Virginia Commonwealth University
VCU Scholars Compass

Theses and Dissertations

Graduate School

2014

Benthic Macroinvertebrate Subsampling Effort and Taxonomic Resolution for Bioassessments of Streams in the James River Watershed of Virginia

Laurel Williams
Virginia Commonwealth University

Follow this and additional works at: http://scholarscompass.vcu.edu/etd

Part of the Biology Commons

© The Author

BENTHIC MACROINVERTEBRATE SUBSAMPLING EFFORT

AND TAXONOMIC RESOLUTION FOR BIOASSESSMENTS OF

STREAMS IN THE JAMES RIVER WATERSHED OF VIRGINIA

A Thesis submitted in partial fulfillment of the requirements for the degree of Master of Science at Virginia Commonwealth University.

by

LAUREL CARY WILLIAMS
BACHELOR OF SCIENCE, VIRGINIA COMMONWEALTH UNIVERSITY, 2004
MASTER OF SCIENCE, VIRGINIA COMMONWEALTH UNIVERSITY, 2014

Director: LEONARD A. SMOCK, PHD
PROFESSOR, DEPARTMENT OF BIOLOGY

Virginia Commonwealth University
Richmond, Virginia
May 2014

Acknowledgements

I am honored to acknowledge the people without whom this thesis would not be possible. I would like to thank Dr. Leonard Smock, whose belief in me made this thesis a reality. I would like to thank Drew Garey, my lab mentor and friend, for his indefatigable patience and willingness to help. I would also like to thank my family, most especially my two beautiful daughters, for their patience and their unwavering support. I would also like to thank my committee members, Dr. Donald Young, Dr. Steven McIninch and Dr. Daniel McGarvey. Lastly, I would like to thank the members of the Aquatics Lab for sharing their space, as well as their stories.

Table of Contents

List of Tables

List of Figures

Page

vi

Abstract

BENTHIC MACROINVERTEBRATE SUBSAMPLING EFFORT AND TAXONOMIC
RESOLUTION FOR BIOASSESSMENTS OF STREAMS IN THE JAMES RIVER
WATERSHED OF VIRGINIA

By Laurel Cary Williams, B.S.

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science
in Biology at Virginia Commonwealth University.

Virginia Commonwealth University, 2014

Major Advisor: Leonard A. Smock, Ph.D.
Professor, Department of Biology

Benthic macroinvertebrate diversity influences stream food web dynamics, nutrient
cycling and material exchange between the benthos and the water column. Stream bioassessment
has moved to the forefront of water quality monitoring in terms of benthic macroinvertebrate
diversity in the recent past. The objectives of this study were to determine optimum subsample
size and level of taxonomic resolution necessary to accurately and precisely describe
macroinvertebrate diversity in streams flowing in the Piedmont province of the James River
watershed in Virginia. Forty-nine sampling sites were selected from streams within the Piedmont
Physiographic Province of the James River watershed. Ten sites were randomly selected to have
all macroinvertebrates in the sample identified to the genus level whenever possible. Optimum
subsampling intensities and Virginia Stream Condition Index (VSCI) metrics and scores were
determined. For samples with the total number of individuals at less than 500, the genus level of

taxonomy provided lower overall optimum subsampling intensities. However, for samples with total individuals over 1000, optimum subsampling intensities at the genus level of taxonomy were higher than the family level for more than 50% of the metrics. For both family and genus levels of taxonomy, the majority of optimum subsampling intensities were well over 50% of the total individuals in the sample, with some as high as 100% of the individuals. While optimum subsampling intensities were valuable in comparing family and genus level taxonomy, they are not reasonable for stream bioassessment protocols; the cost:benefit ratio would be highly unbalanced. A minimum subsample size of 200 individuals is optimum for determining VSCI scores, while optimum taxonomic resolution is dependent on several factors. Thus, the level of taxonomic resolution for a particular study should be determined by the study objectives, level of site impairment and sample size.

Introduction

Biodiversity in aquatic ecosystems has been an important focus of ecological studies for the past several decades, including efforts to determine the level of impact from anthropogenic stressors and how impaired waters can be managed and restored (Gleick 2003; Cook et al. 2005; Giller 2005). Historically, freshwater aquatic ecosystems, and particularly their benthic macroinvertebrate communities, have been negatively affected by agricultural practices and residential, industrial and municipal development. Changes in flow regimes due to the control of water levels, elimination of riparian buffers and creation of reservoirs has reduced the ability of aquatic ecosystems to manage nutrient retention and pollution inputs, creating a need for restoration of these systems (Sondergaard 2007). Aquatic ecosystems include their often extensive and interconnected riparian zones within environmental gradients that cross spatio-temporal scales (Ward 1998). These systems have become inextricably tied to anthropogenic practices and livelihoods, and the resulting decline in biodiversity within them is mainly attributed to the increase in regulation and diversion of freshwater resources, as well as the increase in anthropogenic pollution and invasive species (Kingsford 2011).

Benthic macroinvertebrate diversity and the specific taxa present, particularly in streams, influence food web dynamics, nutrient cycling and material exchange between the benthos and the water column. While describing biodiversity in terms of ecological patterns and biological controls has long been central to the study of ecology, currently freshwater aquatic ecosystems have moved to the forefront of these studies mainly due to the magnitude and extensive scope of anthropogenic impacts in aquatic systems. In low-diversity aquatic systems, particularly impaired systems, biodiversity is crucial to the ability of the system to function (Dodds 2002). In

the face of increasing anthropogenic pressures, the biological assessment of these systems has become necessary to monitor and assess their ecological conditions.

Aquatic ecosystem bioassessment has moved to the forefront of water quality management and preservation in the recent past (Karr and Chu 1999). Accurate and cost-effective methods for acquiring high-quality biodiversity data are crucial for bioassessment at the federal and state level because that information is used to quantify the condition of aquatic ecosystems and to predict changes in their responses to anthropogenic impacts (Pfrender et al. 2010). For water quality monitoring, the cost:benefit ratio is defined by the balance between the time and resource expenditures associated with sampling, processing and analysis (costs) and the accuracy and precision of the assessment (benefits). Simplified field sampling techniques and sample processing in the laboratory are necessary tools for applicable bioassessments (Oliviera et al. 2011). Different sampling techniques for the collection, sorting and identification of stream macroinvertebrates have been and continue to be tested for their reliability in bioassessment programs (Haase et al. 2004). Well-standardized protocols generate results that reflect the reliability of more complex protocols, as long as optimal effort is used without compromising the ecological validity of the assessment (Nichols and Norris 2006).

Since the determination that taxa have different sensitivities to environmental and anthropogenic stressors, many methods have been developed to describe and quantify the ecological condition of streams, summarizing the composition of organism assemblages based on the relative sensitivities of the species present (Chessman 2012). Resulting multimetric indices (MMI's), created to quantify stream condition based on the relative tolerances to disturbances and pollution, are typically unitless values that incorporate abundance, richness and diversity of the taxa present, with the magnitude of the index being indicative of the level of environmental

or anthropogenic stress (Chessman 2012). Accurate and precise MMI's should yield repeatable

results between samples, while remaining sensitive to changes in anthropogenic stressors (Clarke

2006). The Virginia Stream Condition Index (VSCI) is the MMI currently used by the Virginia

Department of Environmental Quality (VADEQ) to determine the level of impairment in non-

coastal Virginia streams using benthic macroinvertebrates (Burton and Gerristen 2003). The use

of MMIs such as the VSCI relies heavily on the accuracy and precision of all aspects of the

methodology associated with the sampling program, including the thoroughness that a site is

sampled and the accuracy of macroinvertebrate identifications (Pfrender et al. 2010).

In the field of stream bioassessment, two subsampling methods have emerged as the

recognized methods of choice: fixed-count and area-based methods (Vinson and Hawkins 1996).

Stream geomorphology and study objectives often dictate which method should be used. For the

fixed-count method, a predetermined subsample size, typically with a minimum of 100

specimens, is used to describe the condition of a stream. With the area-based method, all

individuals in a sample, collected from a given sampling area, are included in the analysis and

thus this method typically requires considerably more time for processing of the sample while

providing more information on the macroinvertebrate assemblage present (Hering et al. 2004).

The simplicity of the fixed-count method allows for the cost and time of the assessment to be

much lower than for the area-based method. Ideally, however, characteristics of the streams and

the study objectives should determine the method employed (Oliviera et al. 2011).

Subsamples must represent the physical heterogeneity and macroinvertebrate diversity of

the stream, and thus of the total sample from which the subsample was drawn. The biological

metrics calculated from data derived from the subsample to determine stream condition must

also adequately reflect that heterogeneity and diversity. In that these metrics also may be

influenced by subsample size, they also are an important factor in determining which subsampling method is used. The U.S. EPA's Rapid Bioassessment Protocols for Use in Streams and Wadeable Rivers (RBP Protocols; Barbour et al. 1999) acknowledges that the minimum subsampling effort of macroinvertebrates may differ across geomorphologically different stream study sites. While arguments have been presented for the use of both area-based subsampling (Courtemanche 1996) and fixed-count subsampling (Vinson and Hawkins 1996, Barbour and Gerritsen 1996), the fixed-count subsampling method tends to be preferred for use with the RBP Protocols.

Besides subsampling methodology, the level of taxonomic resolution used in a study can affect the determination of the ecological condition of streams. The required level of taxonomic resolution is dependent upon the objectives and the geographic range of the study (Resh and McElravy 1993). The larger the differences in macroinvertebrate community composition and structure between sampling sites, the less taxonomic resolution that is required. Family level assessments were adequate for assessing impairment in a variety of studies along a physiographic gradient and environmental gradients (Bowman and Bailey 1998, Bailey et al. 2001, Lenat and Resh 2001, Pond and McMurray 2002, Chessman et al. 2007). In other studies, however, coarse resolution at the family level was not effective in detecting anthropogenic effects on macroinvertebrate diversity (Pfrender et al. 2010). Genus- or species-level identification may be necessary when small differences between sampling sites need to be resolved. The RBP Protocols state that genus-level taxonomic resolution provides more accurate data for bioassessments, as borne out by a number of studies (Guerold 2000, Hawkins et al. 2000, Lenat and Resh 2001, Arscott et al. 2006).

Environmental regulators and the regulated community often disagree on the severity of impairment discerned by different levels of taxonomic resolution and by the metrics used to interpret compliance to water quality regulations (Pond et al. 2008). Virginia currently uses family-level taxonomy and a minimum fixed-count subsampling of 100 specimens in its programs for the biological assessment of stream impairment. Streams flowing in the Piedmont Province of the James River watershed in central Virginia are characterized by diverse geology, variable gradients, and variable substrates. Gradient in particular has been shown to have an overall important effect on the benthic macroinvertebrate assemblages of these Piedmont streams (Marques 1998, Gotelli 2001). Given that changes in the environmental characteristics of streams can affect macroinvertebrate community composition and structure, determining the appropriate subsampling effort and taxonomic resolution required for a given stream is critical for accurate bioassessment of stream condition. The objectives of this study thus were to determine the optimum subsample size and level of taxonomic resolution necessary to accurately and precisely describe macroinvertebrate diversity in streams flowing in the Piedmont province of the James River watershed in Virginia.

Methods

Macroinvertebrate Field Sampling, Subsampling and Identification

Forty-nine sampling sites were selected from streams within the Piedmont Physiographic Province of the James River watershed (Figure 1). All sites had been previously sampled as part of Virginia Commonwealth University's INSTAR stream sampling program (www.instar.vcu.edu), which provides information on the macroinvertebrate and fish assemblages of streams throughout that and other physiographic provinces of Virginia. Collection of macroinvertebrates at those 49 sites followed protocols detailed by *EPA's Rapid Bioassessment Protocols for Streams and Wadeable Rivers* for multiple-habitat sampling (RBP Protocols; Barbour et al. 1999). Macroinvertebrates were collected systematically by kicking the substrate and jabbing with a D-frame dip net in all available major habitat types. Twenty jabs or kicks were taken in each habitat type over the entire 100-m reach, resulting in approximately 3.1 m² of habitat being sampled. The sub-samples from each habitat were placed into one composite multi-habitat sample for each of the 49 sampling sites (RBP Protocols; Barbour et al. 1999). The composite samples were preserved in 70 percent alcohol with Rose Bengal stain.

Approximately 200 individuals were removed in the laboratory from each of the 49 samples. These are hereafter referred to as "200-count samples." Then, 10 of the samples were randomly selected to have all remaining macroinvertebrates in those samples removed. These 10 samples are hereafter referred to as "total count samples." Macroinvertebrates in each sample were then identified to the genus level, or for some taxa to the lowest possible taxonomic level of resolution. Identifications were verified for accuracy by Andrew Garey, VCU Aquatics Lab Manager.

The focus of the analysis used eight macroinvertebrate bioassessment metrics selected by the VADEQ for their VSCI MMI stream bioassessment program (Burton and Gerristen 2003). The VSCI uses eight metrics that have been shown to respond predictably to increases in impairment or disturbance at the watershed level (Table 1). The eight VSCI core metrics comprise two richness metrics ("Total Taxa" and "EPT Taxa"), five proportion metrics ("% Ephemeroptera", % Plecoptera+Trichoptera less Hydropsychidae", "% Scrapers", "% Chironomidae" and "% Top 2 Dominant"), and the Hilsenhoff Biotic Index ("HBI"). The richness metrics in the category "Taxonomic Richness/Diversity" count distinct taxa. "Total taxa" is a count of all distinct taxa in the sample at the specified taxonomic level. "EPT Taxa" is a count of all distinct taxa of the orders Ephemeroptera, Plecoptera and Trichoptera. The proportion metrics in the category "Taxonomic Composition" are proportions of the total number of individuals in the entire sample that belong to specific taxonomic groups. "% Ephemeroptera" is the proportion of individuals of the order Ephemeroptera in the sample. "% Plecoptera + Trichoptera less Hydropsychidae" is the combined proportion of the number of individuals of the orders Plecoptera and Trichoptera, minus the tricopteran family of Hydropsychidae. "% Chironomidae" is the proportion of the individuals in the family Chironomidae in the sample. "% Top 2 Dominant" is the proportion of the number of individuals of the two taxa that are most abundant and next most abundant in the sample. The category "Functional Feeding Groups" comprises metrics based on the primary mode of feeding for a taxon. "% Scrapers" is the proportion of the number of individuals in the sample that are classified as scrapers. Finally, the Hilsenhoff Biotic Index, "HBI", a "Degree of Tolerance" metric, is a weighted score for taxa at the family level that measures the tolerance of each taxon to exposure to pollution.

To calculate the standardized individual VSCI metric scores for a sample, the eight metrics are divided into two groups: those that decrease with stress and those that increase with stress. For the metrics that decrease with stress ("Total taxa", "EPT taxa", "% Ephemeroptera", "% Plecoptera+Trichoptera-Hydropsychidae" and "% Scrapers"), the VSCI score is calculated using Equation 1, where metric values can range from 0 to 100 and with higher values indicating less impaired conditions.

Equation 1: Score = 100 x (X/S)

where X = metric value and S = standard (best value) (Table 1)

For the metrics that increase with stress ("% Chironomidae" and "% Top 2 Dominant"), the VSCI score is calculated using Equation 2, where metric values can range from 0 to 100 and with higher values indicating more impaired conditions.

Equation 2: Score = 100 x [(100 – X)/(100 – S)]

where X = metric value and S = standard (best value) (Table 1)

A variation of Equation 2 is used to calculate the HBI, as values range from 0 to 10.

Equation 2v: Score = 100 x [(10 – X)/(10 – S)]

where X = metric value and S = standard (best value) (Table 1)

The final VSCI score is calculated by taking an average of the eight standardized metric scores, with a maximum score of 100 being used for any standardized metric score over 100. VSCI scores are then used to determine the level of impairment at a specific site. Streams with scores over 80 are considered as having "reference" conditions, whereas streams with scores at 60-79 are considered as being "similar to reference". Streams with scores below 60 are considered "impaired".

Data Analysis for Subsampling Effort and Taxonomic Resolution Comparisons

The initial focus of the data analysis was to determine if there were significant differences in metric values if the sample processing protocol called for removing either 100 or 200 individuals (i.e. 100-individual or 200-individual subsamples). Data analysis was conducted at the family level of resolution, conforming to VDEQ protocols. Since most 200-count samples did not consist of exactly 200 individuals, subsamples were produced by randomly selecting exactly 200 individuals from each of the 49 samples using the Vegan package in R (R Core Team, 2012; Oksanen et al. 2012). Since no samples were processed at a 100-individual level, 100-individual subsubsamples were also produced.   This process was iteratively repeated to produce 100 replicates for each subsample at both the 100- and 200-individual level.

Each of the eight core VSCI metric values were calculated for both the 100-individual and 200-individual subsamples from each study site.   Pairwise T-tests (100-individual and 200-individual subsamples paired by study site) were conducted to determine if significant differences ($p=0.05$) occurred between the two levels of subsampling intensity for each of the eight metrics.

