IDENTIFYING LINEAR RELATIONSHIPS BETWEEN STREAMFLOW METRICS AND BENTHIC MACROINVERTEBRATE METRICS IN MINNESOTA

A THESIS
SUBMITTED TO THE FACULTY OF THE
UNIVERSITY OF MINNESOTA
BY

KARA FITZPATRICK

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF SCIENCE

ADVISOR
DR. JOE MAGNER

AUGUST 2018

Acknowledgements

This project was made possible thanks to the data collection and dissemination efforts of many people and groups, including the Minnesota Pollution Control Agency, the U.S. Geological Survey, and Minnesota Department of Natural Resources. Thank you Jeffrey Ziegeweid and Aliesha Krall of the USGS for your invaluable assistance with project design and analysis. I am especially grateful to my co-workers, Ben Lundeen and Kevin Stroom for their consult on macroinvertebrate biology, Mark Evenson for his consult on hydrologic modeling, Greg Johnson for guidance on project direction, and especially Chuck Johnson for his daily guidance and encouragement on many topics.

I express my gratitude to my peers, who served as role models and helped me survive and finish the wonderful graduate school experience: Emily Deering, Lu Zhang, Leah Hall, Masha Guzner, Melissa Fitzpatrick, Petra Kranzfelder, Julia Beni, and Julie Johnston. I am ever grateful to my graduate advising committee, Dr. Joe Magner, Dr. Bruce Wilson, and Dr. Raymond Newman, for their assistance with this project, as well as providing me with exceptional instruction and guidance through rigorous coursework that built my knowledge, skills, and abilities. These researchers have also inspired my love for the art that is environmental science and teaching.

Ultimately, this project was made possible by, and is dedicated to, the citizens of Minnesota, who see the everlasting value in preserving their world-class freshwaters.

Dedication

The completion of my graduate program is dedicated to my best friend and sister, Melissa Fitzpatrick. Thank you for always having my back. Let's go see.

ABSTRACT

Understanding how the flow of a stream affects the aquatic organisms that depend on it will help users and managers of freshwater protect those organisms. Developing flow-biology relationships is a complicated task, wrought with problems of scale and natural variability. The purpose of this project was to determine if linear relationships could be identified between flow metrics and benthic macroinvertebrate metrics within a watershed at two different analytical scales, using both modeled and observed flow data. By using stepwise linear regression, a variable selection method, to select the most predictive independent variables from a large pool of flow metrics, inferences can be drawn about which components of streamflow have the greatest effect on the benthic macroinvertebrate communities in those streams. Two analytical scales were studied, referred to as "large" and "headwater" watershed scales. The large scale analysis used a more generalized biological variable, the index of biotic integrity, while the headwater scale analysis used more specific community measures, such as richness of clinger taxa. Additionally, the large watershed scale used observed flow measurements, while the headwater analysis used modeled streamflow. Significant ($p<0.10$) linear regression equations were developed using flow metrics to predict biological variables at both the large and headwater scales, with the models' adjusted coefficients of determination ($\hat{R}^2$) ranging from zero to 0.93. The results of this project represent one small step toward the goal of defining flow-biology relationships in Minnesota that will help meet the Clean Water Act demands to "restore and maintain the chemical, physical, and biological integrity of the Nation's waters."

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AIC	Akaike Information Criterion
CWA	Clean Water Act
ELOHA	Ecological Limits of Hydrologic Alteration
EPT	Ephemeroptera, Plecoptera, Trichoptera
EPTr	Ephemeroptera, Plecoptera, Trichoptera (taxa) richness
HLU	Hydrologic Landscape Unit
HSPF	Hydrologic Simulation Program Fortran
HUC	Hydrologic Unit Code
IBI	Index of Biotic Integrity
LQ	Lac Qui Parle
MDNR	Minnesota Department of Natural Resources
MIBI	Macroinvertebrate Index of Biotic Integrity
MPCA	Minnesota Pollution Control Agency
O/E	Observed/Expected
RBI	Richards-Baker Index (of flashiness)
$\hat{R}^2$	Adjusted R^2 (coefficient of determination)
USEPA	United States Environmental Protection Agency
USGS	United States Geological Survey
WY	Water Year

CHAPTER 1. INTRODUCTION

CONTEXT

Defining flow-biology relationships is crucial to the management of freshwater quantity and quality. Though the theories and concepts of flow-biology are largely understood and agreed upon (Poff & Zimmerman 2010), flow-biology relationships are difficult to quantify. This project came about in response to the Minnesota Pollution Control Agency's (MPCA's) Intensive Watershed Monitoring results, which point to flow alteration as a primary stressor to the fish and benthic macroinvertebrates in Minnesota's rivers and streams. The term flow alteration can refer to any change in the five major components of a hydrologic system: magnitude, timing, duration, frequency, and rate of change. The causes of flow alteration can range from anything as complex as climate change, to something as simple as a dam. All of these alterations affect the integrity of aquatic life and can be hydrologically measured. The purpose of this study was to detect the response of benthic macroinvertebrate communities to flow metrics across Minnesota streams.

A multi-agency, Ecological Flows research project began in 2017 to determine how human induced hydrologic alterations can be measured throughout the state, with the next step being to determine if those measures can be related to biological data. This thesis project was an auxiliary, exploratory effort to investigate relationships between benthic macroinvertebrates and measures of flow *a priori*, and then observe *a posteriori* if significant relationships were found with the same flow metrics that were determined to have changed over time because of human alterations of the hydrologic system. Ultimately, the information drawn from this thesis project will inform the MCPA's stressor identification process, which largely uses symptomatic measures of flow alteration (e.g. channel stability) to diagnose effects as a stressor.

Thanks to the effort and guidance from MPCA modeling and biological staff, this project evolved from using solely large scale flow gage data and generalized biological measures (e.g. index of biotic integrity scores), to smaller watershed scales and more detailed biological information (e.g. percent predators). Analyses at the smaller analytical scale were also pursued because of the possible "dilution" of signals in the data at larger scales with more generalized biological measures. Additionally, many of Minnesota's

aquatic life use impairments are in headwater drainages, so relationships at that scale may be more pertinent to stressor identification studies and effectiveness monitoring.

PROJECT SUMMARY

This thesis project used benthic macroinvertebrate community data rather than fishes because a general lack of mobility presumably makes them more vulnerable to extreme hydrologic events. Benthic macroinvertebrates have been studied as responders to various types of human activity, such as current land use (Anderson & Vondracek 1999), historical land use (Warwick 1980; Harding et al. 1998), point source discharge (Ferrington & Crisp 1989), and erosion control practices (Lenat 1984; Raymond & Vondracek 2011). Some studies have found them to be reliably associated with stream stability (Bond & Downes 2003; Markos et al. 2016). Many regulatory agencies use benthic macroinvertebrate samples to assess the Nation's waters for their support of aquatic life use, as required in the Clean Water Act of 1972 (33 U.S.C. § 1251 seq.), often summarizing the community into an Index of Biotic Integrity (IBI) score (Karr 1981; Miller *et al.* 1988). Minnesota has a large IBI dataset, and hundreds of additional metrics associated with each IBI score. At the start, this study used only the IBI scores to examine flow-biology relationships. With the help of MPCA biologists, more descriptive data was made available and some guidance was provided on which of the other metrics might be most likely to provide a detectable flow response. Besides IBI, a suite of 29 benthic macroinvertebrate community metrics were chosen for evaluation as dependent variables in a stepwise linear regression analysis on independent variables of streamflow metrics.

The hydrologic data used in this study included both observed and modeled data. Observed flow data were taken from U.S. Geological Survey (USGS) stream gaging stations that had daily mean flow values for each day of the year and were not affected by a significant impoundment. Modeled data were calculated with a Hydrologic Simulation Program Fortran (HSPF) model, built by private consultants and MPCA staff. The modeled data was retrieved using the HSPF-Scenario Application Manager software interface, which is publicly available online. All flow data were analyzed using the EflowStats package (Mills & Blodgett 2017) in R statistical software (R Core Team 2017). This software quickly calculates 171 hydrologic metrics that represent all five components of the flow regime (magnitude, timing, duration, frequency, and rate of

change). The EflowStats package was also used in the aforementioned Ecological Flows project, attempting to define human induced streamflow alterations in Minnesota.

The research question, "**Can benthic macroinvertebrate metrics be predicted by streamflow metrics?**" was tested by regressing 31 different biological metrics on 172 streamflow metrics. Multiple linear regression equations were built for each dependent (biological) variable using a variable selection technique called stepwise linear regression. Most equations were highly significant ($p<0.001$), using anywhere from one to eighteen predictor variables in the equation. The fit of the models varied widely. Ultimately, it was concluded that yes, measures of benthic macroinvertebrate communities can be predicted by measures of streamflow using linear regression at both the large and headwater watershed scales.

LITERATURE REVIEW: Water rights of the environment

The need to quantify flow-biology relationships is dictated, in large part, by the law, and has resulted in an increasing number of research projects in the field of ecological flows (also called environmental flows in some literature). Without significant effort to maintain a balance between the freshwater resource demands of humans and wildlife, the outcome will be as undesirable as it is predictable. Some researchers believe we are presently in the Earth's sixth mass extinction event (Barnosky et al. 2011; Thomas et al. 2004; Novacek & Cleland 2001), wherein the rate of biodiversity decline over the last century would have taken 800-10,000 years, using conservative estimates of historical extinction rates (Ceballos et al. 2016). Additionally, the decline in biodiversity of freshwater organisms may be occurring much more rapidly than that of marine or terrestrial ecosystems (Dudgeon et al. 2006; Strayer and Dudgeon 2010).

In 2016, the USEPA and USGS released a technical report on protecting aquatic life from the detrimental effects of hydrologic alteration (Novak *et al.* 2016), citing the objective of the Clean Water Act to "restore and maintain the chemical, physical, and biological integrity of the Nation's waters (Section 101(a))." The physical integrity of a water body is often equated with physical habitat integrity, while some propose that it should be equated with channel stability (Asmus et al. 2009). However, even channel stability, which does determine habitat quality, is itself determined by the flow regime to which that channel is and has been subjected, which is the basis of the hypothesis of this study and many others (Bunn & Arthington 2002; Poff & Zimmerman 2010; Carlisle *et al.* 2011; Markos *et al.* 2016).

Linkages between flow and biological data may be less useful to water managers than direct measures of habitat. However, the allure of this investigation is that flow data models are arguably more advanced and available than habitat models at this time. Additionally, flow data already have been modeled for the entire state of Minnesota for 1995 through 2012, and through 2015 in most watersheds. Thus, defensible relationships drawn between Minnesota's modeled flow database and biological database would be incredibly useful for the stressor identification process and other efforts to protect the physical and biological integrity of Minnesota's freshwater.

To assist water managers and regulators in meeting the CWA's call for protection of physical integrity, as well as state laws to protect wildlife and satisfy human water demands, an international group of researchers developed the Ecological Limits of Hydrologic Alteration (ELOHA) (Poff *et al.* 2009). The ELOHA framework is a methodology for assessing environmental flow needs for rivers and streams. Some states, Tribes, and other regulatory bodies have already implemented laws and ordinances protecting an environmental flow. For example, the South Carolina Board of Health and Environmental Control required a hydroelectric dam project to meet flow requirements that would protect the endangered shortnose sturgeon before they would grant a license renewal (Novak *et al.* 2016). In the context of water withdrawal permits, Minnesota Rule 6115.0630 Subp. 12 defines protected flow as, "the amount of water required in the watercourse to accommodate instream needs such as water-based recreation, navigation, aesthetics, fish and wildlife habitat, water quality, and needs by downstream higher priority users located in reasonable proximity to the site of appropriation." The Minnesota Department of Natural Resources (MDNR) is required to establish and protect this flow while evaluating and granting water withdrawal permits (Minn. R. part 6115.0670, subp.3.B.6).

In 2010, the North Carolina Department of Environment and Natural Resources was directed, by state law, to establish flows that protect the ecological integrity of rivers and streams. The law defined ecological integrity as,

"…*the ability of an aquatic ecosystem to support and maintain a balanced, integrated, adaptive community of organisms having a species composition, diversity, and functional organization comparable to prevailing ecological conditions and, when subjected to disruption, to recover and continue to provide natural goods and services that normally accrue from the system*." -North Carolina General Assembly Session Law 2010-143

In response, a collaborative research effort ensued and published their findings in the *Journal of American Water Resources Association* (Phelan *et al.* 2017). Ultimately, quantile regression was used to develop twelve flow-biology models, relating six metrics of flow alteration to one metric of fish diversity (Shannon-Weaver diversity of riffle-run guild) and one metric of benthic macroinvertebrate richness (EPTr[1]). The six flow alteration metrics used were: five ecodeficits (annual, winter, spring, summer, and fall), and the reduction in average annual 30-day minimum flow. EPTr was generally more sensitive to changes in flow than the fish metric, as it responded more strongly to four of the six flow metrics, being most sensitive to spring ecodeficit. However, the greatest observed decrease in either biological measure was the fish riffle-run guild response to summer ecodeficit. Because the relationships developed in the study were linear, they did not point to a quantitative threshold of impact beyond which the biological integrity starts to decrease. Rather, species richness decreases incrementally with every incremental increase in hydrologic alteration. So by definition, there was no level of flow alteration that could maintain biological integrity as defined in the state statute.

Some of the earlier ecological flows research includes a study by N. LeRoy Poff & J. David Allan (1995), which showed that hydrologically variable streams in Minnesota and Wisconsin can be expected to support more generalist fish species, whereas stable streams host a higher proportion of specialist species. The authors used data from 34 stream gage stations across the two states to calculate eight different hydrologic metrics, then related those metrics to functionally defined fish groups (e.g. omnivorous) using Canonical discrimination analysis and regression. The hydrologic metrics used were measures of variability and predictability of daily flow, spates, baseflow, and low flows. It was determined that “stable” sites had more benthic invertivores, with fewer omnivores and generalist invertivores. Fish communities from “variable” sites fostered more species that were typical of small streams and lentic waters, with fewer medium to large riverine species than the stable sites. Although, many earlier studies were done investigating flow-biology relationships (Horwitz 1978; Meffe 1984; Bain *et al.* 1988; Poff & Ward 1989; Jowett & Duncan 1990), the Poff and Allan (1995) paper confirmed what had largely been only theorized: that species in hydrologically variable environments were mostly resource generalists.

[1] Ephemeroptera, Plecoptera, and Trichoptera richness

Similar conclusions were drawn by Carlisle *et al.* (2011) in their multi-regional analysis of the conterminous U.S. The researchers tested and confirmed that, 1) flow alteration predicted biotic integrity better than eight other measures, 2) biotic impairment likelihood would increase with flow alteration, and 3) the presence of streamflow alteration could be detected by the functional traits of the streams' biological communities. The researchers used ratios of observed/expected (O/E) conditions for both the biology and hydrology to normalize the variables and control for natural covariates. Among other findings, they concluded that streams with diminished magnitudes fostered more species that are commonly associated with lentic environments and fine substrates. Additionally, the increasing severity of diminished flow magnitudes was associated with a doubling in the likelihood of biotic impairment.

In the Phelan *et al.* (2017), Poff & Allan (1995), and Carlisle *et al.* (2011) studies, data from vastly different ecoregions were compared. Although each study used a different study design and statistical analysis, the same operative theory was applied in all studies. There are predictable rules of community assembly (see Keddy 1992) that allow for functional, rather than taxonomic, comparisons of biological communities across large geographic scales, which enables the study of comparative community ecology (Karr *et al.* 1986; Vannote *et al.* 1980). A stable environment can be typified by strongly interactive species that rely on stable resources, while a variable (or frequently disturbed) environment is typified by weakly interactive, generalist species that are opportunistic and have strategies for using changing resources.

Notably, a large portion of the ecological flows literature is based on dryland river ecosystems, such as those found in the western United States and Australia (e.g. Gasith & Resh 1999; Walters & Post 2011; Kiran *et al.* 2011; Ruhí *et al.* 2016). However, the quantity and geographic distribution of research on the topic seems to increase dramatically after the year 2000. As discussed in Arthington *et al.* (2006), much of society is recognizing "the environment" as a legitimate "user" of water resources, and many are seeking to answer questions of how. How should environmental flows be measured, established, protected, and prioritized? In March 2018, the journal *Environmental Management* published an entire special issue (Volume 61 Number 3) on Adaptive Management of Environmental Flows.

This study fits into the context of the environment having water rights by taking some first steps in understanding how to protect aquatic life from flow alterations, namely,

investigating which component of the flow regime is most associated with the biotic integrity of benthic macroinvertebrates in Minnesota.

CHAPTER 2. METHODS AND MATERIALS

OVERVIEW

Two different watershed scales were used to find flow-biology relationships. In the first part of the study, a large watershed scale was used with observed flow data and generalized biological metrics. The second part used a smaller watershed scale with modeled flow data and more specific biological metrics. At both scales, the biological metrics (dependent variables) were regressed on a nearly identical suite of streamflow metrics (independent variables) using stepwise linear regression. Some flow metrics in the EflowStats package first required the calculation of the 1.67-year flow event; these metrics were used only in large watershed scale analysis.

