Draft Report

Comparability Analysis of Benthic Macroinvertebrate Methods in Montana

Prepared for:

Montana Department of Environmental Quality
Helena, Montana

Prepared by:

Tetra Tech, Inc.
400 Red Brook Blvd., Suite 200
Owings Mills, Maryland 21117-5159

Technical support and direction provided by
U.S. Environmental Protection Agency
EPA Task Order Manager: Danette Quick
EPA Technical Reviewer: Tina Laidlaw

July 11, 2005

TABLE OF CONTENTS

1. Introduction

EPA advocates use of biological criteria and development of numeric biocriteria to assist states in decision-making and management for 305(b) reporting, 303(d) lists, TMDL development, and watershed restoration. Biological assessments use characteristics of the structure and function of biological communities as indicators of stream quality, and different quantitative thresholds as the basis for water quality management decisions. Bioassessments and biocriteria can be used to list and de-list waterbodies on CWA §303(d) lists to assess the effectiveness of TMDL control measures to prioritize streams and watersheds for restoration and/or protection, and to evaluate the effectiveness of stressor control activities.

Montana's Department of Environmental Quality (DEQ) has collected biological data (periphyton and macroinvertebrates) for more than twenty years. Multimetric indices were originally developed for various regions of the state using relatively unstructured approaches or best professional judgment (citations?). Later, the Foothill/Valley/Plains index was revised based on discriminant analysis results (Rhithron Biological Associates 199x). Given the application of biological data in the 303(d) and TMDL process and the statewide datasets now available, a critical evaluation of the State's biological sampling, analysis, and assessment methods has identified several programmatic needs.

Through the 2004 workload planning process, EPA and DEQ found several biomonitoring and bioassessment elements that, once addressed, will enhance the State's ability to improve impairment determinations. These elements are related to recognition of natural variability, reduction of sampling error, and identification of equipment and protocols that will either allow more precise measurements or that will yield comparable (similar) results. This report addresses DEQ's consistency issues related to impairment decisions, and will support interpretation of biological data, suggest opportunities for protocol/equipment improvement, and recommend techniques for combining or segregating data sets for analyses.

Two studies were designed to answer questions related to sampling equipment and sampling protocols. The first study analyzed the effects of sampling equipment mesh size on assessment techniques. The second study considered several of the sampling protocols that have been used throughout Montana; methods that use different equipment, target different habitat types, and specify different laboratory processing. The overall goal of these analyses is to determine the extent to which data collected by these mesh and protocol variants can be combined within a single dataset and subjected to biological assessment procedures.

The three primary purposes of this project are to:

1) Determine the quality and variability of the data generated by the various sampling equipment and protocols, so that sources of variability and error can be reduced in future sampling efforts.

2) Determine the comparability of data generated by the various sampling equipment and protocols, so that assessments and analyses can proceed with either composite or distinct data sets, depending on the outcome of the comparison, and

3) Determine data manipulations that can be made to adjust for bias and to allow comparisons among datasets collected with different equipment or protocols (e.g., rarefaction to a standard target size for sub-samples of various sizes).

Although side-by-side sampling in identical stream locations (as in this project) can provide useful information on inter-protocol comparability, it is not as important as knowledge of data quality, or performance characteristics (Stribling and Diamond 2004). Two or more methods can be judged the same (or even identical) based on evaluation of written standard operating procedures (SOP); knowledge of SOP is critical for representing how a method is supposed to be applied. How that method is actually applied is described by precision, accuracy, bias, representativeness, and completeness. Side-by-side designs for comparability analyses provide only a limited suite of potential outcomes: methods can be judged the same and have similar results or different results, methods can be judged different and likewise have similar or different results. It is, thus, inadequate to rely on whether or not the same, site-specific final assessments are attained with two or more protocols.

Data quality is defined as the magnitude of error associated with a dataset (Peters 1988). To be able to make any judgments on the comparability of multiple protocols, methods, or the datasets they produce, the magnitude of error associated with each must be known. Individual components of protocols are potential error sources that contribute to the overall error (or, variability) of a dataset. In these analyses, we partition the biological assessment process into a series of six methods: field sampling, laboratory sorting and subsampling, taxonomic identification and enumeration, data entry, metric calculation, and site assessment. For each of the methods we examine a series of five performance characteristics, including precision, accuracy, bias, representativeness, and completeness (Table 1).

Precision is defined as the nearness of multiple measures of the same property (Taylor 1988, Taylor and Kuyatt 1994); it is a characterization of consistency or repeatability and can be accomplished in two manners. First, a property can be measured by a single method repeatedly; or a property can be measured by multiple methods. For either approach, differences are characterized by standard error, coefficient of variability, relative percent difference, or other algorithms. Accuracy is the nearness of a measurement to a true value, or analytical truth (Taylor 1988, Clark and Whitfield 1994, Taylor and Kuyatt 1994). Any statement about the accuracy of a measurement system requires that the analytical truth be defined; this can be one of the most difficult aspects of defining the performance of biological assessment protocols. In spite of this conceptual conundrum, it can be addressed by defining the analytical truth in terms of the overall objectives of the initial indicator development process and the monitoring and assessment program. Bias is an indication of error that exists even when a method is consistently applied (Smith et al. 1988, Clark and Whitfield 1994) and it is defined as method error caused by systematically favoring some outcomes over others. Completeness is a measure of the number of valid data points gathered relative to the number of planned data points (Smith et al. 1988). For biological assessments, this performance characteristic is usually evaluated as the number of samples meeting methods specifications and that can be used for analyses. However, depending on the specific component method and the scale of its use, how this value is determined can vary.

Representativeness is typically a qualitative statement about the property a measurement is intended to depict, and whether and how well it actually does represent it.

In this study, our focus is on evaluation of the comparability of benthic macroinvertebrate data produced by two different mesh sizes (500 and 1200μm), and four different field protocols (traveling kick, Surber sampler, EMAP riffle samples, EMAP reachwide samples).

2. Methods

Methods are described below in two parts. The first part covers the performance characteristics that generally apply in the second part, where the specifics of the comparison studies are presented. Performance characteristics are tested and described in each data set using multiple analysis techniques and qualitative criteria. Relevant MQOs are presented with the performance characteristics. The second part of Section 2 includes descriptions of the two study designs. One study examines effects of mesh size on samples and biological assessments and the other examines the effects of sampling protocols.

2.A. Analysis of Performance Characteristics

Comparisons among sampling equipment, protocols, and laboratory processes were made in the context of methods-based performance characteristics. These characteristics include precision, accuracy, bias, representativeness, and completeness (Table 1). Most of the analytical focus is on precision and bias because these are the characteristics that are quantifiable and are applicable throughout several components of the biological monitoring protocols, and for which we have sufficient analytical data. The representativeness of data and information produced by each method is qualitatively characterized by consideration of what the method is intended to do or depict. Thus, for representativeness of samples (or data) produced by a particular field gear or mesh size, the question is asked "What is the sample intended to represent, and are those expectations met?" Data comparability and recommendations regarding data pooling are based on a weight of evidence approach, where several analyses and performance characteristics are considered simultaneously.

2.A.i. Precision

Precision, agreement of multiple measures, was primarily measured in the context of analysis of variance (ANOVA). Variance within groups was assessed to discern the magnitude of variance that is meaningful between groups. We set up the ANOVA such that groups are defined by the sites for which multiple measures were recorded. The mean square error (MSE) term from ANOVA is an estimate of variance within the groups. From MSE, standard deviation within groups can be calculated as the root MSE (RMSE), which is lower when measures are more precise. Likewise, the coefficient of variation (CV) can be calculated, which standardizes variability on the mean of measures (CV = RMSE/mean), allowing comparison of relative precision among metrics and among treatment groups (Diamond et al. 1996).

Using the RMSE and CV, precision of metrics and indices can be quantified within and among mesh sizes and protocols. We expect that variability within groups using the same mesh size or protocol would be less than the variability among the groups. The difference in variability (within vs. among) indicates positive effects (of mesh size or protocol) when the difference is of higher magnitude and consistently higher with all metrics and indices.

CV can be calculated for metrics of sample pairs of different mesh sizes using ANOVA. This CV can be compared to the CV for replicates using the same mesh size to determine whether samples collected with different mesh sizes result in metric variabilities that are greater than those expected based on sample replication alone.

Detectable Difference

The detectable difference (DD) is the likely range around the observed mean that the true mean is likely to fall. For this analysis the significance level of the DD was chosen to be 90%, that is, the range within which the true mean is likely fall, 90% of the time. DD is calculated from RMSE using the equation:

$$DD = \pm RMSE \times t\alpha$$

where $t\alpha$ is the 90% DD value (i.e., $p = 0.10$) from a standard t-table (Zar 1999), which in this analysis equals 1.64. A smaller DD for a method indicates more precise data.

Cluster Analysis and Ordination

Cluster analysis was conducted on the raw and relativized macroinvertebrate taxonomic abundance data of the mesh study sites. The cluster analysis used flexible beta linkage ($\beta = -0.25$) (McCune and Grace 2002, Hawkins and Norris 2000) and Bray-Curtis distance measure. The cluster analysis was interpreted for the cluster groups that best described the relationship of the communities of the different mesh sizes or protocols within streams and between streams of differing ecoregions. We expect that replicate samples (same site and treatment) will be adjacent in the cluster diagram, with shorter linkage distances between them than the distances between samples collected from different sites or with different protocols or equipment.

In ordination diagrams, samples with similar compositions of taxa will plot close together. Those with dissimilar compositions will be further from each other. Precision can thus be assessed qualitatively by observing the position of replicates from the same site and treatment relative to samples from different sites and treatments. We expect that samples collected at the

same site and with the same methods or mesh sizes will plot closer in ordination space than samples from different sites or collected with different protocols or mesh sizes.

The location of samples in the ordination diagram is related to the axes, which are also related to taxa presence and abundance. Samples with a predominance of taxon X will generally plot together and in one region of the diagram. Thus, ordination can be used to detect both precision and bias in the types of organisms that are in the samples. Additional details of the ordination analysis are presented in Section 2.A.iii, Bias.

Percent Difference in Enumeration

Final specimen counts for samples are dependent on the taxonomic identifications, not the rough counts obtained during the initial sorting activity. Comparison of counts uses Percent Difference in Enumeration (PDE) (Stribling et al. 2003), calculated as:

$$PDE = \left(\frac{Lab1 - Lab2}{Lab1 + Lab2} \right) \times 100$$

Lab1 and Lab 2 refer to the counts of organisms identified by each taxonomist. The MQO for PDE is less than or equal to 5%. Samples with PDE ≥ 5% should be further examined for patterns of error.

Percent Taxonomic Disagreement

In each study, 10% of the samples were re-identified by a second laboratory, independent of the first. Side-by-side comparison between the taxonomic results delivered by the two labs was performed. The process entailed examination of the taxa list for each sample and the number of organisms each lab identified for each taxon. For each sample, the number of disagreements was determined, divided by the number of comparisons, and multiplied by 100 to give percent taxonomic disagreement (PTD) (Stribling et al. 2003). PTD was calculated as:

$$PTD = \left[1 - \left(\frac{comp_{pos}}{N} \right) \right] \times 100$$

where $comp_{pos}$ is the number of agreements, and N is the total number of organisms in the larger of the two counts. The lower the PTD value, the more similar are sample taxonomic results, and the greater is the overall taxonomic precision. The MQO for taxonomic disagreement is that the PTD should be ≤ 15%. Differences in individual sample comparisons are evaluated for patterns of disagreement. In cases of consistent disagreement, the causes of disagreements were determined and corrective actions were taken.

2.A.ii. Accuracy

Accuracy, or nearness of a measure to a known, specified analytical truth, requires specification of an analytical truth (Taylor 1988, Clark and Whitfield 1994). For many of the measures we

considered, truth is unknown. In the field and some laboratory processes, we are only working with samples, not a complete census, so we do not know the make-up of the true macroinvertebrate assemblage. A census is not feasible (or necessary) for routine biological monitoring activities. In the case of field samples and some other protocol components, we can only assess accuracy qualitatively.

For taxonomic identifications, analytical truth is defined by 1) the most up-to-date technical literature and keys, 2) an identified reference collection verified by specialists in different taxonomic groups, or 3) specimen by specimen comparison with museum-based type specimens. In this project, a primary laboratory performed all laboratory processing and identification and a secondary laboratory performed re-identification on 10% of samples. All taxonomy in this project was completed using current technical literature. There is currently no reference collection for this project, though the individual labs performing taxonomy have their own reference collections. Type comparison is not feasible, nor considered necessary, for routine monitoring programs.

Accuracy can be assessed for routine metric calculations and for data entry. Metric calculations were checked by hand calculating all the metrics in a random subset of the samples. In this study, data were transferred electronically from spreadsheets provided by the taxonomic laboratory into the database used for metric calculation (EDAS). Accuracy of the data transfer was spot check on a taxon-by-taxon basis and simple metric calculations were compared between the source and the database.

2.A.iii. Bias

Bias is the tendency to systematically favor one outcome over another. Bias was assessed in the entire benthic macroinvertebrate community through ordination, clustering, and indicator analysis using comparisons of sample taxa lists. Bias was also assessed at the metric and index level, where summary measures of the community were calculated based on taxa attributes and compared among samples collected using different mesh sizes or protocols.

For the community level comparisons, PC-ORD (MjM Software Design 2002) was used for ordination, cluster, and indicator species analyses. Because taxonomic resolution was not completely consistent, final taxonomic units for these analyses were decided based on the predominance of identifications in the samples ("data stuffing"). Some taxa were agglomerated (lumped) at higher levels and others were disregarded at the higher level, depending on the process that would result in the least loss of data.

Ordination Analysis

Non-metric multidimensional scaling (NMS) was used to discern patterns in assemblage composition in relation to sites and treatments (mesh size or protocol). NMS has been shown to be the most generally effective method of ordination for community data (McCune and Grace 2002). The ordination was performed using Bray-Curtis distance measure on the relative abundance data and Jaccard's similarity measure for ordination of presence/absence data (McCune and Grace 2002). The number of dimensions to view the data in ordination space was

determined by evaluating the NMS stress of various solutions (McCune and Grace 2002). There has not been a fixed statistical criterion developed for selecting the appropriate number of dimensions (Kruskal and Wish 1978), but it has been shown that stress of 20 or below indicates a stable resulting solution (McCune and Grace 2002).

The results of the ordination were assessed to determine whether samples were more similar within sites than they were within treatments. If a shift occurs between treatments, a consistency in direction of the shift within site groups in ordination space would indicate bias due to the treatment.

Indicator Species Analysis

Indicator species analysis (ISA) was conducted on the groups defined by treatments to find significant indicators taxa that occur more frequently and with greater relative abundance in one or another mesh size or protocol. ISA calculates the frequency and relative abundance of a particular species within each group (Dufrene and Legendre 1997). ISA highlights significant indicator species, based on Monte Carlo simulations of frequency and relative abundance of taxa in community groups. ISA will list these species as being a part of the group, but not a significant indicator of the group. The significance of indicators is determined through Monte Carlo simulations. Any significant indicators (with $p<0.05$ unless otherwise noted) were noted from the ISA as indicators of that particular treatment.

Metric and Index Bias

Bias in metrics was assessed by plotting metric values obtained using one protocol or mesh size against values obtained using the other. Consistent bias is revealed in these plots when a large percentage of points fall to one side or the other of unity (the 1:1 line). We can estimate statistical significance of perceived bias using a Chi-square test of those points above and below unity (Manly 2001). Because we can measure precision using multiple measures within protocols or mesh sizes, we can discount differences between protocols or mesh sizes that are less than the detectable difference (and can be attributed to sampling error and natural micro-site variability). The range around an observation in which we expect to find the true mean in 90% of cases is determined by the RMSE of repeated measures (see Precision: Detectable Difference). We bracket the unity line with the 90% detectable difference and consider those points within the offsets as "ties" when assessing protocol or equipment bias. In the Chi-square test, ties are evenly distributed among the above and below groups.

Percent Sorting Efficiency

In laboratory sample processing, all specimens should be removed from the field sub-sample residue. The accurate enumeration of a sample depends on the completeness of sorting. Bias of sub-sampling is evaluated using a measure similar to percent recovery used in analytical chemistry laboratories, called percent sorting efficiency (PSE) (Barbour et al. 1999, Hill et al. 2004). After the initial sorting effort, all sort residue was rechecked by qualified/experienced sorters. The number of missed organisms recovered in the sort residue from the initial sorting was used to calculate sorting efficiency, using the following formula:

$$PSE = \left(\frac{a}{a+b} \right) \times 100$$

where *a* is the number of organisms originally sorted and *b* is the number of organisms recovered in the QC check. The laboratory sorting/subsampling MQO is to have a dataset where 10% of the samples have a PSE of ≤ 90%. An independent laboratory performed sort residue re-checks on a randomly selected 10% of the samples.

2.A.iv. Representativeness

Representativeness, or ability to collect a representative sample of a population, was only addressed qualitatively in this report. We considered the purpose of the sampling program and the intention of the sample: What is the sample intended to represent? If the samples were intended to represent similar components of the benthic assemblage, from similar habitats, during similar conditions, etc., then we would expect the samples to be comparable among equipment or protocols. Representativeness of the sample relates to program design and can thus address elements of the sampling design that are difficult to test experimentally. For instance, we can examine the intended stream type (size, gradient, predominant substrate) among sampling programs to determine the applicability of all sampling protocols at any particular stream. Likewise, comparability of sampling season, taxonomic resolution, or sub-sample target size can be examined qualitatively to determine similarities or differences among the protocols.

2.A.v. Completeness

Completeness is measured as the degree to which a sampling protocol was followed within a data set. For instance, field sampling data would be considered complete if one hundred percent of the sampling effort was utilized in each of the streams sampled. Data entry would be complete if all the information collected in the field was transferred into electronic format in the database. Completeness should be at or near 100% for all procedures specified in the sampling protocols.

2.B. Study Designs: Mesh Size and Protocols

2.B.i. Mesh Size Study

A study was designed to determine the effect of D-net mesh size on benthic macroinvertebrate samples. At 15 sites in central Montana, samples were collected using the Traveling Kick net protocols (see protocol details below, Section 2.B.ii). At least two samples were collected at each site, one with a 500 μm D-net and one with a 1200 μm D-net. The samples were collected in adjacent stream segments, where site conditions were similar among segments. In a subset of the sites (seven), replicate samples were collected for each mesh size.

The sites were in both the Foothill/Valley/Plains (FVP) ecoregion and the Mountain (Mtn) ecoregion (Table 2). Analysis allowed for investigation into effect differences among ecoregion, though this was not the emphasis of the study and analysis commonly used pooled data. The

sites were of comparable quality when considered along a stressor gradient, being neither high-quality (unimpacted) nor low quality (severely impacted).

2.B.ii. Protocol Study

The goal of the protocol comparison was to determine if the different protocols provided similar results using DEQ macroinvertebrate assessment tools. Multiple protocols were employed in this survey, including those used in the DEQ Fixed Station Network (Hess), Montana Rapid Bioassessment Protocols (traveling kick), and EPA EMAP (Environmental Monitoring and Assessment Program). Because detailed descriptions of methods are provided in the Quality Assurance Project Plan (QAPP) and in Standard Operating Procedures (SOPs, MDEQ 2004), only specific critical protocol information is presented below and in Table 3.

Metrics and indices calculated for samples collected using the different protocols were evaluated at 11 sites (Table 2). Replicate samples were collected at two sites. Ten sites were in the Foothill/Valley/Plains (FVP) ecoregion and one was in the Northwest Glaciated Plains Ecoregion. The study design was completed for three protocols (traveling kick, EMAP reachwide, and EMAP targeted riffle), nearly completed for the Surber protocol (8 of 11 sites), and was not completed for jab and Hess protocols.

Traveling Kick

The Traveling Kick protocol is used in lentic conditions in sites assessed through the DEQ Reassessment & TMDL Monitoring programs. In lotic conditions, a jab protocol is used. Lotic conditions were not the focus of this study and jab collected samples were not included in the comparison analysis. The traveling kick protocol was developed by DEQ as a modification of early EPA guidance for sampling benthic macroinvertebrates (Plafkin et al. 1989, Bukantis 1998). The D-frame net is used for this method. Mesh size has varied in the past and current SOPs prescribe a 1200 μm mesh. The sample is collected from a single riffle within the sampling reach. The sampler disturbs substrate upstream of the D-frame net, capturing the debris and dislodged organisms, for a timed period not less than one minute. Movement proceeds upstream and diagonally across the riffle in order to sample a variety of locations within the riffle. This method focuses on the productive riffle habitats, which are assumed to support the most responsive benthic assemblage and can therefore indicate stresses occurring throughout the stream reach. The traveling kick protocol is a quick method that produces a sample representative of the productive habitats and that is standardized on the time spent collecting the sample. In the laboratory, samples are subsampled to 300 orgainsims ($\pm$ 10%) using a Caton grid. Organisms are identified to taxonomic levels outlined in the DEQ SOP (MDEQ 2004).

EMAP Reachwide and Targeted Riffle

The EMAP sampling method has two nested components, a reachwide sample and a targeted riffle sample (Lazorchak et al. 1998). These two components of the EMAP samples were analyzed as separate samples in this study, although they could be used in concert to determine differences in assemblage characteristics among the entire reach and the targeted riffle sampling

locations. For both components, organisms are counted (500 organism subsample) and identified to the finest practical taxonomic level in the laboratory.

The reachwide sample is collected from 11 transects in the stream reach. The first transect is randomly located using GIS coordinates. Ten more transects are evenly spaced upstream, each separated by a distance of one stream width. At each transect, one square foot of substrate is disturbed and the floating debris and organisms are collected in a D-frame net (500 μm mesh). The sample location within each transect is either at the left bank, center stream, or right bank. The individual collections from each transect are composited to create the sample. This sample thus represents the benthic assemblage that occurs in the entire stream reach, including riffle, run, and pool habitats.

The targeted riffle sample is collected from eight locations within riffles that occur within the stream reach. The sampler selects locations within riffles that are assumed to be most productive (based on substrate, flow, and other channel features). At each location, one square foot of substrate is disturbed and the floating debris and organisms are collected in a D-frame net (500 μm mesh). The individual collections from each location are composited to create the sample. This sample represents the benthic assemblage that occurs in the riffle habitats of the streams reach. Riffles and the selected locations within them are presumed to be the most productive habitats of the reach.

Surber samples

Surber samplers have been used occasionally by DEQ to sample streams. The methods in the DEQ SOP explain the proper use of this sampling device (MDEQ 2004). In the laboratory, samples are subsampled to 300 organisms (± 10%) using a Caton grid. Organisms are identified to taxonomic levels outlined in the DEQ SOP (MDEQ 2004).

The Fixed Station Network

The Fixed Station Network was developed by DEQ to monitor the status and trends of several streams across Montana (Bukantis 1998). This is a comprehensive survey that includes benthic macroinvertebrate samples, water quality, and habitat information. The benthic macroinvertebrate samples were collected using a Hess Sampler. The Hess sampler is a good choice for this study because it provides a quantitative area that other sampling methods used in this survey cannot match. A complete explanation of the Hess Sampler can be found in Bukantis (1998). The goal with this sampling device is to collect a macroinvertebrate sample that is quantitative and allows a more accurate estimate of macroinvertebrate populations in a given riffle. In the laboratory, samples are subsampled to 300 orgainsims (± 10%) using a Caton grid. Organisms are identified to taxonomic levels outlined in the DEQ SOP (MDEQ 2004).

2.C. Data Entry

All data were entered into EDAS (Ecological Data Application System, version 3.2, MS Access 2000) (Tetra Tech 1999). Data types entered included station and sample information, comments, and taxonomic data for each sample and replicate. Taxonomic data included taxa attribute assignments so that metrics could be calculated based on taxonomic hierarchy,

functional feeding groups, mode of locomotion (habit), and pollution tolerance. The MQO for data entry accuracy is 100% after QC checks have occurred and before data is used for subsequent analysis. After QC personnel checked data, all errors were noted and corrected in the database.

2.D. Data Reduction/Metric Calculation

The EDAS database was used to calculate 68 metrics based on sample taxa lists and taxonomic attributes. A subset of metric values was hand-calculated and then compared to those that resulted from the EDAS queries. Cells within a site by metric matrix were randomly selected for hand calculation. Approximately three metrics per site were checked. The purpose of this QC activity was to ensure that the EDAS metric calculation queries performed correctly. The MQO for metric calculation accuracy was 100% after QC checks.