Rarefaction curves were generated for each of the eight metrics, whereby random 10-individual subsamples were drawn from each of the 49 samples, without replacement, along an interval from 10 individuals to the highest integer multiple of 10 possible for each sample. Each of the eight metric values was then calculated for each of these subsampling intervals and plotted on the curve.   Rarefactions and metric calculations were replicated 100 times for each sample.

For the eight metrics, true metric values (TMVs) were calculated at both the family and genus levels for the ten total-count samples. Mean metric values were calculated by determining

the mean values of a metric at each 10-individual interval, for the 100 replicates, for the total number of individuals in the each sample. True metric values are considered to be the mean metric values at the highest 10-individual interval of each sample.

The second focus was to determine if significant differences occurred in metric values when data at the family versus genus level were used. Using data from the ten total-count study sites, optimum subsampling intensity (OSI) was determined for each metric as the minimum number of individuals at which at least 95 of replicate subsamples had metric values that met or exceeded the true metric value. Pairwise t-tests ($p \leq 0.05$) were then used to determine if significant differences occurred between optimum subsampling intensities calculated from the rarefaction curves generated from family-level or genus-level taxonomy. This analysis was conducted on three metrics, "Total taxa", "% Top 2 Dominant" and "HBI", that can vary based on taxonomic resolution. The other five metrics were not analyzed in this manner because by their nature they are not responsive to differences in taxonomic resolution.

The eight metrics for the total-count samples were calculated at 100-individual and 200-individual counts and compared at family-level and genus-level taxonomy to determine the departure from the true metric value for the sample. The 100-individual and 200-individual count metric values were compared to the true metric values as a percent of the true metric value to determine if the current subsampling protocols accurately represent the true metric values of the samples.

Results

Only seven of the 49 sampling sites had VSCI scores that met or exceeded the VDEQ threshold score of 60, indicating their level of impairment as being "similar to reference" conditions. Five of those seven sites were included among the 39 200-count sampling sites and two were included among the 10 total-count sites (Fig. 2 and 3).

Statistically significant differences were found in three of the VSCI metrics between the 100-count and 200-count subsamples. The metrics "% Top 2 Dominant", "Total taxa" and "EPT taxa" were the metrics most affected by the different subsampling protocols. All other metrics showed no statistically significant difference in their values between the 100- and 200-count subsampling protocols (Table 2, Figure 4).

For the 10 total-count samples, the total number of individuals in each sample ranged from a minimum of 210 to a maximum of 1640. Four samples contained over 1000 individuals, while five samples contained less than 500 individuals. The true metric values, at both the family and genus levels of taxonomy, varied based on these differences in numbers of individuals in the sample and the variability of the metric values at each interval (Tables 3 and 4).

The optimum subsampling intensity (OSI) for each metric was determined for sites with less than 500 individuals (Figure 5) and sites with more than 1000 individuals (Figure 6). The OSI, calculated as a percent of the total number of individuals in the sample (Table 5), varied widely across each sample. For the samples where the OSI was 100% of the total individuals, it was reached in the final 10-individual interval or it was never reached. In the sites with over 1000 total individuals, only three sites had a single metric not reach the OSI and the overall patterns of OSIs tended to show the highest OSIs in the richness metrics, with the lowest in each

20

sample being "HBI" (Figure 5). In sites with less than 500 individuals, each site had at least two metrics not reach the OSI, with one site not reaching OSI for six of the eight metrics. The overall OSI patterns tended to show highest OSIs in the richness metrics and "% Scrapers", while the "HBI" metric had the lowest OSIs for all samples (Figure 6). .

Rarefaction curves for each metric were then used to compare the optimum subsampling intensity at the family- and genus-levels of taxonomy. The rarefaction curves for sampling sites JM42_02 and JM52_01 were chosen to represent the baseline curves for the comparisons for each of the 10 total-count sampling sites. These two sites were chosen as the baseline because they had the highest VSCI scores of the 10 sampling sites, with values at or near 60 and thereby identifying these sites as "similar to reference" conditions. Also, site JM42_02 had a total of 470 individuals and thus its rarefaction curve was used for comparisons with other sites with fewer than 500 individuals in the sample (Figure 7). Site JM52_01 had a total of 1640 individuals and thus its curve was used for comparisons with sites with greater than 1000 individuals (Figure 15). This allowed for a demonstration of comparisons at the family and genus levels of taxonomy for two different sized samples at the same relative level of impairment (Figures 7-22).

The rarefaction curves for the proportion metrics showed high variability, with a leveling off as the total number of individuals in the samples was approached.   The OSI at site 42_02 was higher at the family level than at the genus level for five metrics ("EPT Taxa", "% Ephemeroptera", "% Plecoptera + Trichoptera less Hydropsychidae", % Scrapers", and "HBI") and was the same as the genus level for the remaining three metrics ("Total taxa", "% Chironomidae", and "% Top 2 Dominant") (Figures 7-14). This is representative of the majority of sites with less than 500 individuals, as these sites tended to favor genus level taxonomy overall for optimum subsampling intensity.

The family level OSI at site 52_01 was lower for five metrics ("% Ephemeroptera", "% Plecoptera + Trichoptera less Hydropsychidae", "% Chironomidae", "% Top 2 Dominant", and "HBI") and higher for three metrics ("Total taxa", "EPT Taxa", and "% Scrapers") compared to the family level (Figures 15-22). This is representative of the majority of sites with more than 1000 individuals, as these sites tended to favor family level taxonomy overall for optimum subsampling intensity. Pairwise T-tests showed no significant differences (p≤0.05) in the optimum sampling intensity between the family or genus levels taxonomy for any of the three metrics analyzed ("Total Taxa", "Top 2 Dominant" and "HBI").

The accuracy at which the 100-count and 200-count subsamples represented the total macroinvertebrate composition in the samples at the family and genus levels varied depending on the metric (Figures 23 and 24). Ideally, metrics would score between 95% and 100% for an accurate representation of the true metric value, whereas metrics scoring over 100% would be overestimated in the VSCI score. For both the 100-individual and 200-individual counts at the family level of taxonomy, the metric values were at least 95% of the true value for the metrics "% Ephemeroptera", "% Chironomidae", "% Top 2 Dominant" and "HBI" for all replicates. Both the "% Plecoptera plus Trichoptera less Hydropsychidae" and "% Scrapers" metrics showed 70% of the samples scoring at or above 95% of the true metric values. The "% Top 2 Dominant" was the only metric at the family level that scored at over 100% of the true metric value for all ten samples. Neither the "Total taxa" or "EPT taxa" metrics were scored at over 100% of the total metric value; both metrics did have one sample at over 95% of the true metric value (Figure 23). At the genus level of taxonomy, four metrics ("% Ephemeroptera", "% Chironomidae", "% Top 2 Dominant", "HBI") were scored at over 95% of the true metric values for both 100-individual and 200-individual subsamples, As with the family level comparison,

22

both "Total taxa" and "EPT taxa" had only one sample each scoring over 95% of the true metric

value (Figure 24).

Discussion

Significant differences in three of the eight metrics occurred when comparing all sites subsampled at both 100 and 200 individual counts.   The optimum sub-sample size at either 100 or 200 individual counts varied depending on the metric being examined. The differences to note were for the "Total taxa", "EPT taxa", and "% Top 2 Dominant" metrics, these being two richness metrics and a composition metric, respectively. These differences indicated that the values of these metrics are being underestimated when a 100-individual subsample protocol is used, compared to using a 200-individual subsample. In general, richness metrics were more sensitive to the effects of counting 100 versus 200 individuals from a sample than were the proportion metrics. These results are consistent with other studies that showed that richness metrics are the metrics most affected by changes in subsample size (Doberstein et al. 2000, Lorenz et al. 2004, Clarke et al. 2006, Oliviera et al. 2011).

The eight VSCI metrics, as used in bioassessment programs, are averaged together with equal weight to create a VSCI score. Each metric has some unknown extent of inaccuracy, to some degree depending on the subsampling protocol and the level of taxonomy employed. Metrics that are overestimated with a given protocol would need to balance underestimated metric scores for the VSCI index to provide an accurate representation of the macroinvertebrate community and thus condition of the stream. Since the current Virginia protocols require 100-individual subsamples, there would be a benefit to increasing the minimum subsample size to 200 individuals, as this would decrease the overestimation by the proportion metrics and decrease the underestimation by the richness metrics, thereby generating a more accurate VSCI score. Other studies have also shown that a 200-count provides greater metric stability and

functionality and an overall more accurate representation of a stream's macroinvertebrate assemblage than does a 100-count protocol (Lorenz et al. 2004).

For the purposes of this analysis, optimum range was determined to be metric values greater than 95%, but not exceeding, the true metric value. The majority of the samples had proportion metric values that were within the optimum range when taxonomy was at the family level. Richness metrics were particularly sensitive to differences in sample count, rarely achieving the optimum range and being consistently underestimated at nearly all sampling sites. The metrics "HBI", "% Chironomidae", "% Top 2 Dominant" and "% Ephemeroptera", at both levels of taxonomy, had 100% of the samples achieving the optimum range. However, these metrics differed between taxonomic levels on the percentage of sites with metric values exceeding the optimum range. "HBI" and "% Chironomidae" had less sites exceed the optimum range at the family level, while "% Top 2 Dominant" and "% Ephemeroptera" had less sites exceed the optimum range at the genus level. For the metrics "% Plecoptera + Trichoptera less Hydropsychidae" and "% Scrapers", less sites exceeded the optimum range at the family level and less sites achieved the optimum range at the family level. Thus, it is unclear for these two metrics which taxonomic level is optimal. In general, richness metrics responded equally to both levels of taxonomic resolution, neither adequately estimating the sample richness. The proportion metrics varied, but the majority favored family level over genus level taxonomy.

The cost:benefit ratio of stream bioassessment requires that the cost of sampling, processing and analyzing the samples be considered relative to the need for accuracy and precision in the final VSCI score for each sampled site. To provide the highest level of precision and accuracy, all individuals in a sample would need to be removed and included in the data set. This is generally not possible given agency budgets and time constraints, with the cost per

sample exceeding allowable costs within the cost:benefit ratio balance. Optimum subsampling

intensities thus were calculated for each metric, at both the family and genus levels of taxonomy,

under the assumption that a lower optimum subsampling intensity would provide a more

appropriate cost:benefit ratio.

Optimum subsampling intensities for samples with fewer than 500 individuals and

identified at the genus level of taxonomy were lower or the same for all metrics compared to the

family level of taxonomy. However, for samples with greater than 1000 individuals, the

optimum subsampling intensities at the genus level of taxonomy were higher than for the family

level for the majority of the metrics. Richness metrics and percent scraper metric had lower

optimum subsampling intensities at the genus level than the family level of taxonomy, most

likely due to rarer species causing more variability at the family level. For the majority of the

proportion metrics, the family level of taxonomy provided the lower optimum subsampling

intensity for the majority of the proportion metrics when processing the larger samples.

The majority of the optimum subsampling intensities, at both the family and genus level

of taxonomy, were well over 50% of the total sample size, with some as high as 100%. While

the optimum subsampling intensities were valuable in comparing family and genus level

taxonomy, they are not reasonable for stream bioassessment protocols, as the cost:benefit ratio

would be highly unbalanced. Therefore, it would be difficult to determine the level of taxonomic

resolution better suited for subsampling without first knowing the total number of individuals in

the sample. The required level of taxonomic resolution should be dependent on the study

objectives, whereas family level resolution was deemed adequate for determining impairment

levels in a variety of studies along physiographic and environmental gradients (Resh and

McElravy 1993, Bowman and Bailey 1998, Bailey et al. 2001, Lenat and Rash 2001, Pond and

McMurray 2002, Chessman et al. 2007), other studies determined that genus level resolution was more effective in determining anthropogenic effects on macroinvertebrate diversity (Pfrender et al. 2010).

In conclusion, a minimum subsample size of 200 individuals is optimum for determining VSCI scores, while the optimum taxonomic resolution is dependent on several factors. Thus, the level of taxonomic resolution for a particular study should be determined by the study objectives, level of site impairment and sample size.

References

Arscott D.B., Jackson J.K., and Kratzer E.B. 2006. Role of rarity and taxonomic resolution in a regional and spatial analysis of stream macroinvertebrates. Journal of North American Benthological Society, 25: 977-997.

Bailey R.C., Norris R.H., Reynoldson T.B. 2001. Taxonomic resolution of benthic macroinvertebrate communities in bioassessments. Journal of North American Benthological Society, 20: 280-286.

Barbour M.T., Gerritson. J. 1996. Subsampling of benthic samples: a defense of the fixed-count method. Journal of North American Benthological Society, 15: 386-391.

Barbour M.T., Gerritson J., Snyder D., and Stribling J.B. 1999. Rapid Bioassessment Protocols for use in Streams and Wadeable Rivers; Periphyton, Benthic Macroinvertebrates, and Fish. 2nd Edition. EPA 841-B-99-002. Washington, D.C.: Office of Water, U.S. Environmental Protection Agency.

Barbour M.T., Plafkin J., Bradley B.P., Graves C.G. and Wisseman R.W. 1992. Evaluation of EPA's rapid bioassessment benthic metrics: metric redundancy and variability among reference stream sites. Environmental Toxicology and Chemistry, 11: 437-449.

Bowman M.F., and Bailey R.C. 1998. Does taxonomic resolution affect the multivariate description of the structure of freshwater benthic macroinvertebrate communities? Canadian Journal of Fisheries and Aquatic Sciences, 54: 1802-1807.

Burton, J. and Gerristen, J. 2003. A stream condition index for Virginia non-coastal streams. Tetra Tech, Inc. Owing Mills, MD. Report prepared for USEPA Office of Science and Technology, Office of Water, Washington, DC; USEPA Region 3 Environmental Services Division, Wheeling, WV; Virginia Department of Environmental Quality, Richmond, VA.

Cao Y., Hawkins C.P. & Vinson M. 2003 Measuring and controlling data quality in assemblage surveys a special reference to stream macroinvertebrates. Freshwater Biology, 48, 1898–1911.

Cao Y., Williams D.D. & Larsen D.P. 2002 Comparison of ecological communities - the problem of sample representativeness. Ecological Monographs, 72, 41–56.

Cao Y., Hawkins C.P. & Storey A.W. 2005 A method for measuring the comparability of different sampling methods used in biological surveys: implications for data integration and synthesis. Freshwater Biology, 50, 1105-1115.

Carter J.L. & Resh V.H. (2001) After site selection and before data analysis: sampling, sorting, and laboratory procedures used in stream benthic macroinvertebrate monitoring programs by USA state agencies. Journal of the North American Benthological Society, 20, 658–682.

Chessman B., Williams S., and Besley C. 2007. Bioassessment of streams with macroinvertebrates: effect of sampled habitat and taxonomic resolution. Journal of North American Benthological Society, 26: 546-565.

Chessman, Bruce; McEvoy, Paul. 2012. Insights into Human Impacts on Streams from Tolerance Profiles of Macroinvertebrate Assemblages. Water, Air and Soil Pollution, 223: 1343-1352.

Clarke, R. T., Lorenz, A., Sandin, L., Schmidt-Kloiber, A., Strackbein, J., Kneebone, N. T., et al. (2006).Effects of sampling and sub-sampling variation using the STAR-AQEM sampling protocol on the precision of macroinvertebrate metrics. Hydrobiologia, 566, 451–459.

Clarke R.T., Murphy J.F. 2006. Effects of locally rare taxa on the precision and sensitivity of RIVPACS bioassessment of freshwaters. Freshwater Biology, 51: 1924-1940.

Coffman W.P. and Ferrington Jr., L.C. 1984. "Chironomidae", Chapter 25, in R.W. Merritt and K.W. Cummins (eds) An Introduction to the Aquatic Insects of North America, 2nd edn., Kendall/Hunt Publishing Company, Dubuque, Iowa. 722.

Courtemanche, D.L. 1996. Commentary on the subsampling procedures used for rapid bioassessments. Journal of the North American Benthological Society. 15:381-385.

Doberstein, C. P., Karr, J. R., & Conquest, L. L. (2000). The effect of fixed-count subsampling on macroinvertebrate biomonitoring in small streams. Freshwater Biology, 44, 355–371.

Dodds, Walter K. 2002. Freshwater Ecology Concepts and Environmental Applications. Academic Press, San Diego, CA.

Gotelli, Nicholas J. and Colwell, Robert K. 2001. Quantifying Biodiversity: procedures and pitfalls in the measurement and comparison of species richness. Ecology Letters, 4, 379-391.

Guerold F. 2000. Influence of taxonomic determination level on several community indices. Water Research 34:487-492.

