
		Genus									
	HES	Family									
		Genus							-3.5714**		
# Metabolic conformers	LES	Family									-1.4690*
		Genus									-1.4666*
	HES	Family									-2.2993**
		Genus									-2.4073**

Table 7. Regression model statistics of each indicator metric for a 2x2 factorial combination of sampling effort and taxonomic resolution.

Dependent variable	Sampling effort	Taxonomic resolution	R^2 (%)	SE	df	F-ratio	# variables
# Taxa	LES	Family	30.2	0.2375	48	6.92	3
		Genus	42.0	0.2369	48	11.6	3
	HES	Family	46.1	0.2077	47	10.1	4
		Genus	46.3	0.2371	47	10.1	4
# EPT	LES	Family	67.6	0.3964	47	24.6	4
		Genus	72.0	0.4055	47	30.3	4
	HES	Family	77.4	0.3701	47	40.1	4
		Genus	79.4	0.3947	47	45.3	4
MBI	LES	Family	64.6	0.1150	44	11.5	7
		Genus	64.5	0.1212	44	11.4	7
	HES	Family	70.4	0.0919	45	17.9	6
		Genus	72.8	0.1004	44	16.8	7
# Sensitive	LES	Family	46.0	0.5298	49	20.8	2
		Genus	50.2	0.5959	48	16.1	3
	HES	Family	65.4	0.4635	47	22.2	4
		Genus	66.4	0.5286	47	23.2	4
# Metabolic conformers	LES	Family	69.0	0.4172	46	20.5	5
		Genus	71.6	0.4230	46	23.2	5
	HES	Family	80.7	0.3553	46	38.5	5
		Genus	78.9	0.4118	46	34.3	5

Table 8. Means and medians of normalized scores of each indicator metric for a 2x2 factorial combination of sampling effort and taxonomic resolution (n=52).

Dependent variable	Sampling effort	Taxonomic resolution	Mean	Median	SD	Min	Max
Composite scores	LES	Family	-1.3416	-1.4928	1.2001	-4.7534	1.1087
		Genus	-1.4779	-1.6224	1.1578	-4.9025	1.1418
	HES	Family	-1.5320	-1.4265	1.2612	-5.2163	1.6173
		Genus	-1.6343	-1.5955	1.3041	-5.1286	1.2520
# Taxa	LES	Family	-0.4735	-0.1610	1.0514	-3.5215	1.0075
		Genus	-0.5702	-0.5229	1.0737	-3.7266	1.3049
	HES	Family	-0.8651	-0.9288	1.1625	-4.0128	1.6570
		Genus	-0.9620	-0.9370	1.1985	-3.9022	1.7295
# EPT	LES	Family	-1.6818	-1.6701	1.5392	-6.1423	1.0753
		Genus	-1.7855	-1.9218	1.5612	-6.1947	1.1394
	HES	Family	-1.7032	-1.4982	1.5202	-6.5538	1.1108
		Genus	-2.0439	-1.8316	1.6631	-7.0725	0.6143
MBI	LES	Family	-1.6268	-1.8564	1.4815	-4.7396	1.8116
		Genus	-1.8779	-1.9659	1.2827	-4.5700	1.3842
	HES	Family	-1.7429	-1.6826	1.4569	-4.6326	2.8630
		Genus	-1.7808	-1.7533	1.4714	-4.4455	2.8327
# Sensitive	LES	Family	-1.1291	-0.9788	1.2547	-5.2310	1.5407
		Genus	-0.9667	-1.0201	1.1924	-3.4670	1.3605
	HES	Family	-2.5342	-2.7110	1.9739	-8.1283	1.2287
		Genus	-2.1857	-2.3141	1.7768	-5.5911	1.7262
# Metabolic conformers	LES	Family	-1.5846	-1.6210	1.4601	-5.2418	1.0606
		Genus	-1.6779	-1.6964	1.5067	-5.5479	1.1405
	HES	Family	-1.8168	-1.6229	1.5779	-6.0788	1.6385
		Genus	-1.7507	-1.4350	1.5442	-5.7839	1.0239