LARGE SCALE ANALYSIS

-Streamflow data

In a separate project, data from 47 long-term USGS flow gages were used to find metrics that could be indicative of altered hydrology across the state of Minnesota (Ziegeweid *et al.* in prep). In an effort to produce comparable results, gages from that list were chosen for this part of the study, and the same hydrologic metrics were calculated. From those 47 gages, a subset was chosen that met the following criteria: 1) Flow at the gage was unimpacted by nearby impoundments 2) The drainage area contained at least ten biological data points that were not separated from the gage by a dam, lake, or wetland complex 3) The drainage area contained no more than one HUC8. If an impoundment was located between a biological station and the flow gage but it was not holding back enough water to change the channel width, as viewed from aerial imagery, the dam was disregarded and the biological data point was used. Applying those criteria resulted in fifteen USGS-operated flow gaging stations and 14-73 total biological data points per watershed that were used for the large scale analysis of flow-biology relationships (Table 1 and Figure 1).

Daily average discharge, in cubic feet per second (cfs), was used to calculate 172 different flow metrics. The years of record used included water year (WY) 1996 through 2015, using the water year definition of October 1st – September 30th, where the water year is defined by the calendar year in which it ends. Therefore, the entire record analyzed was 10/1/1996 - 9/30/2015. One exception was the Lac Qui Parle River

station, where there was a gap in data from 1999-2002. To resolve this, the metrics were calculated for WY1996-1998 and again for WY2003-2015, and the results were averaged.

Of the 172 flow metrics, 171 of them were calculated using the EflowStats package in R statistical software (Mills & Blodgett 2017; The R Core Team 2017). The various EflowStats metrics represent all five components of the flow regime: magnitude, timing, duration, frequency, and rate of change, and usually are normalized by a mean flow. Most of the EflowStats metrics are named according to which component they measure. For example, “dl22” is a measure of the duration of low flows. Often, the second letter is “h” or “l” because it stands for “high” or “low,” and is followed by a number. The numbers are just an identifier because the package measures duration of low flows in 21 other ways, for example. A description of the EflowStats metrics discussed in this report is given where applicable (Table 3), while some are also defined in-text for discussion purposes. A description of all 171 EflowStats metrics used in this analysis can be found in Appendix 5 of Henriksen *et al.* (2006).

The one, non-EflowStats flow metric used in this study was the Richards-Baker Index (RBI) (Baker *et al.* 2004). The RBI was developed as a measure of streamflow flashiness, which is a commonly observed stressor of aquatic life in Minnesota. The RBI was included in the analysis because the input data requirements and calculation are rather simple, resulting in an easily used tool with many potential applications.

Table 1. USGS flow gaging stations used in the analysis.

USGS site number	USGS site name	Drainage area [mi^2]	# Biological data points
05304500	Chippewa River nr Milan, MN	1880	60
05131500	Little Fork River at Little Fork, MN	1700	73
05132000	Big Fork River at Big Falls, MN	1480	45
05078500	Clearwater River at Red Lake Falls, MN	1380	48
05286000	Rum River nr St. Francis, MN	1360	48
05107500	Roseau River at Ross, MN	1090	24
05244000	Crow Wing River at Nimrod, MN	1030	25
05062000	Buffalo River nr Dilworth, MN	975	54
05300000	Lac Qui Parle River nr Lac Qui Parle, MN	960	43
05313500	Yellow Medicine River nr Granite Falls, MN	666	44
05316500	Redwood River nr Redwood Falls, MN	629	15
04010500	Pigeon River at Middle Falls, MN	609	14
05275000	Elk River nr Big Lake, MN	559	42
05069000	Sand Hill River at Climax, MN	460	15
05353800	Straight River nr Faribault, MN	435	45

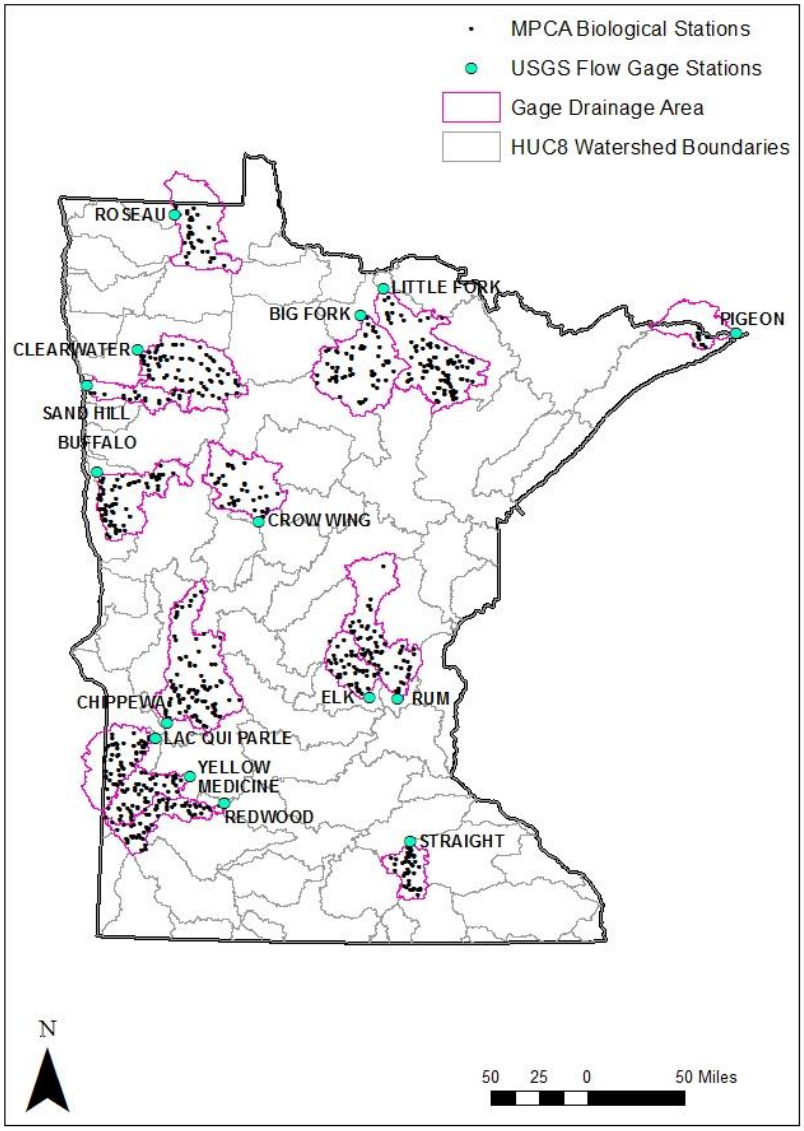

Figure 1. Map of USGS flow gage stations used in the analysis.

-Biological data

The MPCA's macroinvertebrate index of biotic integrity (MIBI) scores are the sum of ten biological metrics (MPCA 2014). Exactly which ten metrics are used to calculate the MIBI depends on characteristics of the stream from which the macroinvertebrate community was sampled. Minnesota's MIBI tool was calibrated to a human disturbance

score, which measures land-use, feedlot density, point-source density, reach channelization, watershed channelization, impervious surfaces, road density, and riparian conditions, among other things. Additionally, to be used in the final MIBI calculation, metrics had to pass a range test, signal to noise test, responsiveness test, and redundancy test. Furthermore, some metrics were transformed to eliminate their correlation with natural gradients. This procedure resulted in nine MIBI classes, differentiated by region, drainage area, stream gradient, and thermal regime. In doing so, researchers ensured that the tool would measure macroinvertebrate response to anthropogenic disturbance rather than natural phenomena.

Scores from the different MIBI classes are compared against different numeric thresholds. The thresholds are numeric MIBI standards, below which the benthic macroinvertebrate community is deemed “impaired,” and to be exhibiting a stress response. In addition to MIBI, a “Ratio” metric was used in this study, which was calculated by dividing the raw MIBI score by the numeric threshold of its class. Based on communication with MPCA biologists, using the Ratio was a plausible way to make the MIBI scores more comparable across the different classes. A similar approach was used by Carlisle *et al.* (2011) in their multi-regional analysis of flow-biology relationships, where the observed biotic measure was divided by an expected value (O/E), resulting in a binary division of the data, impaired (O/E<1) and unimpaired (O/E>1).

Biological information collected outside the years 1996-2015 was excluded from the study. If a station had more than one year of data, the MIBI scores were considered as two separate data points. If a station had two valid samples from within the same year, those MIBI scores were averaged and considered as a single data point. The MIBI and Ratio scores of all the biological data points were summarized into a mean and median value for each of the fifteen gaged watersheds.

HEADWATER SCALE ANALYSIS

-Streamflow data

In the second part of this study, modeled hydrologic data were used to seek flow-biology relationships at a smaller watershed scale. It was necessary to use a model due to a lack of gage stations at the desired scale with year-round data (required by the EflowStats package) and several years of record. Two HUC8 watersheds were chosen for this analysis, the Chippewa River (HUC 07020005) and Hawk Creek-Yellow Medicine River (HUC 07020004) (hereafter referred to as the Yellow Medicine) (Figure 2). These

two watersheds were also represented by a gage station in the large scale part of the study. The Chippewa and Yellow Medicine watersheds are similar in land use, geology, and geographic location. Both watersheds are in the Minnesota River-Headwaters basin. Using the headwater information from both of these HUC8s was intended to increase sample size (n) while introducing as little additional variability as possible.

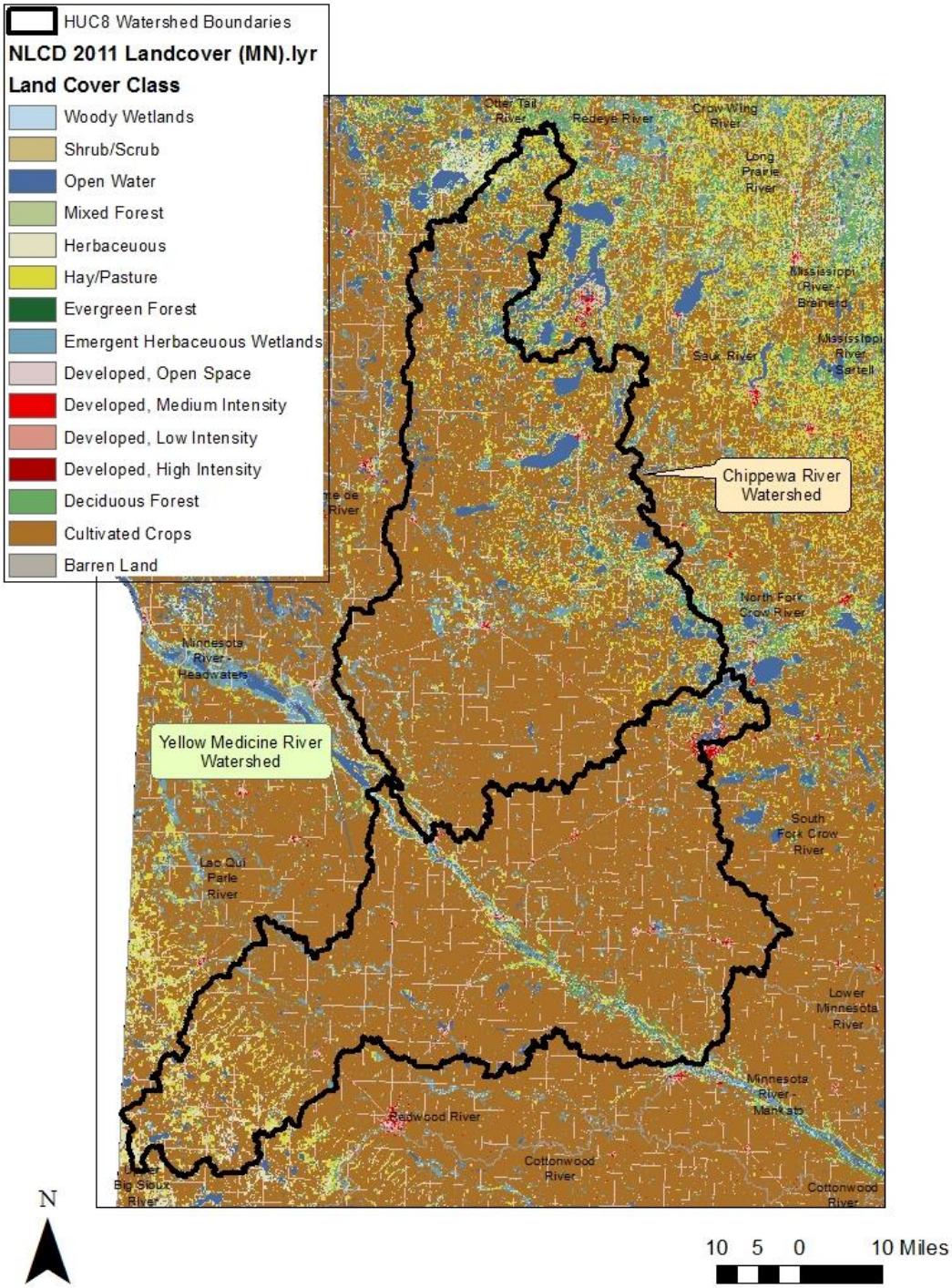

Figure 2. Land use in the Chippewa and Hawk Creek-Yellow Medicine HUC8 watersheds, according to the USGS 2011 National Land Cover Dataset.

A Hydrologic Simulation Program-Fortran (HSPF) model was used, the methodology for which was originally developed in the 1960's and has been modified by the USEPA, USGS, and others over time. The most recent iteration is Version 11,

hosted on the USGS website (Bicknell *et al.* 1997). The model uses several meteorological, hydrologic, geologic, and climatic data inputs to compute data points on hourly timesteps. The hydrologic metrics used in this analysis were computed based on average daily discharge values (in cfs). The modeled record available for the Chippewa and Yellow Medicine watersheds goes from 1996 through 2012. Water years, rather than calendar years, were used in this study, so to maximize the available data, the modeled hydrologic record used was 10/1/1996 - 9/30/2012.

In each HUC8, the “headwaters” were determined by mapping the average annual discharge volume from each HUC12 watershed. Then, the HUC12s were classified into five categories according to their magnitude of total discharge using the Jenks natural breaks classification system (Figure 3), which clusters data by minimizing within-class average deviation from the mean and maximizing the intra-class deviation from the means (Jenks 1967). The class of HUC12s chosen as “headwaters” were those in the first (smallest) of the five discharge categories. This method allowed for identifying the smallest possible contributing drainage areas that were also large enough to have been sampled for invertebrates. Some HUC12 headwaters did not contain any biological stations and were excluded from the study. In total, 48 HUC12s were used in the headwaters scale analysis, 19 from the Chippewa River watershed and 29 from the Yellow Medicine River watershed.

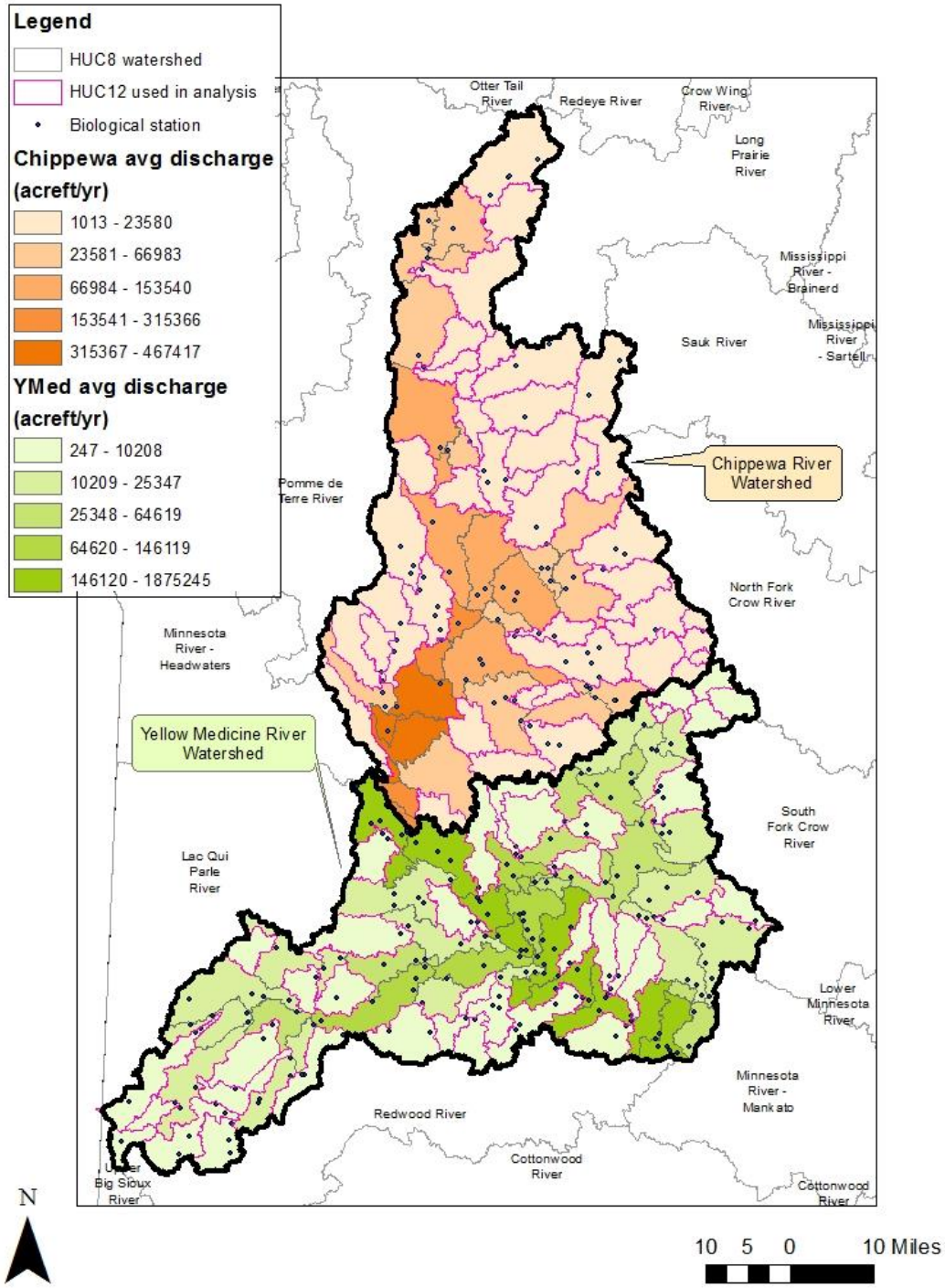

Figure 3. Map of watershed discharge classes used to define “headwaters.”