Haase P., Lohse S., Pauls S., Schindehutte K., Sundermann A., Rolauffs P. 2004. Assessings streams in Germanywith benthic invertebrates: Development of a practical standardised protocol for macroinvertebrate sampling and sorting. Limnologica 34:349-365.

Hawkins C.P., Norris R.H., Hogue J.N. and Feminella J.W. 2000. Development and evaluation of predictive models for measuring the biological integrity of streams. Ecological Applications, 10: 1456-1477.

Heck K.L. Jr, Van Belle G. & Simberloff D. 1975. Explicit calculation of the rarefaction diversity measurement and the determination of sufficient sample size. Ecology, 56, 1459–1461.

Hering, D., Moog, O., Sandin, L., & Verdonschot, P. F. M. (2004). Overview and application of the AQEM assessment system. Hydrobiologia, 516, 1–20.

Karr J.R. and Chu E.W. 1999. Restoring life in running waters: Better biological monitoring p.206. Washington D.C.: Island Press.

Kingsford R.T., Biggs H.C., and Pollard S.R. 2011. Strategic adaptive management in freshwater protected areas and their rivers. Biological Conservation, 144: 1194-1203.

Lenat, D.R. (1993). A biotic index for the southeastern U.S.: derivation and list of tolerance values with criteria for assigning water quality ratings. Journal of North American Benthological Society, 12: 279-290.

Lenat D.R., Resh V.H. 2001. Taxonomy and stream ecology - The benefits of genus-and species-level identifications. Journal of North American Benthological Society, 20: 287-298.

Li J., Herlihy A., Gerth W., Kaufmann P., Gregory S., Urguhart S., and Larsen D. 2001. Variability in stream macroinvertebrates at multiple spatial scales. Freshwater Biology, 46: 87-97.

Lorenz, A., Kirchner, L., & Hering, D. (2004). Electronic subsampling of macrobenthic samples: How many individuals are needed for a valid assessment result? Hydrobiologia, 516, 299–312.

Macdonald E.E. and Taylor B.R. 2006. Incidence of mentum deformities in midge larvae (Diptera: Chironomidae) from Northern Nova Scotia, Canada. Hydrobiologia, 563: 277-287.

Marques, D.M. 1998. Comparison of macroinvertebrate stream communities across the Southeast Plains ecoregion of Virginia. Thesis. Virginia Commonwealth University.

Martinez E.A., Moore B.C., Schaumloffel J., and Dasgupta N. 2004. Effects of exposure to zinc- and lead-spiked sediments on mouthpart development and growth in *Chironomus tentans*. Environmental Toxicology and Chemistry, 23: 662-667.

Mebane, Christopher A. 2001. Testing Bioassessment Metrics: Macroinvertebrate, Sculpin, and Salmonid Responses to Stream Habitat, Sediment, and Metals. Environmental Monitoring and Assessment 67:293-322.

McCune B., and Grace J.B. 2002. Analysis of ecological communities. MjM Software Design, Gleneden Beach, Oregon.

Mykra H., Aroviita J., Kotanen J., Hamalainen H., Muotka T. 2008. Predicting the stream macroinvertebrate fauna across regional scales: influence of geographical extent on model performance. Journal of the North American Benthological Society, 27: 705-716.

Nagel D., B. J. 2006. Comparison of Methods for Estimating Stream Channel Gradient Using GIS. Boise, Idaho: USDA Forest Service, Rocky Mtn. Research Station, Boise Aquatic Sciences Lab.

Nichols S.J. and Norris R.H. 2006. River condition assessment may depend on the sub-sampling method: Field live-sort versus laboratory subsampling of invertebrates for bioassessment. Hydrobiologia 572:95-213.

Oksanen J., Blanchet F.G., Kindt R., Legendre P., Minchin P.R., O'Hara R.B., Simpson G.L., Solymos P., Henry M., Stevens H., and Wagner H. 2012. Vegan: Community Ecology Package. R package version 2.0-4. URL http://CRAN.R-project.org/package=vegan.

Oliviera R., Mugnai R., Castro C. and Baptista D. 2011. Determining subsampling effort for the development of a rapid Bioassessment protocol using benthic macroinvertebrates in streams of Southeastern Brazil. Environmental Monitoring and Assessment, 175: 75-85.

Pfrender, Michael et al. 2010. Assessing macroinvertebrate diversity in freshwater ecosystems: advances and challenges in DNA-based approaches. The Quarterly Review of Biology, 85: 319-340.

Plafkin J.L., Barbour M.T., Porter K.D., Gross S.K. and Hughes R.M. 1989. Rapid Bioassessment Protocols For use in streams and rivers. In:Benthic Macroinvertebrates and Fish. EPA|444|4-89|001. Office of Water Regulations and Standards, U.S. EPA, Washington D.C.

Pond G.J., Passmore M.E., Borsuk F.A., Reynolds L., and Rose C.J. 2008. Downstream effects of mountaintop coal mining: comparing biological conditions using family- and genus-level macroinvertebrate assessment tools. Journal of North American Benthological Society, 27 (3): 717 - 737.

R Core Team. 2012. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-9000051-07-0, URL http://www.R-project.org/.

Rabeni, Charles F. and Wang, Ning. 2001. Bioassessment of Streams Using Microinvertebrates: Are the Chironomidae Necessary? Environmental Monitoring and Assessment, 71:177-185.

Resh V.H., McElravy E.P. 1993. Contemporary quantitative approaches to biomonitoring using benthic macroinvertebrates. In e. D. Resh, Freshwater Biomonitoring and Benthic Macroinvertebrates (pp. 159-194). New York: Chapman and Hall.

Sondergaard, Martin and Erik Jeppesen. 2007. Anthropogenic impacts to lake and stream ecosystems, and approaches to restoration. Journal of Applied Ecology, 44: 1089-1094.

Vanderklift M.A., Ward T.J. and Jacoby C.A. 1996. Effect of reducing taxonomic resolution on ordinations to detect pollution-induced gradients in macrobenthic infaunal assemblages. Marine Ecology Progress Series. 136: 137-145.

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

Ward, J.V. 1998. Riverine landscapes: biodiversity patterns, disturbance regimes, and aquatic conservation. Biological Conservation, 83: 269-278.

Wiley M.J., Kohler S.L. and Seelbach P.W. 1997. Reconciling landscape and local views of aquatic communities: lessons from Michigan trout streams. Freshwater Biology, 37: 133-148.

Appendix

Table 1. Metrics used to calculate the VSCI and analyzed in this study. ER represents the expected response to an increase in disturbance, with + or − indicating that the metric value should increase or decrease with increased disturbance. Equations 1 and 2 refer to the equations used in the VSCI manual for calculating the index. The index is calculated on a unitless 100 point scale, where all metrics contribute equal weight.

VSCI Core Metrics	Metric Category	ER	Standard (best value)	Equation
Total taxa	Taxonomic Richness/Diversity	-	22	1
EPT taxa	Taxonomic Richness/Diversity	-	11	1
% Ephemeroptera	Taxonomic Composition	-	61.3	1
% Plecoptera+Trichoptera less Hydropsychidae	Taxonomic Composition	-	35.6	1
% Scrapers	Functional Feeding Group	-	51.6	1
% Top 2 Dominant	Taxonomic Composition	+	0	2
% Chironomidae	Taxonomic Composition	+	30.8	2
HBI	Degree of Tolerance	+	3.2	2v

Table 2. Mean VSCI metric values for 100-individual count and 200-individual count subsamples with standard errors, for each of the eight metrics. Pairwise t-tests were calculated to determine significant differences between the subsample sizes for each metric at significance level p≤0.05. Highlighted p-values indicate significant differences were found.

VSCI Metric	mean100	se100	mean200	se200	p-value
%chiro	54	4	54	4	0.4715
%ephem	22	2	22	2	0.8900
%pt-h	6	1	7	1	0.0741
%scrapers	8	1	8	1	0.8218
%top2dom	61	2	61	2	0.0000
ept_taxa	7	1	6	0	0.0462
hbi	6	0	7	1	0.1435
richness	13	1	19	1	0.0000

Table 3. True metric values (TMV) calculated at Family level of taxonomic resolution for the ten sites picked in their entirety. TMVs chosen are the metrics used to calculate the Virginia Stream Condition Index (VSCI) score. TMVs were determined using the mean value of the rarefied sample (100 replicates) at the last interval of ten individuals sampled.

	#INDIV	TOTAL_TAXA	EPT_TAXA	%EPHEM	%PT-H	%SCRAPERS	%CHIROS	%TOP2_DOM	HBI
JA07_01	1020	27	8	42.38	9.21	3.91	29.54	65.74	5.51
JA18_01	1570	30	6	3.26	1.50	2.64	80.59	84.30	6.70
JA19_02	840	26	6	3.78	12.33	2.96	55.12	69.32	6.11
JA33_04	430	28	7	20.83	0.70	3.00	50.48	60.31	6.48
JM18_03	1340	31	10	12.27	21.01	7.04	36.71	48.14	5.06
JM37_01	430	23	10	14.57	22.14	5.00	44.87	54.68	5.41
JM42_02	470	22	7	52.17	14.33	3.16	9.04	51.34	3.97
JM52_01	1640	37	9	20.53	17.24	14.57	25.78	43.81	4.84
JM71_05	210	26	6	10.55	2.74	11.91	18.23	44.71	6.78
JM77_02	460	30	9	7.26	1.28	4.70	58.65	65.98	6.36

Table 4. True metric values (TMV) calculated at Genus level of taxonomic resolution for the ten sites picked in their entirety. TMVs chosen are the metrics used to calculate the Virginia Stream Condition Index (VSCI) score. TMVs were determined using the mean value of the rarefied sample (100 replicates) at the last interval of ten individuals sampled.

	#IND	TOT_TAXA	EPT_TAXA	%EPHEM	%PT-H	%SCRAPER	%CHIROS	%TOP2_DOM	HBI
JA07_01	1020	38	18	42.35	9.20	1.37	29.54	46.57	5.78
JA18_01	1570	45	13	3.26	1.51	3.71	80.58	82.53	6.73
JA19_02	840	35	13	3.80	12.30	2.95	55.16	69.35	6.11
JA33_04	430	34	9	20.84	0.69	8.71	50.50	57.58	6.48
JM18_03	1340	42	15	12.26	21.03	16.35	36.70	48.13	5.06
JM37_01	430	30	16	14.61	22.09	6.83	44.84	53.48	5.36
JM42_02	470	28	10	52.24	14.31	34.99	9.04	51.39	4.31
JM52_01	1640	55	19	20.54	17.25	31.78	25.78	43.25	4.53
JM71_05	210	33	7	10.52	2.74	5.45	18.18	44.69	6.73
JM77_02	460	40	15	7.25	1.27	5.11	58.65	65.90	6.52

Table 5. Optimum subsample intensities shown as percent of the true metric value for each of the 10 total-count samples at both the family and genus levels of taxonomic resolution.