Table 9. Independent T-tests of normalized scores for each indicator metric between LES and HES at each taxonomic resolution and between family and genus levels at each sampling effort.

	t-test for Equality of Means				
	t	df	Sig (2-tailed)	Mean Differece	Std. Error difference
LES vs. HES in family level					
Composite ((# T + # E + # M + MBI)/4)	0.788	102	0.432	0.190	0.241
# Taxa (# T)	1.802	102	0.075	0.392	0.217
# EPT (# E)	0.701	102	0.943	0.021	0.300
Macroinvertebrate Biotic Index (MBI)	0.403	102	0.688	0.116	0.288
# Sensitive (# S)	5.005	102	**0.000**	1.597	0.319
# Metabolic conformers (# M)	0.779	102	0.438	0.232	0.298
LES vs. HES in genus level					
Composite ((# T + # E + # M + MBI)/4)	0.647	102	0.519	0.157	0.242
# Taxa (# T)	1.756	102	0.082	0.392	0.223
# EPT (# E)	0.817	102	0.416	0.258	0.316
Macroinvertebrate Biotic Index (MBI)	-0.359	102	0.720	-0.097	0.271
# Sensitive (# S)	4.108	102	**0.000**	1.219	0.297
# Metabolic conformers (#M)	0.243	102	0.808	0.073	0.299
Family vs. Genus level in LES					
Composite ((# T + # E + # S + MBI)/4)	0.589	102	0.557	0.136	0.231
# Taxa (# T)	0.464	102	0.643	0.097	0.208
# EPT (# E)	0.341	102	0.743	0.104	0.304
Macroinvertebrate Biotic Index (MBI)	0.924	102	0.358	0.251	0.272
# Sensitive (# S)	0.367	102	0.714	0.087	0.236
# Metabolic conformers	0.321	102	0.749	0.093	0.291
Family vs. Genus level in HES					
Composite ((# T + # E + # S + MBI)/4)	0.407	102	0.685	0.102	0.252
# Taxa (# T)	0.418	102	0.676	0.097	0.232
# EPT (# E)	1.090	102	0.278	0.341	0.313
Macroinvertebrate Biotic Index (MBI)	0.132	102	0.895	0.038	0.287
# Sensitive (# S)	-0.795	102	0.429	-0.291	0.366
# Metabolic conformers	-0.216	102	0.829	-0.066	0.306

Table 10. Pearson correlations of normalized scores among macroinvertebrate indicator metrics between degrees of sampling effort at each taxonomic resolution and between taxonomic resolution at each level of effort. One star indicates significance at $p \leq 0.05$, and two stars indicate indicates significance at $p \leq 0.01$.