-Biological data

The number of biological stations in each headwater drainage ranged from one to six. For those with multiple stations, the biological metrics from each station were

averaged. Stations that were separated from the HUC12 outlet by an impoundment were excluded from the analysis if the impoundment caused a change in channel width, as viewed from aerial imagery. Some stations had samples from multiple years; each year was treated as an individual data point. If a station had two valid samples from within the same year, those scores were averaged and considered as a single data point. Biological data that were sampled outside the years 1996-2015 were excluded from the study. The MIBI classes represented by the biological dataset were: Class 5-Southern streams riffle-run (18.3%), Class 6-Southern forest streams glide-pool (1.7%), Class 7-Prairie streams glide-pool (89%), and Class 9-Southern coldwater (2.6%) (see MPCA 2014). Twenty-nine biological metrics were calculated for each HUC12 (Table 2).

Table 2. Metrics of benthic macroinvertebrates used in the study.

Metric	Description
Climber	Taxa richness of climbers (excluding chironomids)
ClimberCh	Taxa richness of climbers
ClimberChTxPct	Percent of climber taxa
ClimberPct	Percent of climber individuals
Clinger	Taxa richness of clingers (excluding chironomids)
ClingerCh	Taxa richness of clingers
ClingerChTxPct	Percent of clinger taxa
ClingerPct	Percent of clinger individuals
CollectorFilterer	Taxa richness of collector-filterers (excluding chironomids)
CollectorFiltererCh	Taxa richness of collector-filterers
CollectorGatherer	Taxa richness of collector-gatherers (chironomid and beatid each treated as one taxon)
CollectorGathererCh	Taxa richness of collector-gatherers
DomOneChPct	Percent of most dominant taxa in subsample
DomTwoChPct	Percent of two most dominant taxa in subsample
DomThreeChPct	Percent of three most dominant taxa in subsample
InsectPct	Percent of insect individuals
InsectTxPct	Percent of insect taxa
LongLivedPct	Percent of long-lived individuals
LongLivedChTxPct	Percent of long-lived taxa
NonInsectPct	Percent of noninsect individuals
NonInsectTxPct	Percent of noninsect taxa
PredatorPct	Percent of predator individuals
PredatorChTxPct	Percent of predator taxa
Sprawler	Taxa richness of sprawlers (excluding chironomid and baetid sprawler taxa)
SprawlerCh	Taxa richness of sprawlers
SprawlerPct	Percent of sprawler individuals
Swimmer	Taxa richness of swimmers (excluding chironomids; baetid taxa treated as one taxon)
SwimmerCh	Taxa richness of swimmers
SwimmerPct	Percent of swimmer individuals

REGRESSION ANALYSIS

Stepwise linear regression was performed at the large and small watershed scales with each biological metric (dependent variable) regressed on the suit of streamflow metrics (independent variables). Stepwise linear regression is a variable selection technique that adds and removes predictor variables from the linear model to find the best possible fit while avoiding multicolinearity of the predictor variables. As each predictor variable is added and removed from the equation, the Akaike information criterion (AIC) is minimized, which produces the best possible fit for the number of variables included (Akaike 1998). The stepwise regression function used, called “step,” is part of the R package ‘stats’ (version 3.4.2) (The R Core Team 2017).

The output of each stepwise linear regression analysis (one for each biological metric) were multiple linear regression equations in the form of Equation 1, where: $\hat{y}$ is the predicted dependent variable; b_0 is the intercept; b_j is the coefficient of the independent variable; x_j is the independent variable. Of note, the final model output of the stepwise algorithm ordered the predictor variables by their usefulness, where considered alone, the first independent variable in the equation would give the best fit of all.

$$\hat{y}=b_0+b_1x_1+b_2x_2\ldots+b_jx_j \qquad \textbf{(1)}$$

The Adjusted R^2 ($\hat{R}^2$) is used to discuss the goodness of fit because it accounts for the number of independent variables used in the equation, whereas (unadjusted) R^2 does not. Equation 2 shows that $\hat{R}^2$ will decrease when new parameters are added (because m increases) that do not improve the fit very much (because SSE decreases), where: n is the sample size; m is the number of independent variables; SSE is the sum of squares error/residual; SST is the total sum of squares.

$$\hat{R}^2 = 1 - \frac{\mathrm{SSE}}{\mathrm{SST}} * \frac{\mathrm{n}-1}{\mathrm{n}-\mathrm{m}-1} \qquad \textbf{(2)}$$

Equations with more than ten predictor variables were not considered for further analysis. Those with no more than five predictor variables and a $p<0.05$ were analyzed further by: iteratively removing predictor variables from the equation, assessing the appropriateness of a linear model using residual plots, and/or transforming the dependent variable.

CHAPTER 3. RESULTS

OVERVIEW

Significant ($p<0.05$) linear regression equations were built for all of the macroinvertebrate metrics. The number of independent variables in the equations ranged from one to 47, with an $\hat{R}^2$ range of 0.17 - 0.93. A subset of the best-fitted equations was chosen for further investigation. An explanation of the EflowStats metrics used in these equations was adapted from the R package's documentation (Table 3). For an explanation of all EflowStats metrics, see Appendix 5 in Henriksen *et al.* (2006).

Table 3. Names and descriptions of the flow metrics used as independent variables in the linear regression equations.

Metric name	Description
dh6	Variability of annual maximum daily flows. Compute the standard deviation for the maximum 1-day moving averages. Dh6 is 100 times the standard deviation divided by the mean.
dh7	Variability of annual maximum 3-day moving average. Compute the standard deviation for the maximum 3-day moving averages. Dh7 is 100 times the standard deviation divided by the mean.
dh22	Flood interval. Compute the flood threshold as the flow equivalent for a flood recurrence interval of 1.67 years. Determine the median number of days between flood events for each year. Dh22 is the mean of these values.
dl9	Variability of annual minimum 30-day moving average flow. Compute the standard deviation for the minimum 30-day moving averages. Dl9 is 100 times the standard deviation divided by the mean.
dl17	Low flow pulse variability. Compute the average pulse duration for each year for flow events below the 25th percentile value from the entire record. Dl17 is 100 times the standard deviation of these values, divided by the mean.
fh5	Flood frequency. Compute the number of flow events each year with flows above a threshold equal to the median flow value from the entire record. Fh5 is the mean of these values.
fh10	Flood frequency. Compute the number of flow events each year with flows above a threshold equal to the median of the annual minima for the entire flow record. Fh10 is the mean of these values.

ma18	Mean of monthly flows. Compute the mean of July flows over the entire record.
ma21	Mean of monthly flows. Compute the mean of October flows over the entire record.
mh8	Mean maximum monthly flows. Compute the mean of August maximum flows over the entire record.
mh19	Skewness in annual maximum flow. Use the equation below, where: N is the number of years on record qm is the log of annual maximum flows σ is the standard deviation of annual maximum flows $$\frac{3[(2N\sum(qm3)) - (3N\sum(qm)\sum(qm2)) + (2\sum qm)]}{N(N-1)(N-2)(3\sigma)}$$
ml1	Mean of minimum monthly flows. Compute the mean of minimum January flows over the entire record.
ml12	Mean of minimum monthly flows. Compute the mean of minimum December flows over the entire record.
ml16	Median annual minimum flow. Compute the ratio of the annual minimum to median flow for each year on record. Ml16 is the median of these values.
ra8	Number of reversals. Compute the number of days in each year that the change in flow from one day to the next changes direction. Ra8 is the mean of these values.
tl1	Julian date of annual minimum flow. Determine the Julian date that the minimum flow occurs in each year. Transform the dates to relative values on a circular scale (radians or degrees). Compute the x and y components for each year and average them across all years. Compute the mean angle as the arc tangent of the y-mean divided by the x-mean. Transform the resultant angle back into a Julian date.

LARGE SCALE ANALYSIS

The biological metrics analyzed at the large watershed scale were the mean and median of MIBI and Ratio scores (Figure 4). However, there was little difference between the watershed mean and median for either metric, plus the same independent variables were useful in predicting each, therefore interpretation of the results will focus on the mean only. Most of the independent variables are dimensionless; however, their coefficients are not directly comparable because some of the metric formulas include a multiplicative coefficient (usually 100) and others do not.

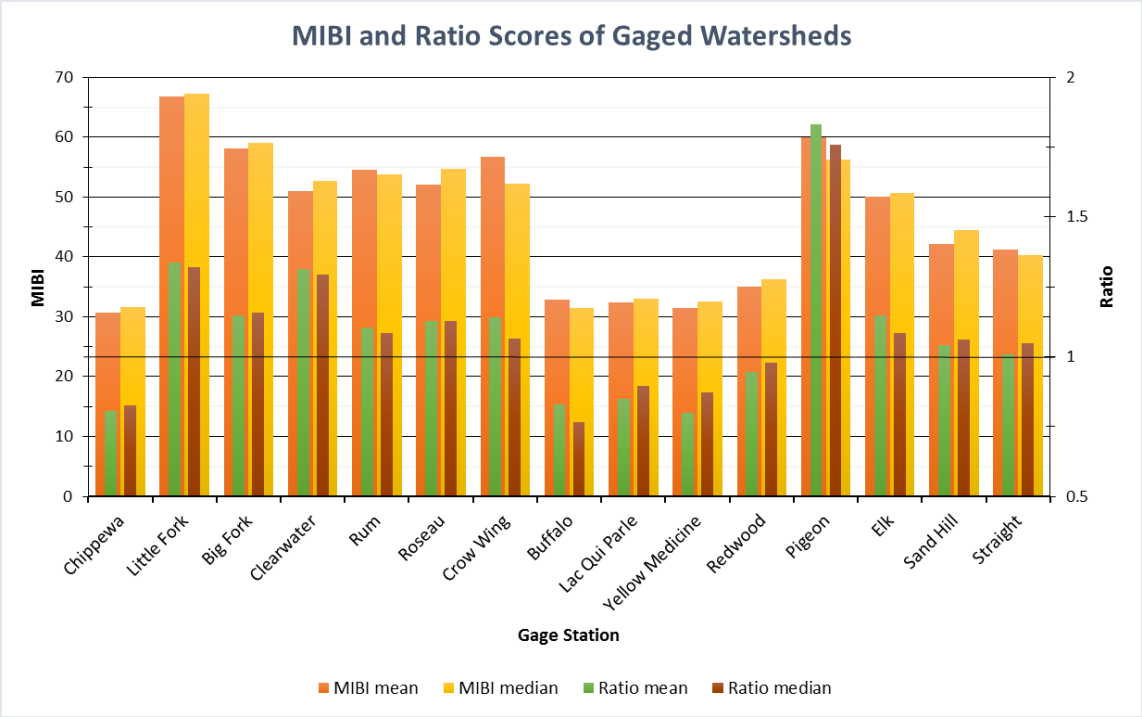

Figure 4. Mean and median MIBI and Ratio scores from biological data points within the flow gage station drainage area. MIBI and Ratio are displayed on different vertical axes. The solid horizontal line drawn at Ratio = 1.0 represents the threshold below which benthic macroinvertebrate communities are considered impaired.

The single most predictive flow metric was dh7 for MIBI and mh19 for the Ratio, with dh22 at a close second. In the interest of selecting well-fit models that use as few predictors as possible, the model statistics are reported for various versions of the MIBI and Ratio models for comparison (Table 4). Ultimately it was decided that the equations in bold were the best models for that biological variable.

Table 4. Summary of simple and multiple linear regression equations. Data from fifteen watersheds (n=15) were used to build all equations. All independent variables are significant at the 95% level. Equations in bold were chosen as the best model for the biological variable.

Dependent variable (n=15)	Intercept	Independent variables (p<0.05)	Adjusted R^2	p-value
MIBI	77.030	(-0.481)dh7	0.667	1.24E-4
	120.269	(-0.847)dh7 + (-72.390)ml16	0.886	8.87E-7
	131.434	**(-0.846)dh7 + (-74.961)ml16 + (-0.037)tl1**	**0.927**	**4.04E-7**
Ratio	1.011	(-0.408)mh19	0.674	4.30E-7
	2.803	(-0.263)mh19 + (-0.006)dh22	0.843	5.98E-6
	3.491	(-0.232)mh19 + (-0.007)dh22 + (-0.004)dl17	0.926	4.30E-7
Sqrt(Ratio)	**2.051**	**(-0.114)mh19 + (-0.003)dh22 + (-0.002)dl17**	**0.912**	**1.15E-6**

-MIBI

In predicting the watershed's mean MIBI score with dh7, $\hat{R}^2$ is improved considerably (≥24%) by adding the second predictor variable, ml16. The $\hat{R}^2$ increases a bit further by adding one more flow metric, tl1. Also, the addition of tl1 yields residual plots that suggests a better linear fit than when just the first one or two independent variables are used (Figure 5).

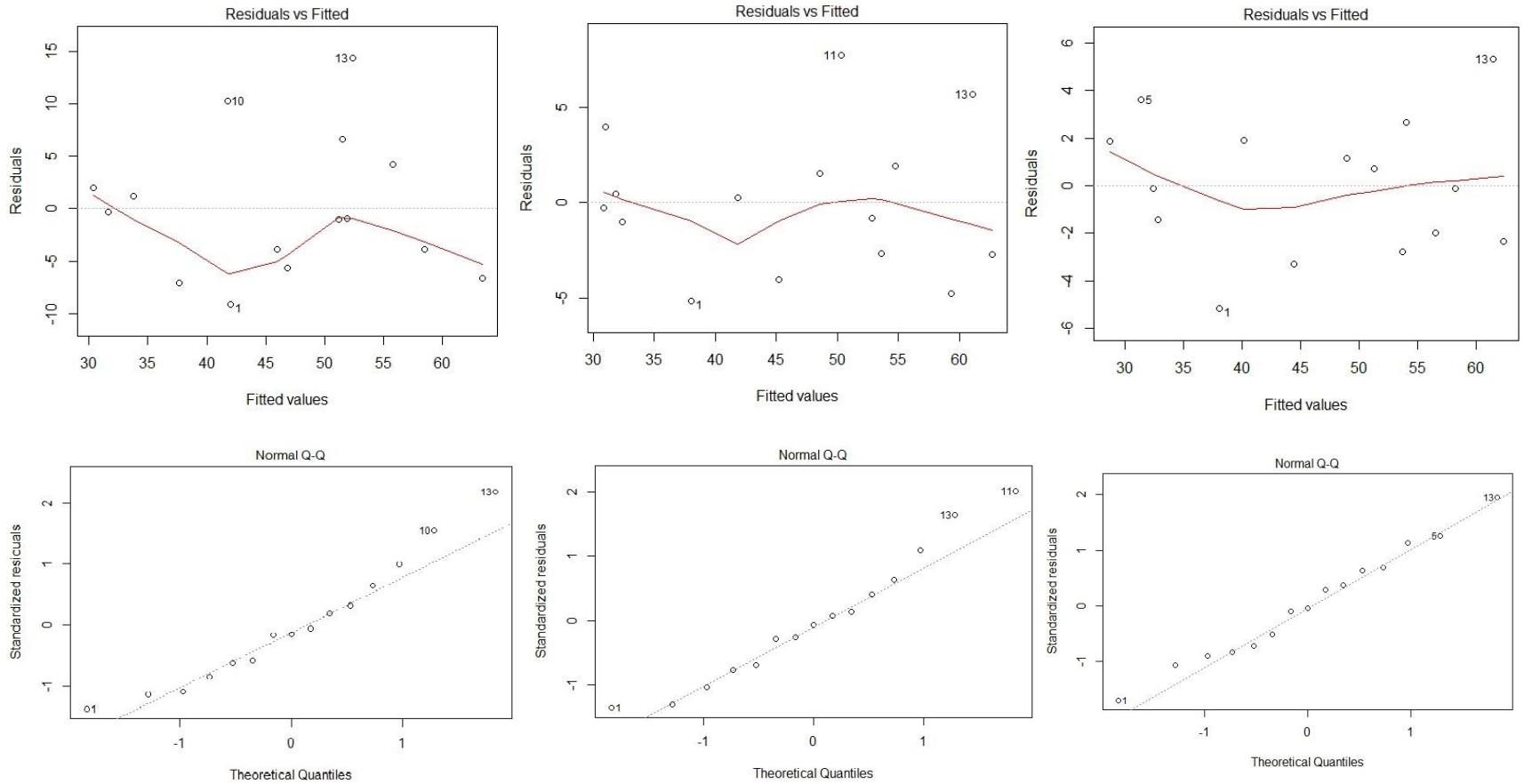

Figure 5. Residual plot matrix of MIBI on dh7 (left column), dh7+ml16 (middle column), and dh7+ml16+tl1 (right column). The red line in the residuals vs. fitted (top row) plots is a locally weighted regression line.