2.E. Site Assessment and Interpretation

Site assessments are based on multimetric indices specific to the physiographic region of Montana. These indices are described in the MDEQ biomonitoring SOPs (MDEQ 2004) and include the Mountain index, the Foothill/Valley/Plains index, and the Prairie Plains index. The Prairie Plains index was not evaluated because streams from that region were not included in the study. The indices are used to assign narrative assessments (e.g., attaining or not attaining aquatic life uses) to a given waterbody based on the metrics calculated from the stream sample. Precision of an index includes an assessment of the narrative result regarding agreement of attainment status among sample collected with various mesh sizes or protocols.

3. Results

3.1 Field Measures: Mesh Size

Ordination Analysis

Macroinvertebrate community ordination suggests that the strongest groupings based on relative abundance data are a result of stream site differences such as ecoregion, gradient and local substrate, rather than by differing mesh sizes (Figure 1). Samples collected from the same site generally grouped together. No groupings were detected based on mesh size. Samples from the same sites and differing mesh sizes did not show a consistent shift within the site group, indicating that mesh size did not **bias** the sample collection at the assemblage scale.

Replicate samples at the same site collected with the same mesh size generally plotted close together, often overlapping with samples from the same site and the alternate mesh size. This suggests that **precision** within mesh size is similar to precision among mesh size.

Most of the study sites showed a clear separation in ordination space by differing ecoregion (Mountain and Montana Valley/Foothills). Some of the Mountain ecoregion sites (Deep Creek

& Thompson Gulch) clustered very closely and distinctly. Classification of macroinvertebrate community structure in streams of different ecoregions has been reported to be significant or non-significant depending on which part of the country the study was conducted (Hawkins & Norris 2000).

Ordination results show that samples cluster according to site, replicate, and ecoregion. Mesh size does not affect the sample to a degree that would cause confusion as to the collection site or the biological condition of the collection site based on the mesh size used to collect the biological sample. This positive evidence of the null effects of mesh size is based on qualitative indications from the ordination that samples are not **biased** by mesh size and **precision** among mesh sizes is not greatly different than **precision** within mesh size.

Cluster Analysis

In 11 of 14 comparisons (79%), the linkage distance between replicate samples of the same size mesh were smaller than the distance between samples collected with the different mesh size in the same site (Figure 2). This indicates that sampling **precision** is greater within samples collected with a single mesh size. Analysis using relative abundance data showed clearer patterns of site similarity when compared to analysis using presence/absence data (not shown).

Indicator Analysis

Based on frequency of occurrence and relative abundance in samples, the following 6 of 233 taxa were significantly different ($p<0.05$) in their capture probabilities among mesh sizes:

Greater abundance and frequency in 500 μm mesh:

- Ostracoda
- Ephemeroptera: Baetidae: *Diphetor hageni*
- Ephemeroptera: Leptophlebiidae: *Paraleptophlebia*
- Diptera: Chironomidae: *Parakieferiella*

Greater abundance and frequency in 1200 μm mesh:

- Coleoptera: Dytiscidae (predaceous diving beetles)
- Hemiptera: Corixidae: *Sigara* (water boatmen)

All have swimming abilities, except for the midge, *Parakieferiella*. Ostracods and *Parakieferiella* are small and it may be that they could pass through the nets with larger mesh size. *Paraleptophlebia* are usually larger mayflies, and it is therefore somewhat surprising that they would be in greater abundance and frequency in the smaller mesh size, though the samples may have captured smaller instars (information on specimen size and maturity not recorded). Dytiscids and *Sigara* are larger swimmers that may have abilities of escaping capture by swimming out of the net. Escape by swimming may be easier with slower currents, as may be caused by smaller mesh sizes. Because 227 of 233 taxa (97%) were captured at similar frequency and abundance in the different mesh sizes, the evidence for mesh size effects from indicator analysis is negative – no **bias**.

Variability Within and Among Mesh Sizes

Of 68 metrics, 48 were less variable within sample pairs collected with the same mesh size as compared to samples collected with different mesh sizes (Table 4). This is perhaps expected because with similar mesh sizes, there is one less source of variability to account for. Only one third (22) of the metrics had CVs more than 10% higher in samples collected with different mesh sizes.

Both indices were less variable within replicates of the same mesh size compared to replicates with different mesh sizes. In the Mountain Index, the CV for samples collected with different mesh sizes was 20.5, while the CV for samples collected with the same mesh size was 13.9. In the FVP Index, the CV for samples collected with different mesh sizes was 11.5, while the CV for samples collected with the same mesh size was 8.7. For both indices, it is evident that there is more variability associated with natural variability and sampling error than there is with mesh size. The additional variability (increase in CV) in the indices for samples collected with different mesh sizes is less than the base level of variability (CV for samples collected with the same mesh size).

Sensitivity to Mesh Size

In 46 metrics (69%), the difference in mean values between mesh sizes is less than half of the standard deviation (RMSE) calculated from replicate sample pairs with the same mesh size (Table 5). In only 3 metrics (% Trichoptera, % sprawlers, and % univoltine) was the difference greater than the standard deviation. This indicates that on average, the **bias** due to mesh size is less than the **precision** of the method.

For the two indices tested, the mean difference between mesh sizes was 22% of the standard deviation for the Mountain Index and 75% of the standard deviation for the FVP Index.

Precision Compared Between Mesh Sizes

Of 67 metrics, somewhat more than half (38) were more **precise** in the replicates collected with 500 µm mesh compared to those collected with 1200µm mesh (Table 6). This was determined by comparing CVs calculated from replicates within each mesh size. Both indices were more **precise** in the 1200µm mesh replicates. These results indicate that while more of the tested metrics are less variable when samples are collected with a smaller mesh size, the indices that have proven sensitive to stress are less variable when samples are collected with a larger mesh size. Because of these counter-indications, we conclude that **precision** is similar among samples collected with either mesh size.

Bias Associated with Mesh Size

Only two metrics showed significant ($p<0.05$) **bias** between the mesh sizes (Table 7, Appendix A). These metrics included % Trichoptera and % univoltine, both of which had higher values in the 1200µm mesh samples. No bias was apparent in either the Mountain or FVP index using either Chi-square analysis (Table 7) or inspection of plotted indices (Figure 3).

3.2 Recommendations Regarding Mesh Sizes

The weight of evidence from the several analyses conducted to find similarities among metrics and indices calculated with different mesh sizes suggests that both 500µm and 1200µm meshes yield bioassessment results that are similar enough to allow direct comparison (Table 8). Five of six analyses show evidence of similarity among samples collected by the two mesh sizes.

While these analyses suggest sufficient similarities for pooling data, any analyses conducted with pooled data should include statements of uncertainties and variabilities associated with each mesh size and the comparisons among them.

3.3 Field Measures: Protocols

Ordination

Ordination results (Figure 4) show samples grouping by site and not by the protocol used to collect samples. Furthermore, the samples did not align with any of the ordination axes by protocol within site groupings, i.e., there was no consistent shift in taxa based on protocol. Groupings of similar samples by site and no consistent shift in taxa within sites indicate that samples were not biased by sampling protocols. An outlier from the Little Blackfoot River collected using Surber protocols was different from other samples from the site because of a low number of invertebrates collected in the sample.

Indicator Analysis

Based on frequency of occurrence and relative abundance in samples, none of the taxa were significantly different ($p<0.05$) in their capture probabilities among the four protocols included in the analysis (traveling kick, EMAP reachwide, EMAP targeted riffle, and Surber). The protocols are not biased to collect any particular taxa.

Variability

Variability was measured as the average standard deviation of metric or index values for samples collected with the same method at the same site. As such, it is an estimate of the average sampling error for all protocols. Data included in the analysis were from two sites, where four sample protocols were duplicated. Composition metrics had the highest CVs, ranging up to 150% for Percent Gastropoda. Three-quarters of metrics had CVs less than 50%. CVs of the indices were 11.6 and 12.3% after rarefaction of component metrics in the mountains and foothill/valley/plains, respectively.

Because the protocols include different subsample target sizes (300 and 500), it was expected that richness metrics (counts of taxa) would be greater with larger subsamples. This source of variability among protocols was reduced through rarefaction. Rarefaction is a systematic re-sampling based on probabilities of selecting a taxon if the subsample was of a smaller size (Hurlbert 1971). The rarefaction of richness metrics generally resulted in lower variability when measured across duplicates within protocols (Table 9). Of 22 richness metrics, 16 CVs decreased as a result of rarefaction.

Sensitivity to Protocols

Sensitivity was measured as the difference in mean metrics or indices by protocol divided by the variability estimated as sampling error (RMSE). The resulting statistic (Mean diff./RMSE) illustrates the number of standard deviations (associated with sampling error) can fit into the difference between protocol means. Larger values reveal metrics or indices that are sensitive to the protocol, relative to sampling error (Table 10).

Traveling Kick vs. EMAP Reachwide Protocols:

In 44 metrics (49%), and the Foothill Valley and Plains index, the difference in mean values between traveling kick and EMAP reachwide protocols was more than one standard deviation (RMSE) based on sampling error. Twenty metrics (21%) were found to have a greater than two standard deviations difference between these protocols. In all but one richness metric (shredder taxa) rarefaction reduced sensitivity.

Traveling Kick vs. EMAP Targeted Riffle Protocols:

Thirty-seven (41%) of the metrics demonstrated a greater than one standard deviation difference between these protocols. Ten (11%) metrics demonstrated a greater than two standard deviation difference between the protocols. The Mountain metric suite also demonstrated a greater than one standard deviation difference between the protocols before rarefaction.

Kick vs. Surber Protocols:

These two protocols showed lower sensitivity to protocol. Eight (9%) of the metrics had differences between protocols greater than one standard deviation. Only one metric (% Tanytarsini) had greater than two standard deviations between means. The Mountain index showed greater than one standard deviation between means for both rarefacted and raw data.

Precision Compared Between Protocols

Replication was insufficient to perform an analysis of precision within any of the protocols. Replicates were collected at only two sites.

Bias Associated with Protocols

Significant bias was found between metrics collected using the traveling kick protocol and those collected using the EMAP protocols (Table 11). Most of the bias was observed in richness metrics when comparing traveling kick to EMAP reachwide protocols. This bias was eliminated through rarefaction in some metrics, but not all. The numbers of EPT, collector, burrower, and swimmer taxa were significantly greater ($p>0.10$) in EMAP reachwide samples compared to traveling kick samples, even after rarefaction. In comparison to EMAP targeted riffle samples, only the swimmer taxa metric was significantly greater than traveling kick after rarefaction. No metrics showed significant bias when traveling kick samples were compared to Surber samples. No significant bias was found in the index comparisons calculated on samples collected using different traveling kick protocols (Table 11, Figure 5).

3.4 Recommendations Regarding Protocols

The weight of evidence from the several analyses conducted to find similarities among metrics and indices calculated with the different protocols (Table 12) suggests that four protocols yield bioassessment results that are similar enough to allow direct comparison, especially after rarefacting large samples to a 300 organism target size.

The four protocols that are similar include the traveling kick, EMAP reachwide (rarefacted), EMAP targeted riffle (rarefacted), and Surber. Comparisons to other protocols (e.g., Hess, jab, REMAP) were not assessed due to insufficient sample sizes in this study. These unassessed methods should not be included in any subsequent analyses of data pooled by protocol.

While these analyses suggest sufficient similarities for pooling of certain data (after rarefaction), any analyses conducted with pooled data should include statements of uncertainties and variabilities associated with each data set and the comparisons among them.

3.5 Measurement Quality Objectives (*move to performance characteristics section*)

3.5.i. Sorting/Subsampling Bias - Results

Five samples were randomly selected from total sample lot, representing approximately 10% of the mesh study (only). Only one of the five samples (Moose Creek, U1062) passed the MQO of >90% (Table 13). Among all five samples, the organisms that were missed were primarily very small midges, occasionally elmid larvae, and a few worms. Corrective actions were that the primary lab sorted through the sort residue of all remaining (non-QC) samples a second time to check for missed specimens. Any recovered specimens were added to the full sample.

Sample number 4 (Beaver Creek, U 1081 MB) came out with a very low PSE of 67.6; this, in large part, resulted from there being only a small number of organisms originally found in the sample. This single sample seems to be an anomaly. All samples from all protocols were treated consistently in this phase of the project; the sorting/subsampling process is not introducing bias into the dataset.

3.5.ii Taxonomic Precision - Results

Results of the taxonomic re-identification comparisons are presented in Table 14**.** All samples met the MQO for enumeration (PDE <5%). None of the samples met the MQO for percent taxonomic disagreement (PTD) at the lowest practical taxonomic level; 3 of the 10 exceeded PTD at the genus level (Table 14 sample numbers 5, 8, and 10). For lowest practical taxonomic level, approximately 76% of the differences were hierarchical, and constitute the primary reason for not meeting the project MQO of 15%; about 10% were straight differences, and 14% due to missing specimens. For the genus level comparison, the 3 samples exceeding the MQO were also primarily due to hierarchical differences (Figures 6**,** 7).

In Sample 5 (U1058-M) the primary difference was hierarchical, with the primary difference being in 132 specimens of mayflies (60 identified as Nixe by T1, and as Heptageniidae by T2; and 72 specimens identified as Baetis tricaudatus by T1 and as Baetis by T2). In Sample 8 (M09LUMPG01), the primary hierarchical differences were 8 mites (identified to Acariformes by T1, and to genus by T2); 8 psychodids (identified to family by T1 and to the genus Pericoma by T2), and 10 nemourids (identified as Nemouridae by T1, and as Zapada cinctipes by T2). Similar to the other two samples in exceedence, differences in Sample 10 (M09TENMC05) were primarily hierarchical, and were with Heptageniidae (14 specimens) and Acarina (6 specimens).

Overall, the number of straight differences was very small, approximately 134 out of a total of 2865 specimens re-identified in the 10 samples (~4.7%). Taxonomic precision (consistency) for this dataset is acceptable at genus level.

4. References

Barbour, M. T., J. Gerritsen, B. D. Snyder, and J. B. Stribling. 1999. Rapid Bioassessment Protocols for Use in Streams and Wadeable Rivers: Periphyton, Benthic Macroinvertebrates, and Fish. Second edition. EPA/841-B-99-002. U.S. EPA, Office of Water, Washington, DC.

Clark, M.J.R. and P.H. Whitfield. 1994. Conflicting perspectives about detection limits and about the censoring of environmental data. Water Resources Bulletin 30(6): 1063-1079.

MDEQ 2004. Sample Collection and Taxonomic Identification of Macroinvertebrates: Standard Operation Procedure. WQPBWQM-009, Rev#: 01. Montana DEQ, Water Quality Planning Bureau.

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

Dufrene, M. and P. Legendre. 1997. Species assemblages and indicator species: the need for a flexible asymmetrical approach. Ecological Monographs 67:345-366.

Hill, C.R., J.B. Stribling, and A.C. Gallardo. 2005. Documentation of Method Performance Characteristics for the Anne Arundel County Biological Monitoring Program. Prepared by Tetra Tech, Inc., Owings Mills, MD for Anne Arundel County Office of Environmental & Cultural Resources. Annapolis, MD. (For further information, contact Christopher Victoria, 410-222-7441 or by email at ***cvictoria@mail.aacounty.org***)

Hurlbert, S.H. (1971). The nonconcept of species diversity: a critique and alternative parameters. *Ecological Monographs* 54, 187–211.

Kruskal, J. B., & Wish, M. (1978). Multidimensional scaling. Beverly Hills, CA: Sage Publications.

Lazorchak, J. M., D. J. Klemm, and D. V. Peck (editors). 1998. Environmental monitoring and assessment program—surface waters: field operations and methods for measuring the ecological condition of wadeable streams. EPA/620/R-94/004F, U.S. Environmental Protection Agency, Washington, D.C.
Web link: http://www.epa.gov/emap/html/pubs/docs/groupdocs/surfwatr/field/ws_abs.html

Manly, B.F.J. 2001. Statistics for Environmental Science and Management. Chapman Hall, New York.

Peters, J.A. 1988. Quality Control Infusion into Stationary Source Sampling. Chapter 22, IN, Lawrence H. Keith (editor), Principles of Environmental Sampling. Pp. 317-333. ACS Professional Reference Book. ISBN 0-8412-1173-6. American Chemical Society. Columbus, Ohio.

Smith, F., S. Kulkarni, L. E. Myers, and M. J. Messner. 1988. Evaluating and Presenting Quality Assurance Sampling Data. Chapter 10, IN, Lawrence H. Keith (editor), Principles of Environmental Sampling. Pp. 157-168. ACS Professional Reference Book. ISBN 0-8412-1173-6. American Chemical Society. Columbus, Ohio.

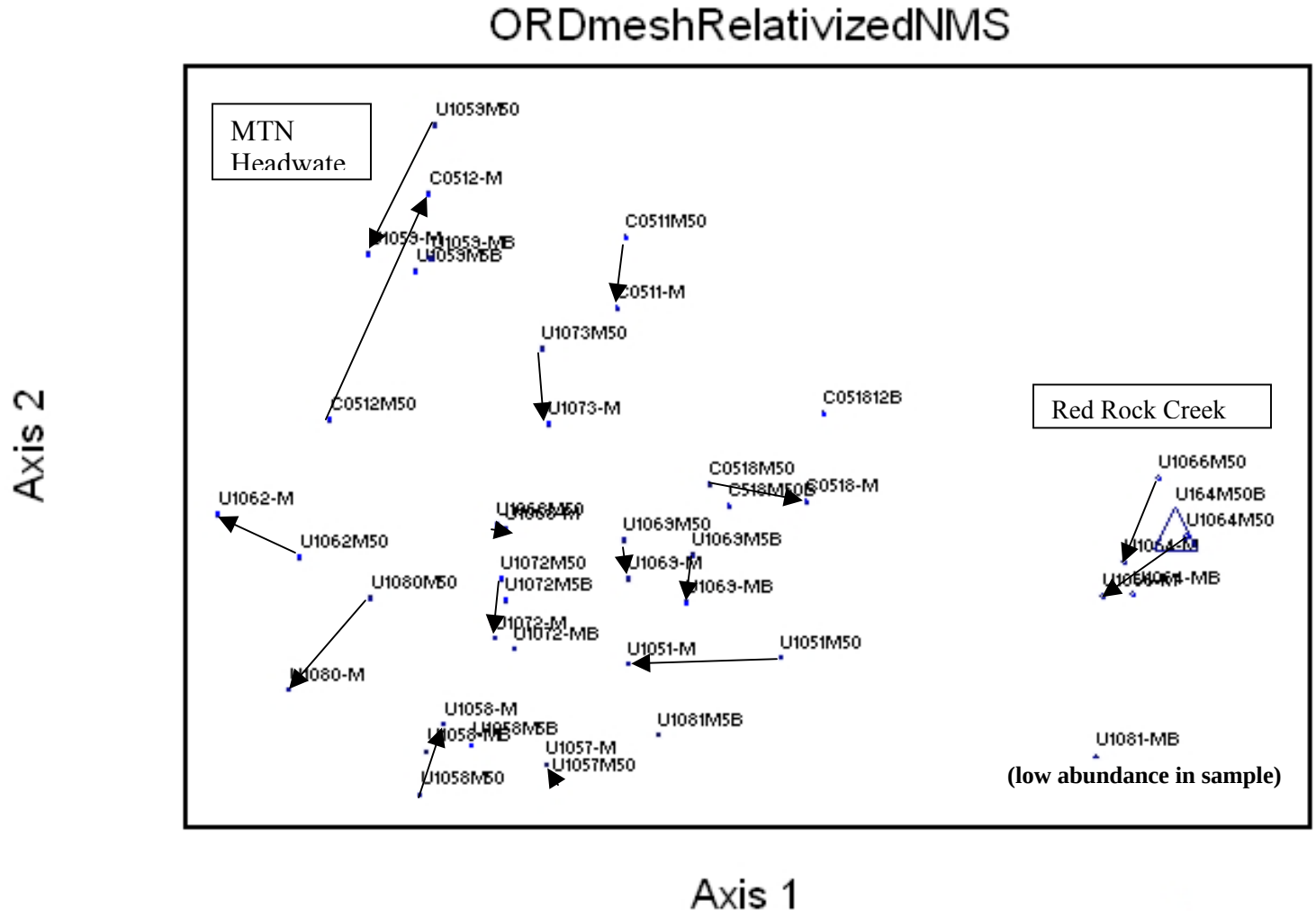
Stribling, J. B. and J. M. Diamond. 2004. The relationship of performance characteristics and data quality to the comparability of biological assessments. Poster presentation and abstract (No. 223). 2004 Conference of the National Water Quality Monitoring Council. Chattanooga, Tennessee. http://water.usgs.gov/wicp/acwi/monitoring/conference/2004/proceedings_contents/13_titlepages/posters/poster_223.pdf

Taylor, J.K. 1988. Defining the Accuracy, Precision, and Confidence Limits of Sample Data. Chapter 6, pages 102-107, *In* Lawrence H. Keith (editor), *Principles of Environmental Sampling*. ACS Professional Reference Book. American Chemical Society. Columbus, Ohio.

Taylor, B. N., and C. E. Kuyatt. 1994. Guidelines for evaluating and expressing the uncertainty of NIST measurement results. NIST Technical Note 1297. National Institute of Standards and Technology, U.S. Department of Commerce. Washington, DC.

Zar, J.H. 1999. Biostatistical Analysis. Fourth Edition. Prentice Hall, Upper Saddle River, New Jersey 07458. 663 pp.

Figure 1. NMS ordination diagram of relativized macroinvertebrate abundance. Arrows are drawn from 500 µm mesh samples to corresponding 1200 um mesh samples from the same site. Sample replicates (indicated with a "B" at the end of the site code) do not have arrows. One obvious outlier had low abundance in the sample, which may affect relative abundance of taxa within the sample.



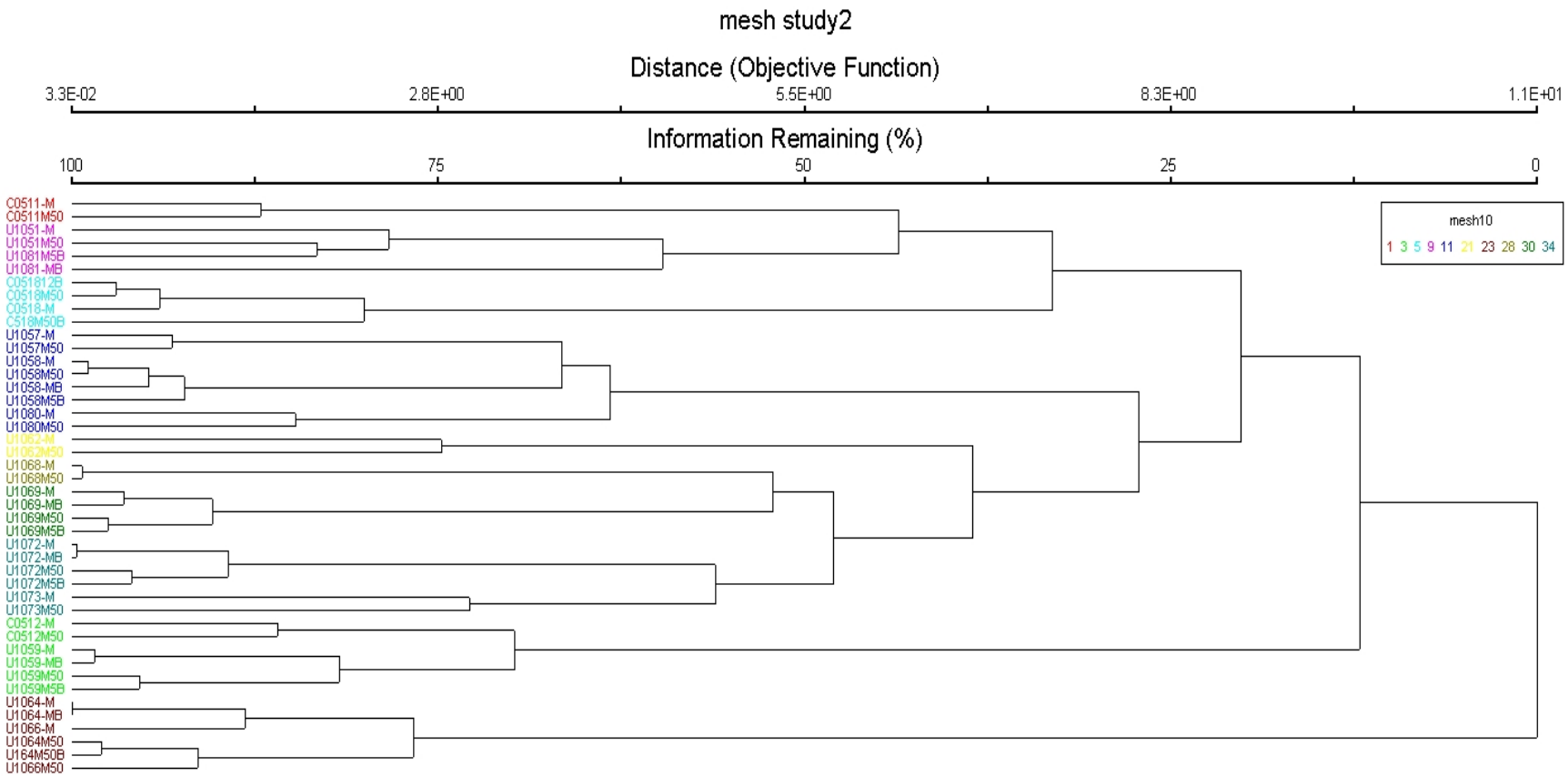


Figure 2. Relativized macroinvertebrate data dendrogram run using Cluster Analysis. Replicates are indicated with a "B" at the end of the site code.