	Family level, HES					
	Com	# Taxa	# EPT	MBI	# S	# M
Family level, LES						
Composite (Com)	**0.787****	0.465**	0.739**	0.744**	0.726**	0.773**
# Taxa (# T)	0.329*	**0.372****	0.215	0.372**	0.403*	0.228
# EPT (# E)	0.758**	0.433**	**0.777****	0.655**	0.670**	0.752**
Macroinvertebrate Biotic Index (MBI)	0.765**	0.357**	0.719**	**0.779****	0.714**	0.771**
# Sensitive (# S)	0.475**	0.263	0.365**	0.602**	**0.621****	0.419**
# Metabolic conformers (# M)	0.773**	0.442**	0.726**	0.698**	0.668**	**0.803****
	Genus level, HES					
	Com	# Taxa	# EPT	MBI	# S	# M
Genus level, LES						
Composite (Com)	**0.798****	0.526**	0.733**	0.768**	0.765**	0.766**
# Taxa (# T)	0.374**	**0.417****	0.271	0.410**	0.381**	0.256
# EPT (# E)	0.789**	0.488**	**0.791****	0.683**	0.748**	0.783**
Macroinvertebrate Biotic Index (MBI)	0.699**	0.395**	0.602**	**0.797****	0.696**	0.646**
# Sensitive (# S)	0.673**	0.429**	0.635**	0.657**	**0.722****	0.628**
# Metabolic conformers (# M)	0.774**	0.478**	0.730**	0.682**	0.713**	**0.809****
	Genus level, LES					
	Com	# T	# E	MBI	# S	# M
Family level, LES						
Composite (Com)	**0.977****	0.640**	0.949**	0.739**	0.728**	0.936**
# Taxa (# T)	0.587**	**0.967****	0.480**	0.213	0.390**	0.436**
# EPT (# E)	0.929**	0.516**	**0.982****	0.652**	0.674**	0.914**
Macroinvertebrate Biotic Index (MBI)	0.869**	0.358**	0.806**	**0.906****	0.732**	0.809**
# Sensitive (# S)	0.721**	0.681**	0.605**	0.581**	**0.748****	0.610**
# Metabolic conformers (# M)	0.930**	0.500**	0.920**	0.667**	0.660**	**0.980****
	Genus level, HES					
	Com	# Taxa	# EPT	MBI	# S	# M
Family level, HES						
Composite (Com)	**0.976****	0.763**	0.929**	0.808**	0.830**	0.934**
# Taxa (# T)	0.720**	**0.974****	0.658**	0.426**	0.485**	0.560**
# EPT (# E)	0.922**	0.622**	**0.964****	0.695**	0.791**	0.931**
Macroinvertebrate Biotic Index (MBI)	0.846**	0.552**	0.710**	**0.939****	0.814**	0.770**
# Sensitive (# S)	0.831**	0.614**	0.752**	0.803**	**0.897****	0.756**
# Metabolic conformers (# M)	0.921**	0.613**	0.902**	0.731**	0.782**	**0.967****

Table 11. Impairment classification based on normalized composite scores from four combinations of sampling effort and taxonomic resolution.

	Genus level, LES					
	Exceptional	Good	Threatened	Poor	Very poor	Total
Family level, LES						
Exceptional	**1**	2	0	0	0	3
Good	0	**8**	4	0	0	12
Threatened	0	0	**4**	0	0	4
Poor	0	0	2	**13**	3	18
Very poor	0	0	0	0	**15**	15
Total	1	10	10	13	18	**52**
	Genus level, HES					
	Exceptional	Good	Threatened	Poor	Very poor	Total
Family level, HES						
Exceptional	**2**	0	0	0	0	2
Good	0	**6**	2	0	0	8
Threatened	0	1	**4**	1	0	6
Poor	0	0	1	**16**	1	18
Very poor	0	0	0	2	**16**	18
Total	2	7	7	19	17	**52**
	HES, family level					
	Exceptional	Good	Threatened	Poor	Very poor	Total
LES, family level						
Exceptional	**0**	3	0	0	0	3
Good	2	**4**	2	**4**	0	12
Threatened	0	0	**2**	2	0	4
Poor	0	**1**	2	**9**	6	18
Very poor	0	0	0	3	**12**	15
Total	2	8	6	18	18	**52**
	HES, genus level					
	Exceptional	Good	Threatened	Poor	Very poor	Total
LES, genus level						
Exceptional	**0**	1	0	0	0	1
Good	2	**4**	1	**3**	0	10
Threatened	0	1	**4**	5	0	10
Poor	0	**1**	2	**8**	2	13
Very poor	0	0	0	3	**15**	18
Total	2	7	7	19	17	**52**