| | JA07_01 | | JA18_01 | | JA19_02 | | JA33_04 | | JM18_03 | |
	family%	genus%	family%	genus%	family%	genus%	family%	genus%	family%	genus%
TOTAL TAXA	94%	91%	95%	94%	93%	95%	95%	95%	96%	87%
EPT TAXA	96%	99%	97%	101%	96%	100%	70%	79%	99%	98%
% EPHEM	71%	68%	99%	98%	99%	98%	98%	95%	92%	90%
%PT-H	93%	87%	100%	100%	94%	94%	98%	100%	78%	84%
%SCRAP	99%	100%	100%	92%	98%	100%	100%	100%	96%	84%
%CHIRO	75%	78%	18%	22%	58%	65%	84%	81%	69%	66%
%TOP2	44%	65%	31%	22%	42%	50%	72%	70%	49%	59%
HBI	13%	8%	3%	1%	12%	32%	16%	16%	16%	25%
	JM37_01		JM42_02		JM52_01		JM71_05		JM77_02	
	family%	genus%	family%	genus%	family%	genus%	family%	genus%	family%	genus%
TOTAL TAXA	98%	100%	96%	96%	95%	95%	100%	100%	98%	96%
EPT TAXA	98%	100%	100%	98%	98%	91%	100%	100%	100%	100%
% EPHEM	100%	98%	77%	72%	80%	81%	100%	100%	100%	100%
%PT-H	95%	95%	96%	94%	80%	87%	100%	100%	100%	100%
%SCRAP	100%	100%	100%	85%	82%	68%	100%	100%	100%	100%
%CHIRO	77%	79%	98%	98%	71%	76%	100%	100%	72%	72%
%TOP2	72%	72%	74%	74%	50%	52%	95%	95%	65%	63%
HBI	28%	44%	38%	30%	15%	16%	10%	19%	17%	9%
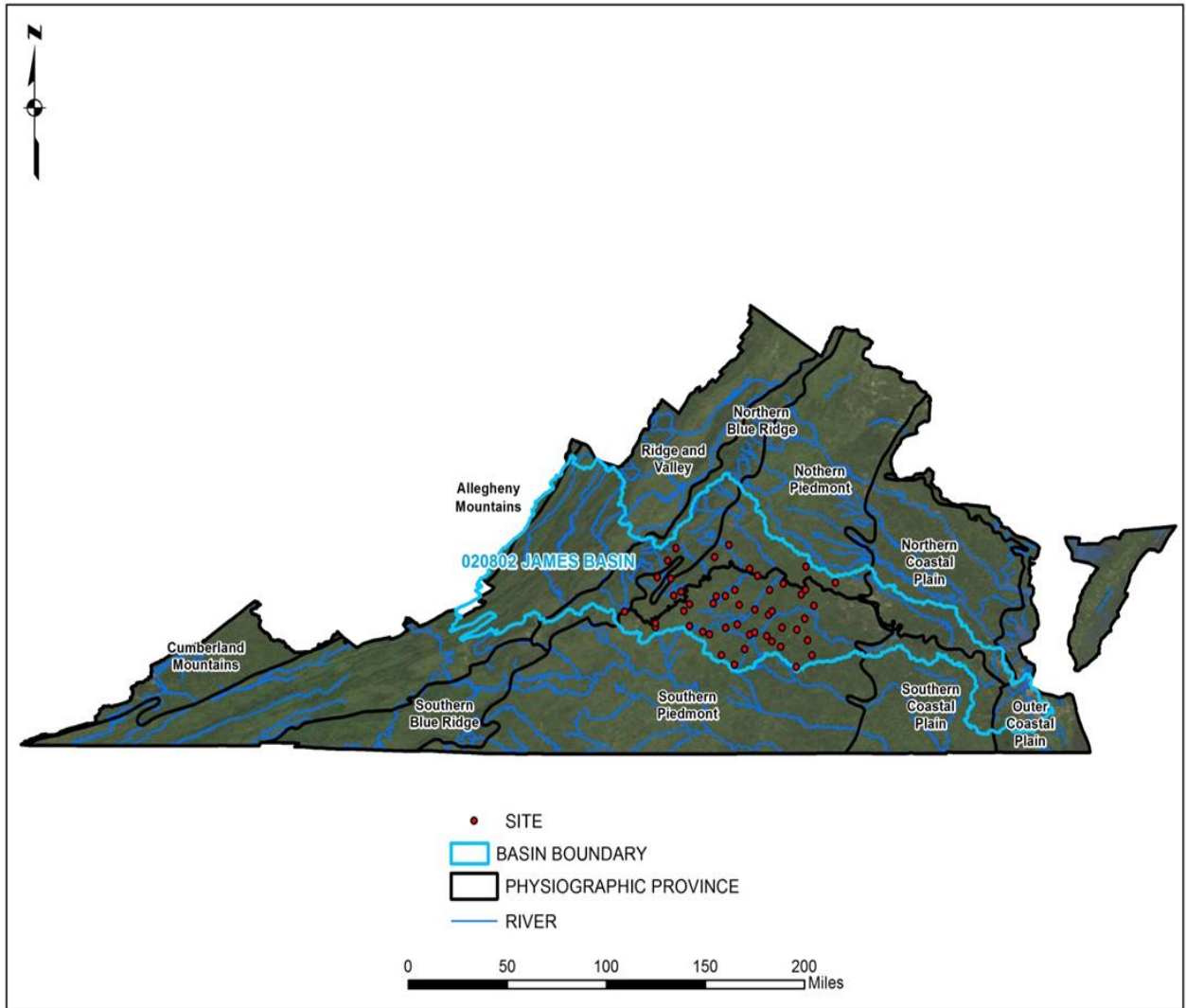
Figure 1. Forty-nine sampling sites shown on a map of Virginia, within the Piedmont physiographic province and James River watershed.
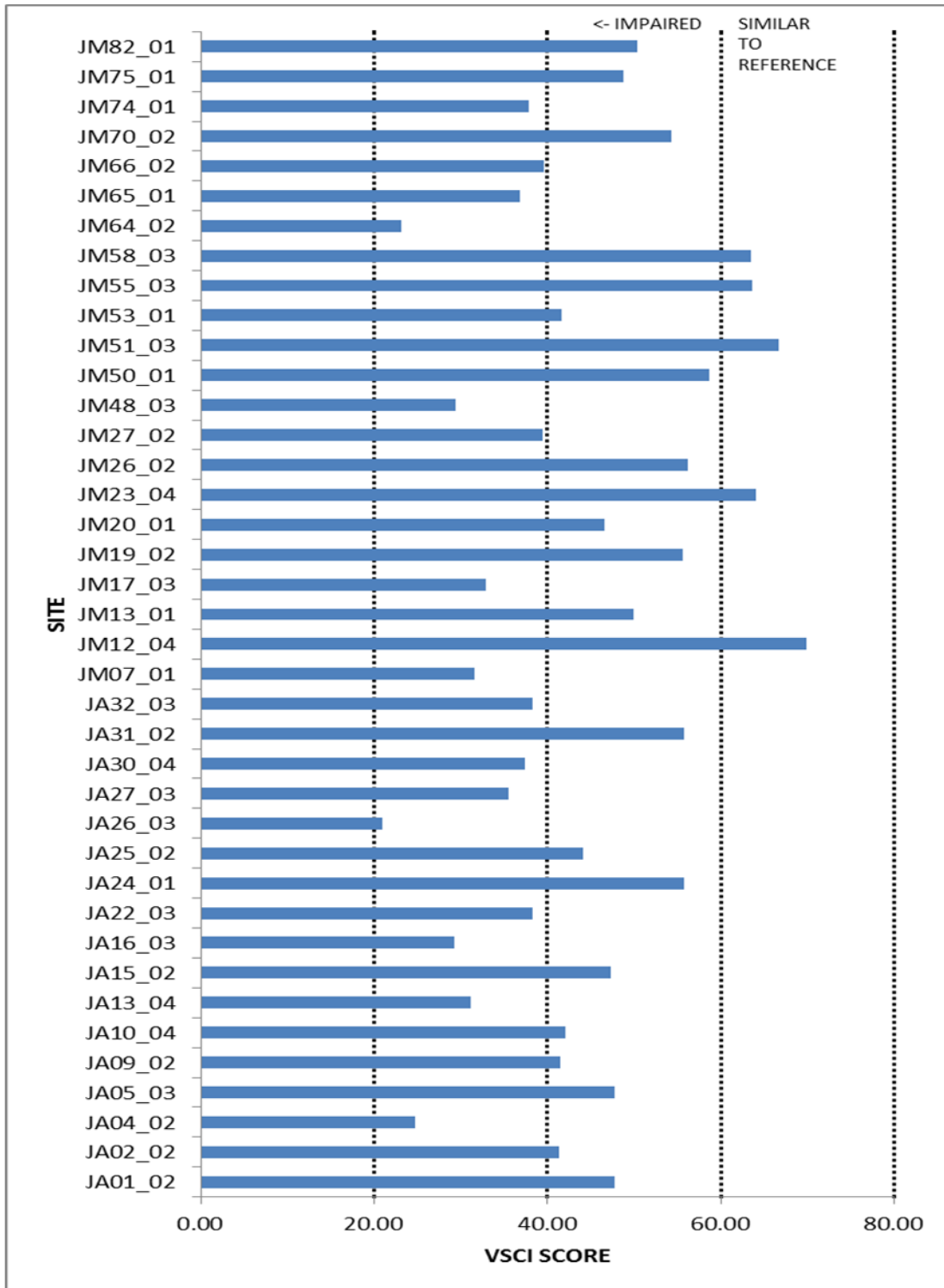
Figure 2. Calculated Virginia Stream Condition Index (VSCI) scores determined for 39 sites at the Family level of taxonomic resolution for 100 individuals. VSCI scores were calculated using the eight mean metric values at the 100 individual count, each metric having equal weight. The classification threshold of note is a VSCI score of 60, indicating scores above are "similar to reference" sites While scores below are considered "impaired" waters.
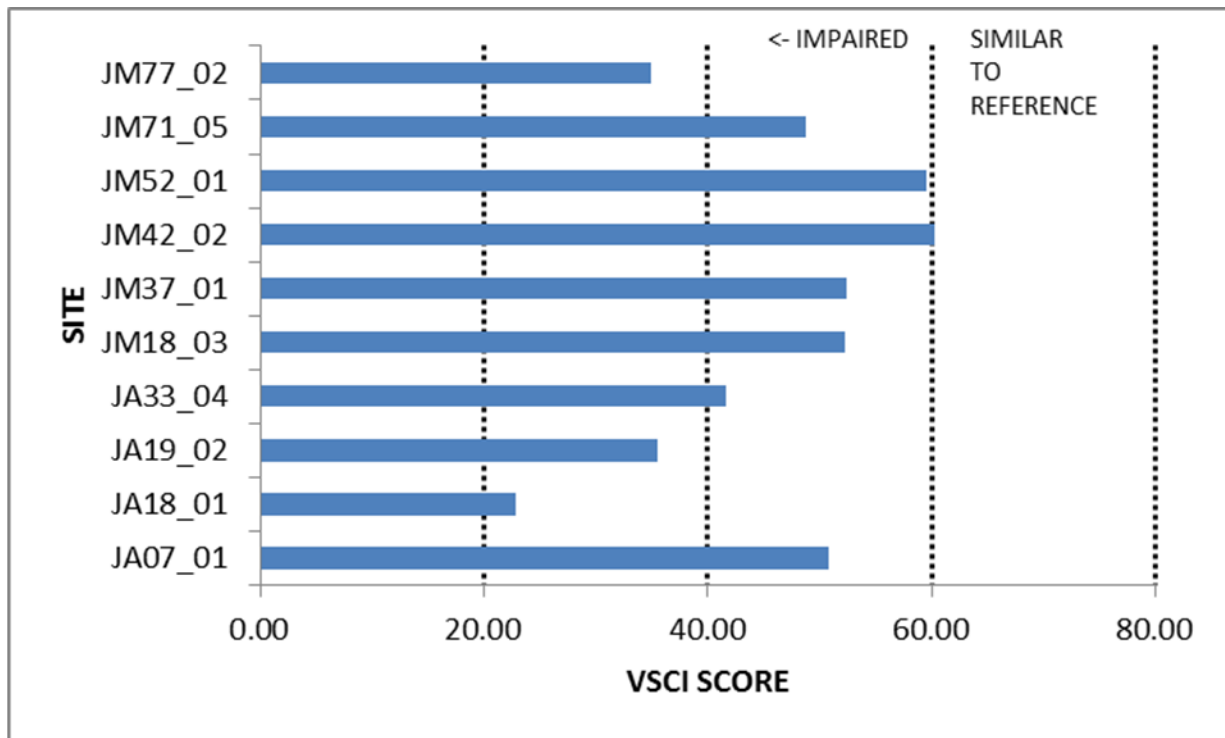
Figure 3. Calculated Virginia Stream Condition Index (VSCI) scores determined for 10 sites picked in entirety, at the Family level of taxonomic resolution for 100 individuals. VSCI scores were calculated using the eight mean metric values at the 100 individual count, each metric having equal weight. The classification threshold of note is a VSCI score of 60, indicating scores above are "similar to reference" sites While scores below are considered "impaired" waters.
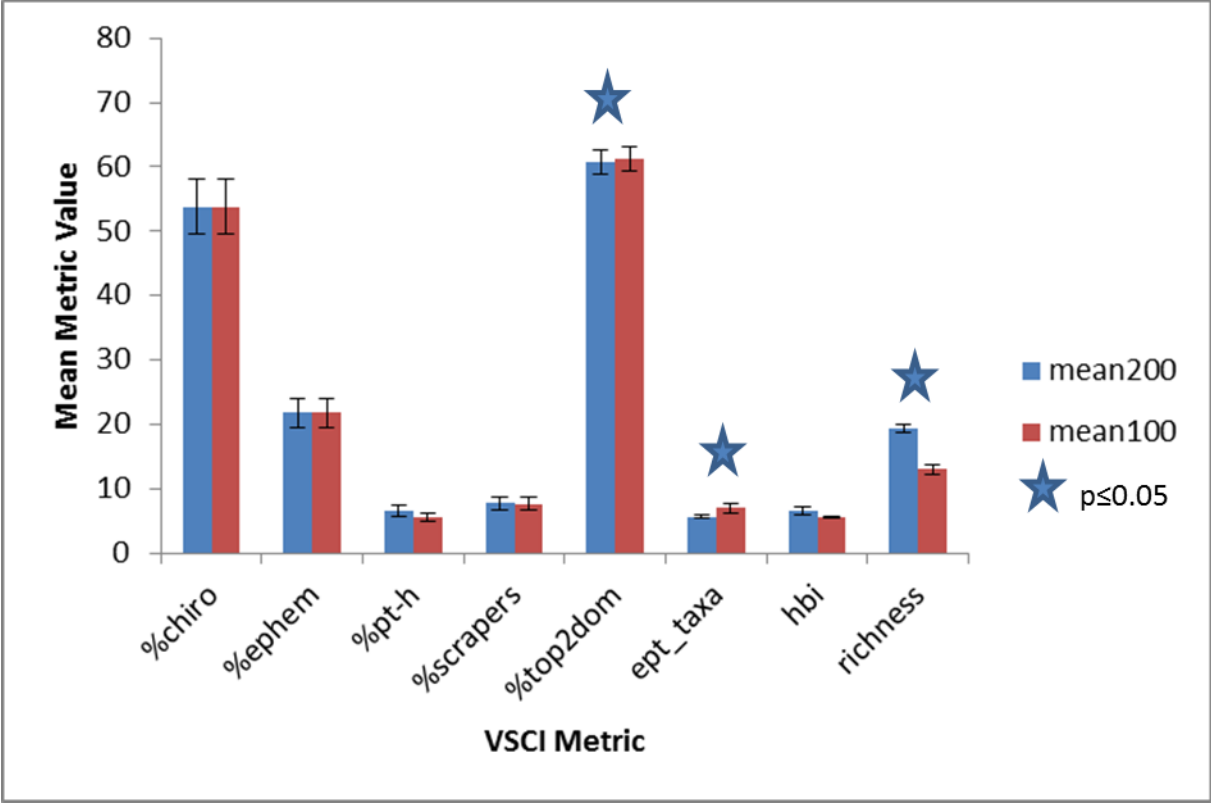
Figure 4. Calculated Virginia Stream Condition Index (VSCI) mean metric values for each metric at 100-individual and 200-individual subsample sizes. Pairwise t-tests were calculated to determine if significant differences exist between the two subsample sizes for each metric at significance level of p≤0.05. Stars above indicate when a significant difference is present.
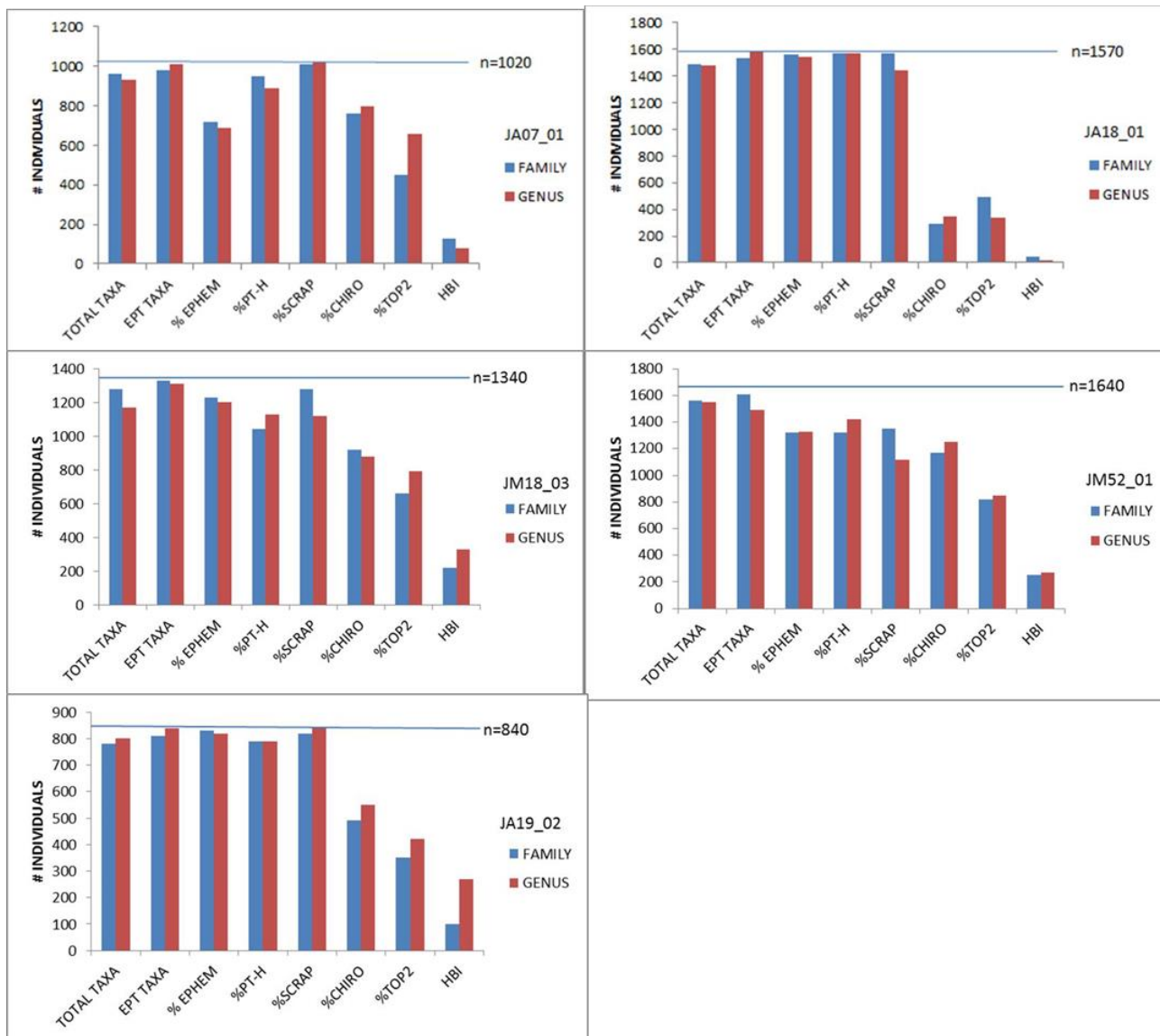
Figure 5. Optimum Subsampling Intensity for sampling sites with greater than 1000 individuals, for both Family and Genus levels of taxonomic resolution, for each of the eight VSCI metrics. The optimum subsampling intensity value was calculated by determining 95% of the true metric value, with less than 5% error (variation)among the 100 replicates for each ten-individual interval. The y=n line represents the total number of individuals in the sample (n). The metrics where the optimum subsampling intensity is equal to the total number of individuals in the sample, either did not meet the criteria or met in the last possible ten-individual interval. The site JA19_02 was included due to the total number of individuals being greater than 500.
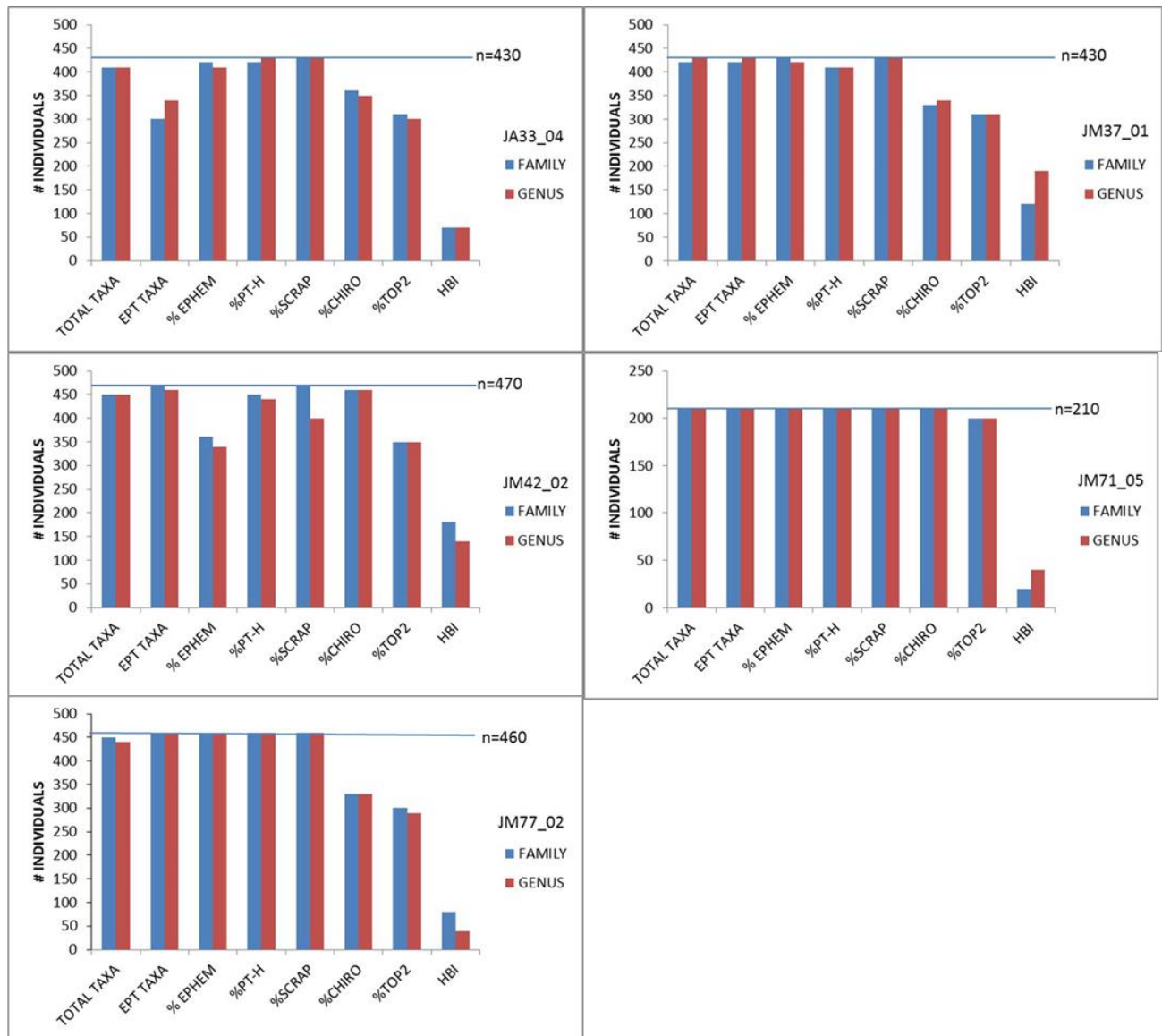
Figure 6. Optimum Subsampling Intensity for sampling sites with less than 500 individuals, for both Family and Genus levels of taxonomic resolution, for each of the eight VSCI metrics. The optimum subsampling intensity value was calculated by determining 95% of the true metric value, with less than 5% error (variation) among the 100 replicates for each ten-individual interval. The y=n line represents the total number of individuals in the sample (n). The metrics where the optimum subsampling intensity is equal to the total number of individuals in the sample, either did not meet the criteria or met in the last possible ten-individual interval.