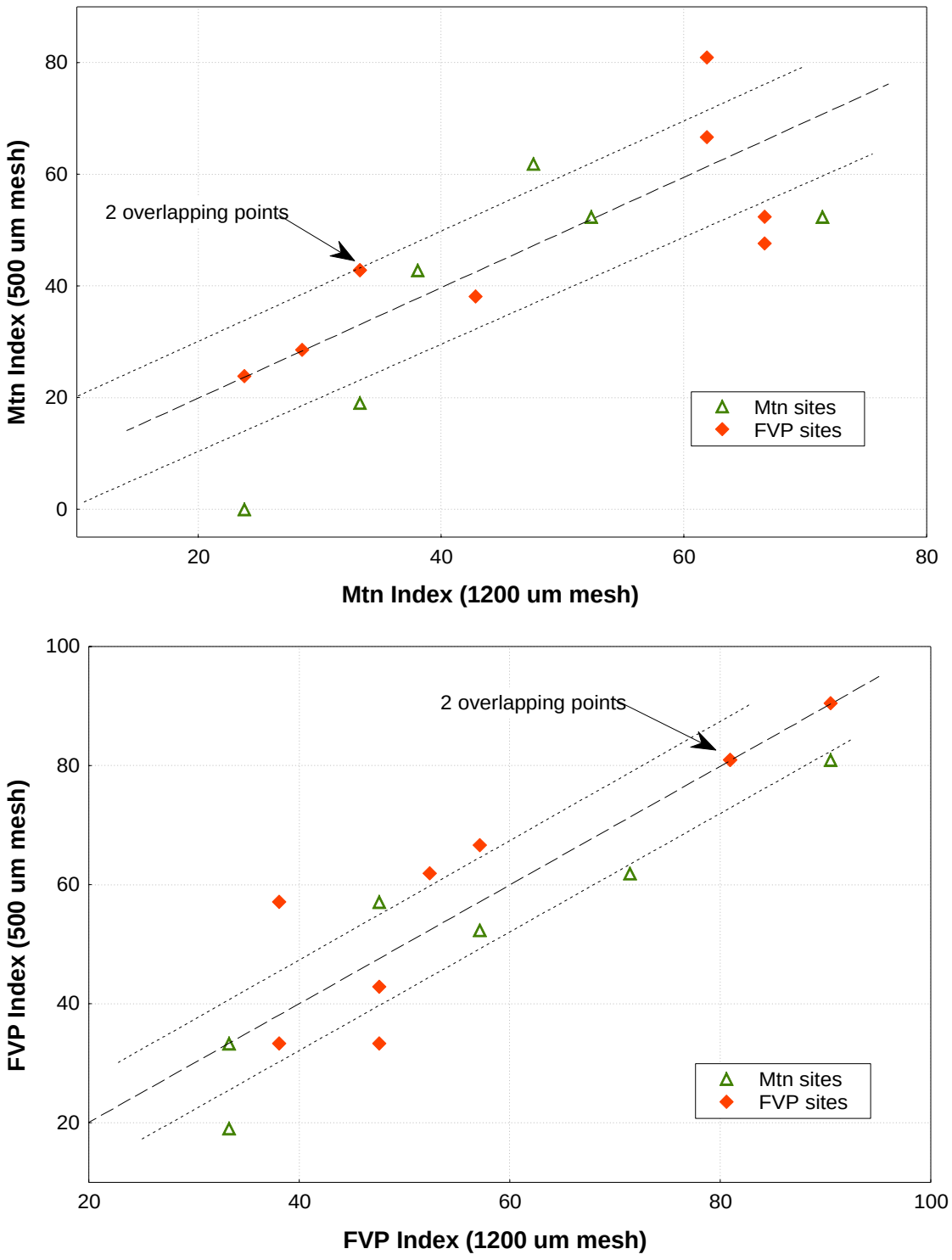


Figure 3. Plots illustrate index values calculated from samples collected with 1200µm mesh on the *x*-axis against those collected with 500µm mesh on the *y*-axis. The unity (1:1) line is shown with the 90% detectable difference on either side.

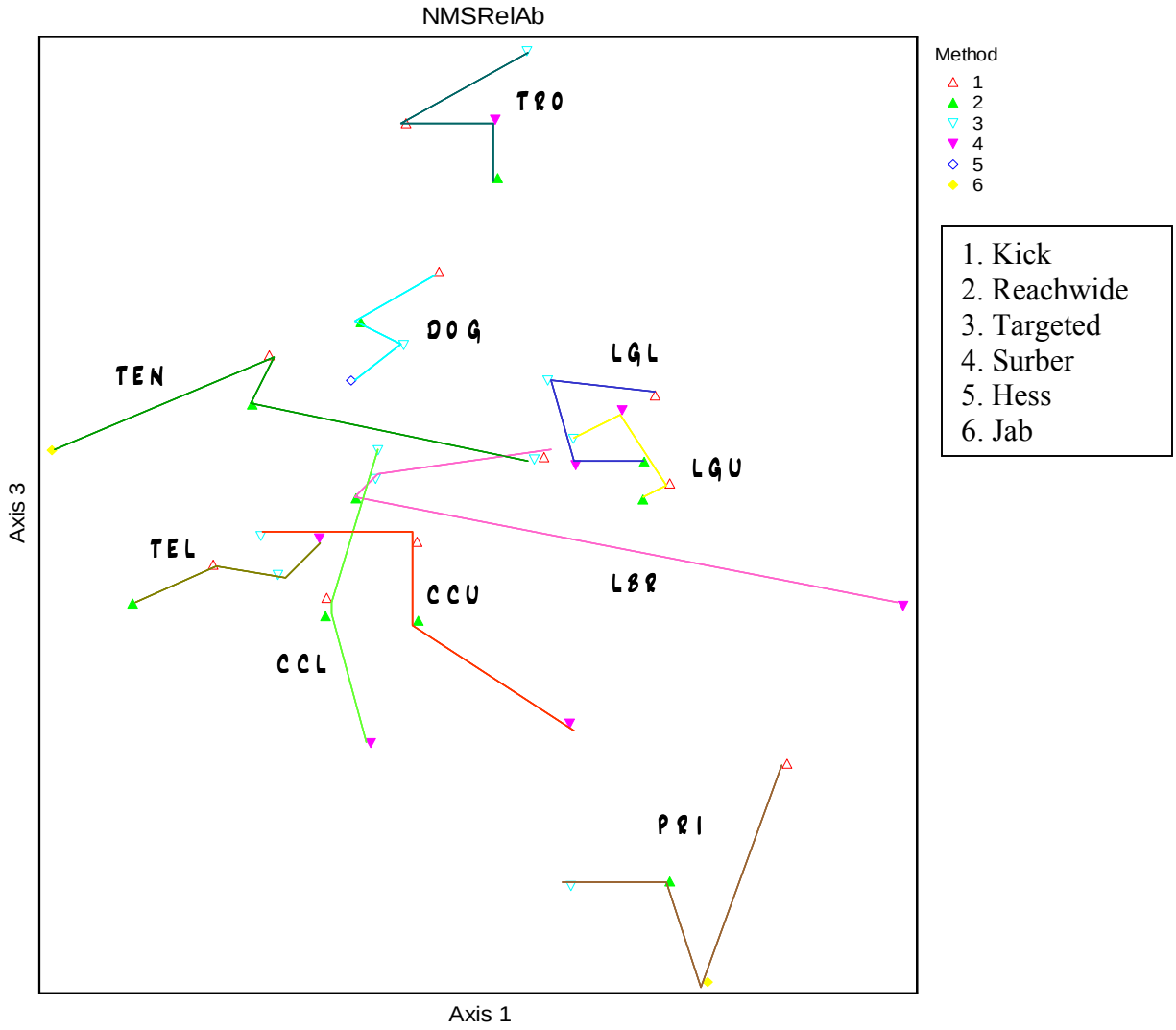


Figure 4. Ordination diagram of samples in taxa space. Three-letter abbreviations are site codes that correspond to the closest grouping of linked samples. Symbols distinguish the protocol used to collect the sample.

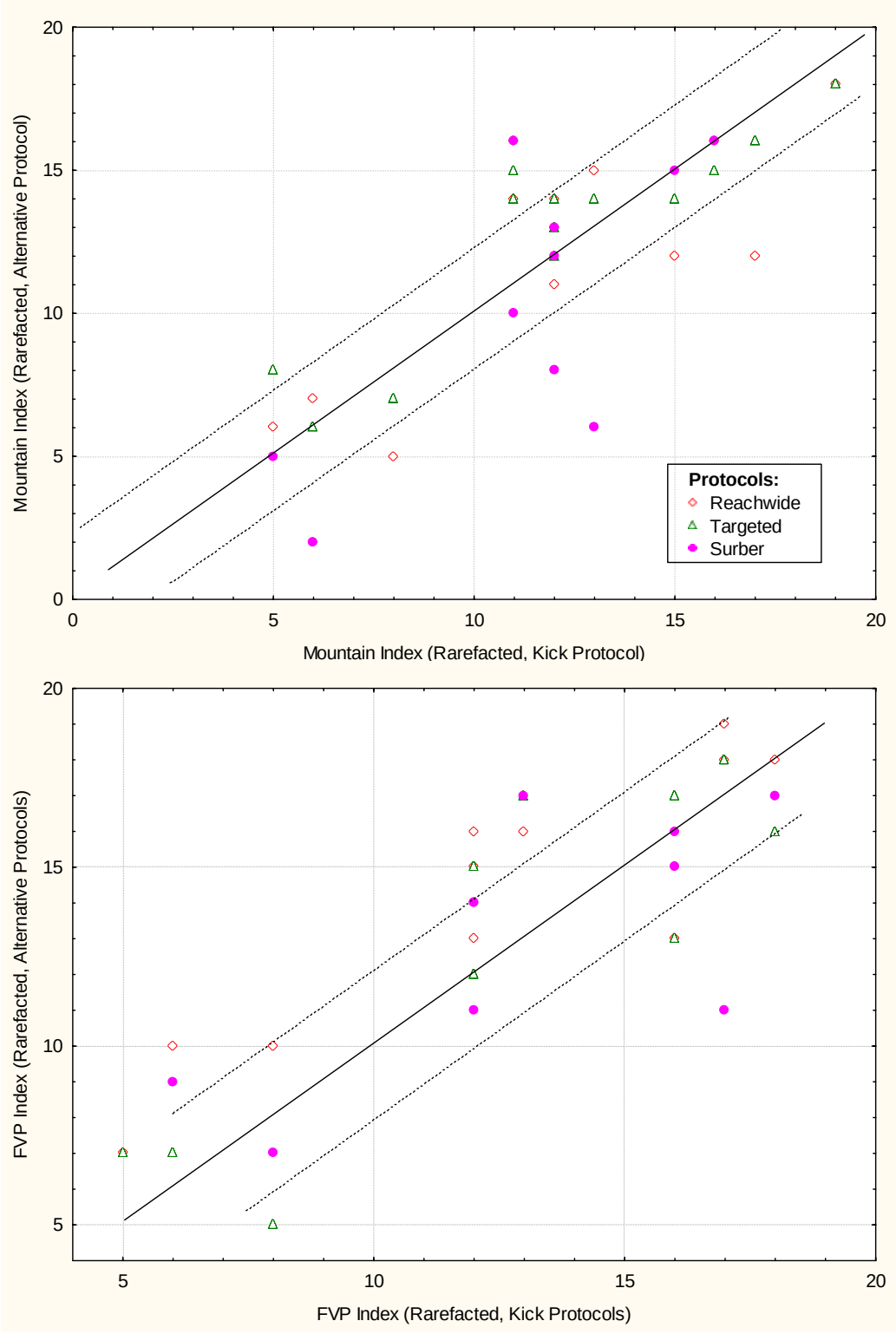


Figure 5. Plots illustrate rarefacted index values calculated from samples collected with traveling kick protocols on the *x*-axis against the alternative protocols on the *y*-axis. The unity (1:1) line is shown with the 90% detectable difference on either side.

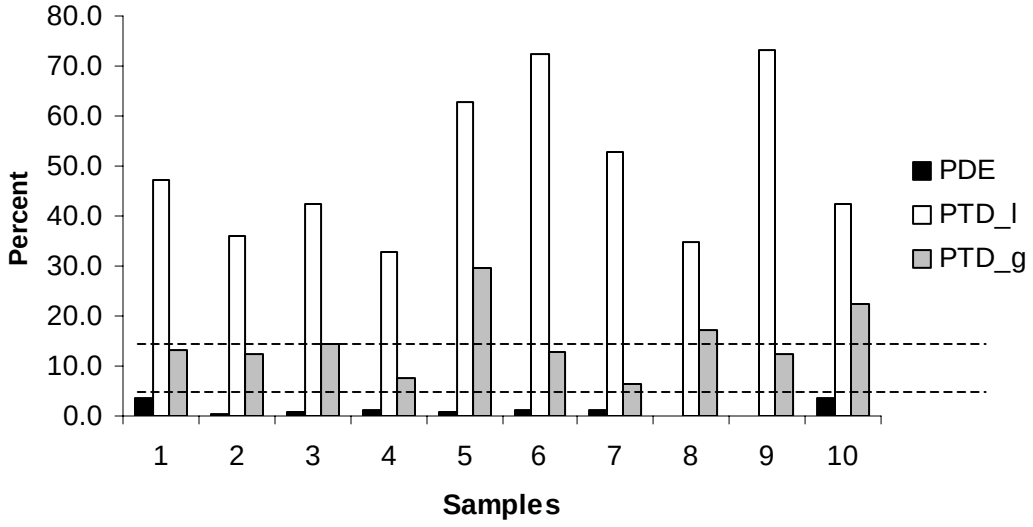


Figure 6. Comparison results of taxonomic re-identifications. Dashed lines represent the MQO for percent difference in enumeration (PDE<5%) and for percent taxonomic disagreement (PTD<15%). PTD_l is calculated at lowest practical taxonomic level, and PTD_g primarily at genus level.

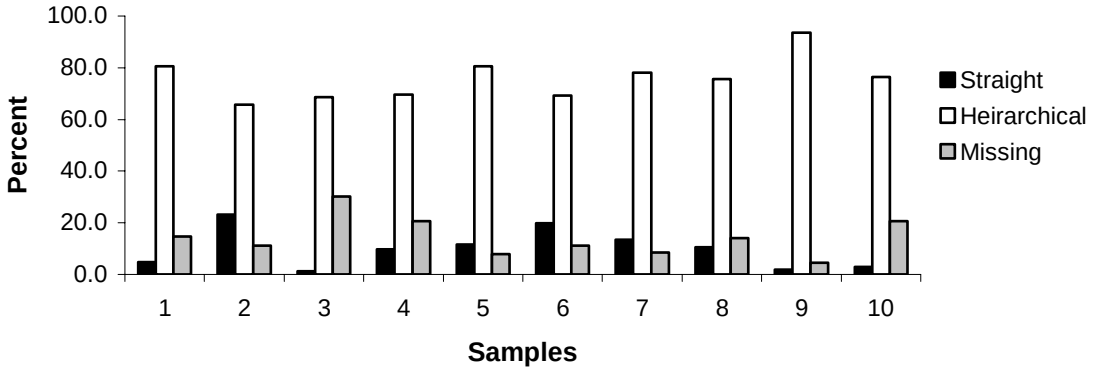


Figure 7. Types of differences (%) causing taxonomic disagreements at lowest practical taxonomic level (lptl) in this dataset.

Table 1. Error partitioning framework for biological assessment protocols. Performance characteristics may be quantitative (QN), qualitative (QL), or not applicable (na). Green = Ben's unsure – discuss with Sam.

	Performance Characteristics				
Component Method or Activity	**Precision**	**Accuracy**	**Bias**	**Representativeness**	**Completeness**
1. Field sampling	QN	na	QN	QL	QN
2. Laboratory sorting/subsampling	na	na	QN	QL	na
3. Taxonomy	QN	QL	QL	na	na
4. Enumeration	QN	QL	QL	na	na
5. Data entry	na	QN	na	na	na
6. Metric calculation (e.g., data reduction)	na	QN	na	na	na
7. Site assessment	QN/QL	na	na	QL	QN

Table 2. Sites sampled in the two comparison studies. Ecoregions include the Mountains (Northern and Middle Rockies, MTN), the Montana Foothills and Valley Prairies (MVFP), and the Northwest Glaciated Plains (NGP).

Stream Name	Site ID	Reps?	BasinID	Ecoregion	County	Lat_Dec	Long_Dec
Mesh Study							
Ten Mile Creek	C0511	N	17010201	MTN	Granite	46.76221	-113.37038
Deep Creek	C0512	N	17010201	MTN	Granite	46.79803	-113.29928
Bloody Dick Creek Upstream	U1073	N	10020001	MTN	Beaverhead	45.06921	-113.42140
Bloody Dick Creek Mid	U1072	Y	10020001	MTN	Beaverhead	45.01658	-113.49782
Red Rock River Mid	U1064	Y	10020001	MTN	Beaverhead	44.64280	-112.03622
Red Rock River Upstream	U1066	N	10020001	MTN	Beaverhead	44.65744	-111.98227
Barnes Creek	C0518	Y	17010202	MVFP	Granite	46.61089	-113.16017
Medicine Lodge Creek UpStr	U1068	N	10020001	MVFP	Beaverhead	44.75140	-113.03617
Medicine Lodge Creek Mid	U1069	Y	10020001	MVFP	Beaverhead	44.87056	-113.00747
Moose Creek Downstream	U1062	N	10030103	MVFP	Meagher	46.81136	-110.90415
Newlan Creek	U1057	N	10030103	MVFP	Meagher	46.62837	-110.97917
Thompson Gulch Downstream	U1058	Y	10030103	MVFP	Meagher	46.60046	-111.12523
Thompson Gulch Headwaters	U1059	Y	10030103	MVFP	Meagher	46.51624	-111.21378
Beaver Creek Downstream	U1051	Y	10030103	MVFP	Meagher	46.74395	-111.40850
Beaver Creek Upstream	U1080	N	10030103	MVFP	Meagher	46.75185	-111.19135
Protocol Study							
Crow Creek Lower	CCL	N	10030101	MVFP	Jefferson	46.25070	-111.67390
Little Blackfoot River	LBR	N	17010201	MVFP	Lewis & Clark	46.45717	-112.41901
West Fork Poplar River	WFP	N	10060004	NGP	Daniels	48.69700	-105.83200
Dog Creek	DOG	N	17010201	MVFP	Lewis & Clark	46.65983	-112.38963
Lump Gulch Lower	LGL	N	10030101	MVFP	Lewis & Clark	46.47435	-112.08497
Crow Creek Upper	CCU	N	10030101	MVFP	Jefferson	46.30028	-111.73418
Lump Gulch Upper	LGU	N	10030101	MVFP	Lewis & Clark	46.47450	-112.08568
Telegraph Creek	TEL	N	17010201	MVFP	Powell	46.48125	-112.36128
Tenmile Creek	TEN	N	10030101	MVFP	Lewis & Clark	46.52758	-112.25385
Trout Creek	TRO	Y	10030101	MVFP	Lewis & Clark	46.76700	-111.64918
Prickly Pear Creek	PRI	Y	10030101	MVFP	Lewis & Clark	46.66123	-111.97618

Table 3. Protocols descriptions.

Program	Method	Equipment	Mesh size	Substrate Area	Habitat	Sub-sample Target Size
Reassessment & TMDL Monitoring	Lotic: Traveling Kick: kicking for a minimum of one minute in a diagonal upstream and across a stream. Lentic: 20 jabs, 1/3 square meter each.	D-frame net	1200 um	Lotic: Variable, effort is standardized by sampling time, not stream area. Lentic: 6.7 square meters.	Lotic: In one riffle. Lentic: Banks and snags (productive areas)	300 individ.
Fixed Station Network	Four replicates, composited	Hess sampler	1000 um	0.35 square meter	Riffles (in lentic conditions jabs are used as above)	300 individ.
Proposed/Optional	Four replicates, composited	Surber sampler	500 um	0.36 square meter	Riffles (in lentic conditions jabs are used as above)	300 individ.
EMAP & Reference; Reachwide	11 semi-random transects distributed evenly throughout reach; one replicate per transect, composited to create sample	D-frame net	500 um	1.0 square meter	Multihabitat, as randomly encountered in transects	500 individ.
EMAP & Reference; Targeted Riffle	8 sample locations are selected in productive riffles within the entire reach. 1 square meter of substrate is disturbed at each location.	D-frame net	500 um	8 square feet	Riffle	500 individ.

Table 4. Coefficient of Variability (CV) calculated for metrics and indices within replicates collected using the same mesh size compared to CV among replicates collected with different mesh sizes. The lower CV of each comparison indicates greater precision and is shown in bold.

	Metric	CV within	CV among
Richness	Total taxa	**15.0**	12.8
	EPT taxa	**20.0**	21.5
	Ephemeroptera taxa	31.1	**30.1**
	Plecoptera taxa	**21.7**	39.5
	Trichoptera taxa	**44.8**	47.1
	Coleoptera taxa	41.4	**30.6**
	Diptera taxa	**19.7**	24.1
	Chironomidae taxa	**22.4**	26.4
	Tanytarsini taxa	**39.8**	52.7
	Crustacea and Mollusca	51.7	**29.8**
	Oligochaeta taxa	**70.7**	111.3
Composition	% EPT	**18.0**	23.8
	% Ephemeroptera	**32.3**	35.5
	% Trichoptera	**41.3**	53.3
	% Plecoptera	**61.9**	78.2
	% Diptera	**22.7**	22.9
	% Chironomidae	**24.1**	26.9
	% Coleoptera	45.0	**44.0**
	% non-Insects	**22.0**	36.6
	% Oligochaeta	**123.4**	125.1
	% Amphipoda	**32.0**	88.0
	% Bivalvia	80.9	**64.2**
	% Crustacea and Mollusca	**38.7**	68.2
	% Gastropoda	**34.4**	**71.8**
	% Tanytarsini	**63.8**	**96.5**
	Tanytarsini / Midges	**43.5**	62.3
	Baetidae / Mayflies	61.0	**49.5**
	Hydropsychidae / EPT	**131.3**	224.8
	Hydropsychidae / Caddisflies	**169.6**	218.2
	Cricotopus & Chironomus / Midges	**47.2**	83.5
Evenness	% Dominant	**24.4**	38.5
	Shannon-Wiener Index	**6.8**	10.6
	D	**21.8**	44.4
	Evenness	**10.0**	12.9
	D Margoleff's	13.2	**10.6**
Habit	% Burrowers	**30.7**	34.2
	% Climbers	137.5	**128.0**
	% Clingers	**14.9**	16.3
	% Sprawlers	**15.7**	25.9
	% Swimmers	**36.2**	48.6
	Burrower taxa	32.6	**23.7**
	Climber taxa	**65.7**	70.4
	Clinger taxa	21.0	**14.1**
	Sprawler taxa	**14.3**	22.3
	Swimmer taxa	**28.5**	30.9
Feeding Groups	% Collectors	15.7	**15.6**
	% Filterers	**47.2**	48.7
	% Predators	**41.3**	60.6
	% Scrapers	**52.8**	54.5
	% Shredders	**53.7**	67.1
	% Collectors & Filterers	**14.1**	14.2
	% Scrapers & Shredders	**35.5**	40.9
	Collector taxa	14.1	**10.7**
	Filterer taxa	35.5	**28.3**
	Predator taxa	23.6	**23.6**
	Scraper taxa	37.2	**33.4**
	Shredder taxa	**36.6**	40.4
Lifestage & Tolerance	Intolerant taxa	21.7	**21.4**
	Tolerant taxa	**20.1**	35.7
	% Intolerant	**35.1**	35.3
	% Tolerant	**28.3**	49.9
	MT Sensitive taxa	72.0	**48.8**
	MT % Tolerant	**25.7**	31.8
	Hilsenhoff Biotic Index	**9.5**	10.8
	Beck's Biotic Index	20.9	**19.8**
	% Multivoltine	**17.5**	19.7
	% Univoltine	**24.0**	28.7
Indices			
	Mountains	**13.9**	20.5
	Foothill Valleys & Plains	**8.7**	11.5

Table 5. The difference of mean metric or index values among mesh sizes (500 or 1200 μm) as a function of variability observed within samples collected with the same mesh size within sites (RMSE).