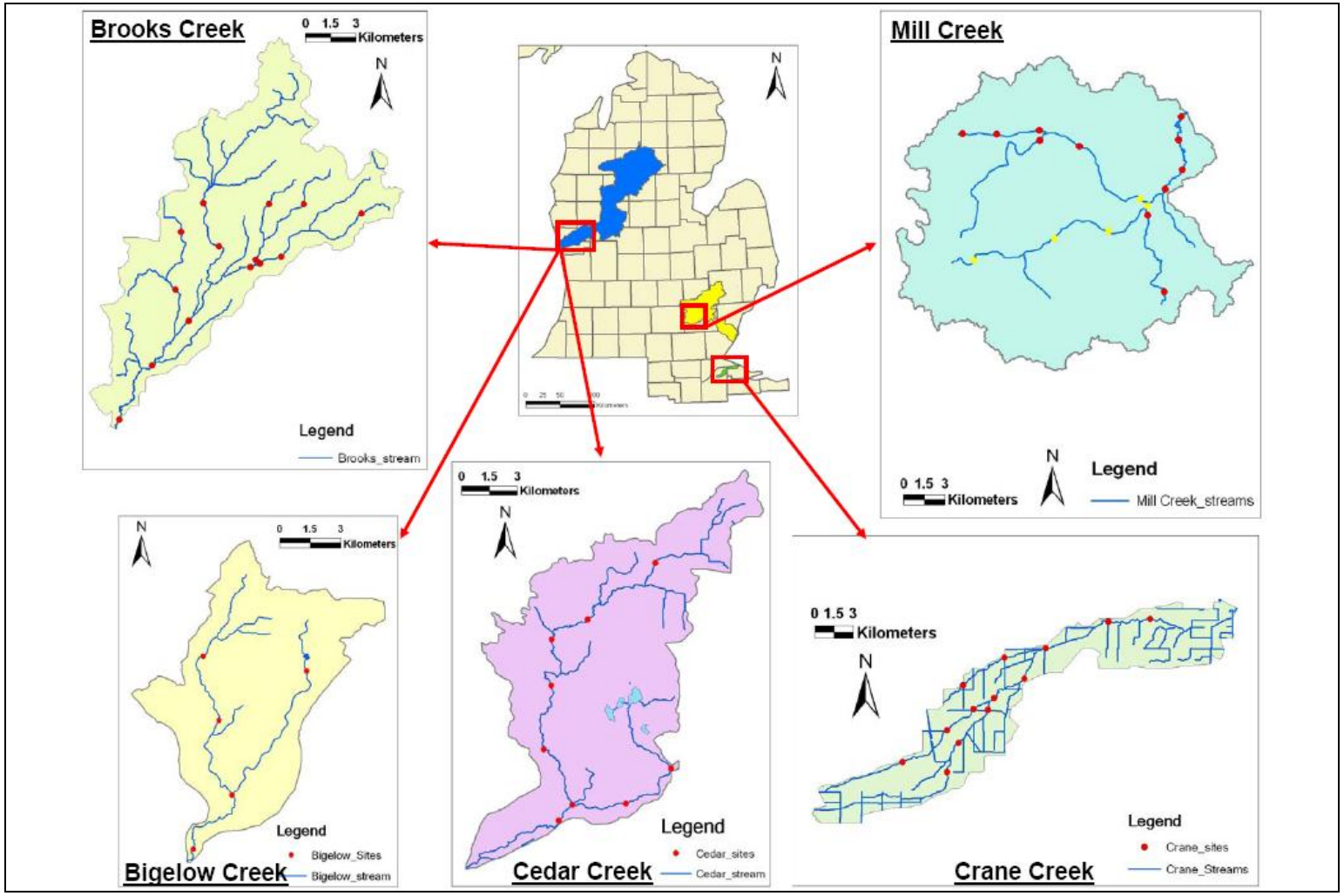

Figure 1. Map showing sampling site locations on Bigelow Creek, Brooks Creek, Cedar Creek, Crane Creek, and Mill Creek.