JM42_02 FAMILY LEVEL (n=470, TMV=22)
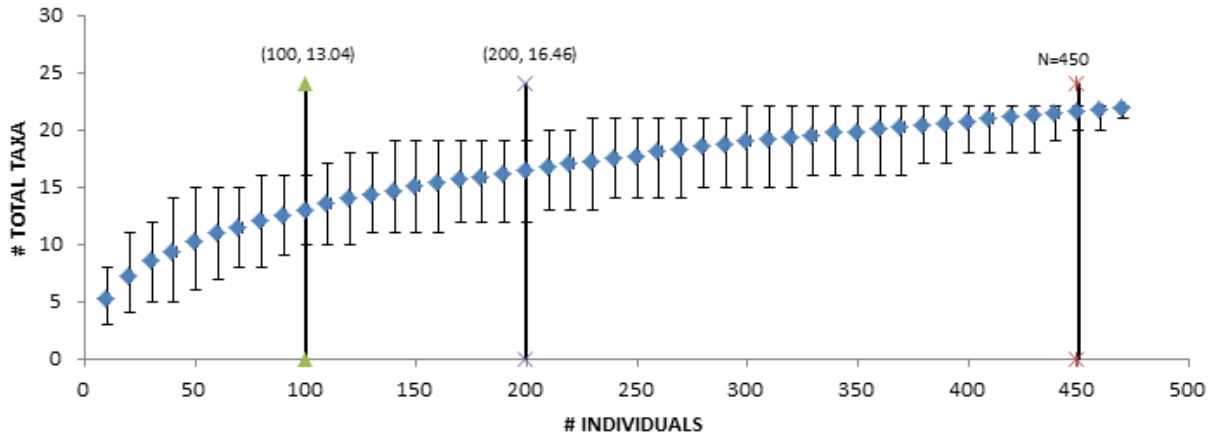
JM42_02 GENUS LEVEL (n=470, TMV=28)
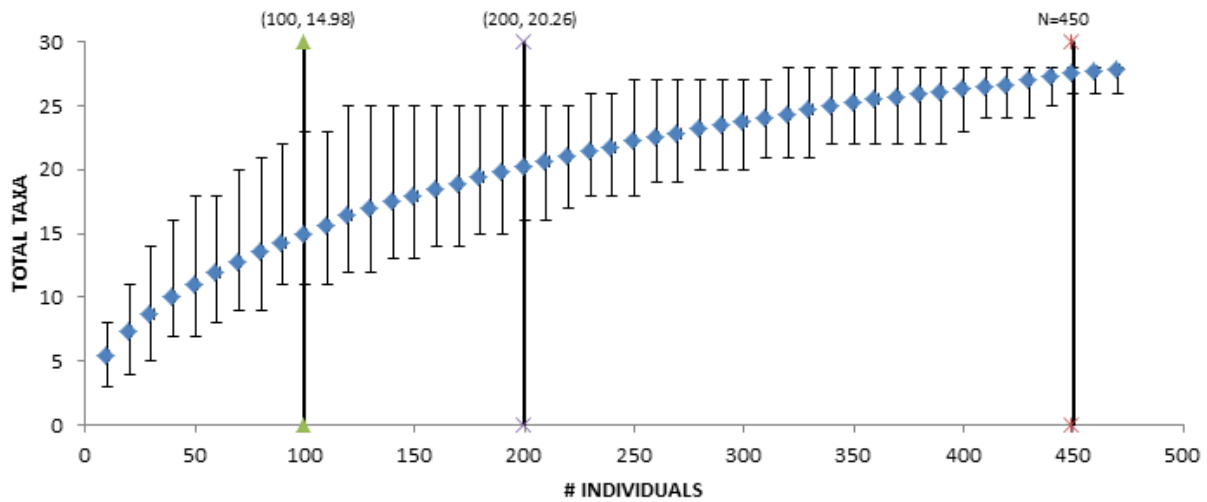
Figure 7. Family and Genus level comparison of the VSCI metric "Total Taxa", calculating the total taxa richness of the sample JM42_02 at ten-individual intervals for the entire sample, where n=sample size and TMV=true metric value. The 100 and 200 individual counts are represented by a vertical line showing the mean metric value at the respective individual count. The optimum subsample size for the metric (95% of TMV at 5% error rate) is represented by "N" and the optimum subsample size is given. For metrics with the optimum subsample size equal to the sample size, (N=n), the criteria were not met or were met in the last ten-individual Interval for the sample. For richness metrics, the error bars represent minimum and maximum values.

JM42_02 FAMILY LEVEL (n=470, TMV=7)

JM42_02 GENUS LEVEL (n=470, TMV=10)
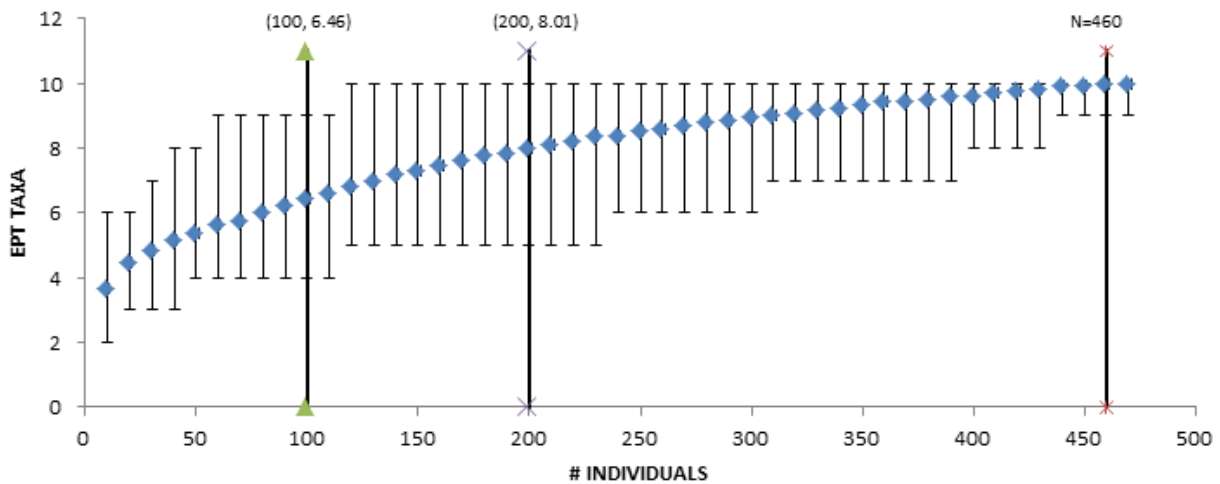
Figure 8. Family and Genus level comparison of the VSCI metric "EPT Taxa", calculating the taxa richness of EPT (Ephemeroptera, Plecoptera and Trichoptera) of the sample JM42_02 at ten-individual intervals for the entire sample, where n=sample size and TMV=true metric value. The 100 and 200 individual counts are represented by a vertical line showing the mean metric value at the respective individual count. The optimum subsample size for the metric (95% of TMV at 5% error rate) is represented by "N" and the optimum subsample size is given. For metrics with the optimum subsample size equal to the sample size, (N=n), the criteria were not met or were met in the last ten-individual Interval for the sample. For richness metrics, the error bars represent minimum and maximum values.

JM42_02 FAMILY LEVEL (n=470, TMV=52.17)

JM42_02 GENUS LEVEL (n=470, TMV=52.24)
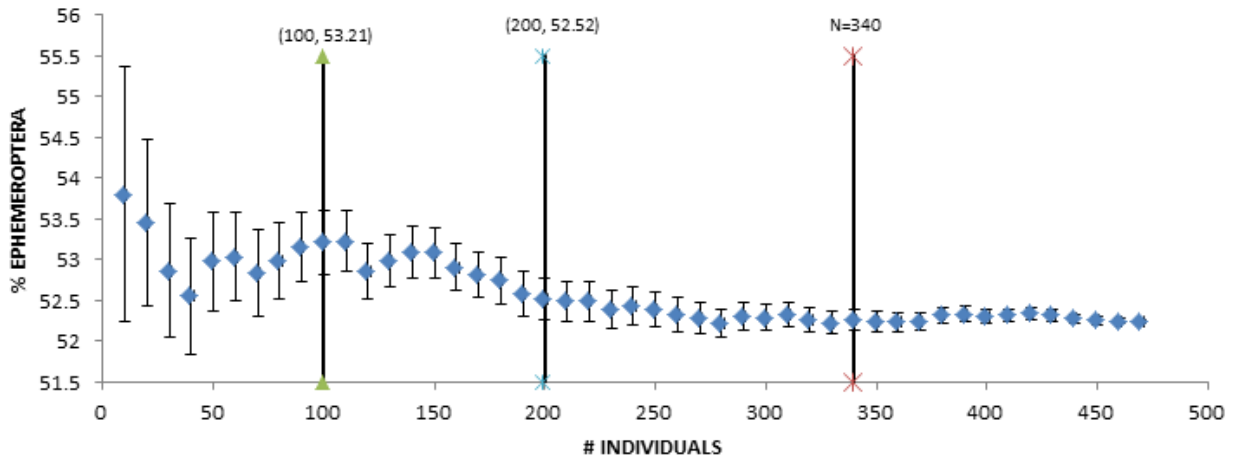
Figure 9. Family and Genus level comparison of the VSCI metric "% Ephemeroptera", calculating the proportion of Ephemeroptera in the sample JM42_02 at ten-individual intervals for the entire sample, where n=sample size and TMV=true metric value. The 100 and 200 individual counts are represented by a vertical line showing the mean metric value at the respective individual count. The optimum subsample size for the metric (95% of TMV at 5% error rate) is represented by "N" and the optimum subsample size is given. For metrics with the optimum subsample size equal to the sample size, (N=n), the criteria were not met or were met in the last ten-individual Interval for the sample. For proportion metrics, the error bars represent standard error.

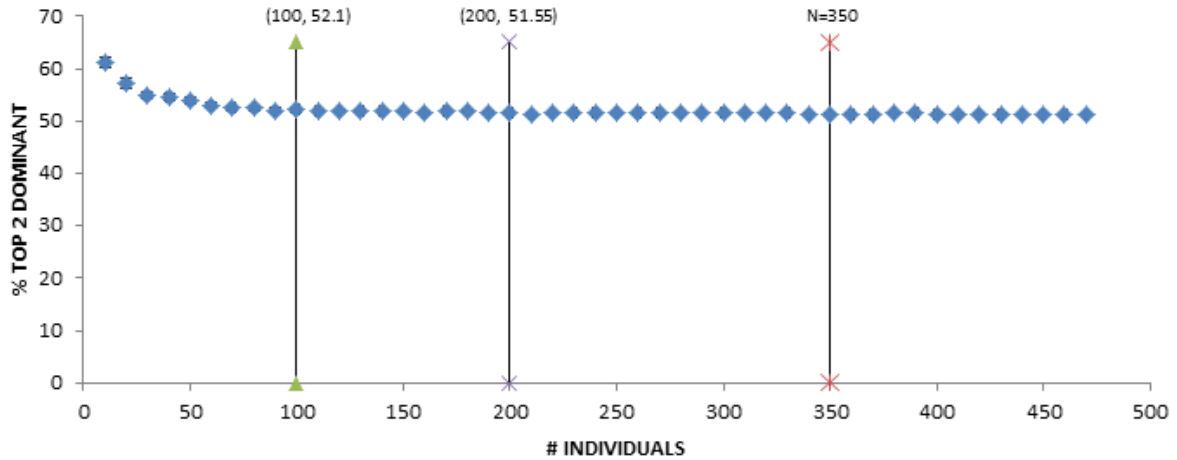## JM42_02 FAMILY LEVEL (n=470, TMV=14.31)

JM42_02 GENUS LEVEL (n=470, TMV=14.33)
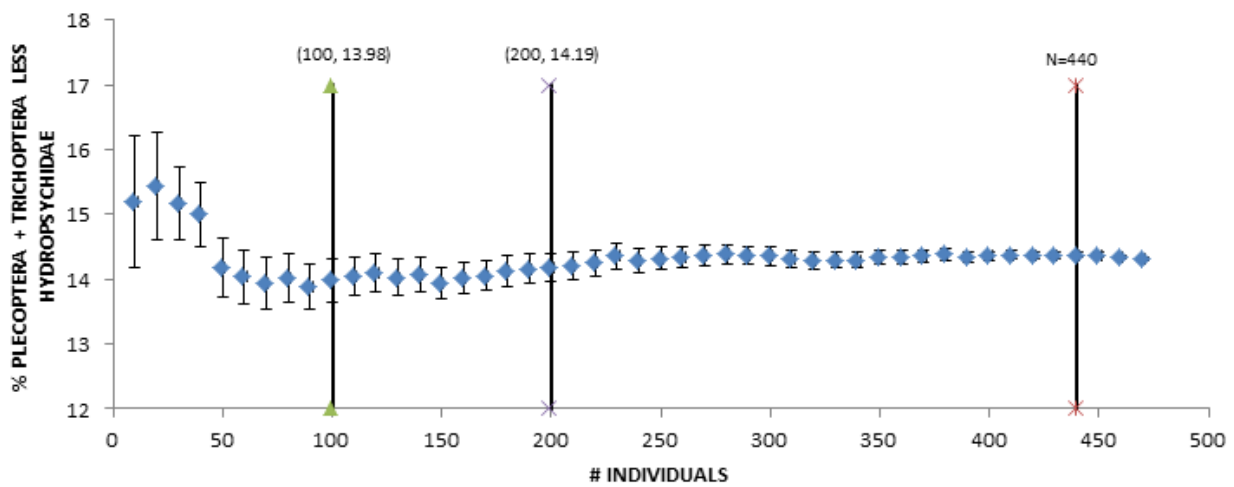
Figure 10. Family and Genus level comparison of the VSCI metric "% Plecoptera plus Trichoptera Less Hydropsychidae" (HP-T), calculating the proportion of HP-T in the sample JM42_02 at ten-individual intervals for the entire sample, where n=sample size and TMV=true metric value. The 100 and 200 individual counts are represented by a vertical line showing the mean metric value at the respective individual count. The optimum subsample size for the metric (95% of TMV at 5% error rate) is represented by "N" and the optimum subsample size is given. For metrics with the optimum subsample size equal to the sample size, (N=n), the criteria were not met or were met in the last ten-individual Interval for the sample. For proportion metrics, the error bars represent standard error.