	Metric or Index	500	1200	RMSE	Diff/RMSE
Indices					
	Mountain	43.2	44.6	6.1	-0.22
	Foothill Valleys and Plains	59.2	55.4	5.0	0.75
Richness	Total taxa	33.6	32.0	4.9	0.33
	EPT taxa	11.5	11.1	2.3	0.16
	Ephemeroptera taxa	4.6	3.9	1.3	0.55
	Plecoptera taxa	2.9	2.4	0.6	0.88
	Trichoptera taxa	4.1	4.9	2.0	-0.43
	Coleoptera taxa	2.4	2.6	1.0	-0.14
	Diptera taxa	13.4	11.6	2.5	0.72
	Chironomidae taxa	10.9	9.2	2.3	0.76
	Tanytarsini taxa	1.6	1.7	0.7	-0.22
	Crustacea and Mollusca taxa	1.8	1.6	0.9	0.16
	Oligochaeta taxa	0.9	1.1	0.7	-0.40
Composition	% EPT	28.0	27.1	5.0	0.17
	% Ephemeroptera	17.5	13.6	5.0	0.78
	% Trichoptera	4.8	7.6	2.6	-1.10
	% Plecoptera	5.7	5.9	3.6	-0.07
	% Diptera	31.2	32.1	7.2	-0.12
	% Chironomidae	28.1	28.9	6.9	-0.12
	% Coleoptera	8.7	10.7	4.4	-0.47
	% non-Insects	30.6	25.0	6.1	0.91
	% Oligochaeta	3.9	4.6	5.2	-0.12
	% Amphipoda	6.3	6.7	2.1	-0.21
	% Bivalvia	1.1	0.8	0.8	0.42
	% Crustacea and Mollusca	20.4	15.6	7.0	0.70
	% Gastropoda	3.1	4.3	1.3	-0.94
	% Tanytarsini	6.3	6.7	4.1	-0.09
	Tanytarsini / Midges	21.2	20.2	9.0	0.12
	Baetidae / Mayflies	53.0	40.5	28.5	0.44
	Hydropsychidae / EPT	2.6	2.9	3.6	-0.08
	Hydropsychidae / Caddisflies	11.2	8.1	16.4	0.20
	Cricotopus & Chironomus / Midges	20.0	18.2	9.0	0.20
Evenness	% Dominant	26.3	23.2	6.0	0.50
	Shannon-Wiener Index	2.7	2.7	0.2	-0.04
	D	0.1	0.1	0.0	0.57
	Evenness	0.5	0.5	0.0	-0.34
	D Margoleff's	5.9	5.7	0.8	0.20
Habit	% Burrowers	19.0	21.6	6.2	-0.42
	% Climbers	1.7	3.6	3.7	-0.52
	% Clingers	54.1	53.1	8.0	0.13
	% Sprawlers	30.9	38.7	5.5	-1.42
	% Swimmers	31.7	22.9	9.9	0.89

	Burrower taxa	6.8	6.9	2.2	-0.06
	Climber taxa	1.5	2.0	1.1	-0.44
	Clinger taxa	17.6	16.3	3.6	0.38
	Sprawler taxa	12.1	11.1	1.7	0.56
	Swimmer taxa	6.4	5.2	1.7	0.73
Feeding Groups	% Collectors	53.3	57.0	8.6	-0.44
	% Filterers	5.9	6.7	3.0	-0.28
	% Predators	12.1	10.1	4.6	0.45
	% Scrapers	13.7	14.1	7.3	-0.05
	% Shredders	10.3	9.6	5.3	0.13
	Collector taxa	15.9	14.6	2.1	0.60
	Filterer taxa	3.2	3.4	1.2	-0.18
	Predator taxa	6.3	6.7	1.5	-0.28
	Scraper taxa	3.6	2.8	1.2	0.72
	Shredder taxa	3.2	2.7	1.1	0.46
	% Collectors and Filterers	59.1	63.8	8.7	-0.53
	% Scrapers and Shredders	24.0	23.7	8.5	0.04
Lifestage and Tolerance	Intolerant taxa	10.0	9.6	2.1	0.17
	Tolerant taxa	5.6	6.0	1.2	-0.37
	% Intolerant	16.1	18.0	6.0	-0.32
	% Tolerant	23.1	30.1	7.5	-0.93
	HBI	5.0	5.3	0.5	-0.50
	Beck's Biotic Index	20.2	19.2	4.1	0.24
	MT % Tolerant	31.6	32.6	8.2	-0.13
	MT Sensitive taxa	1.3	1.3	0.9	0.00
	% Multivoltine	56.1	51.0	9.4	0.55
	% Univoltine	30.1	40.7	8.5	-1.24

Table 6. Coefficient of Variability (CV) compared among mesh sizes in 7 sites with replicate samples (28 replicates total). Lower CV indicates greater precision and is shown as bold in each comparison.

	Metric	CV 500	CV 1200
Richness	Total taxa	15.9	**13.9**
	EPT taxa	**18.2**	21.7
	Ephemeroptera taxa	35.1	**24.0**
	Plecoptera taxa	22.9	**19.6**
	Trichoptera taxa	**19.7**	55.6
	Coleoptera taxa	53.9	**25.5**
	Diptera taxa	**10.5**	27.5
	Chironomidae taxa	**10.7**	32.2
	Tanytarsini taxa	**34.0**	44.1
	Crustacea and Mollusca	65.2	**28.2**
	Oligochaeta taxa	**62.4**	74.0
Composition	% EPT	**12.5**	22.4
	% Ephemeroptera	**16.6**	47.7
	% Trichoptera	42.7	**39.3**
	% Plecoptera	**24.4**	82.5
	% Diptera	24.1	**21.3**
	% Chironomidae	**23.0**	25.1
	% Coleoptera	**44.7**	44.8
	% non-Insects	**19.6**	25.0
	% Oligochaeta	169.4	**71.5**
	% Amphipoda	**26.2**	36.3
	% Bivalvia	**79.0**	81.2
	% Crustacea and Mollusca	**36.3**	41.6
	% Gastropoda	**29.0**	36.2
	% Tanytarsini	89.8	**22.5**
	Tanytarsini / Midges	58.0	**15.9**
	Baetidae / Mayflies	**44.3**	81.0
	Hydropsychidae / EPT	**110.0**	146.0
	Hydropsychidae / Caddisflies	**153.8**	191.0
	Cricotopus & Chironomus / Midges	**36.6**	57.3
Evenness	% Dominant	29.4	**15.7**
	Shannon-Wiener Index	8.3	**4.8**
	D	27.3	**11.2**
	Evenness	**7.8**	11.7
	D Margoleff's	14.9	**11.0**

	Metric	CV 500	CV 1200
Habit	% Burrowers	42.8	**15.9**
	% Climbers	**40.0**	142.4
	% Clingers	15.8	**13.9**
	% Sprawlers	16.8	**14.8**
	% Swimmers	**31.1**	43.3
	Burrower taxa	39.6	**24.1**
	Climber taxa	69.0	**62.7**
	Clinger taxa	23.6	**17.4**
	Sprawler taxa	**13.5**	15.2
	Swimmer taxa	**27.6**	29.4
Feeding Groups	% Collectors	20.5	**9.7**
	% Filterers	**44.9**	48.8
	% Predators	**39.4**	43.6
	% Scrapers	70.8	**25.8**
	% Shredders	**48.7**	58.8
	% Collectors & Filterers	18.1	**9.4**
	% Scrapers & Shredders	44.0	**23.8**
	Collector taxa	16.7	**10.0**
	Filterer taxa	**34.3**	36.6
	Predator taxa	**19.9**	26.4
	Scraper taxa	**30.2**	46.0
	Shredder taxa	**36.2**	36.8
Lifestage & Tolerance	Intolerant taxa	16.7	**10.0**
	Tolerant taxa	**34.3**	36.6
	% Intolerant	**19.9**	26.4
	% Tolerant	**30.2**	46.0
	MT Sensitive taxa	**58.8**	83.1
	MT % Tolerant	**23.2**	27.8
	Hilsenhoff Biotic Index	**8.8**	10.2
	Beck's Biotic Index	**12.2**	27.5
	% Multivoltine	19.3	**15.1**
	% Univoltine	33.2	**16.3**
Indices			
	Mountains	16.4	**11.1**
	Foothill Valleys & Plains	9.6	**7.6**

Table 7. Assessment of index and metric bias attributed to mesh size. The number of 500µm to 1200µm differences that are above or below zero by an amount greater than expected for sampling error alone determines bias. Critical Chi-square values are 3.841 and 2.706 for probabilities of 0.05 and 0.10, respectively. DD = detectable difference calculated from repeated measures within sites and mesh sizes. Differences not exceeding the DD are distributed evenly among groups of the Chi-square analysis.

Metric	DD	Exceeding DD		Chi-square
		500	1200	
Mountain Index	2.1	3	4	0.07
FVP Index	1.7	3	3	0.00
Total taxa	8.06	1	1	0.00
EPT taxa	3.71	2	5	0.60
Ephemeroptera taxa	2.15	1	1	0.00
Plecoptera taxa	0.93	5	4	0.07
Trichoptera taxa	3.31	0	5	1.67
Coleoptera taxa	1.70	1	2	0.07
Diptera taxa	4.05	2	1	0.07
Chironomidae taxa	3.69	3	0	0.60
Tanytarsini taxa	1.07	0	2	0.27
Crustacea and Mollusca taxa	1.45	1	0	0.07
Oligochaeta taxa	1.16	2	2	0.00
% EPT	8.13	2	4	0.27
% Ephemeroptera	8.27	4	3	0.07
% Trichoptera	4.19	0	8	**4.27**
% Plecoptera	5.87	2	1	0.07
% Diptera	11.77	1	3	0.27
% Chironomidae	11.26	1	3	0.27
% Coleoptera	7.17	1	1	0.00
% non-Insects	10.05	6	2	1.07
% Oligochaeta	8.61	3	7	1.07
% Amphipoda	3.42	2	1	0.07
% Bivalvia	1.23	5	0	1.67
% Crustacea and Mollusca	11.42	5	1	1.07
% Gastropoda	2.07	1	4	0.60
% Tanytarsini	6.79	3	1	0.27
% Tanytarsini of Chironomidae	14.73	3	2	0.87
% Baetidae of Ephemeroptera	46.76	5	0	1.13
% Hydropsychidae of EPT	5.95	1	1	0.00
% Hydropsychidae of Trichoptera	26.85	1	0	0.07
Cricotopus & Chironomus of Chironomidae	14.79	3	3	0.00
% Dominant	9.90	5	2	0.60
Shannon-Wiener Index	0.30	2	2	0.00
D	0.04	5	3	0.27
Evenness	0.08	2	3	0.07
D Margoleff's	1.25	0	1	0.07
% Burrowers	10.24	2	3	0.07
% Climbers	6.02	0	2	0.27

% Clingers	13.07	4	2	0.27
% Sprawlers	8.97	1	4	0.60
% Swimmers	16.20	8	1	3.27
Burrower taxa	3.67	1	1	0.00
Climber taxa	1.89	1	4	0.60
Clinger taxa	5.84	1	1	0.00
Sprawler taxa	2.72	4	3	0.07
Swimmer taxa	2.72	4	2	0.27
% Collectors	14.17	2	1	0.07
% Filterers	4.89	1	2	0.07
% Collectors and Filterers	14.2	2	3	0.07
% Predators	7.53	2	2	0.00
% Scrapers	12.02	3	2	0.07
% Shredders	8.75	2	1	0.07
% Scrapers and Shredders	13.9	3	2	0.07
Collector taxa	3.51	1	1	0.00
Filterer taxa	1.94	1	2	0.07
Predator taxa	2.52	1	4	0.60
Scraper taxa	1.96	1	1	0.00
Shredder taxa	1.78	1	2	0.07
Intolerant taxa	3.49	4	7	0.60
MT Sensitive taxa	1.5	1	0	0.07
Tolerant taxa	1.91	1	5	1.07
% Intolerant	9.78	2	3	0.07
% Tolerant	12.35	0	3	0.60
MT % Tolerant	13.5	3	3	0
HBI	0.80	1	4	0.60
Beck's Biotic Index	6.76	1	4	0.60
% Multivoltine	15.37	4	2	0.27
% Univoltine	13.93	0	8	**4.27**

Table 8. Weight of evidence of similarity among samples collected with 500µm and 1200µm mesh. Positive evidence of similarity is denoted with "+", negative evidence with "-".

Analysis	**Result**	**Evidence of similarity among mesh sizes.**	**Supporting Text, Tables and Figures**
Ordination and Clustering	More variability was associated with sites and ecoregions than with mesh sizes.	+	Figures Ordination, Clusters
Indicator Analysis	Only 6 of 233 taxa (2.6%) showed significant differences among mesh sizes.	+	Section Indicator Taxa
Precision within and among mesh sizes	Two thirds of metrics and both indices are more precise among replicates of similar mesh size compared to replicates of different mesh sizes.	-	Table MetricCVs
Sensitivity to Mesh Size	Two thirds of metrics had a mean difference between mesh sizes that is less than half of one standard deviation calculated from sample replicates of the same mesh size.	+	Table Diff/RMSE
Precision compared between mesh sizes	Slightly more than half of metrics were less variable in samples collected with 500 µm mesh. Indices were slightly less variable in samples collected with 1200 µm mesh.	+	Table CVbyMesh
Bias, Chi-square	Only 2 of 67 metrics showed significant bias with mesh size. Indices showed no bias with mesh size.	+	Table ChiSquare, Appendix MeshMetrics, Figures MtnIndex, FVPIndex

Table 9. Variability associated with indices and metrics calculated from replicate samples collected using multiple protocols. Mean square error (MSE) and Root MSE (RMSE) are estimates of variance and standard deviation associated with replicate measures (same site, same protocol). The 90% detectable difference defines the range around the observation where we expect to find the true mean in 90% of the cases. Coefficient of variation (CV) standardizes the standard deviation on the mean.

Metric/Index	MSE	RMSE	DD90	Mean	CV
Mountain Index	0.88	0.94	1.54	9.50	9.8
Mountain Index Percent	19.84	4.45	7.33	45.24	9.8
Mountain Index rarefacted	1.12	1.06	1.74	9.13	11.6
Mountain Index Percent rarefacted	25.51	5.05	8.31	43.45	11.6
FVP Index	2.25	1.50	2.47	10.38	14.5
FVP Index Percent	51.02	7.14	11.75	49.40	14.5
FVP Index rarefacted	1.56	1.25	2.06	10.19	12.3
FVP Index Percent rarefacted	35.43	5.95	9.79	48.51	12.3
Total taxa	14.31	3.78	6.22	25.69	14.7
Total taxa rarefacted	6.27	2.50	4.12	22.41	11.2
EPT taxa	0.94	0.97	1.59	10.06	9.6
EPT taxa rarefacted	0.62	0.79	1.30	9.19	8.6
Ephemeroptera taxa	0.56	0.75	1.23	4.81	15.6
Ephemeroptera taxa rarefacted	0.51	0.71	1.17	4.46	16.0
Plecoptera taxa	0.19	0.43	0.71	2.31	18.7
Plecoptera taxa rarefacted	0.13	0.36	0.59	2.18	16.5
Trichoptera taxa	0.44	0.66	1.09	2.94	22.5
Trichoptera taxa rarefacted	0.20	0.45	0.74	2.55	17.5
Diptera taxa	7.31	2.70	4.45	8.31	32.5
Diptera taxa rarefacted	4.06	2.02	3.32	6.87	29.4
Chironomidae taxa	4.12	2.03	3.34	5.75	35.3
Chironomidae taxa rarefacted	2.01	1.42	2.33	4.70	30.1
Tanytarsini taxa	0.75	0.87	1.42	1.13	77.0
Coleoptera taxa	1.63	1.27	2.10	2.25	56.7
Coleoptera taxa rarefacted	1.06	1.03	1.69	1.92	53.7
Oligochaeta taxa				0.00	
Oligochaeta taxa rarefacted				0.00	
Crustacea and Mollusca taxa	0.44	0.66	1.09	1.56	42.3
Crustacea and Mollusca taxa rarefacted	0.53	0.73	1.20	1.32	55.1
Shannon-Wiener Index	0.06	0.25	0.42	2.04	12.4
% Dominant	73.53	8.57	14.11	44.01	19.5
Evenness	0.00	0.04	0.07	0.34	11.8
D Margoleff's	0.42	0.65	1.07	4.37	14.9
D	0.00	0.07	0.12	0.25	27.8
% EPT	68.00	8.25	13.57	57.98	14.2
% Ephemeroptera	77.60	8.81	14.49	46.09	19.1
% Plecoptera	8.17	2.86	4.70	5.46	52.3
% Trichoptera	27.64	5.26	8.65	6.43	81.8
% Coleoptera	4.05	2.01	3.31	3.28	61.3
% Diptera	65.48	8.09	13.31	11.21	72.2

% Chironomidae	34.64	5.89	9.68	7.16	82.2
Cricotopus/Chironomus of Chironomidae	2.88	1.70	2.79	2.91	58.3
% Tanytarsini	0.96	0.98	1.61	1.66	59.0
Tanytarsini of Chironomidae	107.09	10.35	17.02	16.85	61.4
% non-Insects	76.03	8.72	14.34	27.39	31.8
% Amphipoda	93.10	9.65	15.87	16.62	58.0
% Bivalvia	0.26	0.51	0.83	0.47	107.8
% Crustacea and Mollusca	76.07	8.72	14.35	26.56	32.8
% Gastropoda	0.98	0.99	1.63	0.66	149.4
% Isopoda				0.00	
% Oligochaeta				0.00	
Intolerant taxa	5.94	2.44	4.01	8.44	28.9
Intolerant taxa rarefacted	3.43	1.85	3.04	7.50	24.7
MDEQ Intolerant taxa	0.06	0.25	0.41	0.69	36.4
MDEQ Intolerant taxa rarefacted	0.03	0.18	0.30	0.55	33.6
Tolerant taxa	1.19	1.09	1.79	3.31	32.9
Tolerant taxa rarefacted	0.56	0.75	1.23	2.65	28.1
% Tolerant	2.36	1.54	2.53	5.05	30.4
MDEQ % Tolerant	80.41	8.97	14.75	34.25	26.2
% Intolerant	24.37	4.94	8.12	42.43	11.6
Beck's Biotic Index	12.75	3.57	5.87	16.13	22.1
Beck's Biotic Index rarefacted	7.82	2.80	4.60	14.31	19.5
HBI	0.06	0.24	0.39	3.24	7.3
% Baetidae of Ephemeroptera	136.53	11.68	19.22	26.10	44.8
% Hydropsychidae of EPT	80.77	8.99	14.78	12.58	71.4
% Hydropsychidae of Trichoptera	87.88	9.37	15.42	32.82	28.6
% Collectors	196.44	14.02	23.06	40.13	34.9
% Filterers	44.83	6.70	11.01	7.87	85.1
% Predators	4.28	2.07	3.40	4.98	41.6
% Scrapers	29.49	5.43	8.93	32.97	16.5
% Shredders	5.67	2.38	3.92	3.21	74.1
Collector taxa	8.19	2.86	4.71	10.19	28.1
Collector taxa rarefacted	3.35	1.83	3.01	8.71	21.0
Filterer taxa	0.81	0.90	1.48	3.06	29.4
Filterer taxa rarefacted	0.74	0.86	1.42	2.67	32.3
Predator taxa	1.56	1.25	2.06	4.81	26.0
Predator taxa rarefacted	0.81	0.90	1.48	3.96	22.7
Scraper taxa	1.25	1.12	1.84	3.75	29.8
Scraper taxa rarefacted	1.14	1.07	1.75	3.52	30.3
Shredder taxa	1.25	1.12	1.84	2.00	55.9
Shredder taxa rarefacted	0.96	0.98	1.61	1.74	56.2
% Burrowers	5.10	2.26	3.72	5.09	44.3
% Climbers	3.01	1.74	2.86	3.82	45.4
% Clingers	124.88	11.18	18.38	63.55	17.6
% Sprawlers	168.11	12.97	21.33	26.76	48.4
% Swimmers	114.79	10.71	17.62	35.75	30.0
Burrower taxa	1.75	1.32	2.18	4.88	27.1
Burrower taxa rarefacted	1.13	1.06	1.75	4.09	26.0

Climber taxa	0.38	0.61	1.01	0.75	81.6
Climber taxa rarefacted	0.08	0.29	0.48	0.62	46.6
Clinger taxa	5.19	2.28	3.75	16.19	14.1
Clinger taxa rarefacted	3.27	1.81	2.97	14.38	12.6
Sprawler taxa	3.12	1.77	2.91	7.63	23.2
Sprawler taxa rarefacted	1.00	1.00	1.64	6.39	15.6
Swimmer taxa	0.50	0.71	1.16	5.25	13.5
Swimmer taxa rarefacted	0.56	0.75	1.23	4.73	15.8
% Multivoltine	37.10	6.09	10.02	26.08	23.4
% Univoltine	55.15	7.43	12.22	52.86	14.0

Table 10. The sensitivity of indices and metrics to protocols calculated as the difference in mean values (Kick – alternative) divided by the RMSE (average standard deviation for replicates within sites and protocols). If the difference among protocol means is greater than the average standard deviation, the statistic is bold-typed.

	Mean Diff (Kick)/RMSE		
Index/Metric	Reachwide	Targeted	Surber
Mountain Index	-0.49	**-1.07**	**1.20**
Mountain Index rarefacted	0.00	-0.43	**1.18**
FVP Index	**-1.09**	-0.67	-0.25
FVP Index rarefacted	**-1.09**	-0.65	-0.30
Total taxa	**-2.84**	**-2.33**	0.33
Total taxa rarefacted	**-2.19**	**-1.45**	**1.04**
EPT taxa	**-3.94**	**-3.38**	-0.52
EPT taxa rarefacted	**-2.11**	**-1.72**	0.38
Ephemeroptera taxa	**-1.82**	**-1.70**	0.67
Ephemeroptera taxa rarefacted	-0.92	**-1.10**	0.99
Plecoptera taxa	**-2.52**	**-2.73**	**-1.44**
Plecoptera taxa rarefacted	**-1.28**	**-1.45**	**-1.00**
Trichoptera taxa	**-2.06**	**-1.24**	-0.57
Trichoptera taxa rarefacted	**-1.22**	-0.10	-0.09
Diptera taxa	**-1.55**	**-1.41**	0.42
Diptera taxa rarefacted	**-1.02**	-0.71	0.82
Chironomidae taxa	**-1.25**	-0.98	0.37
Chironomidae taxa rarefacted	-0.73	-0.34	0.89
Tanytarsini taxa	-0.52	-0.73	0.00
Coleoptera taxa	-0.78	-0.50	-0.10
Coleoptera taxa rarefacted	-0.54	-0.26	-0.06
Crustacea and Mollusca taxa	0.14	0.00	**1.13**
Crustacea and Mollusca taxa rarefacted	0.25	0.15	0.97
% Dominant	0.34	0.31	-0.62
Shannon-Wiener Index	-0.66	-0.46	0.59
Evenness	0.29	0.65	-0.15
D	0.18	0.04	-0.64
D Margoleff's	**-2.10**	**-1.43**	0.16
% EPT	**1.28**	0.78	0.36
% Ephemeroptera	0.07	-0.23	0.35
% Plecoptera	-0.11	-0.48	-0.85
% Trichoptera	**1.96**	**1.87**	0.45
% Coleoptera	**-1.26**	**-1.04**	-0.63
% Diptera	-0.19	0.31	0.31
% Chironomidae	0.37	0.97	0.46
CrCh2ChiPct	**-1.96**	0.89	**-1.59**
% Tanytarsini	**1.17**	**1.99**	**2.95**
Tnyt2ChiPct	-0.12	0.18	0.16
% non-Insects	-0.73	-0.80	-0.49
% Amphipoda	0.16	0.05	-0.24
% Bivalvia	0.21	-0.68	-0.02

% Crustacea and Mollusca	-0.27	-0.30	-0.32
% Gastropoda	-0.08	-0.41	0.13
Intolerant taxa	**-1.72**	**-1.53**	-0.15
Intolerant taxa rarefacted	-0.97	-0.69	0.26
MDEQ Intolerant taxa	**-5.82**	**-5.82**	-0.50
MDEQ Intolerant taxa rarefacted	**-3.53**	**-3.56**	-0.02
Tolerant taxa	-1.17	-0.58	0.00
Tolerant taxa rarefacted	-0.99	0.08	0.12
% Intolerant	0.11	**1.18**	**-1.54**
% Tolerant	-0.45	-0.68	0.17
MDEQ % Tolerant	**2.44**	**1.83**	-0.79
Beck's Biotic Index	**-2.39**	**-2.11**	-0.18
Beck's Biotic Index rarefacted	**-1.38**	**-1.10**	0.35
HBI	-2.08	-0.99	-0.14
% Baetidae of Ephemeroptera	**-1.01**	**-1.21**	0.12
% Hydropsychidae of EPT	0.29	0.31	0.51
% Hydropsychidae of Trichoptera	0.38	0.15	0.32
% Collectors	-0.19	-0.30	-0.43
% Filterers	**1.15**	**1.40**	0.25
% Predators	**-2.40**	**-1.68**	0.96
% Scrapers	-0.17	0.02	0.13
% Shredders	**1.68**	**1.25**	-0.28
Collector taxa	**-2.19**	**-1.49**	0.22
Collector taxa rarefacted	**-1.99**	**-1.03**	0.54
Filterer taxa	-0.81	-0.40	0.14
Filterer taxa rarefacted	-0.46	-0.07	0.25
Predator taxa	**-1.96**	**-1.89**	-0.10
Predator taxa rarefacted	**-1.38**	**-1.42**	0.29
Scraper taxa	-0.73	-0.65	0.22
Scraper taxa rarefacted	-0.40	-0.15	0.31
Shredder taxa	-0.08	-0.41	-0.11
Shredder taxa rarefacted	0.43	0.32	0.20
% Burrowers	**-2.17**	**-2.12**	**-1.81**
% Climbers	-0.17	-0.43	0.02
% Clingers	0.53	0.54	0.59
% Sprawlers	0.01	-0.09	-0.29
% Swimmers	-0.72	-0.73	-0.14
Burrower taxa	**-2.75**	**-2.20**	0.28
Burrower taxa rarefacted	**-2.18**	**-1.47**	0.61
Climber taxa	-0.59	-0.30	-0.20
Climber taxa rarefacted	-0.41	-0.26	-0.26
Clinger taxa	**-2.20**	**-1.64**	0.33
Clinger taxa rarefacted	**-1.12**	-0.54	0.84
Sprawler taxa	**-1.95**	**-2.01**	0.49
Sprawler taxa rarefacted	**-1.57**	**-1.72**	**1.28**
Swimmer taxa	**-3.34**	**-2.96**	0.18
Swimmer taxa rarefacted	**-2.40**	**-1.90**	0.11
% Multivoltine	-0.99	-0.29	0.64

% Univoltine	**1.42**	0.88	0.22

Table 11. Assessment of index and metric bias attributed to protocols. The number of index or metric differences that are above or below zero by an amount greater than expected for sampling error alone determines bias. Critical Chi-square values are 3.841 and 2.706 for probabilities of 0.05 and 0.10, respectively. The protocols in the comparisons are listed in the top two lines of each column.