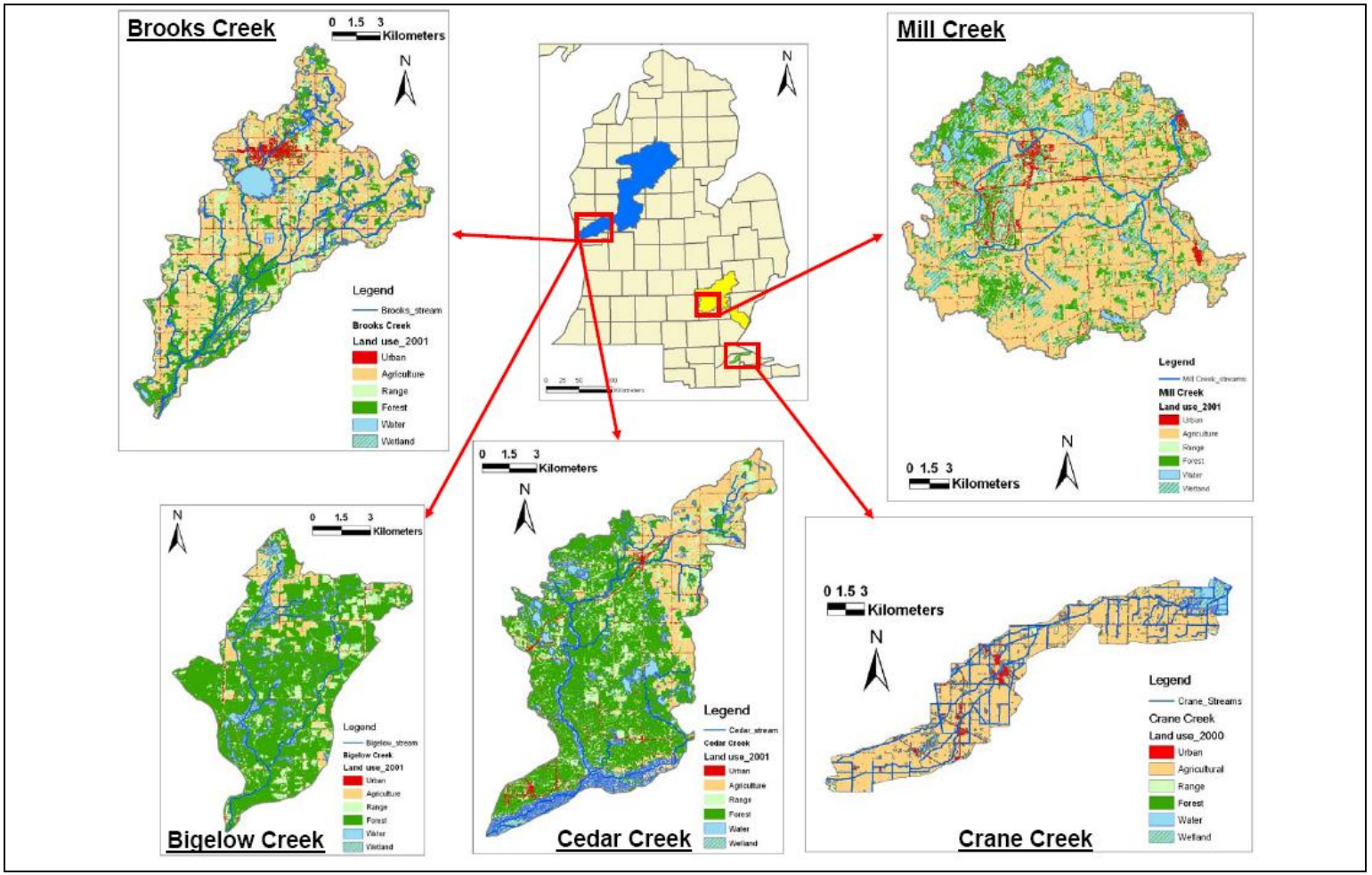

Figure 2. Land uses of each watershed used for this study.