-Ratio

In predicting the watershed's mean Ratio score with mh19, $\hat{R}^2$ is improved considerably (≥20%) by adding the second predictor variable, dh22. The Ratio equation can be improved further by adding dl17. However, the residual plots for all three versions of this model suggest that transformations of some variables may improve the linear fit (Figure 6). Square root and natural log transformations of the dependent variable were tried, and the fit was slightly improved with the square root transformation (Figure 7).

Of note, the Pigeon River watershed had an exceptionally low dh22 score (209.55). Removing Pigeon River from the analysis drops the $\hat{R}^2$ of the final Ratio equation (0.9263) by 11.5% (0.8198).

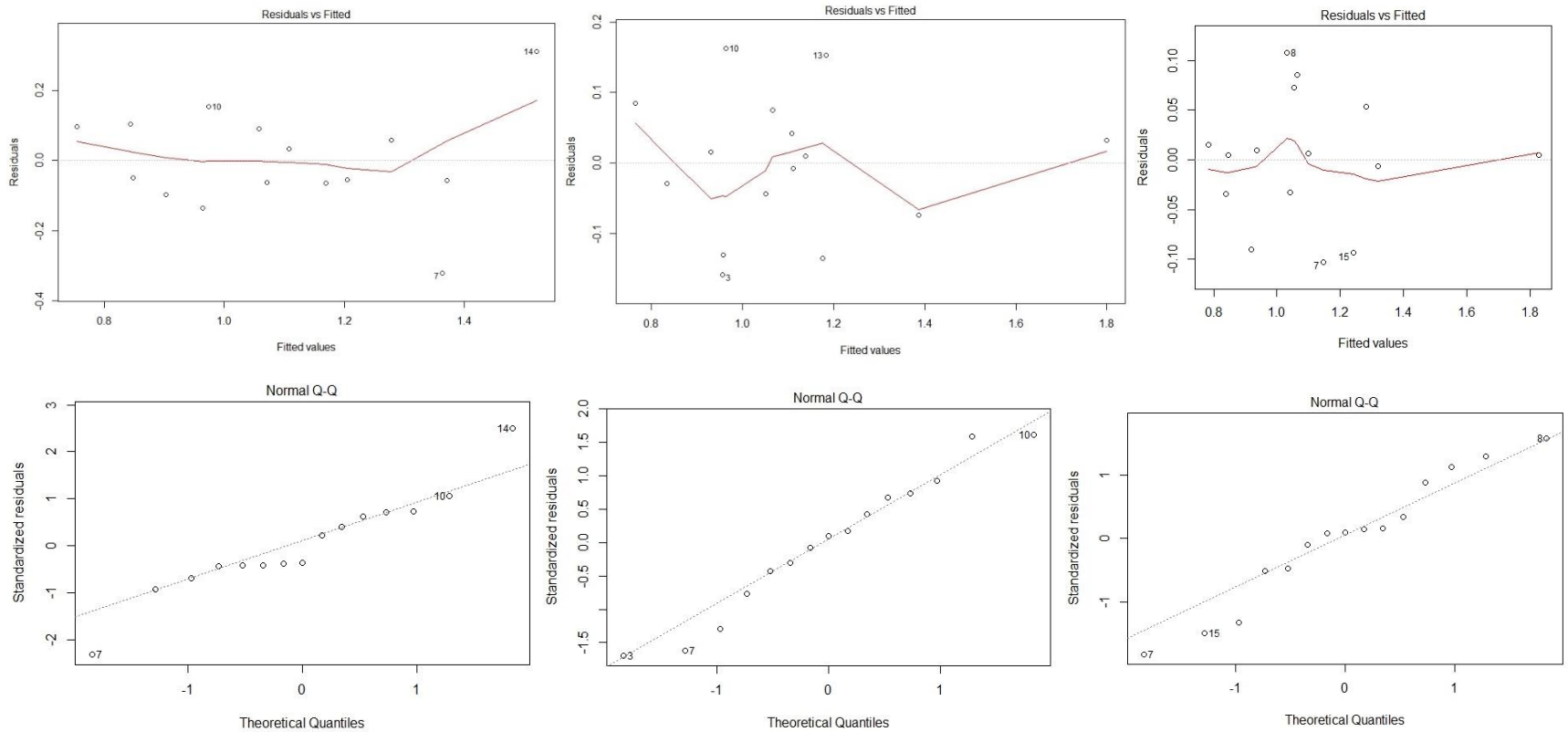

Figure 6. Residual plot matrix of Ratio on mh19 (left column), mh19+dh22 (middle column), and mh19+dh22+dl17 (right column). The red line in the residuals vs. fitted (top row) plots is a locally weighted regression line.

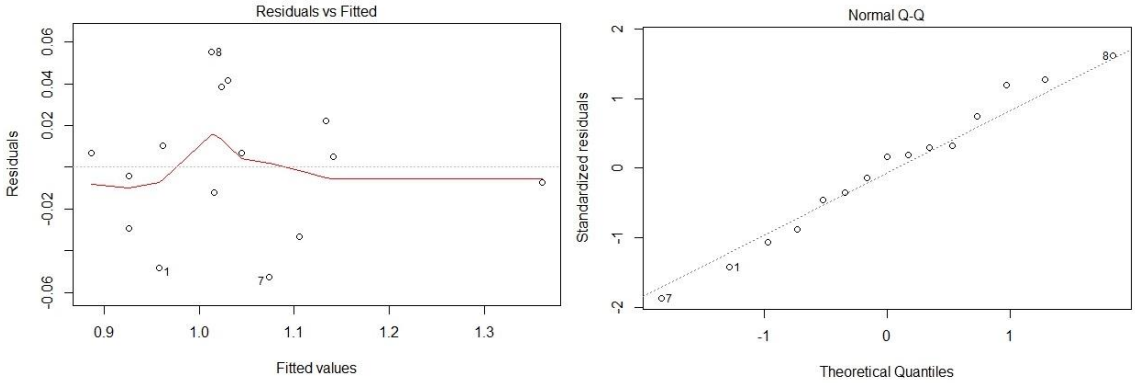

Figure 7. Residual plots of Ratio (square root transformed) on mh19+dh22+dl17. The red line in the residuals vs. fitted (left) plot is a locally weighted regression line.

HEADWATER SCALE ANALYSIS

Stepwise linear regression at the headwater scale yielded significant (p<0.05) linear models for all 29 dependent variables, 28 of which were significant at the 99% level (p<0.01). However, many equations used ten or more independent variables, achieving a range of $\hat{R}^2$ between 0.17 and 0.73. Additionally, some independent variables in the original equations had Pearson correlation coefficients nearing 0.70.

No discernable pattern emerged regarding a specific metric or component of the flow regime that was frequently used among all the equations. A subset of equations was chosen for the focus of this study (Table 5), all of which were modified from the original form outputted by the stepwise algorithm by removing predictor variables. No two independent variables with a Pearson correlation coefficient $\geq$0.40 are used in the same equation. The reader may find all 29 equations in the Appendix in their original form, as output by the stepwise algorithm. All residual plots were examined to evaluate the linear fit and distribution of residuals, though not all plots are shown.

-Clingers

Four different ways of measuring the clinger habitat guild presence in the biological data were analyzed: ClingerPct, ClingerChTxPct, ClingerCh, and Clinger (see Table 2 for more detail). Of all the biological metrics analyzed, the suite of clinger metrics yielded some of the strongest linear relationships found, using only two or three significant predictors. Considering only those equations with three or fewer predictor variables, the largest $\hat{R}^2$ found was 0.479, using three flow metrics to predict ClingerPct (Table 5).

There was a wide range (0-91) of ClingerPct values in the sample, with a relatively even distribution within the range from 0-56. However, 91 was the only data point greater than 56, suggesting that it may be an outlier. Removing this data point from the sample decreases $\widehat{R}^2$ from 0.479 to 0.457. The flow metric fh10 appears in three of the four clinger equations, and ra8 appears in two.

-Swimmers

Considering only equations with three or fewer predictor variables, excluding clinger equations, the next largest $\widehat{R}^2$ found was 0.393, using only two flow metrics to predict SwimmerPct (Table 5). A square root transformation was applied to the dependent variable, and based on residual plots, the transformation improved the linearity of the relationship (Figure 8). Another relatively strong swimmer response was seen in the metric SwimmerCh, which, in its best model, has three predictor variables. No predictor variables were the same between the SwimmerPct and SwimmerCh equations.

Table 5. Summary of multiple linear regression equations. Data from 48 watersheds (n=48) were used to build all equations. Most independent variables are significant at the 90% level. Equations in bold were chosen as the best model for the biological variable.
* Predictor not significant at 90% level

Dependent variable (n=48)	Intercept (p<0.1)	Independent variables (p<0.1)	Adjusted R^2	p-value
ClingerPct	-16.81	(10.45)fh10 + (0.35)mh8	0.337	3.63E-5
	-31.69	**(12.77)fh10 + (0.45)mh8 + (13.15)mh19**	**0.479**	**5.41E-7**
ClingerCh	-13.72	(0.06)dh6 + (0.12)ra8	0.326	5.24E-5
	-17.76	**(0.06)dh6 + (0.13)ra8 + (1.10)fh10**	**0.395**	**1.35E-5**
Clinger	-2.85*	(1.77)fh10 + (0.60)ml12	0.267	0.0003
	-16.03	**(1.89)fh10 + (0.62)ml12 + (0.04)ra8**	**0.378**	**5.24E-5**
ClingerChTxPct	**-14.33***	**(0.21)dh6 + (1.36)fh5**	**0.310**	**8.97E-5**
SwimmerPct	4.39	(4.61)ml1 + (-0.96)ma21	0.393	5.03E-6
Sqrt(SwimmerPct)	**1.61**	**(0.70)ml1 + (-0.12)ma21**	**0.321**	**6.20E-5**
SwimmerCh	1.67	(0.09)ma18	0.330	1.19E-5
	0.68*	(0.11)ma18 + (0.01)dl9*	0.354	2.02E-5
	4.44	**(0.13)ma18 + (0.01)dl9 + (-0.03)th1**	**0.385**	**1.90E-5**

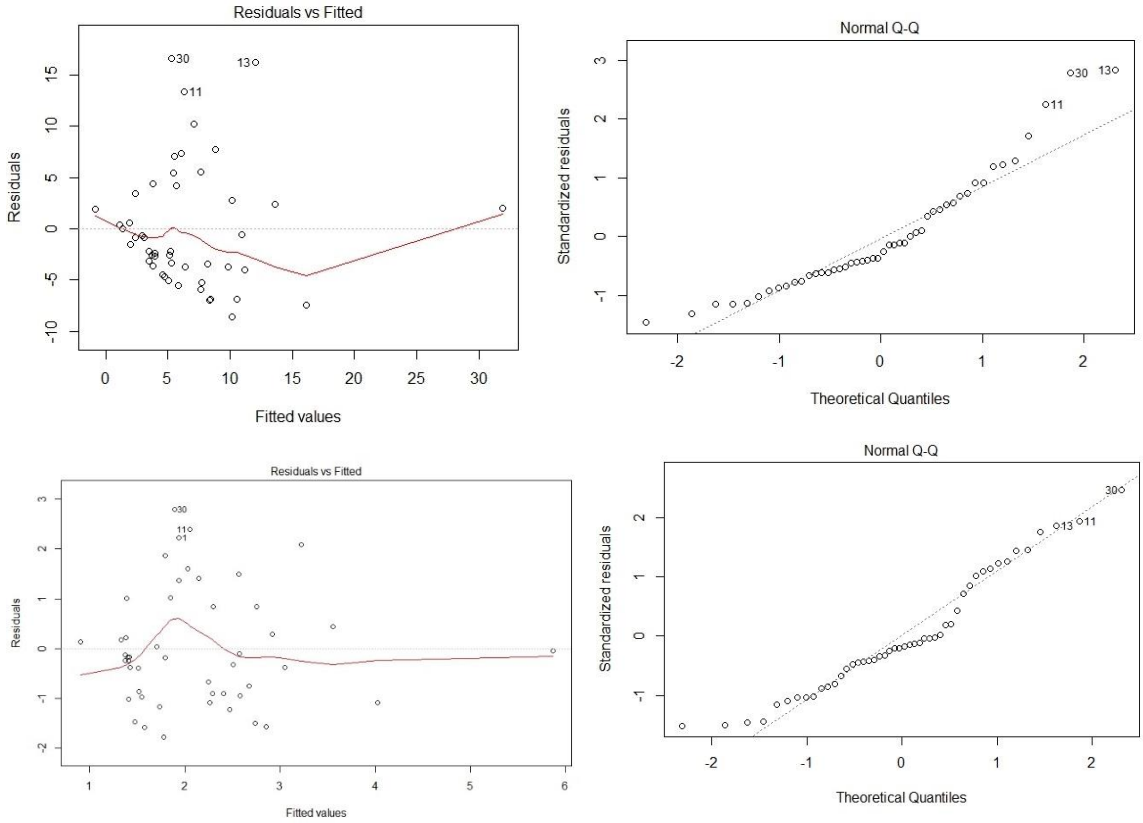

Figure 8. Residual plots of SwimmerPct (top row) and SwimmerPct (square root transformed) (bottom row) on ml1+ma21. The red line The red line in the residuals vs. fitted (left) plot is a locally weighted regression line.

CHAPTER 4. DISCUSSION

OVERVIEW

To deepen the understanding of these flow-biology equations, each dependent variable is discussed in terms of the independent variables. Because other covariates, such as water quality parameters, were not analyzed, speculations can only be made in terms of correlation of occurrence in these specific watersheds, for the time periods analyzed. For details about the metrics discussed here, readers are referred back to Tables 2 and 3, although some explanations are provided in-text.

LARGE SCALE ANALYSIS

The two biological variables analyzed at this scale, MIBI and Ratio, are discussed using specific watershed examples. Much natural variability is to be expected among the fifteen watersheds based on ecoregion differences and species distribution of aquatic organisms. Some research suggests that streams be grouped by a common hydrologic feature, such as gradient (Poff *et al.* 2009). Other studies have effectively detected a biotic response to flow alteration that was not explained by stream type or class (Phelan *et al.* 2017; Carlisle *et al.* 2011).

Phelan *et al.* (2017) discuss the hypothesis that classes of biota behave similarly in different places, and so using biological metrics, such as habitat guilds, can allow for cross-regional comparisons. While species present across the sample of watersheds may not be comparable, the MIBI response to streamflow behavior may be the same. Indeed, in the MIBI documentation (MPCA 2014), the authors cite the rationale of Karr *et al.* (1986) in explaining that some community measures are applicable across a range of conditions because species are replaced by similar species in moving from one ecotype to another.

-MIBI

Of 172 streamflow metrics, MIBI responded most strongly to dh7, a measure of variability in the duration of high flows, and ml16, a ratio of minimum to mean flow; both are negative relationships. The MIBI scores and hydrographs from two watersheds with very different dh7 and ml16 scores, Crow Wing River and Yellow Medicine River, will be examined more closely to further understand these results (Figure 9 and Table 6).