Index/Metric	Kick Reachwide	Kick Targeted	Kick Surber	Surber Reachwide	Surber Targeted	Targeted Reachwide
Mountain Index	0.69	1.92	0.10	3.60	2.50	0.69
Mountain Index Percent	0.69	1.92	0.10	3.60	2.50	0.69
Mountain Index rarefacted	0.08	1.23	0.10	0.90	1.60	1.23
Mountain Index Percent rarefacted	0.08	1.23	0.10	0.90	1.60	1.23
FVP Index	1.23	0.08	0.10	0.10	0.00	0.69
FVP Index Percent	1.23	0.08	0.10	0.10	0.00	0.69
FVP Index rarefacted	0.69	0.00	0.10	0.10	0.00	0.69
FVP Index Percent rarefacted	0.69	0.00	0.10	0.10	0.00	0.69
Total taxa	**6.23**	3.77	0.00	3.60	1.60	0.31
Total taxa rarefacted	1.92	1.92	0.90	**4.90**	2.50	0.69
EPT taxa	**7.69**	2.77	0.40	**6.40**	0.10	1.92
EPT taxa rarefacted	3.77	0.69	0.00	1.60	0.10	1.92
Ephemeroptera taxa	1.92	1.23	0.10	2.50	0.90	0.08
Ephemeroptera taxa rarefacted	0.31	0.69	0.40	0.90	0.90	0.08
Plecoptera taxa	2.77	2.77	0.40	0.00	0.40	0.08
Plecoptera taxa rarefacted	0.31	0.08	0.10	0.10	0.40	0.00
Trichoptera taxa	1.92	1.23	0.00	3.60	0.40	1.23
Trichoptera taxa rarefacted	0.31	0.08	0.10	2.50	0.10	0.31
Diptera taxa	1.92	1.92	0.10	3.60	1.60	0.00
Diptera taxa rarefacted	1.92	0.31	0.40	2.50	1.60	0.31
Chironomidae taxa	1.92	0.08	0.40	1.60	0.90	0.31
Chironomidae taxa rarefacted	0.69	0.31	0.10	0.10	0.40	0.31
Tanytarsini taxa	0.31	0.08	0.00	0.10	0.10	0.00
Coleoptera taxa	0.31	0.08	0.00	0.40	0.40	0.08
Coleoptera taxa rarefacted	0.08	0.31	0.10	0.00	0.10	0.08
Crustacea and Mollusca taxa	0.00	0.08	0.90	0.90	0.90	0.00
Crustacea and Mollusca taxa raref.	0.08	0.08	0.90	0.90	0.40	0.00
% Dominant	0.00	0.08	0.40	0.90	0.90	0.08
Shannon-Wiener Index	0.31	0.69	0.90	1.60	0.90	0.08
Evenness	0.08	0.69	0.40	0.00	0.10	0.31
D	0.00	0.08	0.90	0.90	0.90	0.08
D Margoleff's	2.77	2.77	0.00	**4.90**	2.50	0.31
% EPT	0.31	0.69	0.10	0.40	0.40	0.31
% Ephemeroptera	0.08	0.31	0.10	0.10	0.40	0.08
% Plecoptera	0.08	0.31	0.90	0.10	0.10	0.00
% Trichoptera	0.69	0.31	0.90	0.00	0.00	0.08
% Coleoptera	0.00	1.23	0.00	0.00	0.90	0.08
% Diptera	0.08	0.31	0.90	0.40	0.10	0.00
% Chironomidae	0.69	1.23	0.10	0.10	0.40	0.08
% Tanytarsini	0.08	0.69	0.10	0.10	0.10	0.31

% non-Insects	0.31	0.08	0.10	0.10	0.00	0.08
% Amphipoda	0.08	0.08	0.00	0.10	0.40	0.08
% Gastropoda	0.08	0.08	0.00	0.10	0.10	0.00
% Bivalvia	0.00	1.23	0.00	0.10	1.60	1.92
% Crustacea and Mollusca	0.08	0.00	0.10	0.00	0.10	0.08
Intolerant taxa	1.92	1.23	0.00	1.60	0.90	0.08
Intolerant taxa rarefacted	1.23	0.69	0.00	1.60	0.40	0.31
MDEQ Intolerant taxa	1.92	3.77	0.40	0.90	0.90	0.08
MDEQ Intolerant taxa rarefacted	1.92	1.23	0.40	0.10	0.40	0.08
Tolerant taxa	1.92	0.08	0.10	0.90	0.40	0.69
Tolerant taxa rarefacted	0.69	0.00	0.10	0.90	0.10	1.23
% Intolerant	1.92	1.92	0.40	2.50	2.50	0.69
% Tolerant	0.31	0.31	0.10	0.40	0.10	1.92
MDEQ % Tolerant	0.08	0.00	0.40	0.40	0.90	0.00
Beck's Biotic Index	3.77	3.77	0.00	2.50	1.60	0.31
Beck's Biotic Index rarefacted	1.92	1.23	0.10	1.60	0.90	0.08
HBI	2.77	1.23	0.10	1.60	0.10	2.77
% Baetidae of Ephemeroptera	0.69	0.69	0.10	0.10	0.10	0.08
% Hydropsychidae of EPT	0.31	0.08	0.10	0.00	0.10	0.31
% Hydropsychidae of Trichoptera	1.23	0.00	0.10	0.10	0.40	0.69
% Collectors	0.08	0.08	1.60	3.60	**4.90**	0.08
% Filterers	0.69	0.69	0.00	0.10	0.10	0.08
% Predators	1.23	0.69	1.60	0.90	0.10	0.31
% Scrapers	0.08	0.31	0.00	0.40	1.60	0.00
% Shredders	0.69	0.31	0.40	1.60	0.40	0.31
Collector taxa	2.77	1.23	0.10	2.50	1.60	0.31
Collector taxa rarefacted	2.77	0.31	0.00	2.50	0.90	0.69
Filterer taxa	1.92	0.31	0.00	1.60	1.60	0.08
Filterer taxa rarefacted	0.31	0.31	0.00	0.90	0.90	0.08
Predator taxa	2.77	1.92	0.00	2.50	1.60	0.08
Predator taxa rarefacted	1.92	1.92	0.40	1.60	1.60	0.00
Scraper taxa	0.31	0.31	0.90	3.60	0.40	0.00
Scraper taxa rarefacted	0.00	0.00	0.40	0.40	0.40	0.08
Shredder taxa	0.00	0.69	0.40	0.10	0.10	0.31
Shredder taxa rarefacted	0.08	0.08	0.10	0.00	0.00	0.00
% Burrowers	0.69	2.77	0.10	2.50	1.60	0.00
% Climbers	0.31	0.08	0.00	0.10	0.10	0.00
% Clingers	0.00	0.08	0.10	0.10	0.40	0.08
% Sprawlers	0.00	0.00	0.10	0.10	0.40	0.08
% Swimmers	1.23	0.08	0.10	0.40	0.10	0.08
Burrower taxa	**4.92**	2.77	0.40	**4.90**	3.60	0.69
Burrower taxa rarefacted	3.77	1.92	0.90	3.60	2.50	0.69
Climber taxa	0.08	0.00	0.00	0.40	0.10	0.31
Climber taxa rarefacted	0.69	0.69	0.10	0.90	0.40	0.00
Clinger taxa	**4.92**	1.92	0.00	**4.90**	2.50	0.31
Clinger taxa rarefacted	1.23	0.00	0.10	0.90	0.10	0.31
Sprawler taxa	**4.92**	**6.23**	0.10	3.60	3.60	0.31
Sprawler taxa rarefacted	1.92	1.23	0.40	2.50	1.60	0.08

Swimmer taxa	**4.92**	**4.92**	0.00	**6.40**	3.60	0.08
Swimmer taxa rarefacted	3.77	3.77	0.00	**4.90**	2.50	0.31
% Multivoltine	0.31	0.31	0.40	0.40	0.10	0.08
% Univoltine	1.23	0.31	0.10	0.90	0.10	0.31

Table 12. Weight of evidence of similarity among samples collected with 500µm and 1200µm mesh. Positive evidence of similarity is denoted with "+", ambiguous evidence or lack of evidence with "0".

Analysis	**Result**	**Evidence of similarity among protocols.**	**Supporting Text, Tables and Figures**
Ordination	Samples grouped by site. No consistent shift within sites related to protocols.	+	FigProtNMS
Indicator Analysis	No taxa appeared as significant indicators of protocols	+	Section Indicator Analysis above
Metric Variability	CVs of indices were less than 20%. CVs of 75% of metrics were less than 50%.	0	TableVariability
Sensitivity to Protocol	More than half of the metrics showed mean differences that were less than one standard deviation based on sampling error among traveling kick and the EMAP protocols. The means of metrics calculated from Surber samples were less than one standard deviation away from means calculated from traveling kick methods for 90% of metrics.	0	TabProtDiffRMSE
Precision by protocol	NA	0	NA
Bias	After rarefaction, four measures of taxa counts were significantly biased when comparing traveling kick and EMAP reachwide methods. Only one metric was biased when comparing traveling kick and EMAP targeted riffle protocols and no metrics were biased when comparing traveling kick and Surber samples.	+	TabProtChiSquare, FigIndices

Table 13. Results of interlaboratory sort residue re-check

	Site ID	Stream Name	No. of organisms		PSE
			primary (a)	recovery (b)	
1	U 1062	Moose Creek	330	11	96.8
2	U 1069 MB	Medicine Lodge	317	42	88.3
3	U 1072 M	Bloody Dick Creek Mid	329	42	88.7
4	U 1081 MB	Beaver Creek	48	23	67.6
5	U 1086 M	Medicine Lodge US	318	57	84.8

•Mean PSE = 85.2%

Table 14. Results of taxonomic comparisons. Counts are the total numbers of specimens counted by primary (T1) and the QC taxonomist (T2); PDE is percent difference in enumeration.

No.	Site/sample	Count 1	Count 2	PDE	Taxonomic Disagreement (%)	
					Lowest practical	Genus
1	C0518-M500	301	323	3.5	47.4	13.0
2	U1062-M	318	320	0.3	35.9	12.5
3	U1064-M	324	330	0.9	42.4	14.2
4	C0511-M500	321	314	1.1	32.8	7.6
5	U1058-M	329	335	0.9	63.0	29.6*
6	C01DOGC01	313	306	1.1	72.5	12.8
7	C01LTBLR02	185	180	1.4	53.0	6.5
8	M09LUMPG01	304	305	0.2	34.8	17.4*
9	C01TGRPC01	290	289	0.2	73.1	12.4
10	M09TENMC05	180	168	3.4	42.2	22.2*
Mean				**1.3**	**49.7**	**14.8**

Appendix A

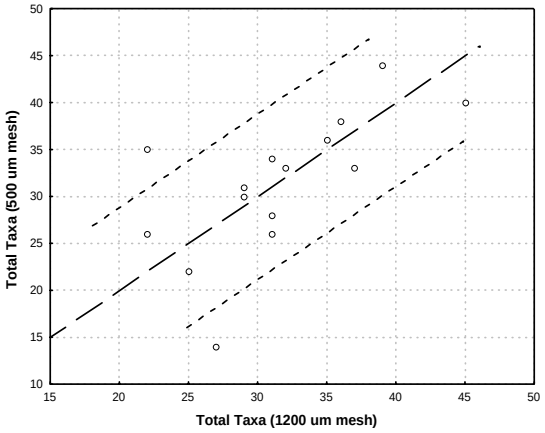
Metric Bias with Mesh Size

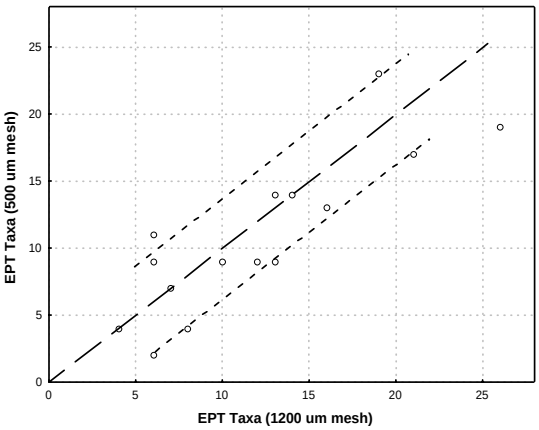
Plots illustrate metric values calculated from samples collected with 1200 μm mesh on the *x*-axis against 500 μm metrics on the *y* axis. The unity (1:1) line is shown with the 90% detectable difference on either side. Those points falling inside of the 90% limits are considered as "ties". They are essentially the same given the performance of the sampling techniques. Those points falling inside of the 90% limits may indicate differences that are due to more than just sampling error and natural variability. These are the points that determine mesh size bias in the chi-square analysis.

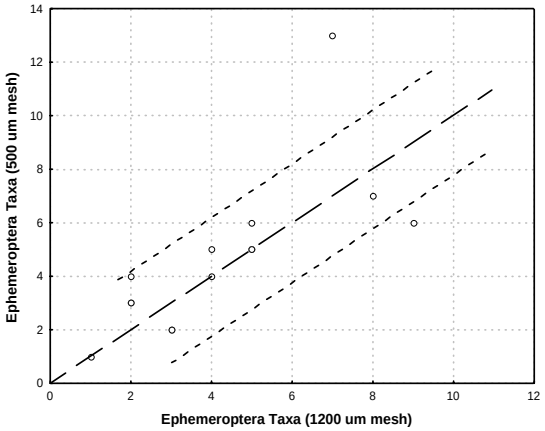
Metrics are shown in the following order, by metric type:

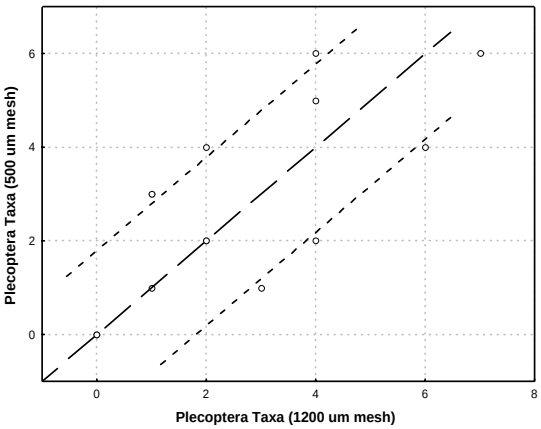
Richness
Diversity
Composition
Pollution Tolerance
Functional Feeding Groups
Habit
Voltinism

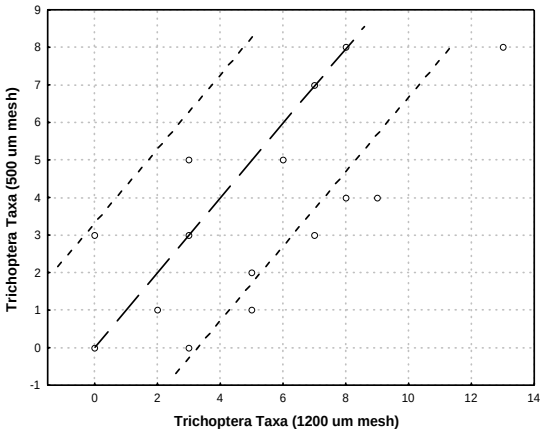
All graph legends use "um" to represent "μm" or microns because the Greek symbol was not available in the graphing software.

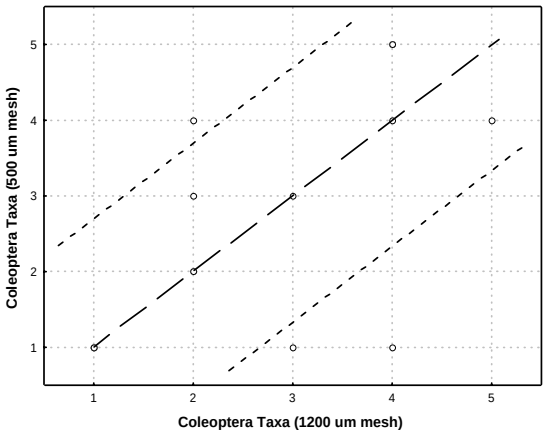

Total Taxa (500 um mesh)
Total Taxa (1200 um mesh)

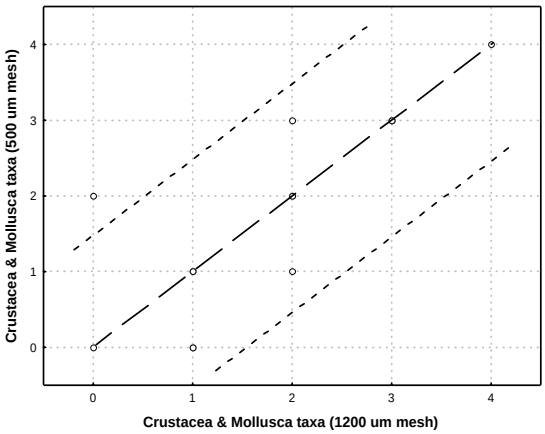

EPT Taxa (500 um mesh)
EPT Taxa (1200 um mesh)

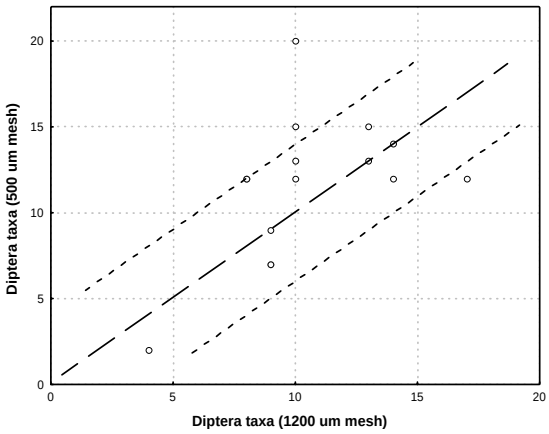

Ephemeroptera Taxa (500 um mesh)
Ephemeroptera Taxa (1200 um mesh)

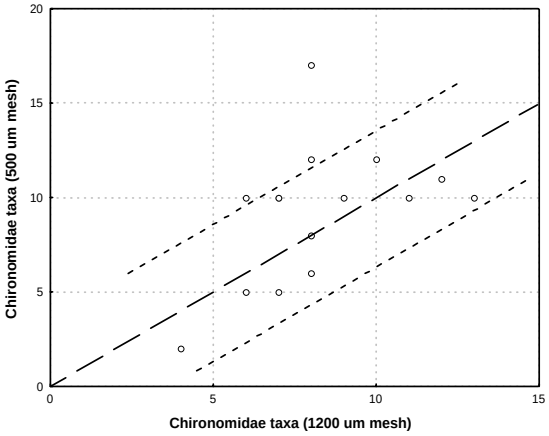

Plecoptera Taxa (500 um mesh)
Plecoptera Taxa (1200 um mesh)

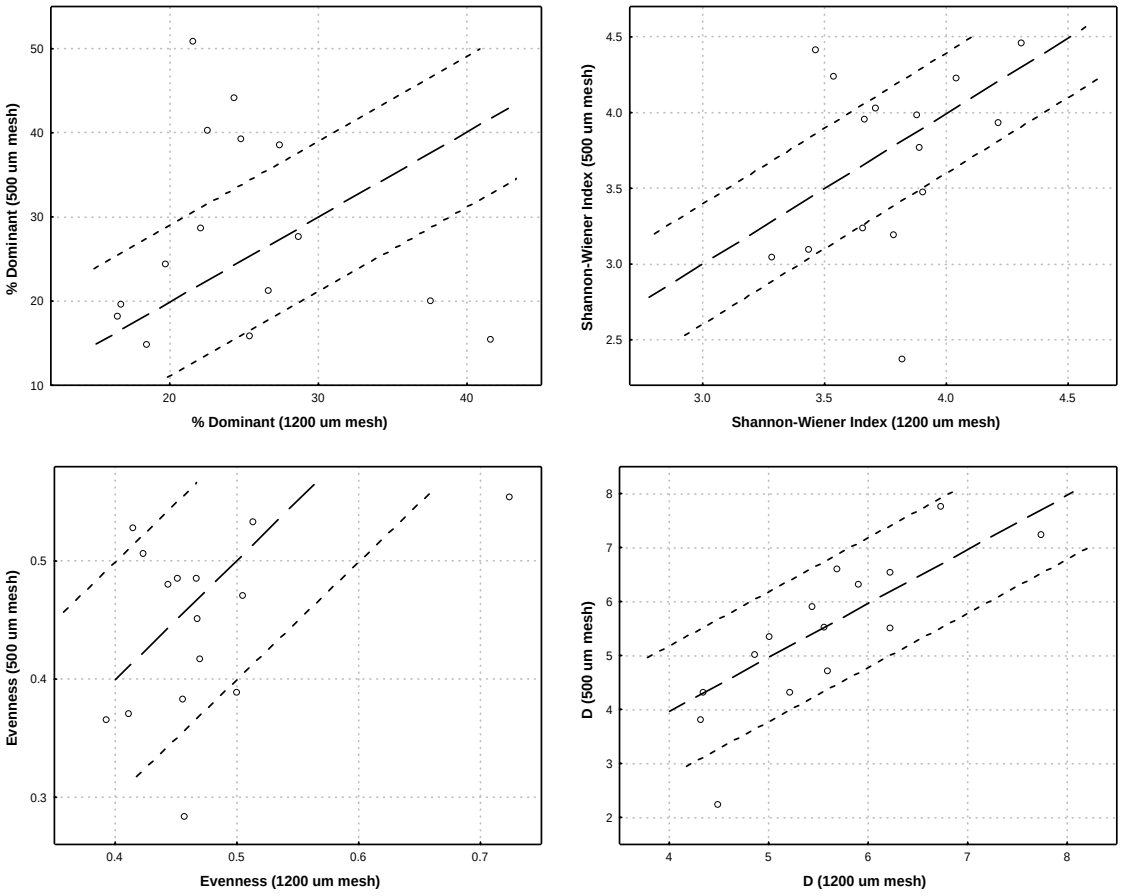

Trichoptera Taxa (500 um mesh)
Trichoptera Taxa (1200 um mesh)