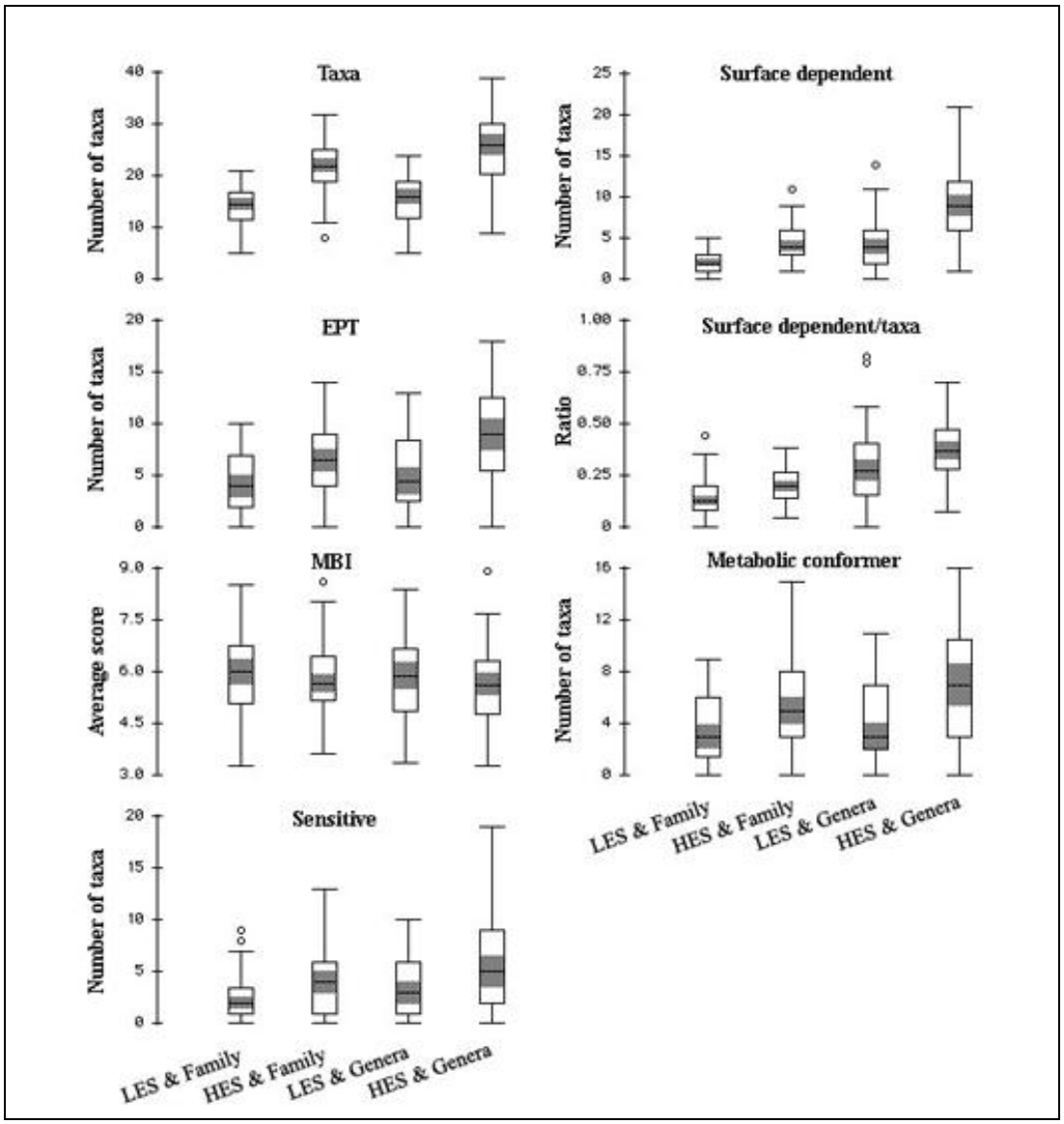

Figure 3. Boxplots show number of taxa for each macroinvertebrate indicator metric collected by different degrees of sampling effort and taxonomic resolution (n=52). Indicator metrics of MBI shows the average MBI scores and # Surface dependent/# Taxa shows relative ratios of number of surface dependent to number of total taxa.