47

JM42_02 FAMILY LEVEL (n=470, TMV=3.16)

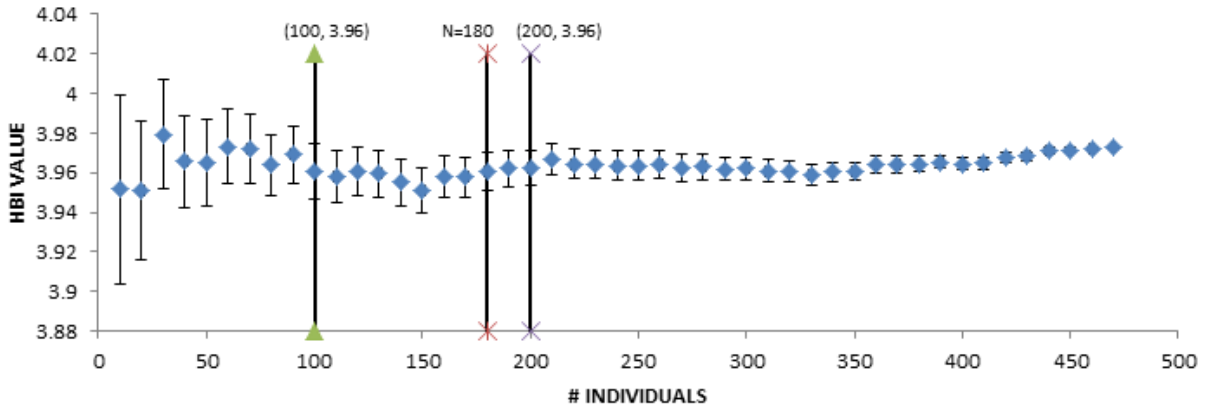
JM42_02 GENUS LEVEL (n=470, TMV=34.99)
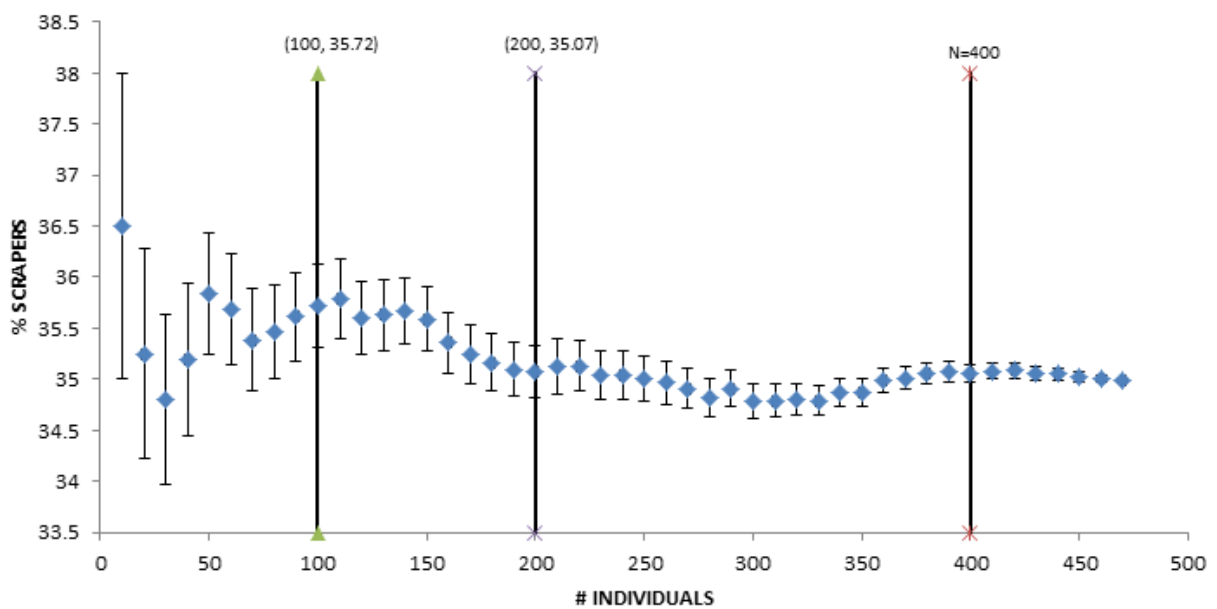
Figure 11. Family and Genus level comparison of the VSCI metric "% Scrapers, calculating the proportion of the functional feeding group "scrapers" in the sample JM42_02 at ten-individual intervals for the entire sample, where n=sample size and TMV=true metric value. The 100 and 200 individual counts are represented by a vertical line showing the mean metric value at the respective individual count. The optimum subsample size for the metric (95% of TMV at 5% error rate) is represented by "N" and the optimum subsample size is given. For metrics with the optimum subsample size equal to the sample size, (N=n), the criteria were not met or were met in the last ten-individual Interval for the sample. For proportion metrics, the error bars represent standard error.

JM42_02 FAMILY LEVEL (n=470, TMV=9.04)

JM42_02 GENUS LEVEL (n=470, TMV=9.04)
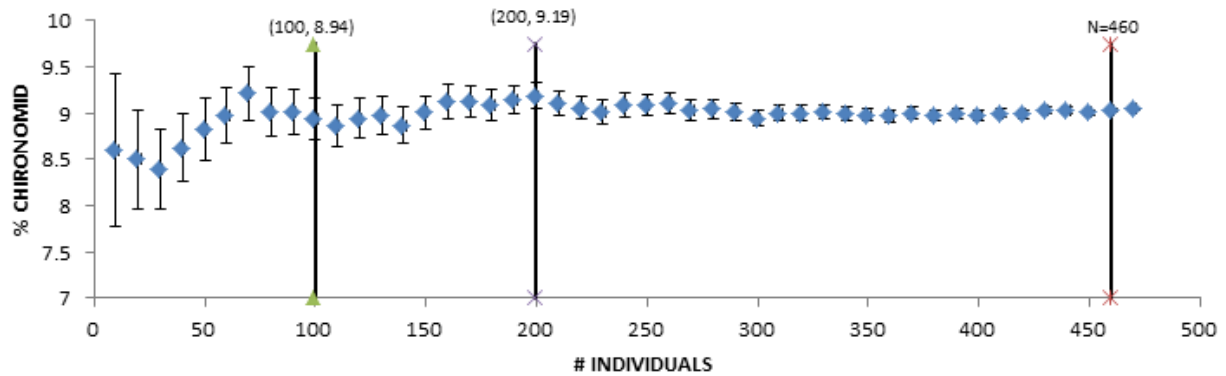
Figure 12. Family and Genus level comparison of the VSCI metric "% Chironomidae," calculating the proportion of Chironomidae in the sample JM42_02 at ten-individual intervals for the entire sample, where n=sample size and TMV=true metric value. The 100 and 200 individual counts are represented by a vertical line showing the mean metric value at the respective individual count. The optimum subsample size for the metric (95% of TMV at 5% error rate) is represented by "N" and the optimum subsample size is given. For metrics with the optimum subsample size equal to the sample size, (N=n), the criteria were not met or were met in the last ten-individual Interval for the sample. For proportion metrics, the error bars represent standard error.

## JM42_02 FAMILY LEVEL (n=470, TMV=51.34)

JM42_02 GENUS LEVEL (n=470, TMV=51.39)
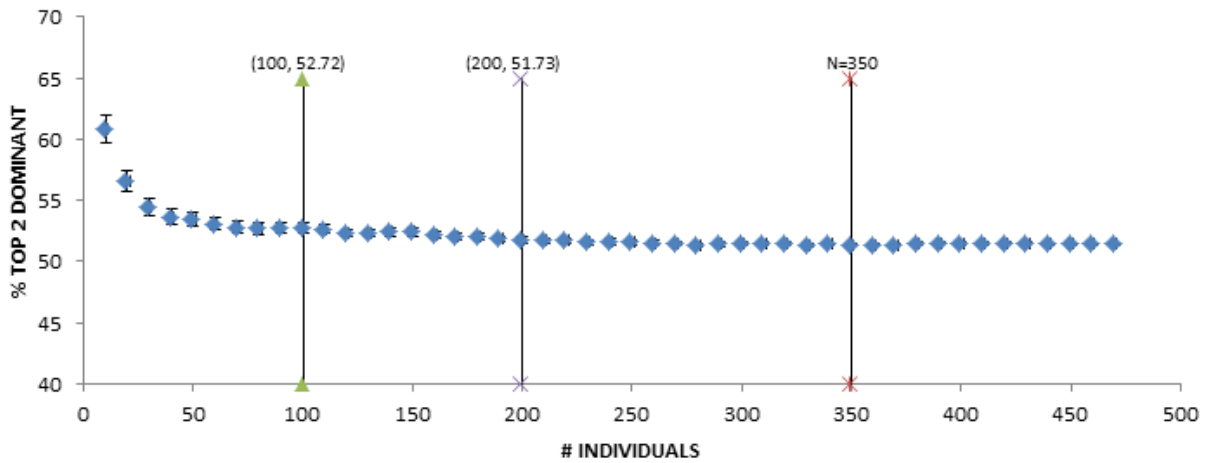
Figure 13. Family and Genus level comparison of the VSCI metric "% Top 2 Dominant", calculating the proportion of the top 2 dominant taxa in the sample JM42_02 at ten-individual intervals for the entire sample, where n=sample size and TMV=true metric value. The 100 and 200 individual counts are represented by a vertical line showing the mean metric value at the respective individual count. The optimum subsample size for the metric (95% of TMV at 5% error rate) is represented by "N" and the optimum subsample size is given. For metrics with the optimum subsample size equal to the sample size, (N=n), the criteria were not met or were met in the last ten-individual Interval for the sample. For proportion metrics, the error bars represent standard error.

JM42_02 FAMILY LEVEL (n=470, TMV=3.97)

JM42_02 GENUS LEVEL (n=470, TMV=4.31)
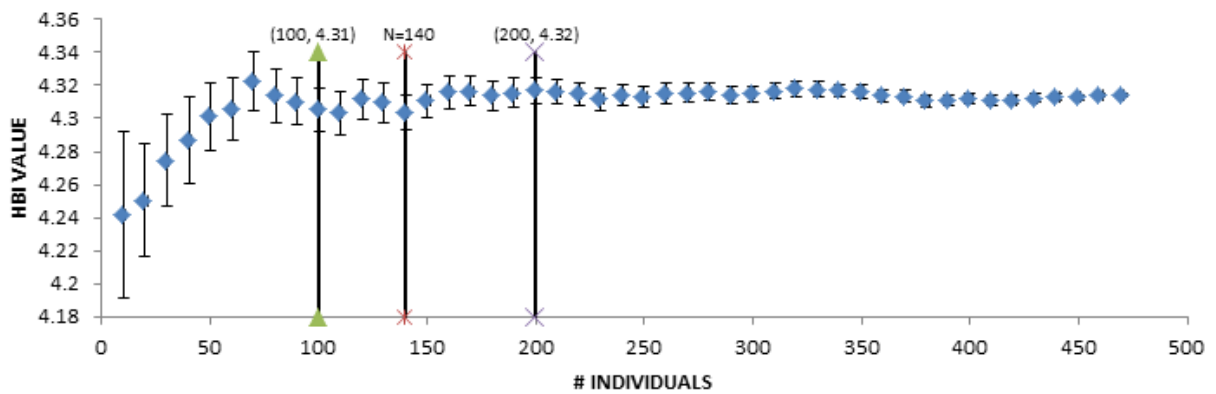
Figure 14. Family and Genus level comparison of the VSCI metric "HBI", calculating the value of the Hilsenhoff Biotic Index for all taxa, based on caluculated tolerance values, for the sample JM42_02 at ten-individual intervals for the entire sample, where n=sample size and TMV=true metric value. The 100 and 200 individual counts are represented by a vertical line showing the mean metric value at the respective individual count. The optimum subsample size for the metric (95% of TMV at 5% error rate) is represented by "N" and the optimum subsample size is given. For metrics with the optimum subsample size equal to the sample size, (N=n), the criteria were not met or were met in the last ten-individual Interval for the sample. For proportion metrics, the error bars represent standard error.

JM52_01 FAMILY LEVEL (n=1640, TMV=37)
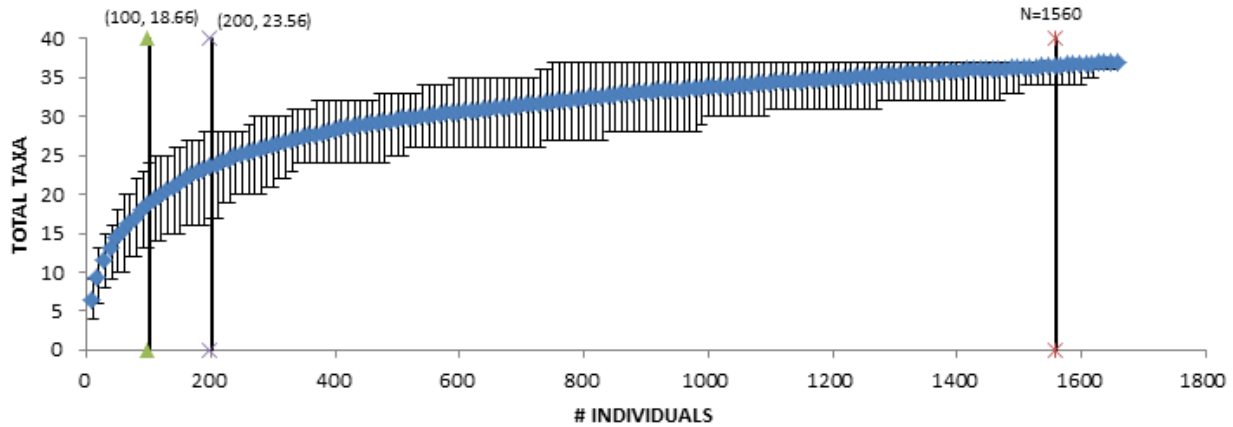
JM52_01 GENUS LEVEL (n=1640, TMV=55)
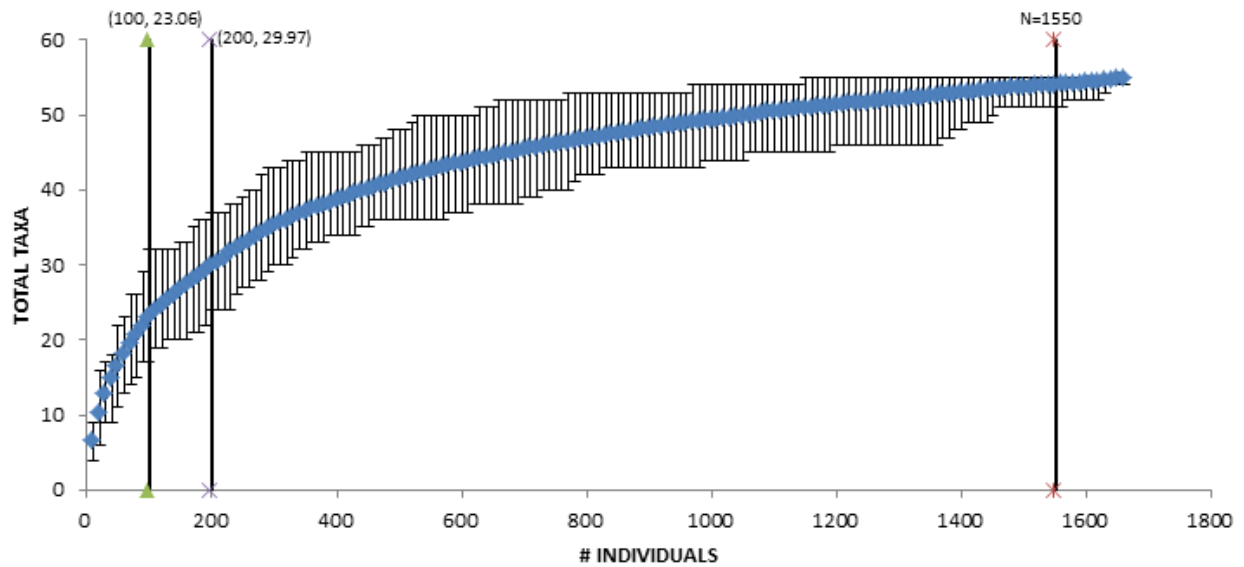
Figure 15. Family and Genus level comparison of the VSCI metric "Total Taxa", calculating the total taxa richness of the sample JM52_01 at ten-individual intervals for the entire sample, where n=sample size and TMV=true metric value. The 100 and 200 individual counts are represented by a vertical line showing the mean metric value at the respective individual count. The optimum subsample size for the metric (95% of TMV at 5% error rate) is represented by "N" and the optimum subsample size is given. For metrics with the optimum subsample size equal to the sample size, (N=n), the criteria were not met or were met in the last ten-individual Interval for the sample. For richness metrics, the error bars represent minimum and maximum values.