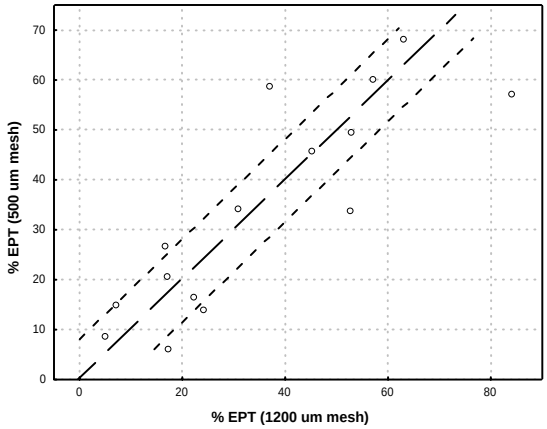

Coleoptera Taxa (500 um mesh)
Coleoptera Taxa (1200 um mesh)

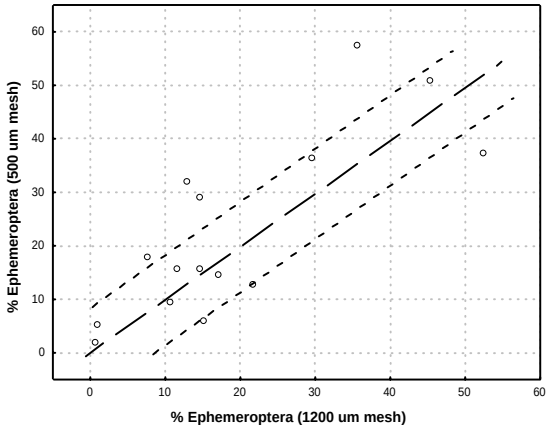

Crustacea & Mollusca taxa (500 um mesh)
Crustacea & Mollusca taxa (1200 um mesh)

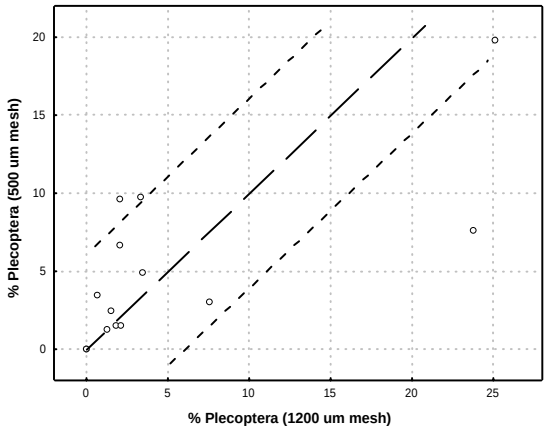

Diptera taxa (500 um mesh)
Diptera taxa (1200 um mesh)

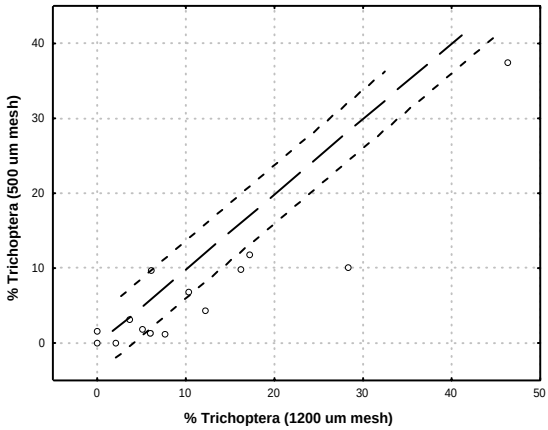

Chironomidae taxa (500 um mesh)
Chironomidae taxa (1200 um mesh)

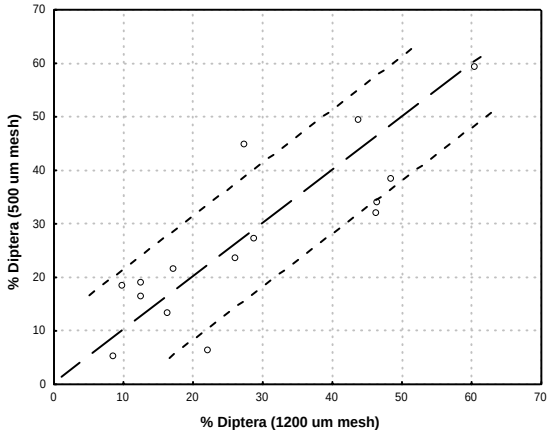

% Dominant (500 um mesh)
% Dominant (1200 um mesh)
Shannon-Wiener Index (500 um mesh)
Shannon-Wiener Index (1200 um mesh)
Evenness (500 um mesh)
Evenness (1200 um mesh)
D (500 um mesh)
D (1200 um mesh)

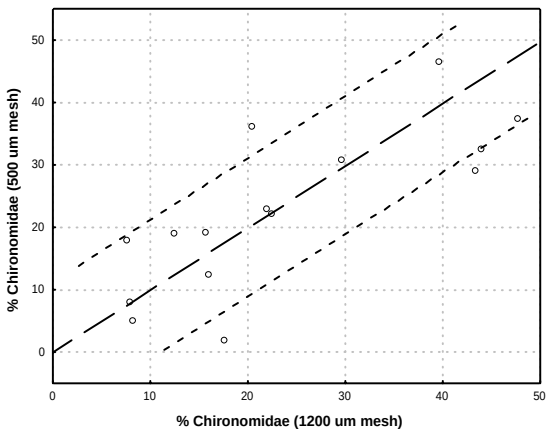

% EPT (500 um mesh)
% EPT (1200 um mesh)

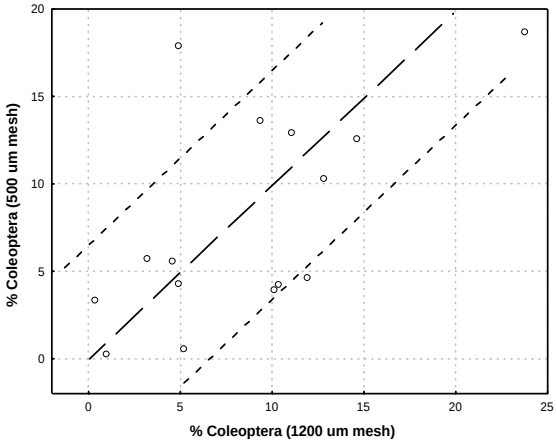

% Ephemeroptera (500 um mesh)
% Ephemeroptera (1200 um mesh)

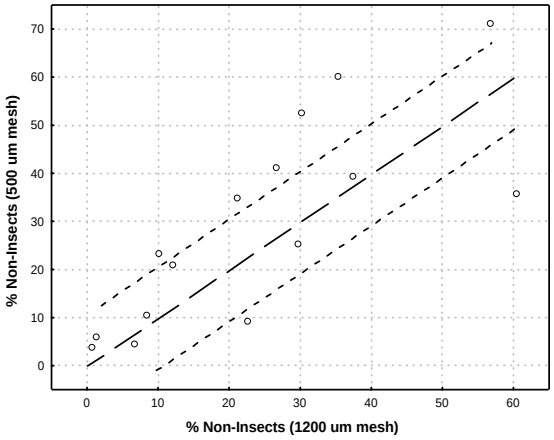

% Plecoptera (500 um mesh)
% Plecoptera (1200 um mesh)

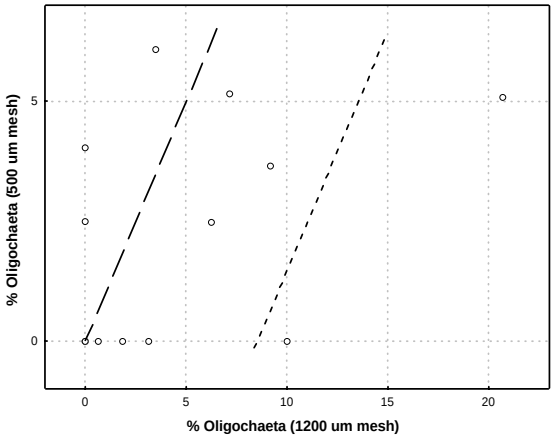

% Trichoptera (500 um mesh)
% Trichoptera (1200 um mesh)

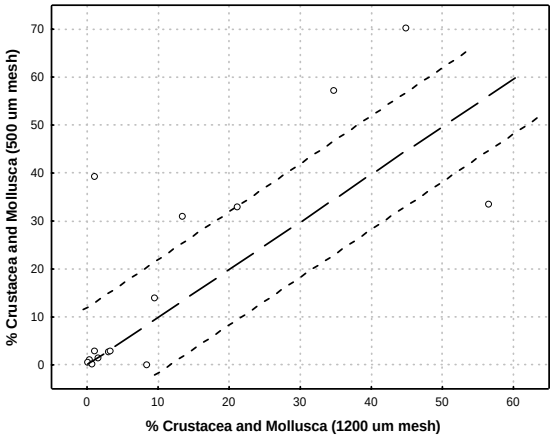

% Diptera (500 um mesh)
% Diptera (1200 um mesh)

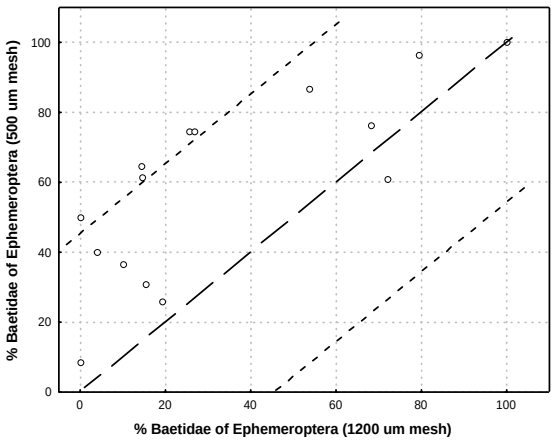

% Chironomidae (500 um mesh)
% Chironomidae (1200 um mesh)

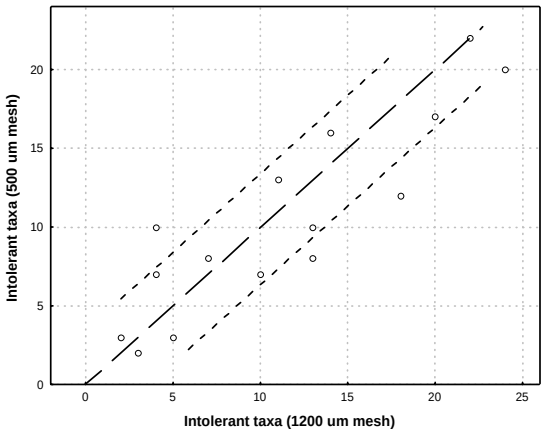

% Coleoptera (500 um mesh)
% Coleoptera (1200 um mesh)
0
5
10
15
20
25

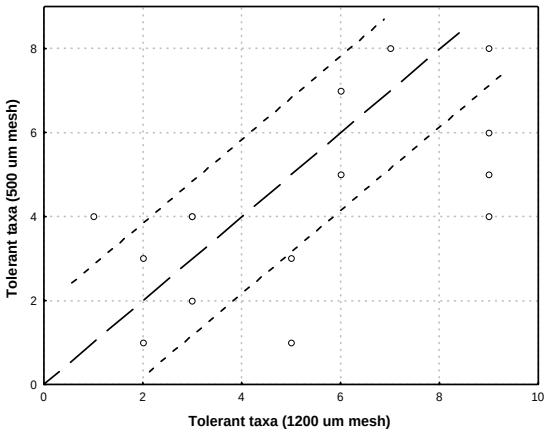

% Non-Insects (500 um mesh)
% Non-Insects (1200 um mesh)
0
10
20
30
40
50
60
70

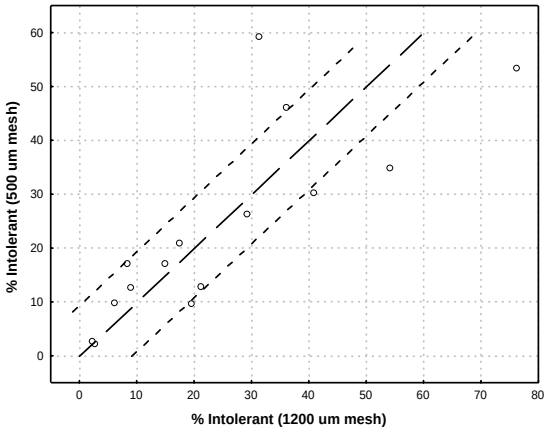

% Oligochaeta (500 um mesh)
% Oligochaeta (1200 um mesh)
0
5
10
15
20

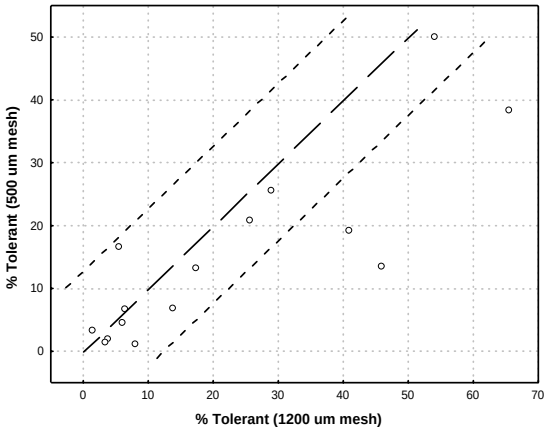

% Crustacea and Mollusca (500 um mesh)
% Crustacea and Mollusca (1200 um mesh)
0
10
20
30
40
50
60
70

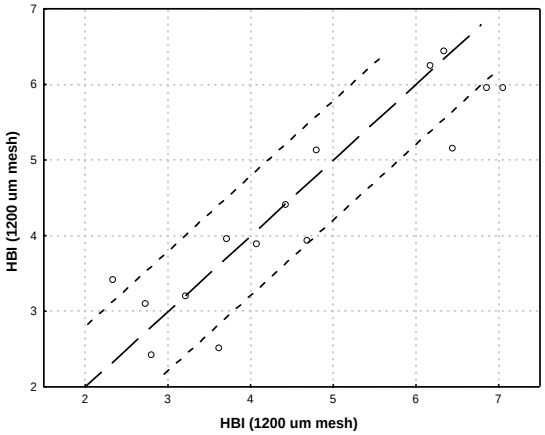

% Baetidae of Ephemeroptera (500 um mesh)
% Baetidae of Ephemeroptera (1200 um mesh)
0
20
40
60
80
100

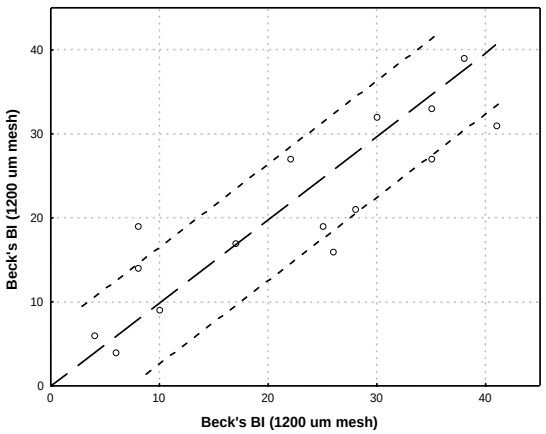

Intolerant taxa (500 um mesh)
Intolerant taxa (1200 um mesh)
0
5
10
15
20
25

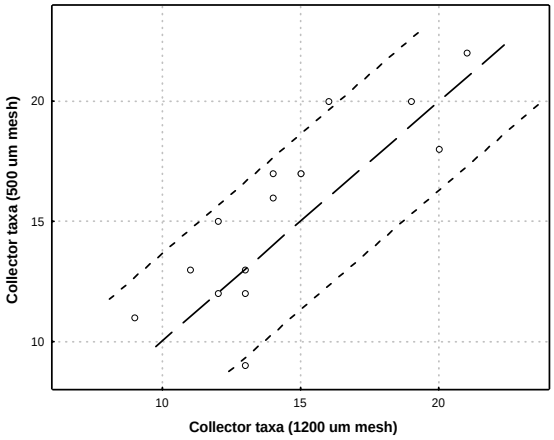

Tolerant taxa (500 um mesh)
Tolerant taxa (1200 um mesh)
0
2
4
6
8
10

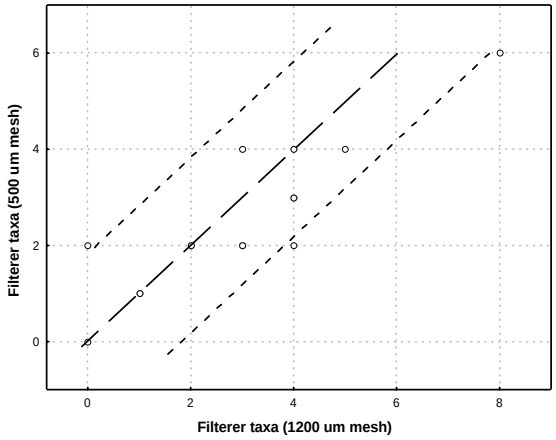

% Intolerant (500 um mesh)
% Intolerant (1200 um mesh)
0
10
20
30
40
50
60
70
80

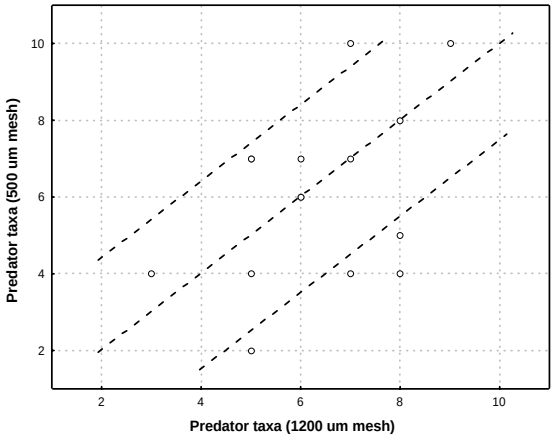

% Tolerant (500 um mesh)
% Tolerant (1200 um mesh)
0
10
20
30
40
50
60
70

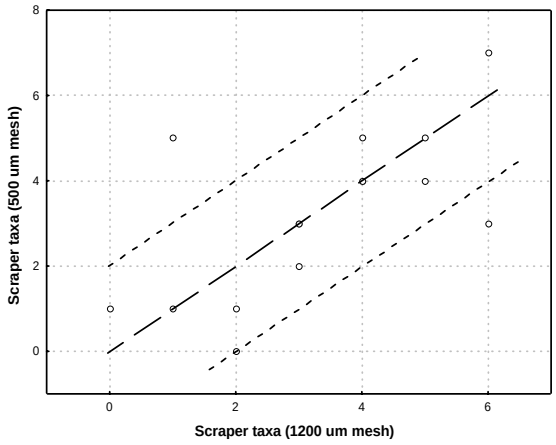

HBI (1200 um mesh)
HBI (1200 um mesh)
2
3
4
5
6
7

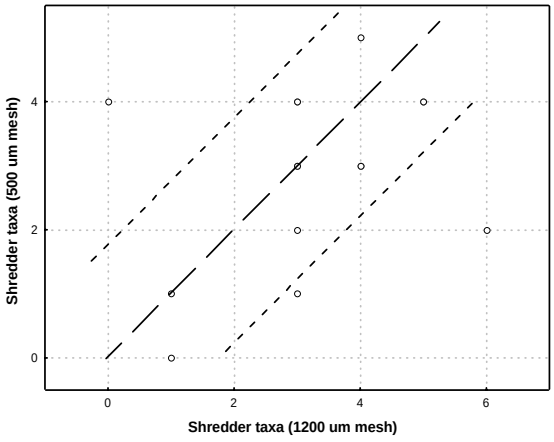

Beck's BI (1200 um mesh)
Beck's BI (1200 um mesh)
0
10
20
30
40

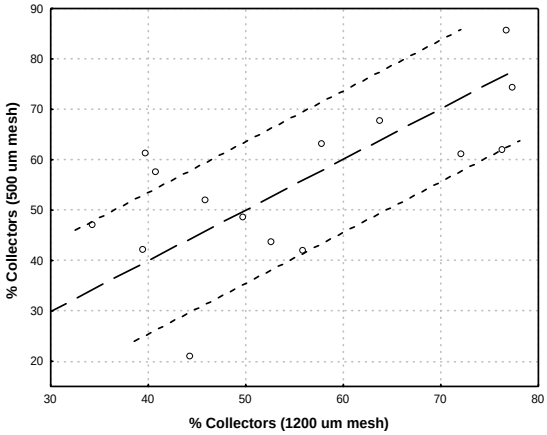

Collector taxa (500 um mesh)
Collector taxa (1200 um mesh)
10
15
20

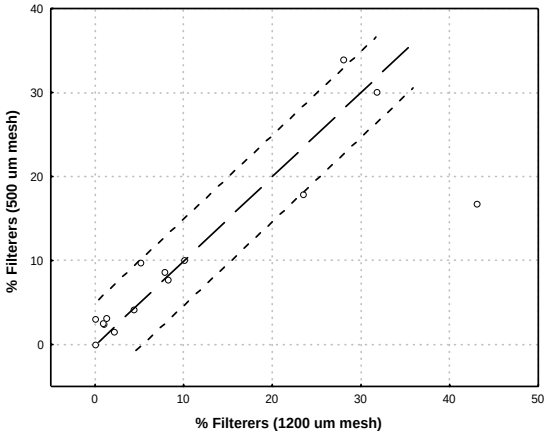

Filterer taxa (500 um mesh)
Filterer taxa (1200 um mesh)
0
2
4
6
8

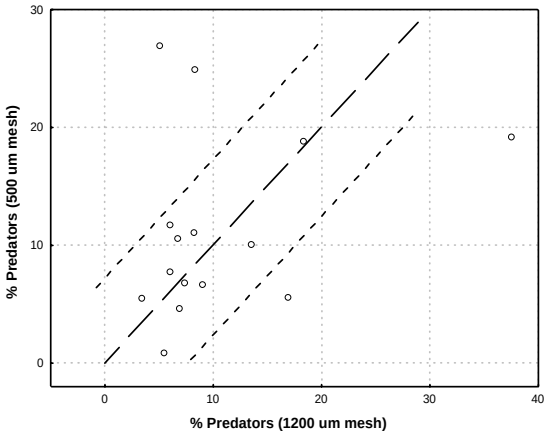

Predator taxa (500 um mesh)
Predator taxa (1200 um mesh)
2
4
6
8
10

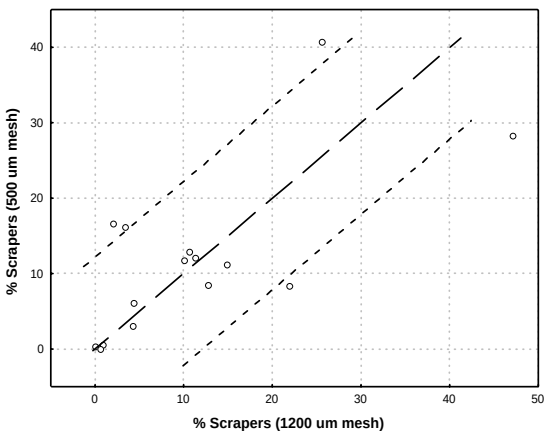

Scraper taxa (500 um mesh)
Scraper taxa (1200 um mesh)
0
2
4
6
8

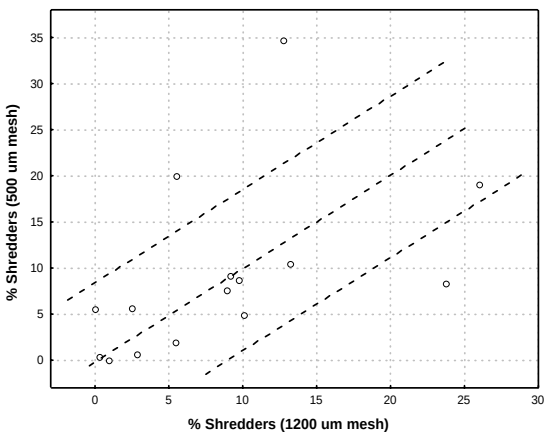

Shredder taxa (500 um mesh)
Shredder taxa (1200 um mesh)
0
2
4
6

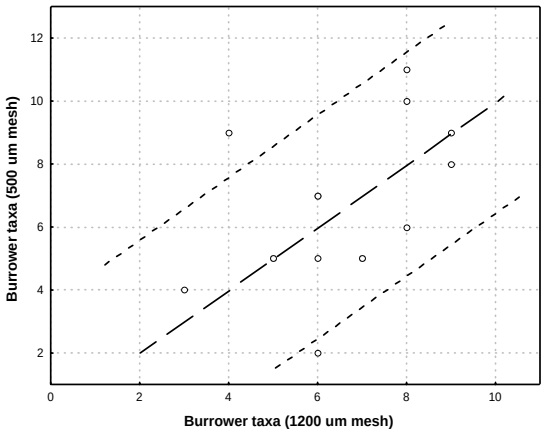

% Collectors (500 um mesh)
% Collectors (1200 um mesh)