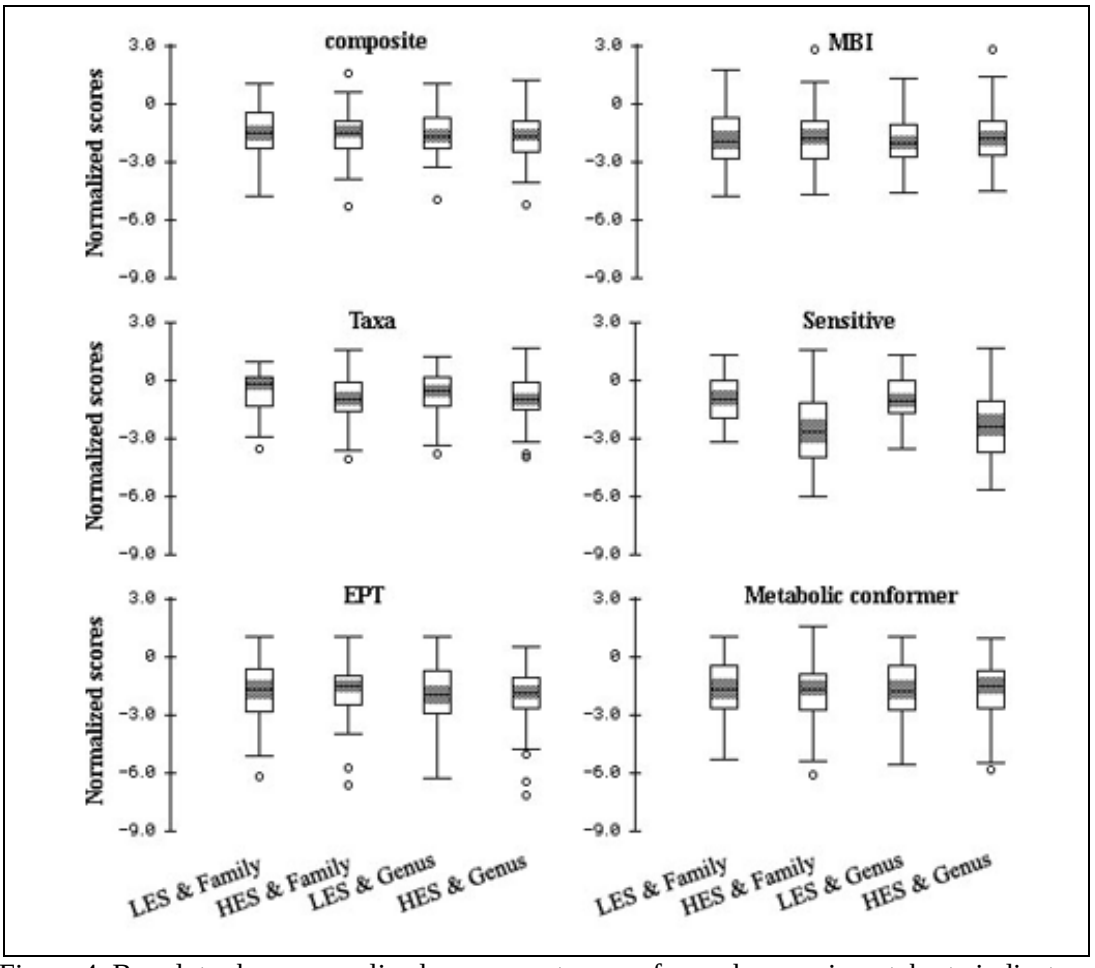

Figure 4. Boxplots show normalized assessment scores for each macroinvertebrate indicator metric collected by different degrees of sampling effort and taxonomic resolution (n=52).