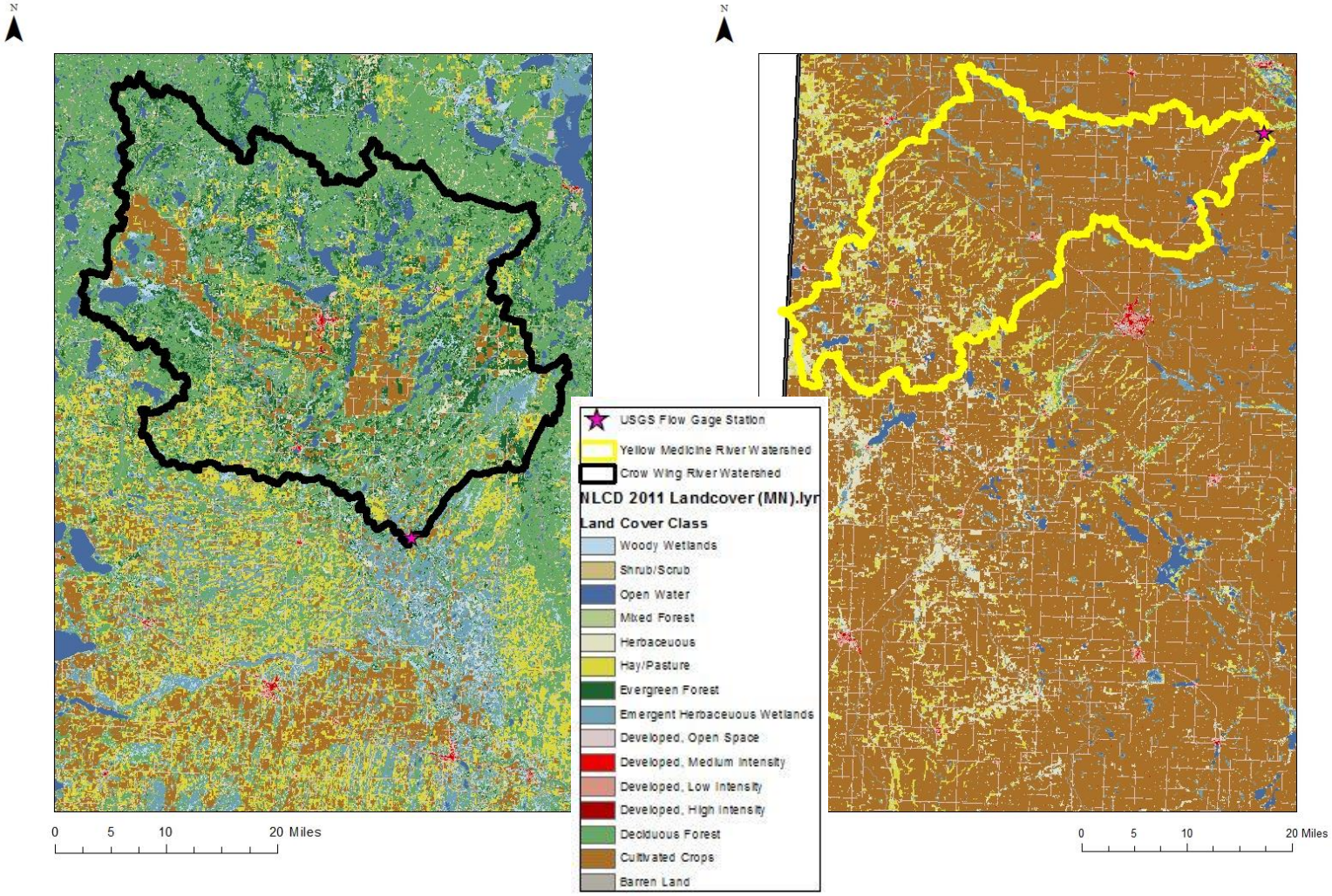

Figure 9. Land use in the Crow Wing River (left) and Yellow Medicine River (right) watersheds, according to the USGS 2011 National Land Cover Dataset.

Table 6. Overview of streamflow and macroinvertebrate metric scores in two different watersheds. The number of biological data points from each MIBI class is also shown.

	Yellow Medicine River	**Crow Wing River**
dh7	94	28
ml16	0.109	0.571
MIBI	31.38	56.67
Ratio	0.798	1.14
Class 1 – Northern forest rivers	-	2
Class 2 – Prairie forest rivers	1	-
Class 3 – Northern forest streams-riffle/run	-	6
Class 4 – Northern forest streams-glide/pool	-	14
Class 5 – Southern streams-riffle/run	12	-
Class 6 – Southern forest streams-glide/pool	-	-
Class 7 – Prairie streams-glide/pool	30	-
Class 8 – Northern coldwater streams	-	2
Class 9 – Southern coldwater streams	-	-
Total biological data points	43	24
Years sampled	91% collected 2010-2015	73% collected 2010-2012

Since these two watersheds have no MIBI classes in common, an analysis of the Ratios is also warranted. The Ratio metric did respond to the same predictor variables as MIBI (dh7 and ml16). A simple linear regression of Ratio on dh7 yields a significant equation ($p<0.01$) with $\hat{R}^2 = 0.40$, increasing to 0.66 when ml16 is added, with both independent variables being significant ($p<0.01$). A 30% difference exists between the Ratio scores of the Crow Wing and Yellow Medicine River watersheds. This suggests that, despite the natural taxonomic variability between these two watersheds, the differences in biotic integrity may be explained, to some extent, by their differences in high flow duration variability (dh7) and the ratio of minimum to median flows (ml16) during WY1996-2015.

Interpreting the ml16 metric results (see Table 6) tells that, generally speaking, the minimum flow is one-tenth the size of the median flow in the Yellow Medicine, whereas the Crow Wing minimum flow is half the size of its median flow. In regards to dh7, looking at the hydrographs from these two flow stations can help illustrate the differences in high flow variability (Figures 10-11).

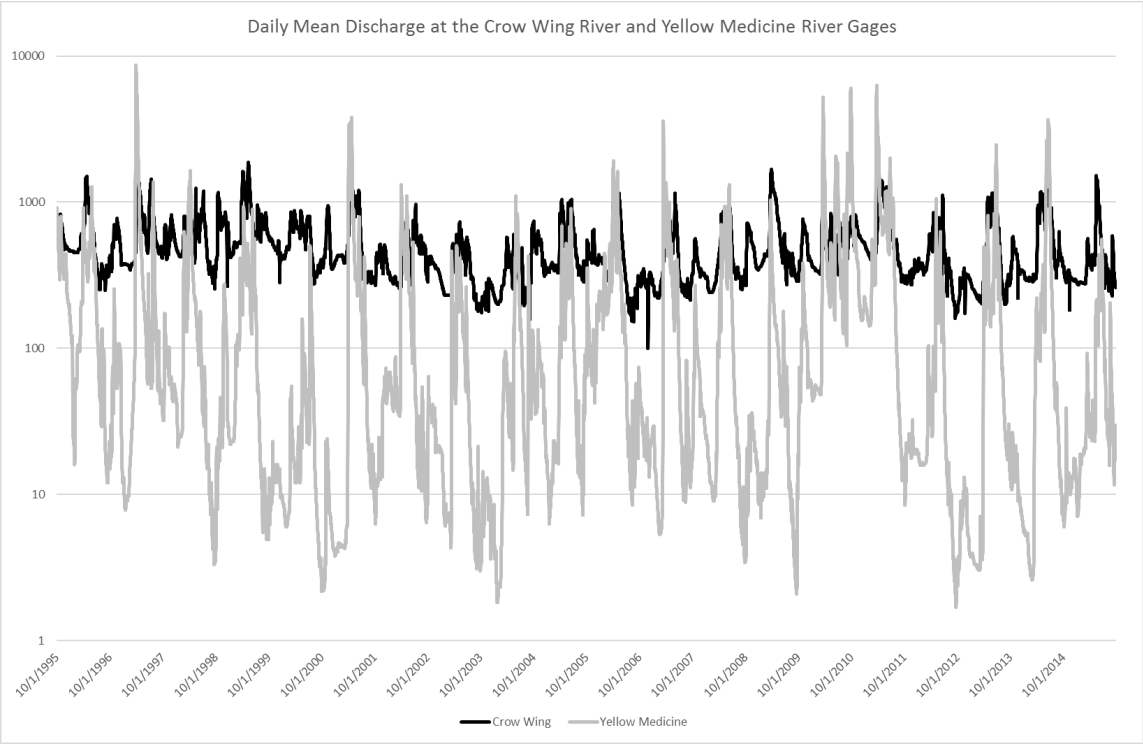

Figure 10. Mean daily discharges of the Crow Wing and Yellow Medicine Rivers as measured at USGS gage stations from 10/1/1995 to 09/30/2015. Values are plotted on a logarithmic scale.

Figure 11. A closer look at the mean daily discharges for water years 2010-2015. Solid Crow Wing) and dashed (Yellow Medicine) horizontal lines follow the annual maximum flow events at the two stations to show the difference in maximum flow variability between sites. Values are plotted on a logarithmic scale.

The difference in high flow variability may be a result of the difference in watershed storage between these systems. Wetlands, floodplains, and backwaters add hydrologic resilience to the system such that peak flows can be tempered and low flows are augmented. Many aquatic organisms need a hydrologic system to maintain its resilience in order to survive extreme events and ecological bottlenecks, as well as be predictably queued for flow-dependent life stages. Additionally, it stands to reason that a hydrologic system with more storage would have a less dramatic difference between minimum and median flows (greater ml16 score) due to the steady release of stored flood water over time.

Ziegeweid and others (2015) detected a non-precipitation driven trend in streamflow in HLU D, using data from gages within the HLU, including the Yellow Medicine River gage used in this study. In contrast, a non-precipitation driven trend in streamflow was not detected in HLU BC, using data from the same Crow Wing River

gage, along with others in the HLU. The authors state that the streamflow trends could be the result of human activity in the watershed, such as cropping patterns, drainage, or wetland construction.

Additional work by Ziegeweid and others (in prep) showed that dh7, calculated using flow gages from HLU D, showed a significant difference between pre- and post-1980 values. It is not known, however, whether the difference is due to a change in the standard deviation of maximum flows (numerator of dh7 metric) or a change in the mean flow (denominator of dh7 metric), or a combination of the two. General speaking, however, mean flow values increased throughout the state (J. Ziegeweid, personal communication, February 2018). The HLU containing the Crow Wing gage (HLU BC) did not exhibit a detectable trend in dh7 values.

It is possible that the MIBI (and Ratio) response to dh7 and ml16 is an artifact of benthic macroinvertebrates being stressed by a lack of hydrologic resilience. In contrast, there are other EflowStats metrics that might measure signals of watershed storage, such as ma6, with which the dependent variables did not show strong linear relationships. A proper space-for-time analysis may help determine if watershed storage differences are indeed responsible for the flow-biology relationship between dh7/ml16 and biotic integrity seen in this sample of fifteen watersheds.

-Ratio

The Ratio scores responded most strongly to a measure of the skewness of annual maximum flows (mh19) and the number of days between a 1.67-year flow event (dh22). The 1.67-year flow event is ecologically relevant because it is considered by many hydrologists to represent the bankfull stage of a channel. The flow associated with bankfull stage, often call the “channel-forming flow,” largely determines the available in-channel habitat because this flow shapes and maintains the channel dimensions and substrate composition (Rosgen 2006).

Both mh19 and dh22 likely are heavily influenced by precipitation patterns. In a map of the flow metric results and Minnesota’s average annual precipitation for 1981-2010, there does not appear to be a spatial relationship between the precipitation data and mh19 or dh22 (Figure 12). However, there does appear to be a clustering spatial pattern, where adjacent watersheds have similar directions (<0 or >0) of mh19 values, and some dh22 values are only days apart. It is possible that other methods of summarizing precipitation may reveal a relationship with the flow metrics.

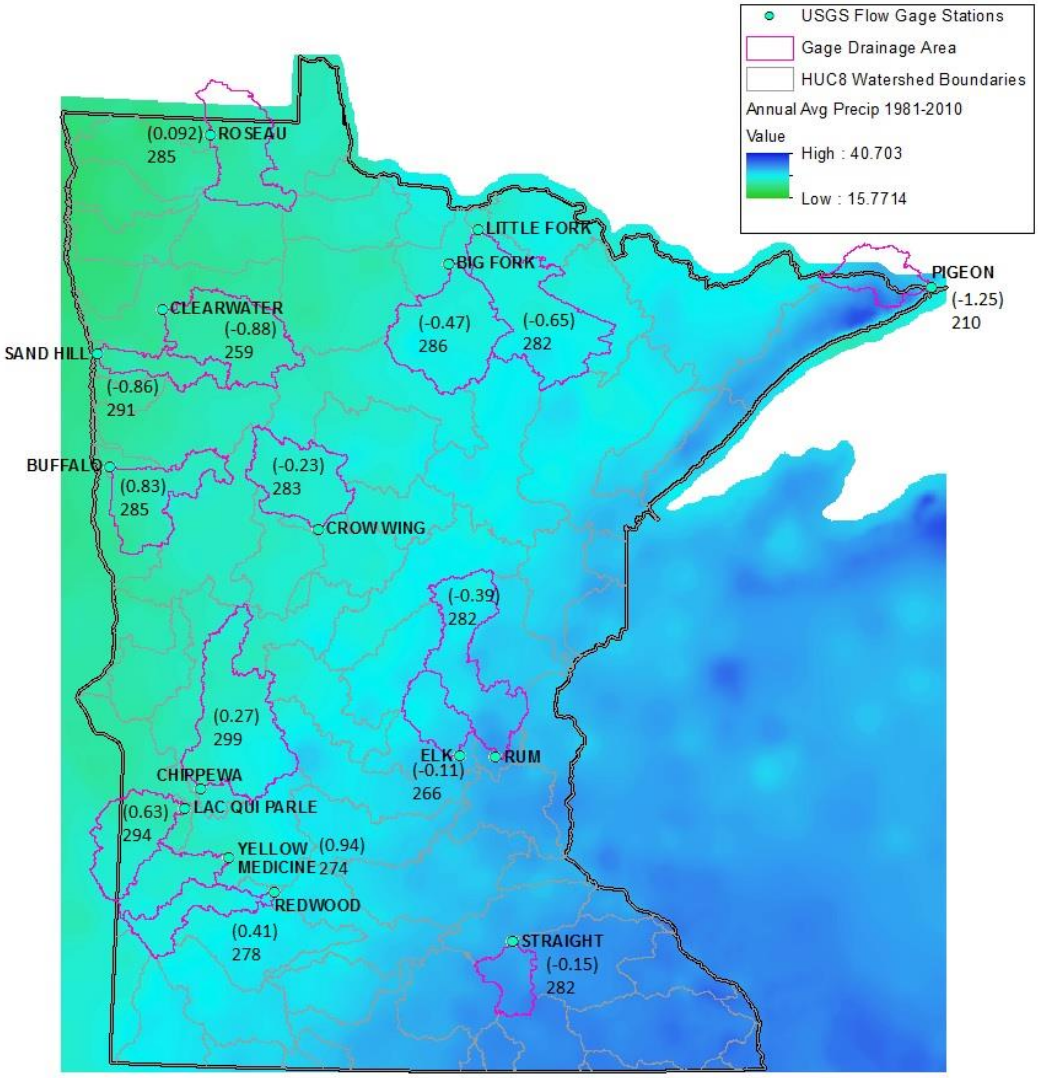

Figure 12. Precipitation map of Minnesota (data from MDNR), including the mh19 (in parenthesis) and dh22 values for each of the watersheds analyzed.

Two watersheds with quite different mh19 and dh22 scores are the Clearwater and Lac Qui Parle (LQ) River watersheds (Table 7). As evidenced by the watersheds' mean Ratio scores, the LQ contains a greater proportion of impaired benthic macroinvertebrate communities than the Clearwater. The negative mh19 value for Clearwater indicates that the distribution of annual maximum flows between WY1996-2015 is asymmetric, weighted toward the low end. In contrast, the positive mh19 value for LQ also indicates asymmetry, but weighted toward the high end. The dh22 values

indicate that, on average, the number of days between the 1.67-year flow event differs between the Clearwater and LQ by about five weeks.

Both of these gages were part of the sample of gages used by Ziegeweid *et al.* (2015) to detect a non-precipitation driven trend in streamflow in HLU A (Clearwater) and HLU D (LQ). Further, mh19, like dh7, was found to show significant differences in HLU D since pre-1980 (J. Ziegeweid, personal communication, March 2018).

In summary, the Clearwater River watershed received its channel-forming flow 35 days more often and generally exhibited more tempered high flows than the LQ from 1996-2015, despite receiving less total precipitation from 1981-2010. This could suggest that the LQ River watershed drainage system delivers precipitation to the watershed outlet more quickly than in the Clearwater, though differences in precipitation frequency and intensity must also be considered. Additionally, unpublished work by Ziegeweid and others suggests that mh19 may be a flow measure indicative of anthropogenic flow alteration.

The hydrologic pathway, from atmosphere to watershed outlet, is complex and ever changing. The results of this large watershed scale analysis may add some clarity in demonstrating that the facets controlling high flow skewness and channel-forming flow frequency seems to differ between the Clearwater and LQ, and correlate with the biotic integrity of benthic macroinvertebrates.

Table 7. Overview of streamflow and biologic metrics in two different watersheds. The number of biological data points from each MIBI class is also shown.

	Clearwater River	**Lac Qui Parle River**
mh19	-0.883	0.632
dh22	259.2	294.5
MIBI mean	50.93	32.34
Ratio mean	1.31	0.85
Class 1 – Northern forest rivers	-	-
Class 2 – Prairie forest rivers	4	8
Class 3 – Northern forest streams-riffle/run	1	-
Class 4 – Northern forest streams-glide/pool	3	-
Class 5 – Southern streams-riffle/run	14	9
Class 6 – Southern forest streams-glide/pool	3	-
Class 7 – Prairie streams-glide/pool	18	23
Class 8 – Northern coldwater streams	5	-
Class 9 – Southern coldwater streams	-	2
Total biological data points	48	42
Years sampled	2005-2015, 56% in 2014	2003-2015, 74% in 2015

HEADWATER SCALE ANALYSIS

The four main groups of macroinvertebrates analyzed at this scale are discussed in terms of the one or two most useful predictor variables used in their regression equations, as well as relevant findings from other studies. The biological metrics analyzed represent only one component of a stream's macroinvertebrate community, and some are not normalized. Therefore, the results may only be comparable to other streams of similar size, gradient, and ecotype, as the community composition is expected to change along those gradients.