JM52_01 FAMILY LEVEL (n=1640, TMV=9)
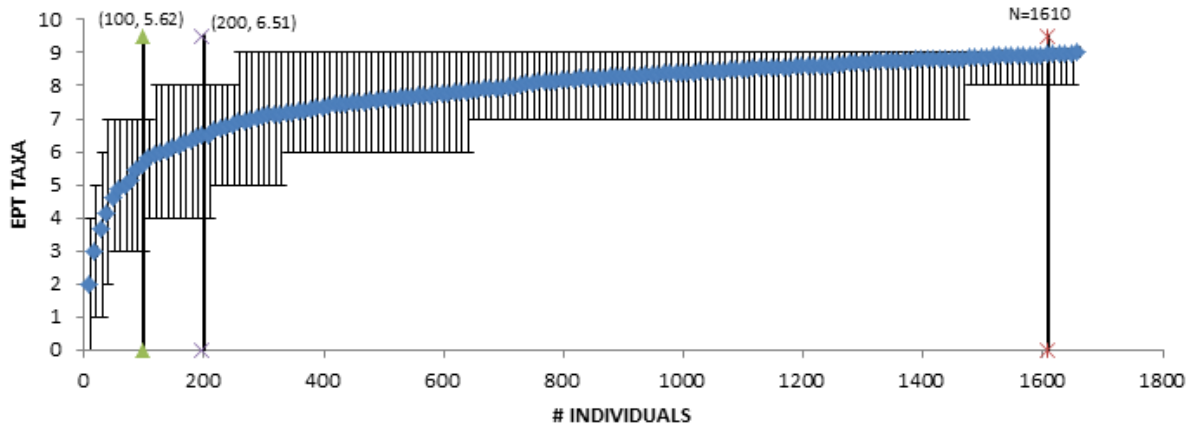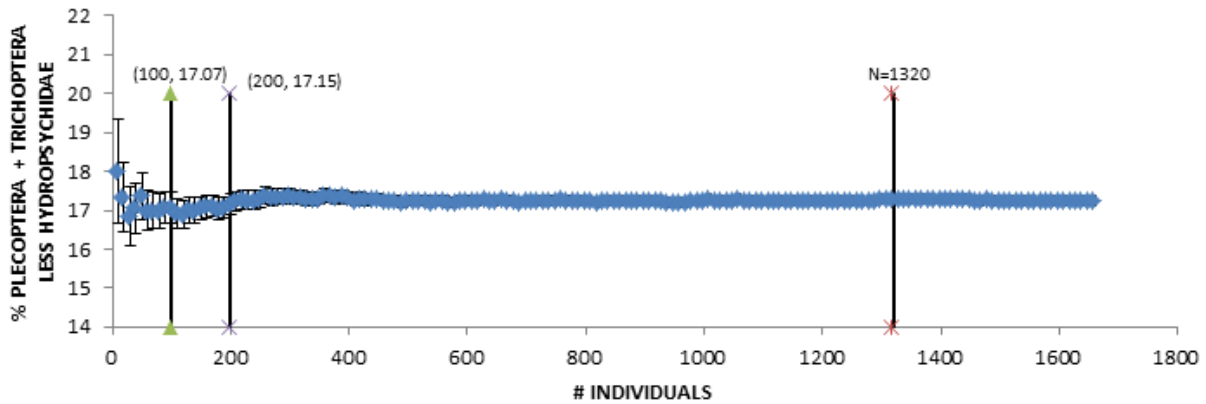

JM52_01 GENUS LEVEL (n=1640, TMV=19)
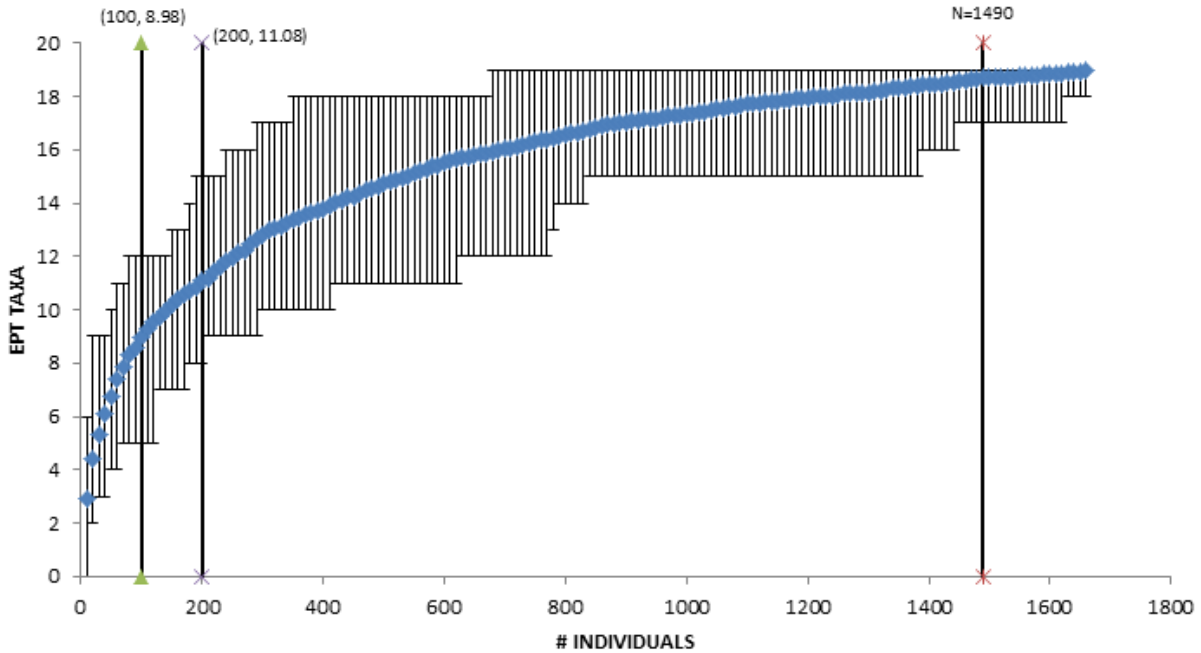
Figure 16. Family and Genus level comparison of the VSCI metric "EPT Taxa", calculating the taxa richness of EPT (Ephemeroptera, Plecoptera and Trichoptera) of the sample JM52_01 at ten-individual intervals for the entire sample, where n=sample size and TMV=true metric value. The 100 and 200 individual counts are represented by a vertical line showing the mean metric value at the respective individual count. The optimum subsample size for the metric (95% of TMV at 5% error rate) is represented by "N" and the optimum subsample size is given. For metrics with the optimum subsample size equal to the sample size, (N=n), the criteria were not met or were met in the last ten-individual Interval for the sample. For richness metrics, the error bars represent minimum and maximum values.

53

JM52_01 FAMILY LEVEL (n=1640, TMV=20.53)

JM52_01 GENUS LEVEL (n=1640, TMV=20.54)
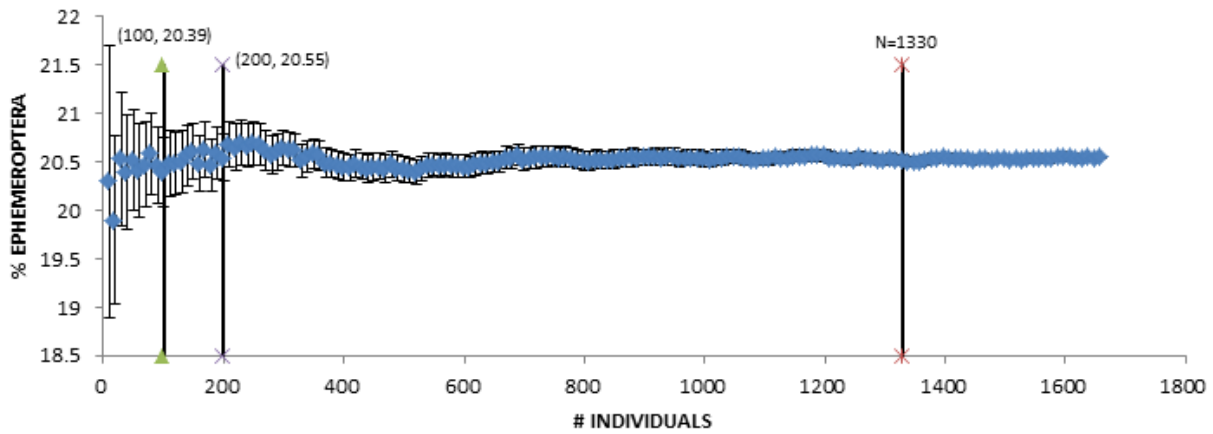
Figure 17. Family and Genus level comparison of the VSCI metric "% Ephemeroptera", calculating the proportion of Ephemeroptera in the sample JM52_01 at ten-individual intervals for the entire sample, where n=sample size and TMV=true metric value. The 100 and 200 individual counts are represented by a vertical line showing the mean metric value at the respective individual count. The optimum subsample size for the metric (95% of TMV at 5% error rate) is represented by "N" and the optimum subsample size is given. For metrics with the optimum subsample size equal to the sample size, (N=n), the criteria were not met or were met in the last ten-individual Interval for the sample. For proportion metrics, the error bars represent standard error.

JM52_01 FAMILY LEVEL (n=1640, TMV=17.24)

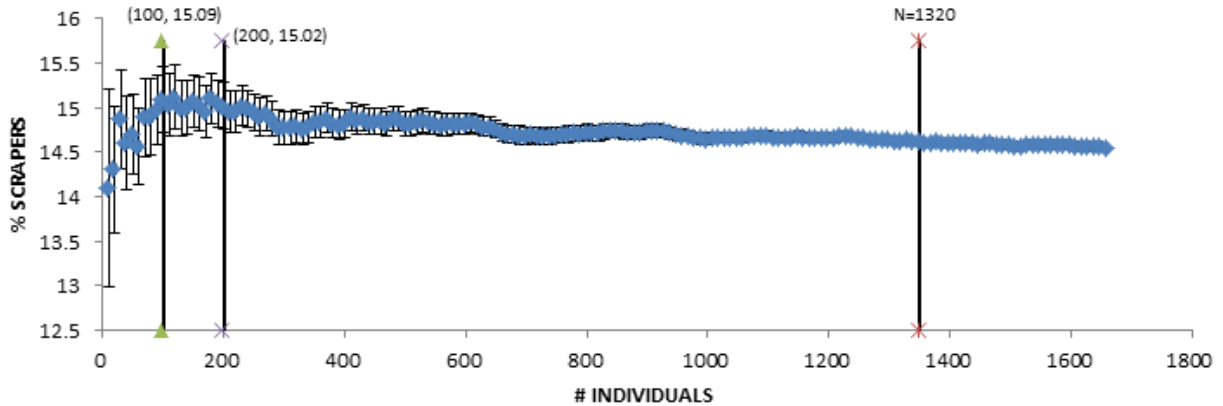**JM52_01 GENUS LEVEL (n=1640, TMV=17.25)**
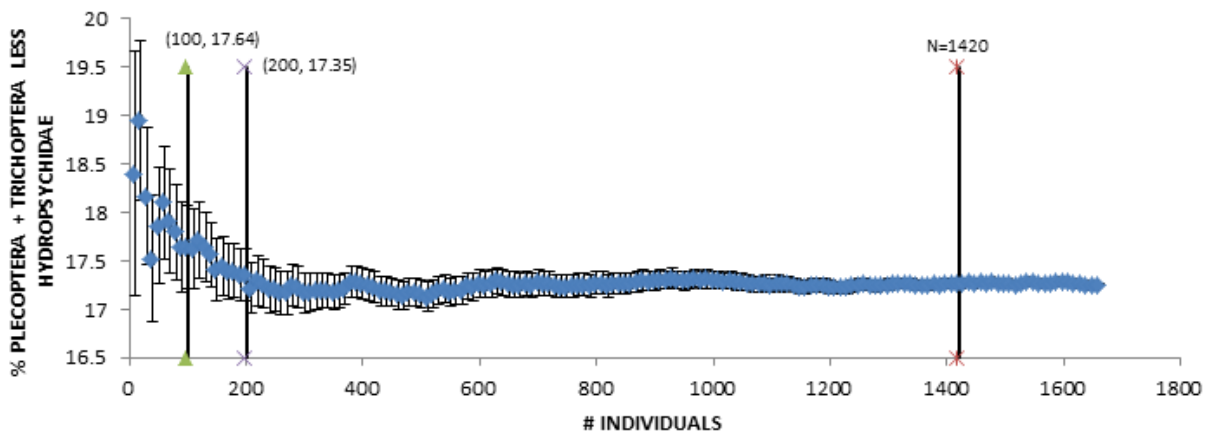
Figure 18 . Family and Genus level comparison of the VSCI metric "% Plecoptera plus Trichophtera Less Hydropsychidae" (HP-T), calculating the proportion of HP-T in the sample JM52_01 at ten-individual intervals for the entire sample, where n=sample size and TMV=true metric value. The 100 and 200 individual counts are represented by a vertical line showing the mean metric value at the respective individual count. The optimum subsample size for the metric (95% of TMV at 5% error rate) is represented by "N" and the optimum subsample size is given. For metrics with the optimum subsample size equal to the sample size, (N=n), the criteria were not met or were met in the last ten-individual Interval for the sample. For proportion metrics, the error bars represent standard error.

JM52_01 FAMILY LEVEL (n=1640, TMV=14.57)

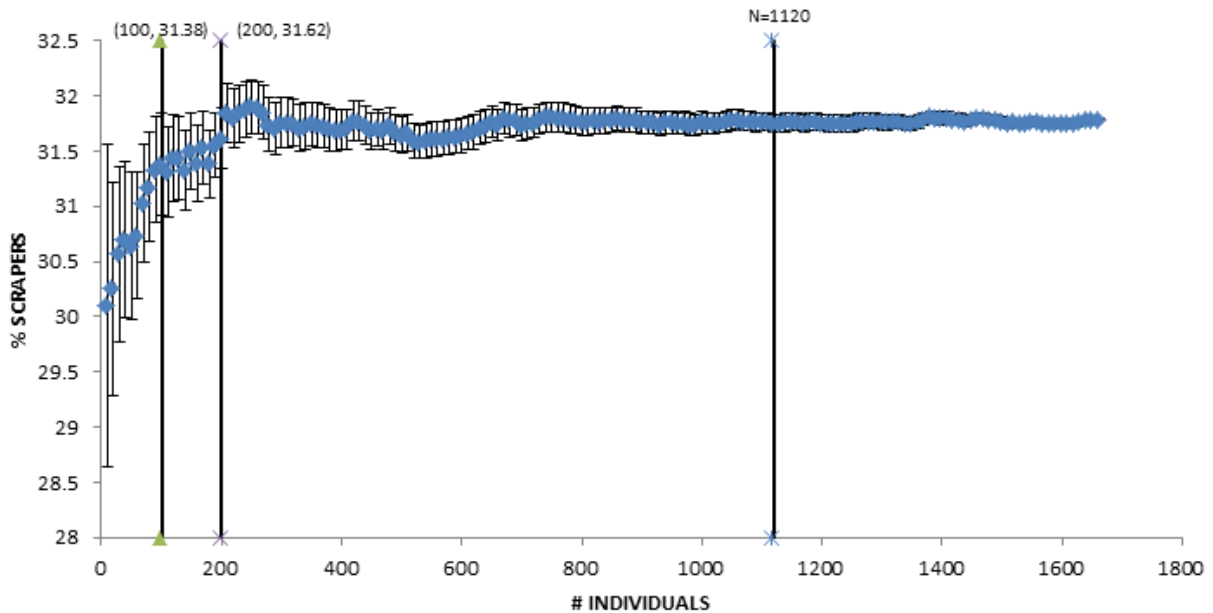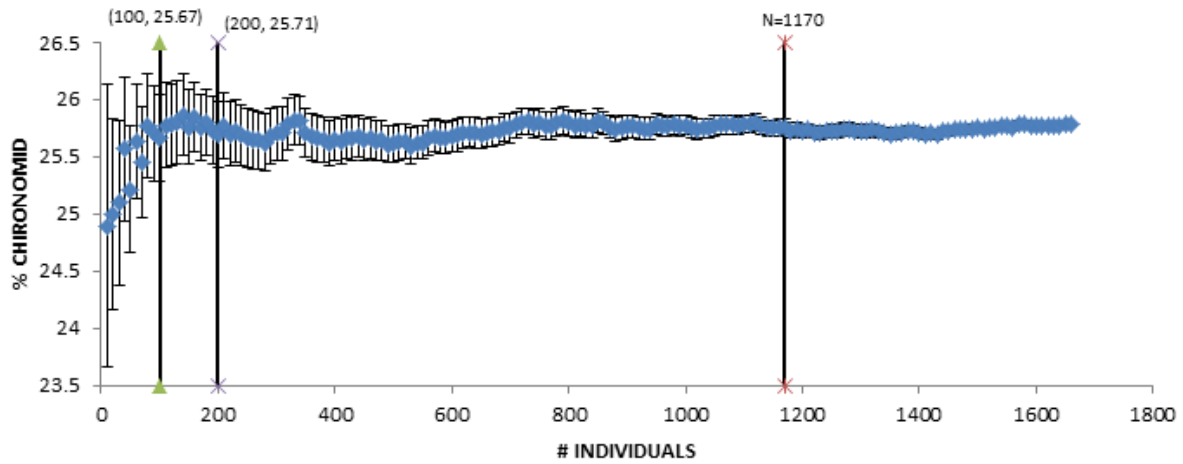JM52_01 GENUS LEVEL (n=1640, TMV=31.78)

Figure 19. Family and Genus level comparison of the VSCI metric "% Scrapers, calculating the proportion of the functional feeding group "scrapers" in the sample JM52_01 at ten-individual intervals for the entire sample, where n=sample size and TMV=true metric value. The 100 and 200 individual counts are represented by a vertical line showing the mean metric value at the respective individual count. The optimum subsample size for the metric (95% of TMV at 5% error rate) is represented by "N" and the optimum subsample size is given. For metrics with the optimum subsample size equal to the sample size, (N=n), the criteria were not met or were met in the last ten-individual Interval for the sample. For proportion metrics, the error bars represent standard error.