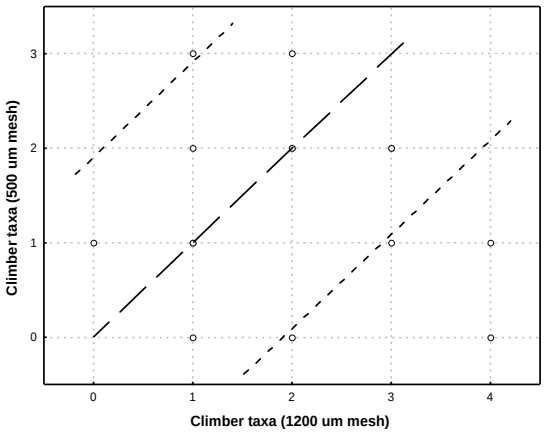

% Filterers (500 um mesh)
% Filterers (1200 um mesh)

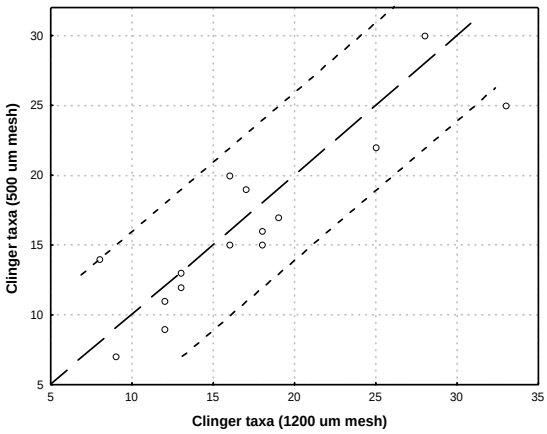

% Predators (500 um mesh)
% Predators (1200 um mesh)

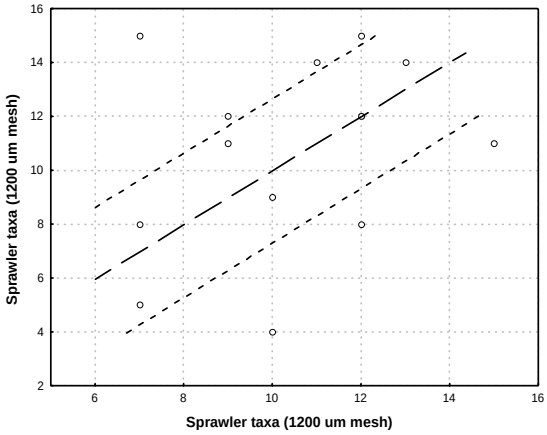

% Scrapers (500 um mesh)
% Scrapers (1200 um mesh)

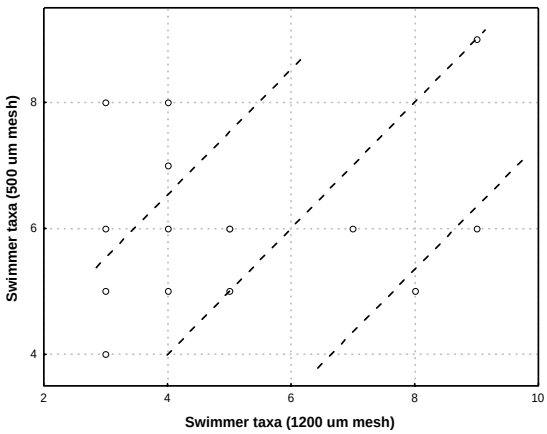

% Shredders (500 um mesh)
% Shredders (1200 um mesh)

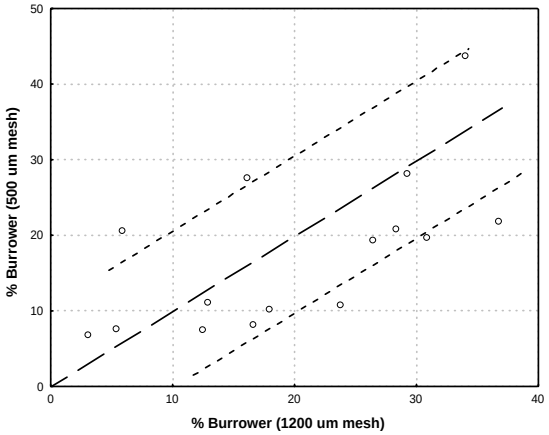

Burrower taxa (500 um mesh)
Burrower taxa (1200 um mesh)

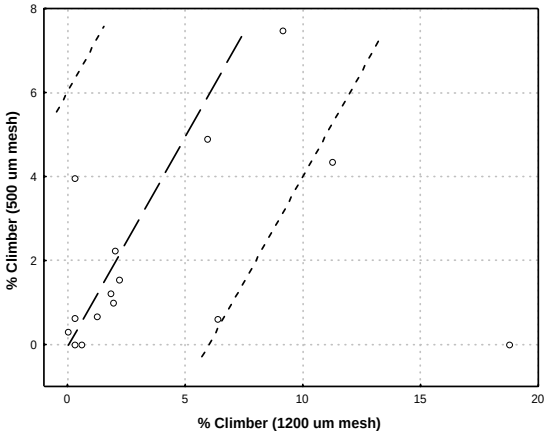

Climber taxa (500 um mesh)
Climber taxa (1200 um mesh)

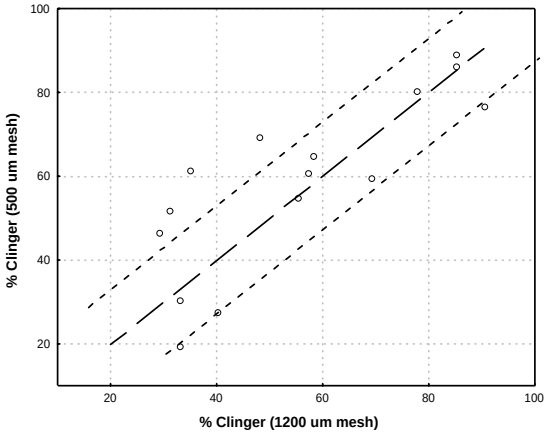

Clinger taxa (500 um mesh)
Clinger taxa (1200 um mesh)

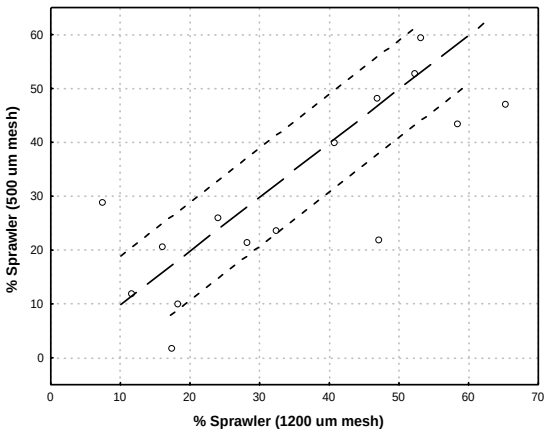

Sprawler taxa (1200 um mesh)
Sprawler taxa (1200 um mesh)

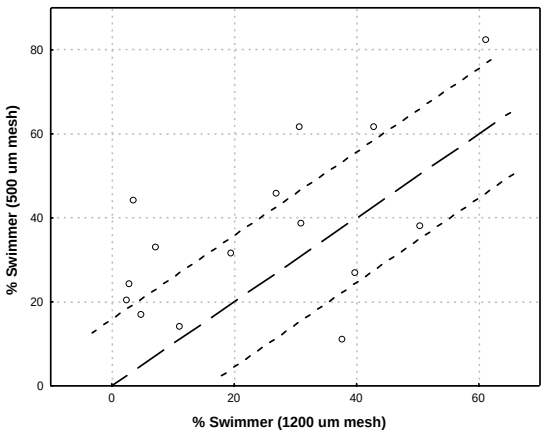

Swimmer taxa (500 um mesh)
Swimmer taxa (1200 um mesh)

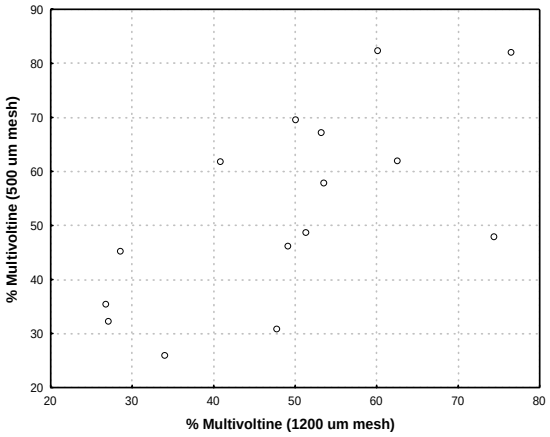

% Burrower (500 um mesh)
% Burrower (1200 um mesh)
0
10
20
30
40
50

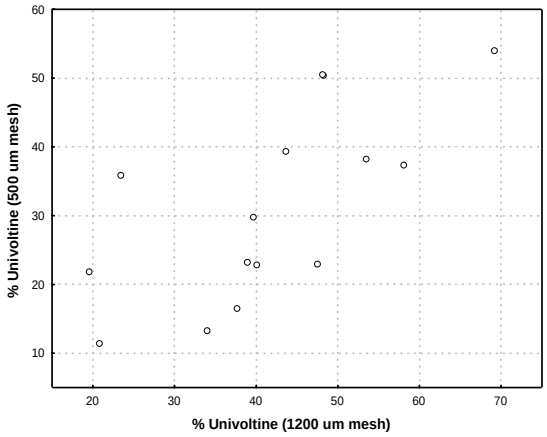

% Climber (500 um mesh)
% Climber (1200 um mesh)
0
2
4
6
8
0
5
10
15
20


% Clinger (500 um mesh)
% Clinger (1200 um mesh)
20
40
60
80
100


% Sprawler (500 um mesh)
% Sprawler (1200 um mesh)
0
10
20
30
40
50
60
70


% Swimmer (500 um mesh)
% Swimmer (1200 um mesh)
0
20
40
60
80


% Multivoltine (500 um mesh)
% Multivoltine (1200 um mesh)
90
80
70
60
50
40
30
20
20
30
40
50
60
70
80


% Univoltine (500 um mesh)
% Univoltine (1200 um mesh)
60
50
40
30
20
10
20
30
40
50
60
70

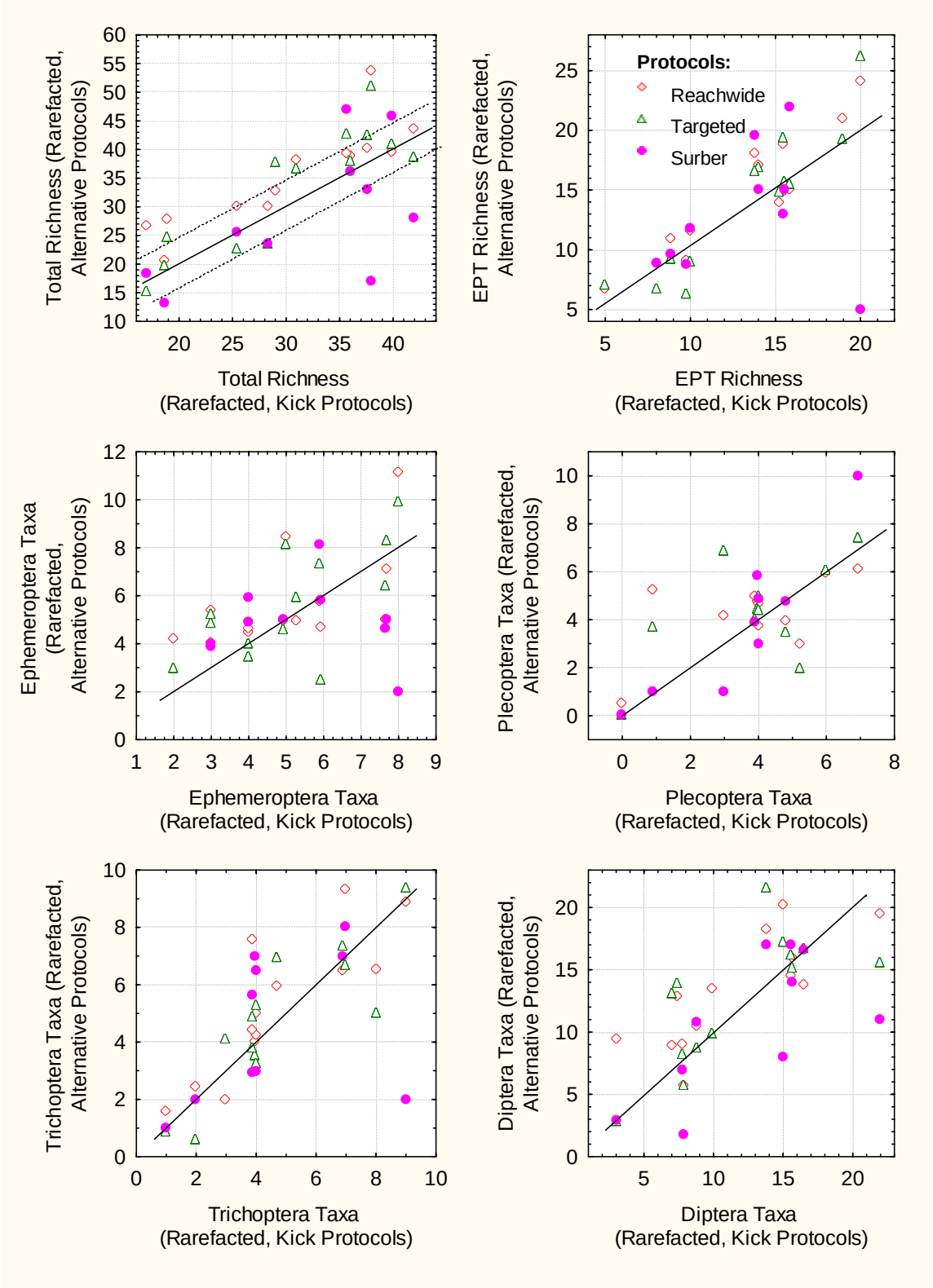
Appendix B

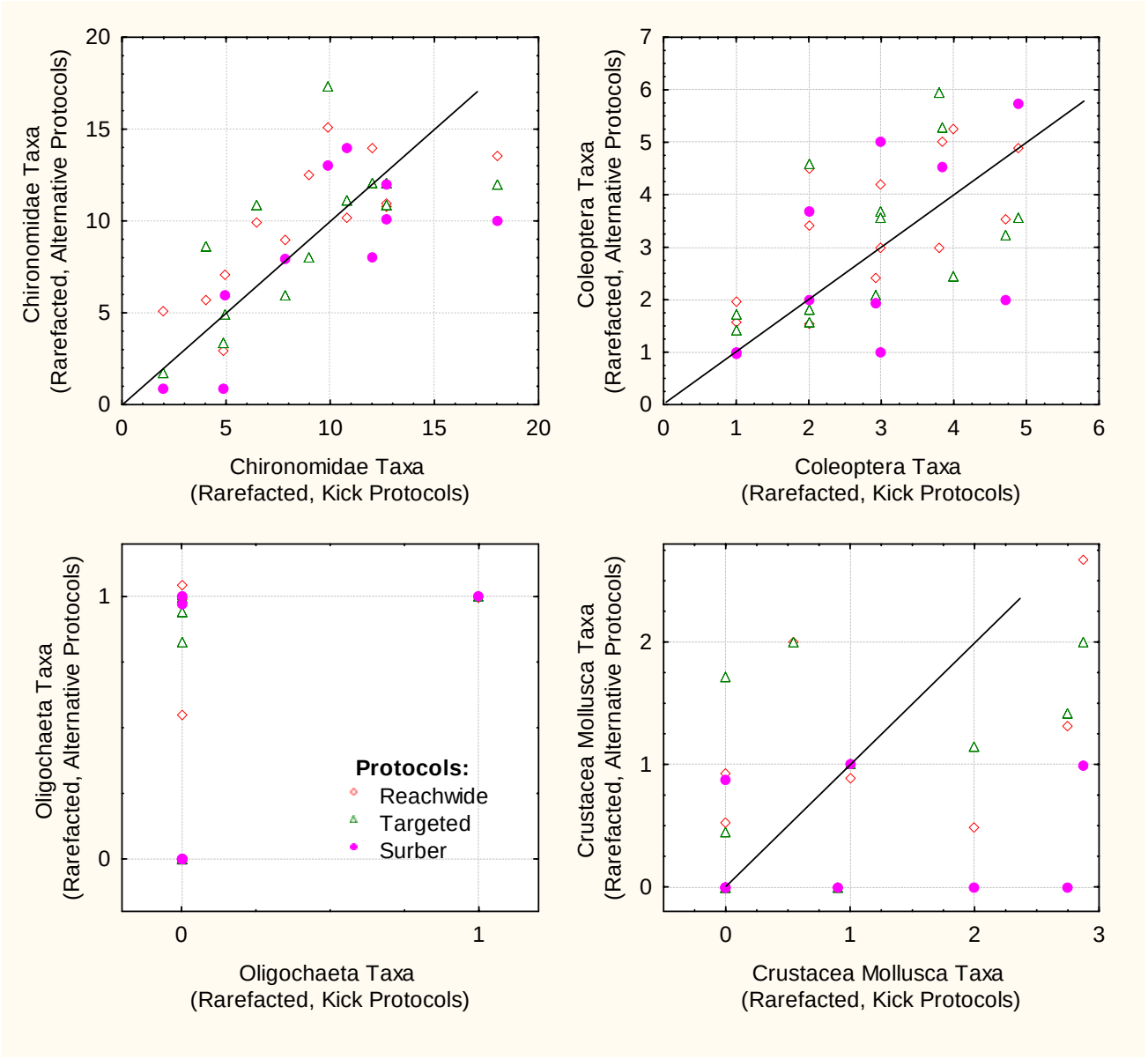
Metric Bias with Protocols

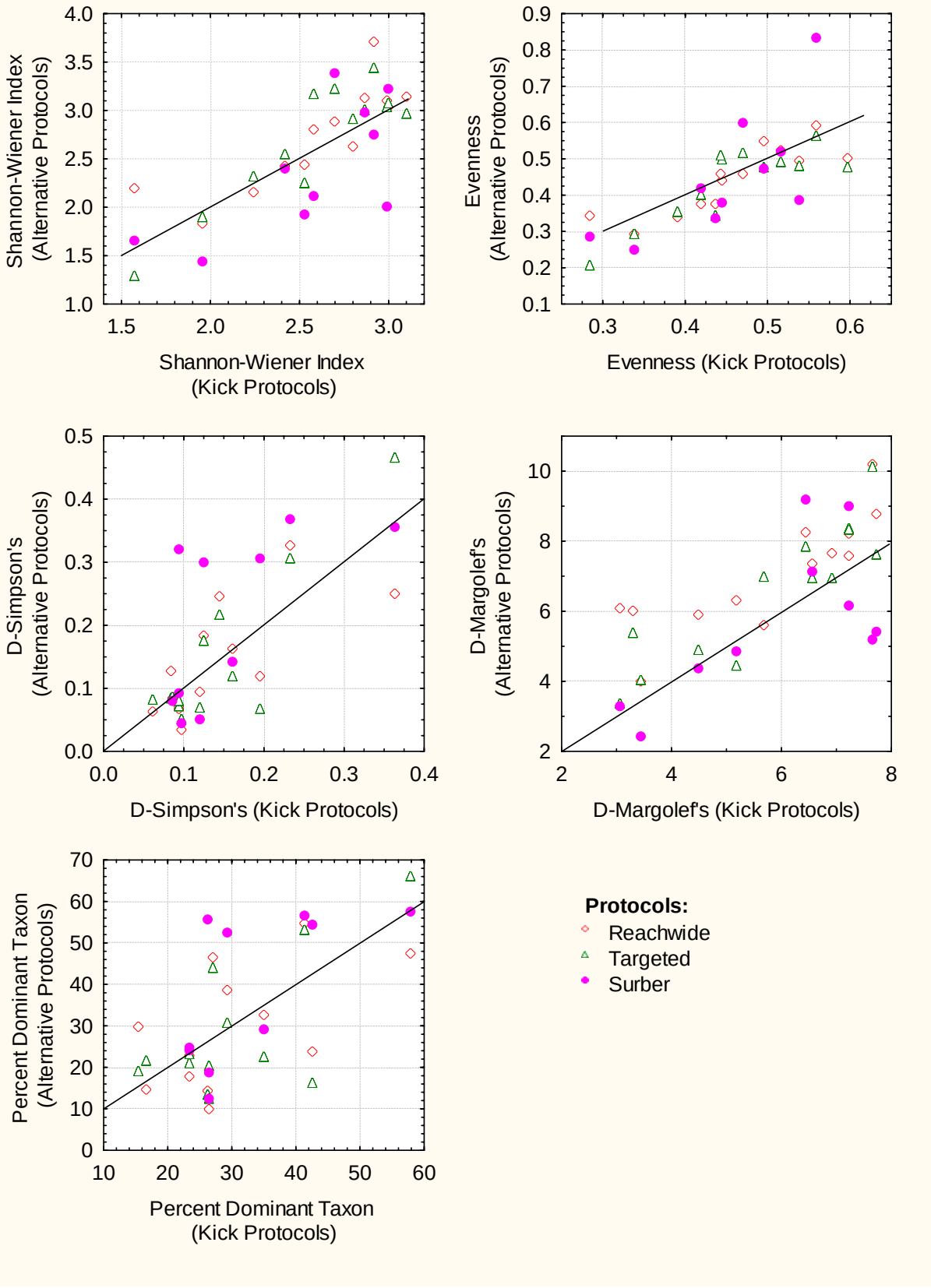
Plots illustrate metric values calculated from samples collected with traveling kick protocols on the *x*-axis against the alternative protocols (EMAP Reachwide, EMAP Targeted, and Surber) on the *y*-axis. The unity (1:1) line is shown.