-Clingers

Clinger metrics were frequently, and always positively, related to fh10, the frequency of flows greater than the median of the annual minimum flow on record. Two

of the clinger equations include dh6, which is essentially the variability in annual maximum flows (one day moving average). Dh6 and dh7 are very similar (Pearson correlation coefficients >0.90), and interestingly, MIBI and clinger metrics have relatively strong, yet oppositely signed relationships with these flow measures (Table 8). Dh7 is the variability in annual maximum three-day moving average.

Table 8. Pearson correlation coefficients between measures of benthic macroinvertebrates and the flow metrics dh6 and dh7. Data shown for two different watershed samples.

	dh6	**dh7**
MIBI (large watershed scale)	-0.828	-0.831
ClingerCh (headwater scale)	0.500	0.455
ClingerChTxPct (headwater scale)	0.487	0.486

Five of the MPCA's nine MIBI classes have one of the clinger metrics in their calculations. Further, all but one (comprising <3% of data points) of the MIBI classes represented in this headwaters dataset do include a clinger metric to calculate MIBI, suggesting that clinger is an especially relevant habitat guild in the analysis of headwater streams of the Chippewa and Yellow Medicine Rivers. It is important to note that not all metrics in the MIBI calculations are weighted evenly, and therefore the clinger richness of a class 7, for example, cannot be presumed to be as useful of a community measure as it would be in a class 5 stream. Nevertheless, it is counter to expectation that the clinger metrics and MIBI would have oppositely signed relationships to the same flow metrics.

Observed opposite relationships may result from the difference in analytical scale (large vs. headwaters) and/or diversity of ecotypes in the sample. An analysis of MIBI scores at the headwaters scale and clinger metrics at the large watershed scale would reveal valuable information on this flow-biology relationship.

It is difficult to find comparable studies regarding this dependent variable specifically; most studies use a biological metric that may include, but is not exclusive to,

clinging benthic macroinvertebrates, such as %EPT[2] or a biotic index that has a clinger component. (e.g. Phelan *et al.* 2017; Carlisle *et al.* 2012). Both fh10 and dh6 are metrics that the aforementioned (unpublished) work by Ziegeweid and others showed significant differences in HLU D from pre- to post-1980. Both the Chippewa and Yellow Medicine watersheds lie in HLU D.

In summary, the clinger habitat guild proved to be a responsive measure of the benthic macroinvertebrate community to flow metrics at the headwaters scale. Based on the aforementioned personal communications with J. Ziegeweid and others, it is possible that the flow metrics used to predict some of the clinger metrics are indicative of human-induced flow alterations. Additionally, it was shown that interpretation of these flow-biology relationships must remain at the analytical scale at which the equations were built.

-Swimmers

Two swimmer metrics showed some of the strongest linear relationships seen of all 29 metrics tested. Further, two of the best predictors of SwimmerCh were ma18 (mean July flow) and dl9 (variability of 30-day minimum moving average), both positive relationships. The positive relationship between richness of swimmer taxa and mean July flows seems opposite to what Carlisle *et al.* (2011) found in their analysis, wherein streams with diminished flow magnitudes had more benthic macroinvertebrate taxa with highly mobile life strategies (e.g. strong swimmers, fast crawlers) replacing other taxa that did not have these traits. The authors suggest this may be a result of mobile traits being favored during ecological bottlenecks of diminished flow. However, whether swimmer richness is increasing along with that of other taxa, or in place of others, was not analyzed in this study. Additionally, this relationship could be better understood if both the independent and dependent variables were normalized (e.g. by drainage area and total richness, respectively).

The fact that swimmer richness also increases with the variability of 30-day minimum flow magnitude suggests that these organisms do have a niche in finding refugia when needed.

The percent of swimmer individuals in a sample (SwimmerPct) increases with January minimum flows (ml1), but is inversely related to October mean flow (ma21).

[2] EPT = Ephemeroptera, Plecoptera, Trichoptera

Again, this relationship might be more telling if the variables were normalized by drainage area or mean flow, especially since many of the HUC12 size streams in the analysis would be completely frozen and/or yield zero flow in January. It is possible that this relationship is driven by groundwater connected streams or streams receiving thermally polluted waters. However, the HSPF model's ability to predict winter flow in headwaters streams is questionable because of the lack of calibration points available at this scale, especially in winter. Flow conditions in October in the Yellow Medicine and Chippewa watersheds are generally representative of baseflow conditions. Thus, the inverse relationship between SwimmerPct and ma21 might actually support the aforementioned findings of Carlisle *et al.* (2011), where lower flows can favor the relative abundance of mobile individuals.

In conclusion, many of the best predictors of the swimmer metrics were measures of non-normalized flow magnitudes, often correlated with drainage area. The ecological reasoning behind these relationships is inconclusive, though reasonable speculations can be made regarding the fitness of swimming benthic macroinvertebrates in variable streamflow environments. Interestingly, no swimmer metrics are used in any of the MPCA's nine MIBI classes.

CHAPTER 5. CONCLUSION

In conclusion, it was determined possible to use linear regression to predict generalized and specific measures of benthic macroinvertebrate communities using flow metrics at both the large and headwater scales. No one specific flow metric was used in many equations. The most frequently used flow metrics in the equations reported were dh6/dh7, mh19, and fh10.

The flow-biology relationships found in this study were stronger at the large watershed scale, as evidenced by the greater $\hat{R}^2$ values. However, this could result from the strength of the MIBI over the other metrics used at the headwater scale. Additionally, care must be taken not to apply the flow-biology relationships found at one scale to other analytical scales.

Because of the project design, using a variable selection method to choose the best predictor from a pool of many predictors, the flow-biology relationships identified in this study can help researchers understand which component of the flow regime may have the strongest effect on benthic macroinvertebrates in Minnesota.

-Future work

There are many ways to improve and expand upon this study, most notably by adding to the number of biological and flow metrics included, as well as normalizing all variables. Also, doing an analysis of relevant covariates, such as stream gradient or ecoregion, would help determine the level of importance of the flow measures on the differences seen in biology.

Additional flow metrics of interest are the ecosurplus/ecodeficit measures, developed by Vogel *et al.* (2007), and modifications of the RBI (Baker *et al.* 2004). Using year-round data to calculate the RBI did not prove more useful than the EflowStats metrics in this study. However, restricting the RBI to one month or one season may prove more useful. Additionally, some EflowStats metrics were not included in the headwaters analysis (those using the 1.67-year return interval) that may prove to be quite useful considering the effect this flow has on maintaining in-stream habitat. Hydrologic measures besides flow, such as percent altered watercourse, could also prove to be quite useful in predicting biotic integrity and possibly linking it with anthropogenic flow alteration.

REFERENCES CITED

Akaike, H. 1998. Information theory and an extension of the maximum likelihood principle. In: Parzen E., Tanabe K., Kitagawa G. (eds) Selected Papers of Hirotugu Akaike. Springer Series in Statistics (Perspectives in Statistics). Springer, New York, NY.

Anderson, D.J. and B. Vondracek. 1999. Insects as indicators of land use in three ecoregions in the prairie pothole region. Wetlands 19(3): 648-664.

Arthington, A.H., S.E. Bunn, N.L. Poff, R.J. Naiman. 2006. The challenge of providing environmental flow rules to sustain river ecosystems. Ecological Applications 16(4): 1311-1318.

Asmus, B., J.A. Magner, B. Vondracek, J. Perry. 2009. Physical integrity: the missing link in biological monitoring and TMDLs. Environmental Monitoring and Assessment 159: 443-463.

Bain, M.B., J.T. Finn, and H.E. Booke. 1988. Streamflow regulation and fish community structure. Ecology 69:382-392.

Baker, D.B., R.P. Richards, T.T. Loftus, and J.W. Kramer. 2004. A new flashiness index: characteristics and applications to Midwestern rivers and streams. Journal of the American Water Resources Association 40(2): 503-522.

Barnosky, A.D., N. Matzke, S. Tomiya, G.O.U. Wogan, B. Swartz, T.B. Quental, C. Marshall, J.L. McGuire, E.L. Lindsey, K.C. Maguire, B. Mersey, and E.A. Ferrer. 2011. Has the Earth's sixth mass extinction already arrived? Nature 471(7336): 51-57.

Bicknell, B.R., Imhoff, J.C., Kittle, J.L., Jr., Donigian, A.S., Jr., and Johanson, R.C. 1997. Hydrological Simulation Program--Fortran, User's manual for version 11: U.S. Environmental Protection Agency, National Exposure Research Laboratory, Athens, GA.

Bond, N.R. and B.J. Downes. 2003. The independent and interactive effects of fine sediment and flow on benthic invertebrate communities characteristic of small upland streams. Freshwater Biology 48: 455-465.

Bunn, S.E. and A.H. Arthington. 2002. Basic principles and ecological consequences of altered flow regimes for aquatic biodiversity. Environmental Management 30(4): 492-507.

Carlisle, D.M., D.M. Wolock, and M.R. Meador. 2011. Alteration of streamflow magnitudes and potential ecological consequences: a multiregional assessment. Frontiers in Ecology and the Environment 9(5): 265-270.

Ceballos, G., P.R. Ehrlich, A.D. Barnosky, A. García, R.M. Pringle, and T.M. Palmer. 2015. Accelerated modern human–induced species losses: entering the sixth mass extinction. Science Advances 1: e1400253.

Dudgeon, D., A.H. Arthington, M.O. Gessner, Kawabata, Z. Kawabata, D.J. Knowler, C. Lévêque, R.J. Naiman, A. Prieur-Richard, D. Soto, M.J. Stiassny, and C.A. Sullivan. 2006. Freshwater biodiversity: importance, threats, status and conservation challenges. Biological Reviews 81(2):163–182.

Ferrington, Jr., L.C. and N.H. Crisp. 1989. Water chemistry characteristics of receiving streams and the occurrence of *Chironomus riparius* and other Chironomidae in Kansas. Acta Biologica Debrecina Oecologica Hungarica 3: 115-126.

Gasith, A. and V.H. Resh. 1999. Streams in Mediterranean climate regions: Abiotic influences and biotic responses to predictable seasonal events. Annual Review of Ecology and Systematics 30: 51-81.

Harding, J.S., E.F. Benfield, P.V. Bolstad, G.S. Helfman, and E.B.D. Jones III. 1998. Stream biodiversity: The ghost of land use past. Proceedings of the National Academy of Sciences 95: 14843-14847.

Henriksen, J.A., J. Heasley, J.G. Kennen, and S. Nisswand. 2006. Users' manual for the hydrologic integrity assessment process software (including New Jersey assessment tools). U.S. Geological Survey Scientific Investigations Report 2006-1093.

Horwitz, R.J. 1978. Temporal variability patterns and the distributional patterns of stream fishes. Ecological Monographs 48: 307-321.

Jenks, G.F. 1967. The data model concept in statistical mapping. International Yearbook of Cartography 7: 186-190.

Jowett, I.G., and M.J. Duncan. 1990. Flow variability in New Zealand rivers and its relationship to in-stream habitat and biota. New Zealand Journal of Marine and Freshwater Research 24: 305-317.

Karr, J.R. 1981. Assessment of biotic integrity using fish communities. Fisheries (6)6: 21-27.

Karr, J.R., K.D. Fausch, P.L. Angermeier, P.R. Yant, and I.J. Schlosser. 1986. Assessing biological integrity in running waters: a method and its rationale. Illinois Natural History Survey Special Publication 5.

Keddy, P.A. 1992. Assembly and response rules: two goals for predictive community ecology. Journal of Vegetation Science 3: 157-164.

Lenat, D.R. 1984. Agriculture and stream water quality: A biological evaluation of erosion control practices. Environmental Management 8(4): 333-344.

Markos, P.D., M.D. Kaller, W.E. Kelso. 2016. Channel stability and the structure of coastal stream aquatic insect assemblages. Fundamental and Applied Limnology 188(3): 187-199.

Meffe, G.K. 1984. Effects of abiotic disturbance on coexistence of predator and prey fish species. Ecology 65: 1525-1534.

Miller, D.L., P.M. Leonard, R.M. Hughes, J.R. Karr, P.B. Moyle, L.H. Schrader, B.A. Thompson, R.A. Daniels, K.D. Fausch, G.A. Fitzhugh, J.R. Gammon, D.B. Halliwell, P.L. Angermeier, and D.J. Orth. 1988. Regional applications of an Index of Biotic Integrity for use in water resource management. Fisheries 13(5): 12-20.

Mills, J. and D. Blodgett. 2017. EflowStats: Hydrologic Indicators and Alteration Stats. R package version 5.0.0.

MPCA. 2014. Development of a macroinvertebrate-based index of biological integrity for assessment of Minnesota's rivers and streams. Minnesota Pollution Control Agency, Environmental Analysis and Outcomes Division, St. Paul, MN.

Novacek, M.J. and E. Cleland. 2001. The current biodiversity extinction event: Scenarios for mitigation and recovery. Proceedings of the National Academy of Sciences 98(10): 5466-5470.

Novak, R., J.G. Kennen, R.W. Abele, C.F. Baschon, D.M. Carlisle, L. Dlugolecki, D.M. Eignor, J.E. Flotemersch, P. Ford, J. Fowler, R. Galer, L.P. Gordon, S.E. Hansen, B. Herbold, T.E. Johnson, J.M. Johnston, C.P. Konrad, B. Leamond, and P.W. Seelbach. 2016. Final EPA-USGS technical report: protecting aquatic life from effects of hydrologic alteration. U.S. Geological Survey Scientific Investigations Report 2016–5164, U.S. Environmental Protection Agency EPA Report 822-R-156-007.

Phelan, J., T. Cuffney, L. Patterson, M. Eddy, R. Dykes, S. Pearsall, C. Goudreau, J. Mead, and F. Tarver. 2017. Fish and invertebrate flow-biology relationships to support the determination of ecological flows for North Carolina. Journal of the American Water Resources Association 53(1): 42-55.

Poff, N.L. and J.D. Allan. 1995. Functional organization of stream fish assemblages in relation to hydrological variability. Ecology 76(2): 606-627.

Poff, N.L. and J.K.H. Zimmerman. 2010. Ecological responses to altered flow regimes: a literature review to inform the science and management of environmental flows. Freshwater Biology 55: 194-205.

Poff, N.L., B.D. Richter, A.H. Arthington, S.E. Bunn, R.J. Naiman, E. Kendy, M. Acreman, C. Apse, B.P. Bledsoe, M.C. Freeman, J. Henriksen, R.B. Jacobson, J.G. Kennen, D.M. Merritt, J.H. O'Keefe, J.D. Olden, K. Rogers, R.E. Tharme, and A. Warner. 2009. The ecological limits of hydrologic alteration (ELOHA): a new framework for developing regional environmental flow standards. Freshwater Biology 55(1): 147-170.

Poff, N.L., and J.V.Ward. 1989. Implications of streamflow variability and predictability for lotic community structure: a regional analysis of streamflow patterns. Canadian Journal of Fisheries and Aquatic Sciences 46:1805-1818.

R core team. 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Raymond, K.L. and B. Vondracek. 2011. Relationships among rotational and conventional grazing systems, stream channels, and macroinvertebrates. Hydrobiologia 669: 105-117.

Rosgen, D. 2006. Watershed Assessment of River Stability and Sediment Supply (WARSSS). Wildland Hydrology, Fort Collins, CO.

Ruhí, A., J.D. Olden, and J.L. Sabo. 2016. Declining streamflow induces collapse and replacement of native fish in the American Southwest. Frontiers in Ecology and Environment 14(9): 465-472.

Strayer, D.L. and D. Dudgeon. 2010. Freshwater biodiversity conservation: recent progress and future challenges. Journal of North American Benthological Society 29(1): 344–358.

Thomas, J.A., M.G. Telfer, D.B. Roy, C.D. Preston, J.J.D. Greenwood, J. Asher, R. Fox, R.T. Clarke, and J.H. Lawton. 2004. Comparative losses of British butterflies, birds, and plants and the global extinction crisis. Science 303: 1879-1881.

Vannote, R.L., G.W. Minshall, K.W. Cummins, J.R. Sedell, and C.E. Cushing. 1980. The river continuum concept. Canadian Journal of Fisheries and Aquatic Sciences 37: 130-137.

Vogel, R.M., J. Sieber, S.A. Archfield, M.P. Smith, C.D. APSE, and A. Huber-Lee. 2007. Relations among storage, yield, and in-stream flow. Water Resources Research 43, W05403, DOI: 10.1029/2006/WR005226.

Walters, A.W. and D.M. Post. 2011. How low can you go? Impacts of a low-flow disturbance on aquatic insect communities. Ecological Applications 21(1): 163-174.

Warwick, W.F. 1980. Chironomidae (Diptera) responses to 2800 years of cultural influence: A paleolimnological study with special reference to sedimentation, eutrophication, and contamination processes. The Canadian Entomologist 112: 1193-1238.