JM52_01 FAMILY LEVEL (n=1640, TMV=25.78)
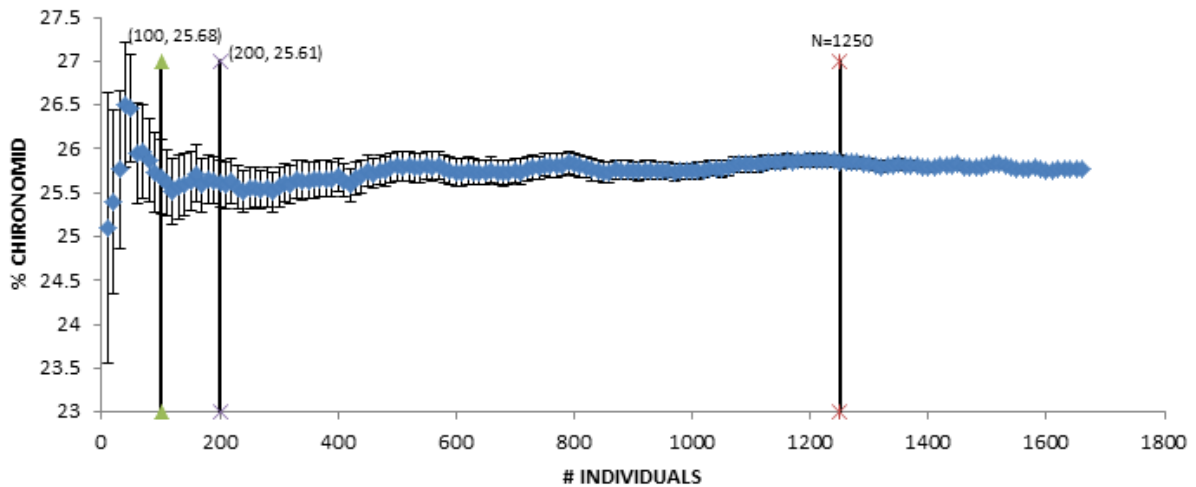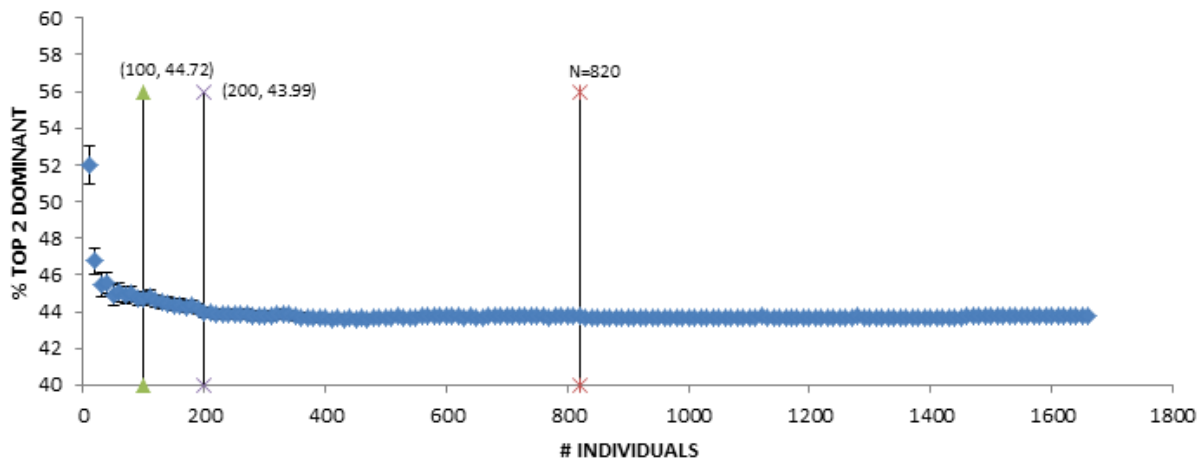

JM52_01 GENUS LEVEL (n=1640, TMV=25.78)

Figure 20 . Family and Genus level comparison of the VSCI metric "% Chironomidae", calculating the proportion of Chironomidae in the sample JM52_01 at ten-individual intervals for the entire sample, where n=sample size and TMV=true metric value. The 100 and 200 individual counts are represented by a vertical line showing the mean metric value at the respective individual count. The optimum subsample size for the metric (95% of TMV at 5% error rate) is represented by "N" and the optimum subsample size is given. For metrics with the optimum subsample size equal to the sample size, (N=n), the criteria were not met or were met in the last ten-individual Interval for the sample. For proportion metrics, the error bars represent standard error.

57

JM52_01 FAMILY LEVEL (n=1640, TMV=43.81)

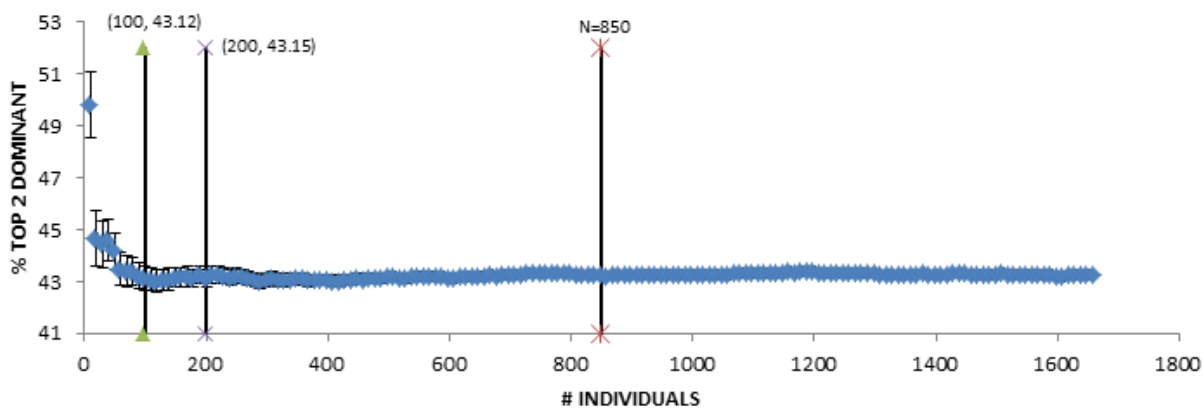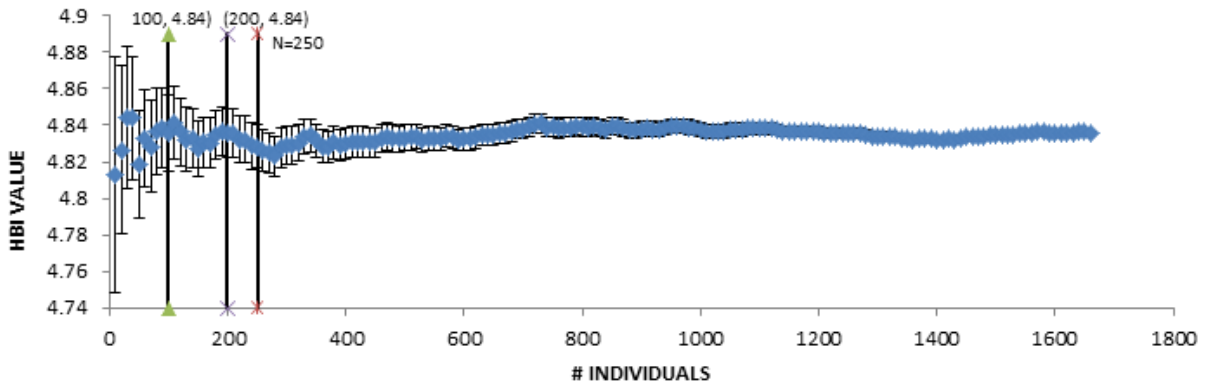JM52_01 GENUS LEVEL (n=1640, TMV=43.25)

Figure 21. Family and Genus level comparison of the VSCI metric "% Top 2 Dominant", calculating the proportion of the top 2 dominant taxa in the sample JM52_01 at ten-individual intervals for the entire sample, where n=sample size and TMV=true metric value. The 100 and 200 individual counts are represented by a vertical line showing the mean metric value at the respective individual count. The optimum subsample size for the metric (95% of TMV at 5% error rate) is represented by "N" and the optimum subsample size is given. For metrics with the optimum subsample size equal to the sample size, (N=n), the criteria were not met or were met in the last ten-individual Interval for the sample. For proportion metrics, the error bars represent standard error.

JM52_01 FAMILY LEVEL (n=1640, TMV=4.84)

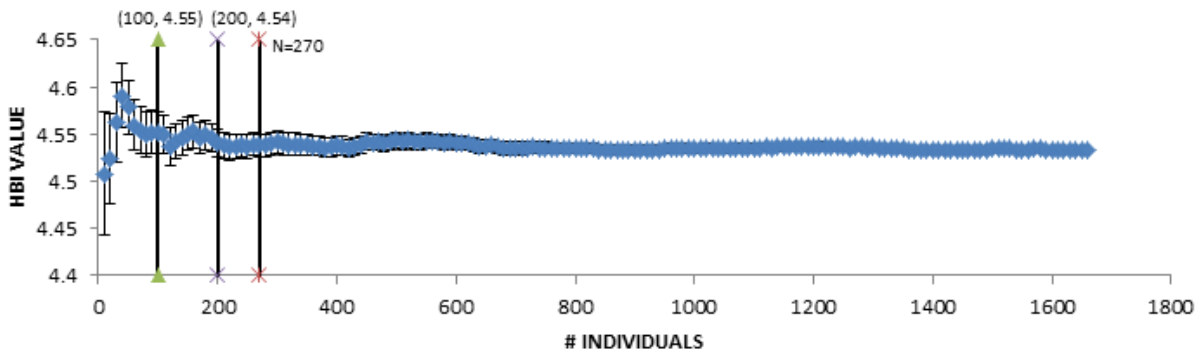## JM52_01 GENUS LEVEL (n=1640, TMV=4.53)

Figure 22. Family and Genus level comparison of the VSCI metric "HBI", calculating the value of the Hilsenhoff Biotic Index for all taxa, based on caluculated tolerance values, for the sample JM52_01 at ten-individual intervals for the entire sample, where n=sample size and TMV=true metric value. The 100 and 200 individual counts are represented by a vertical line showing the mean metric value at the respective individual count. The optimum subsample size for the metric (95% of TMV at 5% error rate) is represented by "N" and the optimum subsample size is given. For metrics with the optimum subsample size equal to the sample size, (N=n), the criteria were not met or were met in the last ten-individual Interval for the sample. For proportion metrics, the error bars represent standard error.
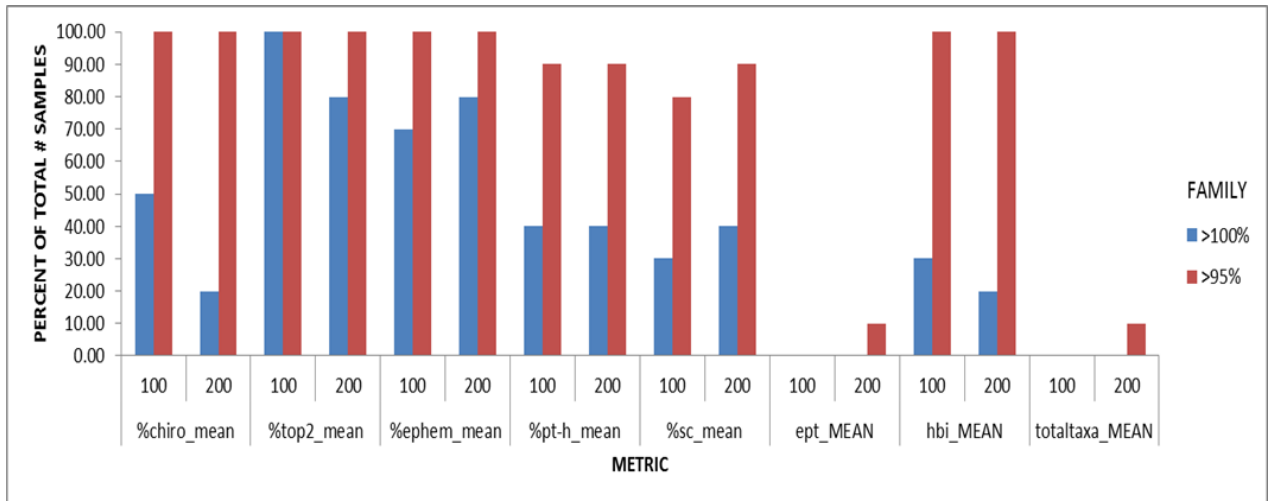
Figure 23. Percent of samples of the 10 samples picked in entirety, at the Family level of taxonomic resolution, where the mean metric values at 100 and 200 individual counts are over 100% of the True Metric Value, and percent of samples are over 95% of the true metric values for each of the eight VSCI metrics. Mean metric values over 100% indicate an overestimate of that metric in the VSCI score.
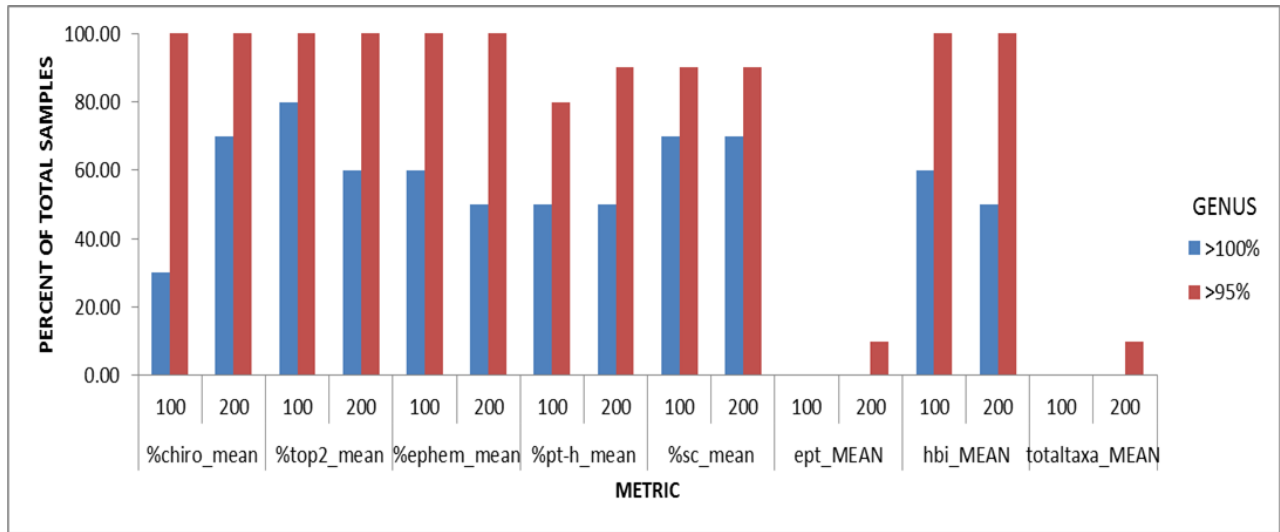
Figure 24. Percent of samples of the 10 samples picked in entirety, at the Genus level of taxonomic resolution, where the mean metric values at 100 and 200 individual counts are over 100% of the True Metric Value, and percent of samples are over 95% of the true metric values for each of the eight VSCI metrics. Mean metric values over 100% indicate an overestimate of that metric in the VSCI score.

Vita

Laurel Cary Williams was born December 8, 1976 in Richmond, Virginia, and is a U.S. citizen. Laurel graduated from Manchester High School in 1994 and received a B.S. in Biology from Virginia Commonwealth University in 2004. After receiving her degree, Laurel was employed at two civil engineering firms, where she worked as an environmental consultant and wetland and stream scientist for five years before entering the M.S. program in Biology at Virginia Commonwealth University.