Ziegeweid, J.R., D.L. Lorenz, C.A. Sanocki, and C.R. Czuba. 2015. Methods for estimating flow-duration curve and low-flow frequency statistics for ungagged locations on small streams in Minnesota. U.S. Geological Survey Scientific Investigations Report 2015-5170.

APPENDIX

Original stepwise algorithm outputs for all 29 biological metrics analyzed at the headwater scale.
Significance codes for all intercepts and independent variables: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

CLIMBER

```
lm(formula = Climber ~ fl3 + ra8 + mh3)
Residuals:
   Min      1Q  Median      3Q     Max
-4.3950 -1.0447 -0.3291  0.8128  4.3138
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 12.277641   3.177128   3.864 0.000363 ***
fl3         -0.337915   0.089168  -3.790 0.000455 ***
ra8         -0.050706   0.027977  -1.812 0.076755 .
mh3          0.013989   0.009648   1.450 0.154163
Residual standard error: 1.843 on 44 degrees of freedom
Multiple R-squared:  0.2856,	Adjusted R-squared:  0.2369
F-statistic: 5.863 on 3 and 44 DF,  p-value: 0.001853
```

CLIMBERCH

```
lm(formula = ClimberCh ~ fl3 + DA)
Residuals:
   Min      1Q  Median      3Q     Max
-5.2694 -0.9757 -0.1917  1.4892  4.2329
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  7.12311    1.24969   5.700 8.69e-07 ***
fl3         -0.23993    0.10540  -2.276   0.0276 *
DA           0.03129    0.02196   1.425   0.1611
Residual standard error: 1.914 on 45 degrees of freedom
Multiple R-squared:  0.2381,	Adjusted R-squared:  0.2042
F-statistic: 7.031 on 2 and 45 DF,  p-value: 0.002203
```

CLIMBERCHTXPCT

```
lm(formula = ClimberChTxPct ~ fh10 + ma19 + ra8 + mh22 + mh19 + ta1 + mh6 + mh15 + mh10 + dh19 +
dh20 + ml3 + dh18 + ml9 + ma17)
Residuals:
   Min      1Q  Median      3Q     Max
-8.8220 -2.3842  0.7991  2.3393  7.1850
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 107.61085   15.30426   7.031 5.69e-08 ***
fh10         -8.13335    1.22793  -6.624 1.81e-07 ***
ma19         -0.75585    0.42952  -1.760  0.08801 .
ra8          -0.61514    0.08808  -6.984 6.50e-08 ***
mh22         -0.41839    0.07606  -5.501 4.63e-06 ***
mh19         -7.28068    2.54596  -2.860  0.00741 **
ta1          41.09769   25.17124   1.633  0.11233
mh6          -0.33707    0.11365  -2.966  0.00566 **
mh15          0.67815    0.14728   4.604 6.25e-05 ***
mh10         -0.14362    0.08329  -1.724  0.09431 .
dh19          2.75286    1.00076   2.751  0.00970 **
dh20         -1.32028    0.51776  -2.550  0.01576 *
ml3           1.97177    0.90999   2.167  0.03780 *
dh18          1.98060    0.98314   2.015  0.05242 .
```

```
ml9          -2.17174    1.27754  -1.700  0.09884 .
ma17          0.81985    0.55610   1.474  0.15017
Residual standard error: 4.148 on 32 degrees of freedom
Multiple R-squared:  0.8195,            Adjusted R-squared:  0.7349
F-statistic: 9.685 on 15 and 32 DF,  p-value: 5.2e-08
```

CLIMBERPCT

```
lm(formula = ClimberPct ~ dh7 + th2 + dh15 + dh20 + ma11 + ma44 + mh7 + fh7 + dl9 + mh16 + dl17 + fh6
+ dh6 + ma39 + ml2 + ma20 + ra5 + fh5 + dh21 + mh1 + ra4 + ma3 + dh14 + dh19 + fh10 + ta1 + mh5 +
mh15 + dh16 + mh4 + fh2 + ra2 + mh17 + ml18 + dh18 + dl14 + ml8 + tl1 + mh13 + ml10 + ml22 + ma36 +
dl15 + ml7 + ma14 + ra1 + DA)
Warning messages:
1: attempting model selection on an essentially perfect fit is nonsense
2: attempting model selection on an essentially perfect fit is nonsense
```

CLINGER

```
lm(formula = Clinger ~ fh10 + ml12 + ra8 + ml2 + ma11 + dh14 + mh8 + mh9 + mh11)
Residuals:
   Min      1Q  Median      3Q     Max
-3.3585 -1.1151  0.1618  1.0485  4.9386
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -10.37064    4.11803  -2.518  0.01612 *
fh10          3.05787    0.52528   5.821 1.00e-06 ***
ml12          1.77817    0.43999   4.041  0.00025 ***
ra8           0.13937    0.02982   4.673 3.66e-05 ***
ml2          -2.79414    0.79360  -3.521  0.00114 **
ma11          0.09326    0.03601   2.590  0.01354 *
dh14         -3.03590    0.68852  -4.409 8.24e-05 ***
mh8           0.04884    0.01731   2.822  0.00755 **
mh9           0.03820    0.01519   2.515  0.01625 *
mh11         -0.12682    0.06881  -1.843  0.07311 .
Residual standard error: 1.854 on 38 degrees of freedom
Multiple R-squared:  0.7219,            Adjusted R-squared:  0.656
F-statistic: 10.96 on 9 and 38 DF,  p-value: 3.573e-08
```

CLINGERCH

```
lm(formula = ClingerCh ~ dh6 + ra8 + fh10 + ml11 + ml9 + dh15 + mh6 + ml5 + mh7 + dl17 + mh19 + ma11)

Residuals:
   Min      1Q  Median      3Q     Max
-4.5116 -1.2552  0.2173  0.9896  3.6033

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -34.20269    8.02683  -4.261 0.000146 ***
dh6           0.04445    0.02043   2.175 0.036472 *
ra8           0.15997    0.04302   3.719 0.000698 ***
fh10          3.53660    0.60209   5.874 1.13e-06 ***
ml11          2.77392    0.54266   5.112 1.14e-05 ***
ml9          -0.85833    0.66631  -1.288 0.206137
dh15          0.45332    0.14597   3.106 0.003752 **
mh6           0.07951    0.02166   3.671 0.000800 ***
ml5          -0.89068    0.21494  -4.144 0.000205 ***
mh7          -0.04498    0.01519  -2.962 0.005466 **
dl17         -0.04428    0.02160  -2.050 0.047926 *
mh19          2.46020    1.32923   1.851 0.072647 .
ma11          0.06947    0.03790   1.833 0.075342 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

```
Residual standard error: 2.011 on 35 degrees of freedom
Multiple R-squared:  0.7771,	Adjusted R-squared:  0.7007
F-statistic: 10.17 on 12 and 35 DF,  p-value: 3.56e-08
```

CLINGERCHTXPCT

```
lm(formula = ClingerChTxPct ~ dh6 + fh5 + fh10 + ma11 + mh8 + ml3 + dl13 + fl1 + mh6 + fh2 + dl14 +
mh14 + ma20 + ra8 + ma18 + ml10)
Residuals:
    Min      1Q   Median      3Q     Max
-11.0901  -2.5048  -0.0285   1.7927  18.3732
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -39.52930   25.02795  -1.579 0.124394
dh6            0.18786    0.05299   3.545 0.001269 **
fh5           -8.62302    1.53181  -5.629 3.54e-06 ***
fh10           3.91155    2.65644   1.472 0.150973
ma11           0.20724    0.11218   1.847 0.074260 .
mh8            0.34003    0.06280   5.414 6.55e-06 ***
ml3           -6.47394    1.25706  -5.150 1.40e-05 ***
dl13         127.75520   28.29364   4.515 8.57e-05 ***
fl1           10.83996    2.04349   5.305 8.97e-06 ***
mh6            0.20148    0.08883   2.268 0.030440 *
fh2            0.43635    0.12764   3.419 0.001781 **
dl14        -166.08204   40.01784  -4.150 0.000240 ***
mh14          -0.38551    0.08665  -4.449 0.000103 ***
ma20           2.44988    0.76214   3.214 0.003048 **
ra8            0.31617    0.15669   2.018 0.052343 .
ma18          -1.35170    0.58861  -2.296 0.028569 *
ml10           3.35901    2.01636   1.666 0.105812
Residual standard error: 5.753 on 31 degrees of freedom
Multiple R-squared:  0.819,	Adjusted R-squared:  0.7255
F-statistic: 8.764 on 16 and 31 DF,  p-value: 1.746e-07
```

CLINGERPCT

```
lm(formula = ClingerPct ~ fh10 + mh8 + mh19 + ta2 + DA + mh3 + mh22 + ml22 + dh21)
Residuals:
   Min      1Q  Median      3Q     Max
-23.140  -4.641  -2.082   6.035  27.597
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  46.35372   21.04089   2.203 0.033729 *
fh10          9.93702    2.48499   3.999 0.000283 ***
mh8           0.48657    0.09680   5.026 1.22e-05 ***
mh19         19.12345    3.98337   4.801 2.47e-05 ***
ta2          -1.70639    0.45807  -3.725 0.000632 ***
DA            0.68135    0.19092   3.569 0.000991 ***
mh3          -0.28727    0.09060  -3.171 0.003004 **
mh22          0.15046    0.04341   3.466 0.001327 **
ml22        209.31788   89.52790   2.338 0.024751 *
dh21         -0.74310    0.25160  -2.954 0.005365 **
Residual standard error: 10.68 on 38 degrees of freedom
Multiple R-squared:  0.7115,	Adjusted R-squared:  0.6432
F-statistic: 10.41 on 9 and 38 DF,  p-value: 6.848e-08
```

COLLECTORFILTERER

```
lm(formula = Collectorfilterer ~ fh10 + ma19 + ra8 + ma13 + mh18 + ml12)
Residuals:
   Min      1Q  Median      3Q     Max
-1.6366 -0.6760 -0.0663  0.4610  2.5497
```

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -3.30475    2.35322  -1.404 0.167745
fh10         0.89537    0.21212   4.221 0.000131 ***
ma19         0.15455    0.07980   1.937 0.059682 .
ra8          0.04187    0.01590   2.634 0.011853 *
ma13        -0.58757    0.17912  -3.280 0.002122 **
mh18        -0.09149    0.04639  -1.972 0.055359 .
ml12         0.27912    0.16784   1.663 0.103945
Residual standard error: 1.007 on 41 degrees of freedom
Multiple R-squared:  0.4615,	Adjusted R-squared:  0.3827
F-statistic: 5.855 on 6 and 41 DF,  p-value: 0.0001773
```

COLLECTORFILTERERCH

```
lm(formula = CollectorfiltererCh ~ ra4 + dh16 + ml13 + ml18 + dl13)
Residuals:
   Min      1Q  Median      3Q     Max
-2.6852 -0.7978  0.0234  0.9250  3.6292
Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.3826601  1.8992115   0.201  0.84129
ra4          0.0008469  0.0004772   1.775  0.08321 .
dh16         0.0142948  0.0076350   1.872  0.06814 .
ml13        -0.0178323  0.0079632  -2.239  0.03049 *
ml18         0.0119785  0.0039236   3.053  0.00392 **
dl13        10.0346738  5.2197327   1.922  0.06135 .
Residual standard error: 1.336 on 42 degrees of freedom
Multiple R-squared:  0.4219,	Adjusted R-squared:  0.3531
F-statistic:  6.13 on 5 and 42 DF,  p-value: 0.0002384
```

COLLECTORGATHERER

```
lm(formula = Collectorgatherer ~ dh6 + mh1 + ma19 + dl17 + ma13 + ma21 + dl10 + th2 + ma20 + mh9 +
mh2)
Residuals:
    Min       1Q   Median       3Q      Max
-2.74513 -0.44257  0.04453  0.44978  1.73143
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  8.037589   1.582505   5.079 1.18e-05 ***
dh6          0.026425   0.005384   4.908 1.99e-05 ***
mh1          1.272785   0.238090   5.346 5.20e-06 ***
ma19        -0.644311   0.107648  -5.985 7.27e-07 ***
dl17        -0.034195   0.013277  -2.576  0.01426 *
ma13        -1.489710   0.645485  -2.308  0.02686 *
ma21         0.904336   0.179795   5.030 1.37e-05 ***
dl10         0.018047   0.009125   1.978  0.05564 .
th2         -0.099481   0.028161  -3.533  0.00115 **
ma20        -0.668426   0.198569  -3.366  0.00182 **
mh9          0.038074   0.015454   2.464  0.01867 *
mh2          0.396146   0.307918   1.287  0.20647
Residual standard error: 1.024 on 36 degrees of freedom
Multiple R-squared:  0.6824,	Adjusted R-squared:  0.5853
F-statistic: 7.031 on 11 and 36 DF,  p-value: 3.475e-06
```

COLLECTORGATHERERCH

```
lm(formula = Collector.gathererCh ~ ml22 + ra8 + ra9)
Residuals:
   Min      1Q  Median      3Q     Max
-8.2057 -2.0431  0.1189  1.6298  6.1653
Coefficients:
```

```
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -3.77847    6.23607  -0.606   0.5477
ml22        32.92340   12.28404   2.680   0.0103 *
ra8          0.09716    0.04647   2.091   0.0424 *
ra9          0.14073    0.09358   1.504   0.1398
Residual standard error: 2.938 on 44 degrees of freedom
Multiple R-squared:  0.2219,	Adjusted R-squared:  0.1689
F-statistic: 4.183 on 3 and 44 DF,  p-value: 0.01088
```

DOMONECHPCT

```
lm(formula = DomOneCHPct ~ ml18 + ra8 + mh16 + ra6 + ma43 + fl2 + dl9 + dh15 + fh5 + mh25)
Residuals:
    Min      1Q  Median      3Q     Max
-15.3481  -4.2332   0.0186   4.0918  15.2949
Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -103.51272   46.48537  -2.227 0.032131 *
ml18          -0.35285    0.05218  -6.763 5.88e-08 ***
ra8           -0.40604    0.19667  -2.065 0.046028 *
mh16          -4.85070    0.93617  -5.181 8.04e-06 ***
ra6          618.44440  105.98204   5.835 1.05e-06 ***
ma43           9.90174    5.69134   1.740 0.090209 .
fl2            0.89996    0.22429   4.012 0.000281 ***
dl9            0.24956    0.07418   3.364 0.001798 **
dh15           2.77445    0.86679   3.201 0.002813 **
fh5            4.01989    1.66626   2.413 0.020916 *
mh25           0.97328    0.46567   2.090 0.043544 *
Residual standard error: 8.27 on 37 degrees of freedom
Multiple R-squared:  0.6922,	Adjusted R-squared:  0.609
F-statistic: 8.321 on 10 and 37 DF,  p-value: 6.923e-07
```

DOMTWOCHPCT

```
lm(formula = DomTwoCHPct ~ ra8 + ml18 + dh14 + mh16 + ra6 + mh15 + dl17 + dh15 + mh14 + dl9)
Residuals:
    Min      1Q  Median      3Q     Max
-21.2563  -5.1920  -0.1165   5.8428  19.2948
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -17.46388   37.73964  -0.463 0.646256
ra8          -0.32869    0.17138  -1.918 0.062864 .
ml18         -0.24261    0.04401  -5.513 2.87e-06 ***
dh14         13.43346    4.74731   2.830 0.007485 **
mh16         -8.29875    2.13193  -3.893 0.000399 ***
ra6         500.76117   96.93183   5.166 8.43e-06 ***
mh15         0.62776    0.24540   2.558 0.014753 *
dl17         0.28190    0.08944   3.152 0.003212 **
dh15         2.03543    0.66021   3.083 0.003860 **
mh14         -0.22007    0.09734  -2.261 0.029738 *
dl9          0.10913    0.07097   1.538 0.132636
Residual standard error: 9.483 on 37 degrees of freedom
Multiple R-squared:  0.6611,	Adjusted R-squared:  0.5694
F-statistic: 7.216 on 10 and 37 DF,  p-value: 3.477e-06
```

DOMTHREECHPCT

```
lm(formula = DomThreeCHPct ~ ml18 + ra6 + ra7 + dl17 + mh16 + mh15 + ml6 + ra9 + fh10 + dl4 + mh14 +
ml9 + ma3 + ma19 + ma40 + dh19)
Residuals:
    Min      1Q  Median      3Q     Max
-11.9193  -4.1728   0.3373   3.2772  21.1661
Coefficients:
```

```
            Estimate Std. Error t value Pr(>|t|)
(Intercept)    9.76404   28.90869   0.338 0.737825
ml18          -0.22226    0.04889  -4.546 7.85e-05 ***
ra6          266.05712  125.24014   2.124 0.041718 *
ra7         -299.86449  177.38167  -1.691 0.100964
dl17           0.62603    0.10884   5.752 2.49e-06 ***
mh16         -11.68312    2.47015  -4.730 4.65e-05 ***
mh15           1.77473    0.51233   3.464 0.001578 **
ml6            7.13470    1.42184   5.018 2.04e-05 ***
ra9           -0.81967    0.65464  -1.252 0.219902
fh10           7.95545    3.34872   2.376 0.023882 *
dl4          -37.73464    9.36046  -4.031 0.000335 ***
mh14          -0.32562    0.13469  -2.418 0.021697 *
ml9           -4.47469    2.19890  -2.035 0.050482 .
ma3            0.29176    0.14715   1.983 0.056311 .
ma19           1.35825    0.90772   1.496 0.144678
ma40         -16.07555    9.58970  -1.676 0.103731
dh19           1.66823    1.34635   1.239 0.224619
Residual standard error: 8.026 on 31 degrees of freedom
Multiple R-squared:  0.7705,	Adjusted R-squared:  0.6521
F-statistic: 6.506 on 16 and 31 DF,  p-value: 4.649e-06
```

INSECTPCT

```
lm(formula = InsectPct ~ th2 + fh10 + mh19 + ra8 + dh7 + mh18 + ma14 + ma8 + mh20)
Residuals:
    Min      1Q   Median      3Q     Max
-26.1231  -7.0138  -0.6122   8.2118  23.1587
Coefficients:
           Estimate Std. Error t value Pr(>|t|)
(Intercept) -63.4957    30.1097  -2.109 0.041608 *
th2          -1.6880     0.2463  -6.854 3.88e-08 ***
fh10         13.3377     2.6814   4.974 1.44e-05 ***
mh19        -46.6339     9.4045  -4.959 1.51e-05 ***
ra8          0.7660     0.2083   3.678 0.000725 ***
dh7          2.2572     0.3640   6.200 3.03e-07 ***
mh18         -5.1767     1.0630  -4.870 1.99e-05 ***
ma14         -0.7570     0.3786  -1.999 0.052752 .
ma8          -0.2680     0.1056  -2.538 0.015374 *
mh20          0.8051     0.4949   1.627 0.112025
Residual standard error: 12.59 on 38 degrees of freedom
Multiple R-squared:  0.7176,	Adjusted R-squared:  0.6507
F-statistic: 10.73 on 9 and 38 DF,  p-value: 4.679e-08
```

INSECTTXPCT

```
lm(formula = InsectTxPct ~ th2 + dh7 + ra5 + mh6 + fh2)
Residuals:
    Min      1Q   Median      3Q     Max
-14.0121  -4.0624  -0.5648   3.8597  10.9505
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  10.92903   20.50365   0.533 0.596823
th2          -0.44386    0.10848  -4.092 0.000190 ***
dh7           0.27719    0.06254   4.432 6.57e-05 ***
ra5         171.48785   46.63610   3.677 0.000664 ***
mh6           0.09376    0.03954   2.371 0.022411 *
fh2           0.17153    0.09719   1.765 0.084839 .
Residual standard error: 5.895 on 42 degrees of freedom
Multiple R-squared:  0.4747,	Adjusted R-squared:  0.4122
F-statistic: 7.592 on 5 and 42 DF,  p-value: 3.734e-05
```

LONGLIVEDPCT

```
lm(formula = LongLivedPct ~ dh6 + ta2 + ml18 + dl17 + ta1 + dh10 + mh13 + dh19)
Residuals:
   Min      1Q  Median      3Q     Max
-7.2221 -1.6633 -0.6118  1.1040 16.0013
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  64.83744   12.60486   5.144 7.95e-06 ***
dh6           0.18600    0.04416   4.212 0.000145 ***
ta2          -2.67061    0.56884  -4.695 3.26e-05 ***
ml18          0.04214    0.01017   4.144 0.000178 ***
dl17         -0.10990    0.03894  -2.822 0.007469 **
ta1         228.44887   53.83989   4.243 0.000131 ***
dh10         -0.22374    0.06447  -3.470 0.001284 **
mh13         -0.04873    0.02278  -2.140 0.038704 *
dh19          0.62751    0.43466   1.444 0.156815
Residual standard error: 3.746 on 39 degrees of freedom
Multiple R-squared:  0.5202,            Adjusted R-squared:  0.4218
F-statistic: 5.286 on 8 and 39 DF,  p-value: 0.000161
```

LONGLIVEDCHTXPCT

```
lm(formula = LongLivedChTxPct ~ dh7 + dl17 + mh5 + mh19 + mh3 + mh14 + ta1 + mh21 + ra8 + mh22 +
mh7 + fh10)
Residuals:
   Min      1Q  Median      3Q     Max
-6.4574 -1.2651  0.0918  1.3373  3.9417
Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -38.518249  10.645216  -3.618 0.000927 ***
dh7           0.370110   0.057491   6.438 2.06e-07 ***
dl17         -0.076810   0.025462  -3.017 0.004739 **
mh5           0.152741   0.028202   5.416 4.55e-06 ***
mh19        -11.073554   2.269880  -4.878 2.31e-05 ***
mh3          -0.068062   0.021591  -3.152 0.003315 **
mh14         -0.144343   0.035547  -4.061 0.000262 ***
ta1         -28.077468   9.438522  -2.975 0.005284 **
mh21          0.120160   0.038933   3.086 0.003947 **
ra8          0.140154   0.057609   2.433 0.020231 *
mh22         -0.051223   0.022450  -2.282 0.028702 *
mh7          0.016383   0.009238   1.774 0.084846 .
fh10         0.821549   0.573803   1.432 0.161083
Residual standard error: 2.471 on 35 degrees of freedom
Multiple R-squared:  0.6311,            Adjusted R-squared:  0.5046
F-statistic:  4.99 on 12 and 35 DF,  p-value: 9.27e-05
```

NONINSECTPCT

```
lm(formula = NonInsectPct ~ th2 + ma19 + ml6 + mh18 + dh7 + mh19 + mh17 + dl16 + mh16 + fh2 + ra5 +
ma11 + tl2 + mh26 + dh16 + dh1 + mh5 + ml17 + fl3 + ml3 + ml15 + mh9 + ma38 + ra6 + mh14 + ma16 +
fh1 + tl1 + dh21 + fh7 + mh13 + ml5 + fl1 + dh10 + ma14 + ma43 + ma37 + fh6 + RBI + ma36 + ml8 + dl3 +
mh23 + mh4 + dl1 + ma41 + DA, data = data)
Residuals:
ALL 48 residuals are 0: no residual degrees of freedom!
Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  5.484e+02         NA      NA       NA
th2         -1.283e-01         NA      NA       NA
ma19         7.220e-01         NA      NA       NA
ml6         -2.688e+01         NA      NA       NA
mh18         5.945e+00         NA      NA       NA
dh7         -1.300e+00         NA      NA       NA
```

```
mh19      -5.999e+00      NA     NA     NA
mh17       2.874e+01      NA     NA     NA
dl16       4.669e+00      NA     NA     NA
mh16      -1.861e+01      NA     NA     NA
fh2       -1.349e+00      NA     NA     NA
ra5       -1.800e+03      NA     NA     NA
ma11       7.743e-01      NA     NA     NA
tl2       -3.219e+00      NA     NA     NA
mh26       6.955e+00      NA     NA     NA
dh16       1.123e-01      NA     NA     NA
dh1       -2.664e-01      NA     NA     NA
mh5        5.463e-01      NA     NA     NA
ml17       3.941e+03      NA     NA     NA
fl3        1.162e+01      NA     NA     NA
ml3        1.787e+01      NA     NA     NA
ml15      -4.650e+03      NA     NA     NA
mh9       -1.283e+00      NA     NA     NA
ma38      -1.971e+01      NA     NA     NA
ra6       -2.136e+02      NA     NA     NA
mh14       6.091e-01      NA     NA     NA
ma16       5.385e+00      NA     NA     NA
fh1        1.076e+01      NA     NA     NA
tl1       -3.438e-01      NA     NA     NA
dh21       8.869e-01      NA     NA     NA
fh7        2.498e+01      NA     NA     NA
mh13      -5.960e-01      NA     NA     NA
ml5        1.099e+01      NA     NA     NA
fl1        7.952e-01      NA     NA     NA
dh10      -4.667e-01      NA     NA     NA
ma14      -1.006e+00      NA     NA     NA
ma43       5.187e+01      NA     NA     NA
ma37       2.705e+01      NA     NA     NA
fh6       -5.099e+00      NA     NA     NA
RBI        2.263e+02      NA     NA     NA
ma36      -4.098e-01      NA     NA     NA
ml8        4.975e+00      NA     NA     NA
dl3       -5.642e+00      NA     NA     NA
mh23      -3.726e-02      NA     NA     NA
mh4        7.751e-03      NA     NA     NA
dl1        7.983e-01      NA     NA     NA
ma41      -3.042e-02      NA     NA     NA
DA        -7.630e-05      NA     NA     NA
Residual standard error: NaN on 0 degrees of freedom
Multiple R-squared:      1,   Adjusted R-squared:    NaN
F-statistic:   NaN on 47 and 0 DF,  p-value: NA
```

NONINSECTTXPCT

```
lm(formula = NonInsectTxPct ~ th2 + dh7 + ra5 + mh10 + dl9 + mh17 + fh2 + ma43 + ma11 + dh1 + dl13 +
dh20 + ma8 + fl3)
Residuals:
   Min      1Q  Median      3Q     Max
-7.9842 -2.5925  0.3704  2.6568  8.1384
Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  9.054e+01  1.716e+01   5.275 8.20e-06 ***
th2          8.378e-01  1.154e-01   7.262 2.48e-08 ***
dh7         -3.696e-01  5.375e-02  -6.875 7.53e-08 ***
ra5         -2.792e+02  5.443e+01  -5.131 1.26e-05 ***
mh10        -4.398e-01  8.460e-02  -5.198 1.03e-05 ***
dl9         -5.508e-02  3.464e-02  -1.590 0.121317
```

```
mh17        7.614e+00  3.349e+00   2.274 0.029629 *
fh2        -2.787e-01  9.266e-02  -3.008 0.005000 **
ma43        1.550e+01  3.527e+00   4.393 0.000109 ***
ma11       -2.073e-01  7.204e-02  -2.877 0.006984 **
dh1         2.171e-02  8.408e-03   2.582 0.014459 *
dl13        4.875e+01  2.414e+01   2.019 0.051647 .
dh20       -6.476e-01  2.773e-01  -2.335 0.025769 *
ma8        -1.078e-01  7.095e-02  -1.519 0.138356
fl3        -1.516e+00  1.035e+00  -1.465 0.152468
Residual standard error: 4.04 on 33 degrees of freedom
Multiple R-squared:  0.7904,	Adjusted R-squared:  0.7015
F-statistic: 8.889 on 14 and 33 DF,  p-value: 1.478e-07
```

PREDATORPCT

```
lm(formula = PredatorPct ~ fh2 + fh1 + ml18 + ma43 + ra8 + ta2)
Residuals:
    Min      1Q   Median      3Q     Max
-12.8997  -3.8143  -0.6783   2.7211  18.8177
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 18.62037   18.86316   0.987  0.32937
fh2         -0.37718    0.11370  -3.317  0.00191 **
fh1         -2.70621    0.91439  -2.960  0.00510 **
ml18         0.04297    0.01590   2.703  0.00996 **
ma43        -8.46124    3.68549  -2.296  0.02687 *
ra8          0.19740    0.11646   1.695  0.09767 .
ta2          0.33721    0.22522   1.497  0.14200
Residual standard error: 6.824 on 41 degrees of freedom
Multiple R-squared:  0.3884,	Adjusted R-squared:  0.2989
F-statistic: 4.339 on 6 and 41 DF,  p-value: 0.001771
```

PREDATORCHTXPCT

```
lm(formula = PredatorChTxPct ~ dh6 + fh2 + fh9 + th1 + fh5 + dl16 + mh6 + ml3 + mh10 + dh19 + mh1 + fl2)
Residuals:
   Min     1Q  Median     3Q    Max
-9.1165 -1.9781  0.1892  1.5036  7.3668
Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 107.83496   25.26640   4.268 0.000143 ***
dh6          -0.13280    0.03017  -4.401 9.64e-05 ***
fh2          -0.57658    0.09656  -5.971 8.43e-07 ***
fh9          -7.79995    1.50769  -5.173 9.49e-06 ***
th1          -0.28999    0.11471  -2.528 0.016140 *
fh5           4.44218    0.86507   5.135 1.07e-05 ***
dl16          0.63127    0.36196   1.744 0.089930 .
mh6          -0.14785    0.04475  -3.304 0.002205 **
ml3           3.23524    1.04127   3.107 0.003738 **
mh10         -0.08530    0.05527  -1.543 0.131732
dh19          0.88746    0.61554   1.442 0.158260
mh1          -1.15851    0.66316  -1.747 0.089414 .
fl2           0.17844    0.11977   1.490 0.145203
Residual standard error: 3.984 on 35 degrees of freedom
Multiple R-squared:  0.6602,	Adjusted R-squared:  0.5437
F-statistic: 5.667 on 12 and 35 DF,  p-value: 2.7e-05
```

SPRAWLER

```
lm(formula = Sprawler ~ ma11 + mh5 + mh6 + ml1 + ml9 + ma21 + ra5 + fl1 + DA + mh8 + ta2 + dl17 + th1
+ ra8)
Residuals:
    Min      1Q   Median      3Q     Max
```

```
-1.54796 -0.30585  0.09135  0.27658  0.91599
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -9.244020   2.495960  -3.704 0.000774 ***
ma11        -0.056374   0.009435  -5.975 1.04e-06 ***
mh5          0.007014   0.004984   1.407 0.168665
mh6         -0.012010   0.006047  -1.986 0.055372 .
ml1          0.445728   0.106676   4.178 0.000202 ***
ml9         -0.730240   0.139827  -5.222 9.57e-06 ***
ma21         0.304320   0.082249   3.700 0.000782 ***
ra5         52.505468   8.829967   5.946 1.13e-06 ***
fl1         -0.544316   0.141590  -3.844 0.000522 ***
DA           0.036575   0.008357   4.376 0.000114 ***
mh8         -0.023890   0.005112  -4.673 4.81e-05 ***
ta2          0.054924   0.020496   2.680 0.011405 *
dl17        -0.013453   0.004936  -2.725 0.010200 *
th1          0.014804   0.009585   1.545 0.131988
ra8         -0.017699   0.012891  -1.373 0.179026
Residual standard error: 0.5219 on 33 degrees of freedom
Multiple R-squared:  0.7986,            Adjusted R-squared:  0.7132
F-statistic: 9.348 on 14 and 33 DF,  p-value: 8.093e-08
```

SPRAWLERCH

```
lm(formula = SprawlerCh ~ ma11 + mh7 + ra5)
Residuals:
   Min      1Q  Median      3Q     Max
-5.7340 -1.1007 -0.2168  1.1620  5.5345
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.115097   3.849401   0.549   0.5855
ma11        -0.103614   0.037668  -2.751   0.0086 **
mh7          0.014577   0.005838   2.497   0.0163 *
ra5         23.654477  13.677947   1.729   0.0908 .
Residual standard error: 2.333 on 44 degrees of freedom
Multiple R-squared:  0.2361,            Adjusted R-squared:  0.184
F-statistic: 4.532 on 3 and 44 DF,  p-value: 0.007459
```

SWIMMER

```
lm(formula = Swimmer ~ dl9 + dh15 + mh10 + ml13 + fh5)
Residuals:
   Min      1Q  Median      3Q     Max
-3.3206 -0.6167 -0.0410  0.7468  2.4003
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -4.176775   2.210845  -1.889 0.065779 .
dl9          0.024703   0.004886   5.056 8.89e-06 ***
dh15         0.232536   0.078361   2.968 0.004939 **
mh10         0.047588   0.012242   3.887 0.000354 ***
ml13        -0.014789   0.004855  -3.046 0.003996 **
fh5          0.222629   0.122318   1.820 0.075878 .
Residual standard error: 1.052 on 42 degrees of freedom
Multiple R-squared:  0.4341,            Adjusted R-squared:  0.3668
F-statistic: 6.445 on 5 and 42 DF,  p-value: 0.000158
```

SWIMMERCH

```
lm(formula = SwimmerCh ~ ma18 + dl9 + th1 + dh21)
Residuals:
   Min      1Q  Median      3Q     Max
-2.9392 -0.9473  0.1231  0.8644  2.7223
Coefficients:
```

```
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  4.854340   2.156627   2.251 0.029559 *
ma18         0.104384   0.028888   3.613 0.000786 ***
dl9          0.012534   0.004856   2.581 0.013341 *
th1         -0.049450   0.020783  -2.379 0.021852 *
dh21         0.040561   0.026256   1.545 0.129717
Residual standard error: 1.327 on 43 degrees of freedom
Multiple R-squared:  0.4547,            Adjusted R-squared:  0.404
F-statistic: 8.963 on 4 and 43 DF,  p-value: 2.347e-05
```

SWIMMERPCT

```
lm(formula = SwimmerPct ~ ml1 + ma21)
Residuals:
   Min     1Q Median     3Q    Max
-8.620 -3.816 -1.854  2.917 16.682
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   4.3933     1.7290   2.541   0.0146 *
ml1           4.6108     1.0114   4.559 3.93e-05 ***
ma21         -0.9629     0.3751  -2.567   0.0136 *
Residual standard error: 6.122 on 45 degrees of freedom
Multiple R-squared:  0.4185,            Adjusted R-squared:  0.3927
F-statistic:  16.2 on 2 and 45 DF,  p-value: 5.03e-06
```