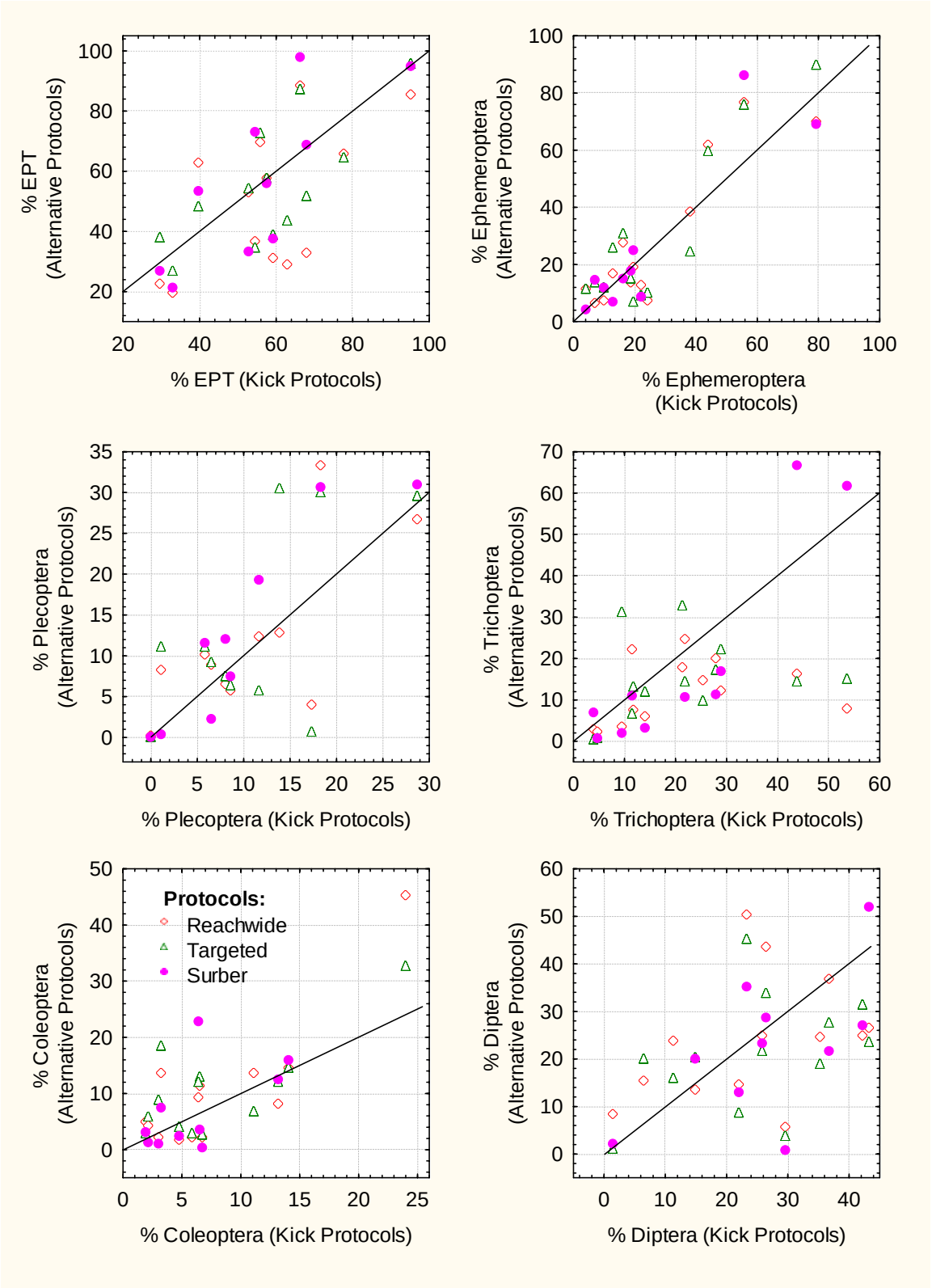
Metrics are shown in the following order, by metric type:

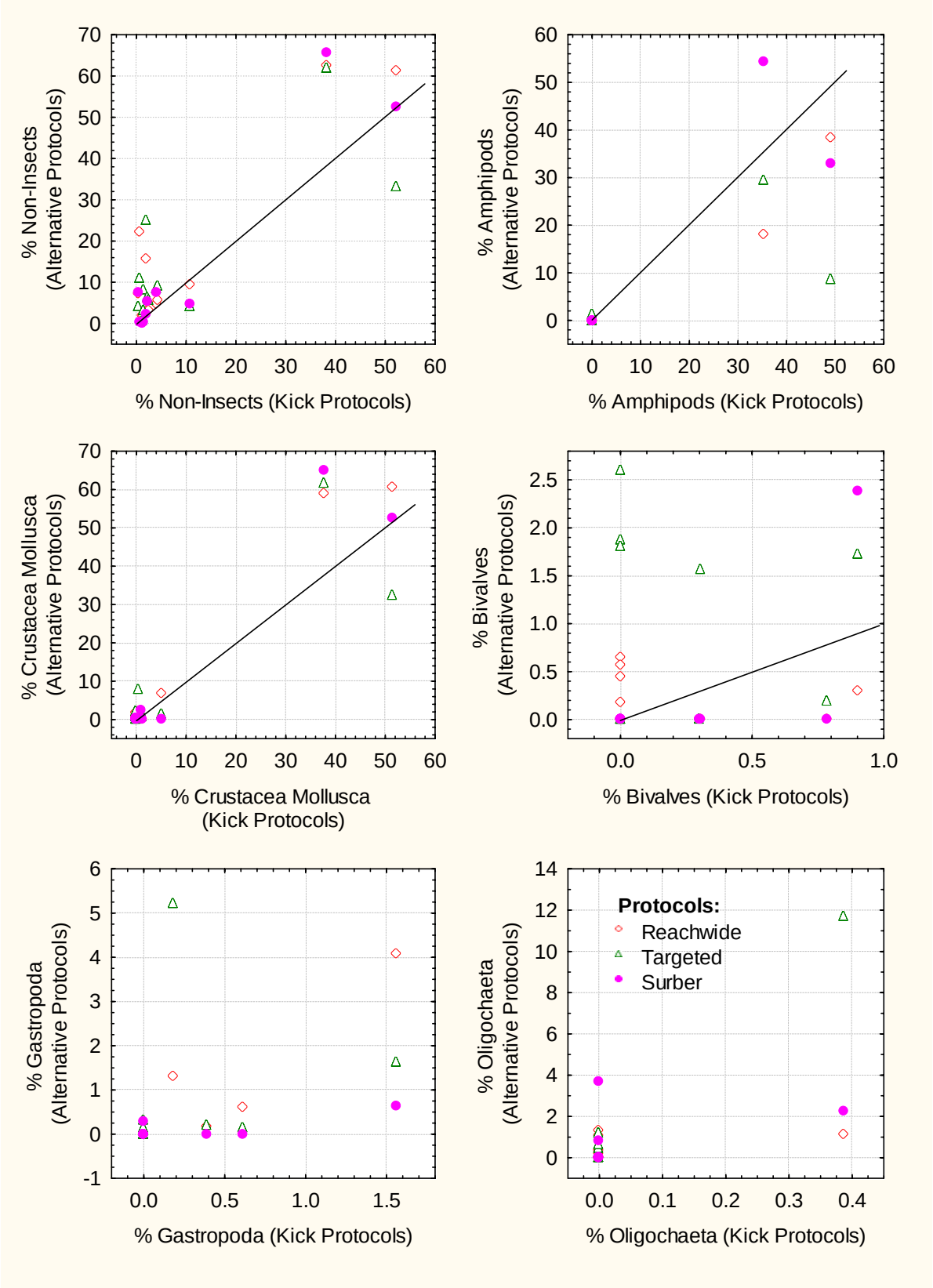
Richness
Diversity
Composition
Pollution Tolerance
Functional Feeding Groups
Habit
Voltinism

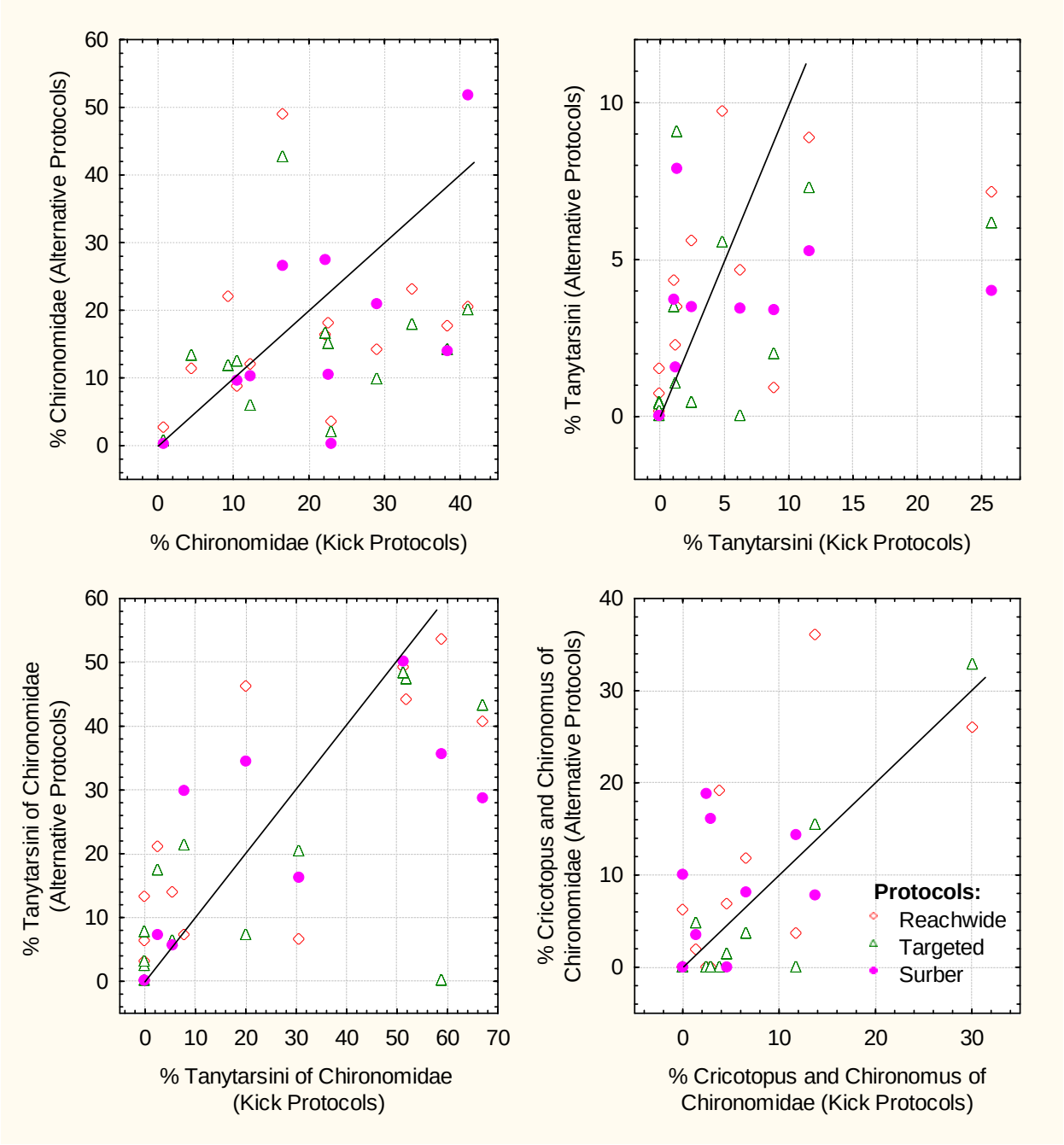

Total Richness (Rarefacted, Alternative Protocols)
Total Richness (Rarefacted, Kick Protocols)
EPT Richness (Rarefacted, Alternative Protocols)
Protocols:
Reachwide
Targeted
Surber
EPT Richness (Rarefacted, Kick Protocols)
Ephemeroptera Taxa (Rarefacted, Alternative Protocols)
Ephemeroptera Taxa (Rarefacted, Kick Protocols)
Plecoptera Taxa (Rarefacted, Alternative Protocols)
Plecoptera Taxa (Rarefacted, Kick Protocols)
Trichoptera Taxa (Rarefacted, Alternative Protocols)
Trichoptera Taxa (Rarefacted, Kick Protocols)
Diptera Taxa (Rarefacted, Alternative Protocols)
Diptera Taxa (Rarefacted, Kick Protocols)

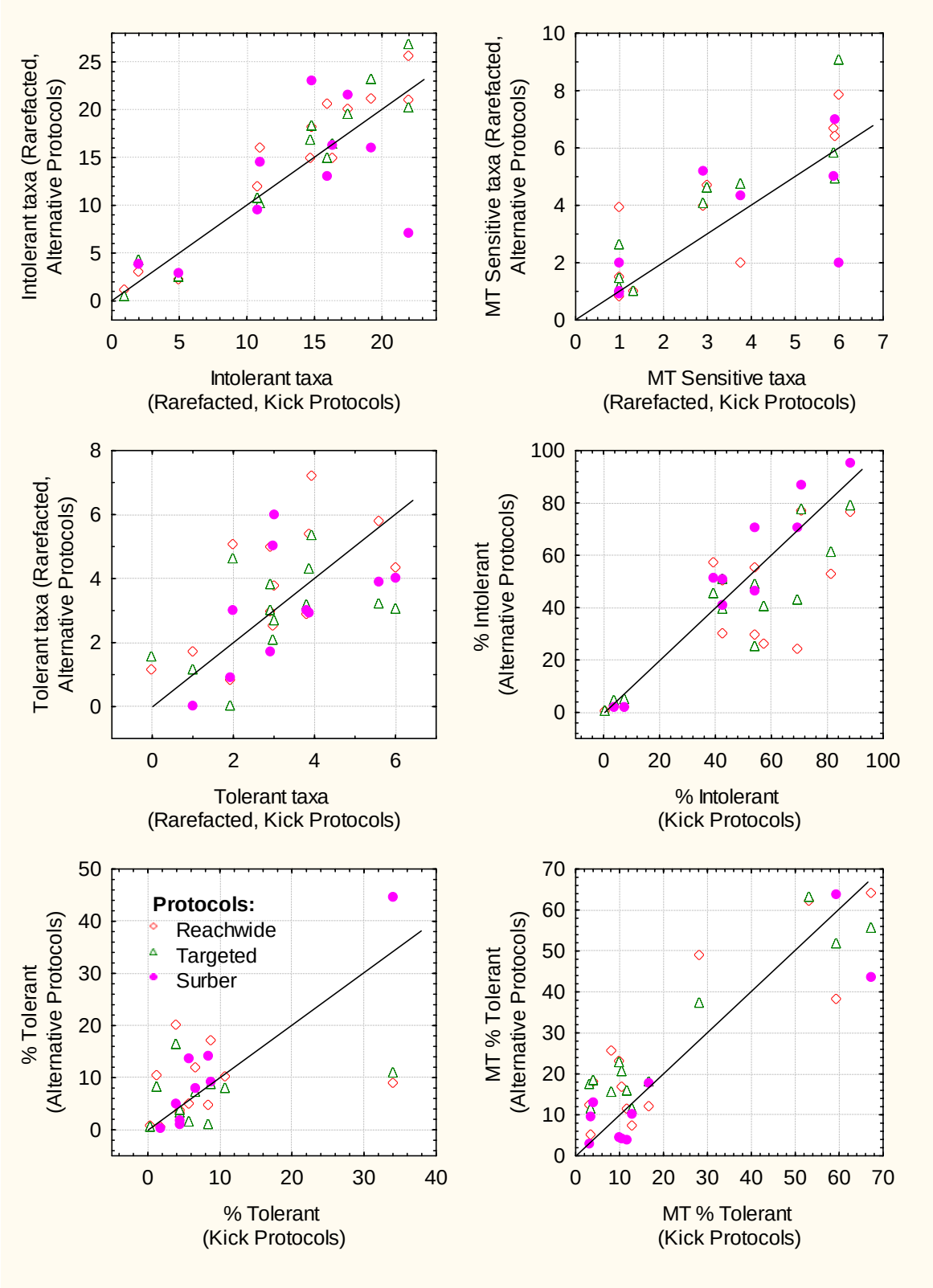

Chironomidae Taxa
(Rarefacted, Alternative Protocols)
Chironomidae Taxa
(Rarefacted, Kick Protocols)
Coleoptera Taxa
(Rarefacted, Alternative Protocols)
Coleoptera Taxa
(Rarefacted, Kick Protocols)
Oligochaeta Taxa
(Rarefacted, Alternative Protocols)
Oligochaeta Taxa
(Rarefacted, Kick Protocols)
Protocols:
Reachwide
Targeted
Surber
Crustacea Mollusca Taxa
(Rarefacted, Alternative Protocols)
Crustacea Mollusca Taxa
(Rarefacted, Kick Protocols)

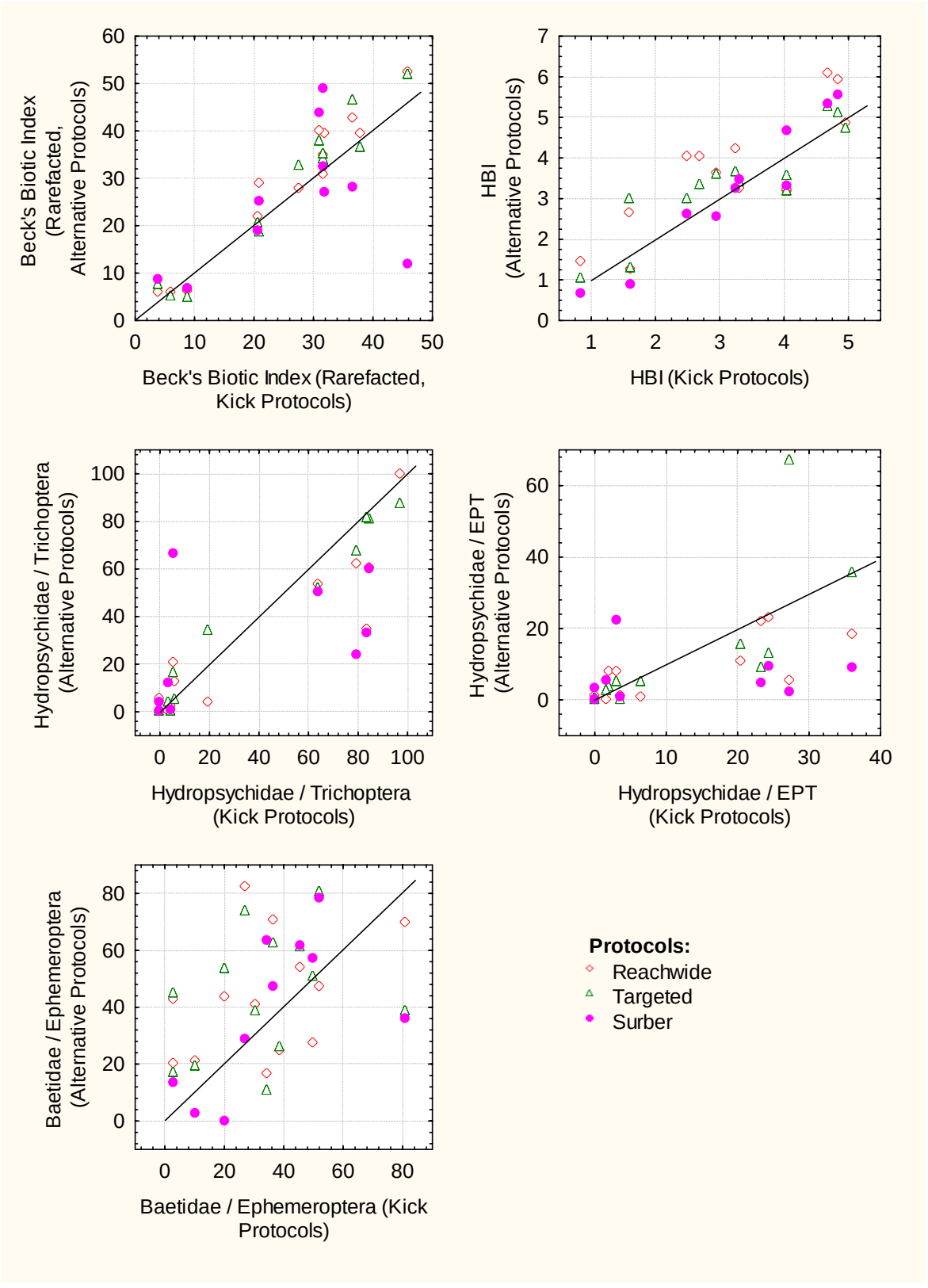

Shannon-Wiener Index (Alternative Protocols)
Shannon-Wiener Index (Kick Protocols)
Evenness (Alternative Protocols)
Evenness (Kick Protocols)
D-Simpson's (Alternative Protocols)
D-Simpson's (Kick Protocols)
D-Margolef's (Alternative Protocols)
D-Margolef's (Kick Protocols)
Percent Dominant Taxon (Alternative Protocols)
Percent Dominant Taxon (Kick Protocols)
Protocols:
Reachwide
Targeted
Surber

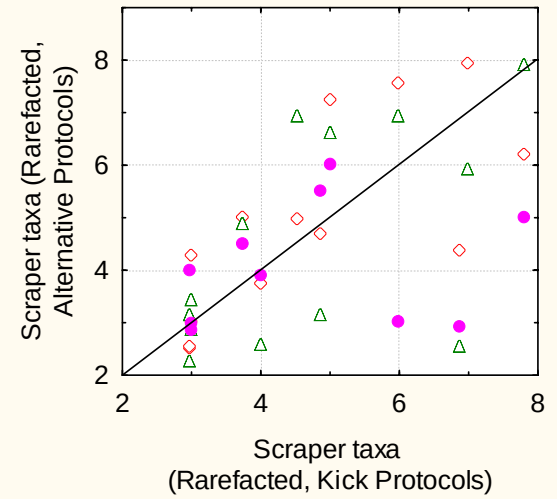

% EPT (Alternative Protocols)
% EPT (Kick Protocols)
% Ephemeroptera (Alternative Protocols)
% Ephemeroptera (Kick Protocols)
% Plecoptera (Alternative Protocols)
% Plecoptera (Kick Protocols)
% Trichoptera (Alternative Protocols)
% Trichoptera (Kick Protocols)
% Coleoptera (Alternative Protocols)
% Coleoptera (Kick Protocols)
% Diptera (Alternative Protocols)
% Diptera (Kick Protocols)
Protocols:
Reachwide
Targeted
Surber

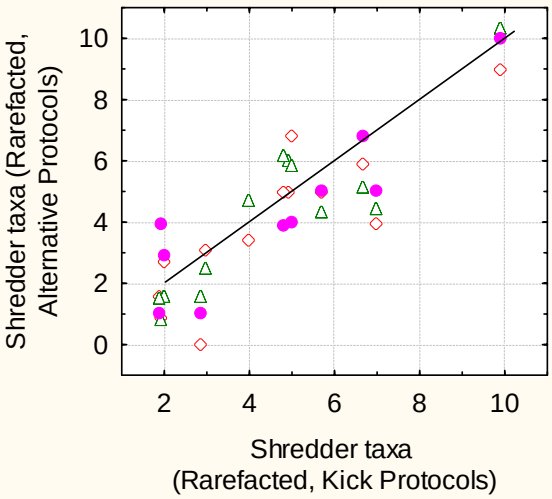

% Non-Insects (Alternative Protocols)
% Non-Insects (Kick Protocols)
% Amphipods (Alternative Protocols)
% Amphipods (Kick Protocols)
% Crustacea Mollusca (Alternative Protocols)
% Crustacea Mollusca (Kick Protocols)
% Bivalves (Alternative Protocols)
% Bivalves (Kick Protocols)
% Gastropoda (Alternative Protocols)
% Gastropoda (Kick Protocols)
% Oligochaeta (Alternative Protocols)
% Oligochaeta (Kick Protocols)
Protocols:
Reachwide
Targeted
Surber

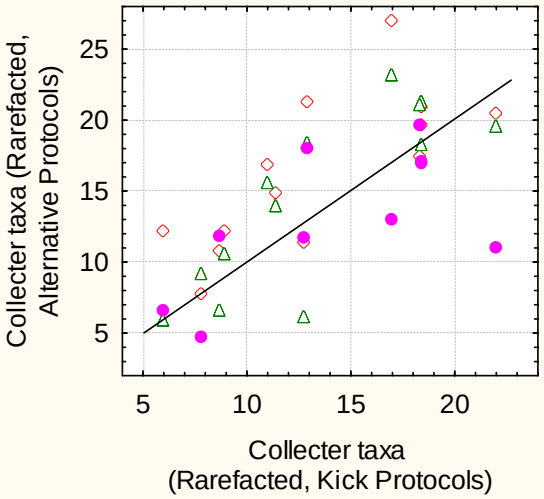

% Chironomidae (Alternative Protocols)
% Chironomidae (Kick Protocols)
% Tanytarsini (Alternative Protocols)
% Tanytarsini (Kick Protocols)
% Tanytarsini of Chironomidae (Alternative Protocols)
% Tanytarsini of Chironomidae (Kick Protocols)
% Cricotopus and Chironomus of Chironomidae (Alternative Protocols)
% Cricotopus and Chironomus of Chironomidae (Kick Protocols)
Protocols:
Reachwide
Targeted
Surber

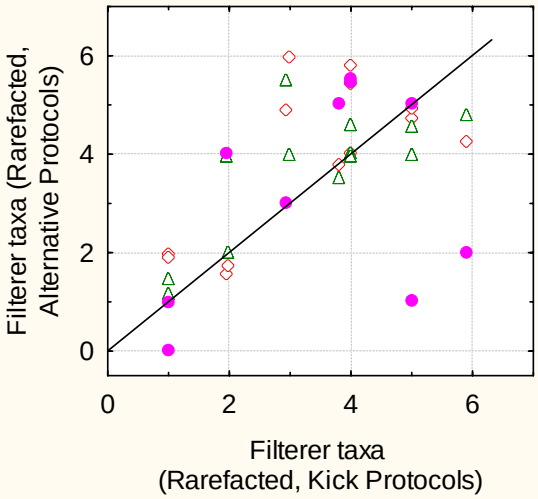

Intolerant taxa (Rarefacted, Alternative Protocols)
Intolerant taxa (Rarefacted, Kick Protocols)
MT Sensitive taxa (Rarefacted, Alternative Protocols)
MT Sensitive taxa (Rarefacted, Kick Protocols)
Tolerant taxa (Rarefacted, Alternative Protocols)
Tolerant taxa (Rarefacted, Kick Protocols)
% Intolerant (Alternative Protocols)
% Intolerant (Kick Protocols)
% Tolerant (Alternative Protocols)
% Tolerant (Kick Protocols)
MT % Tolerant (Alternative Protocols)
MT % Tolerant (Kick Protocols)
Protocols:
Reachwide
Targeted
Surber

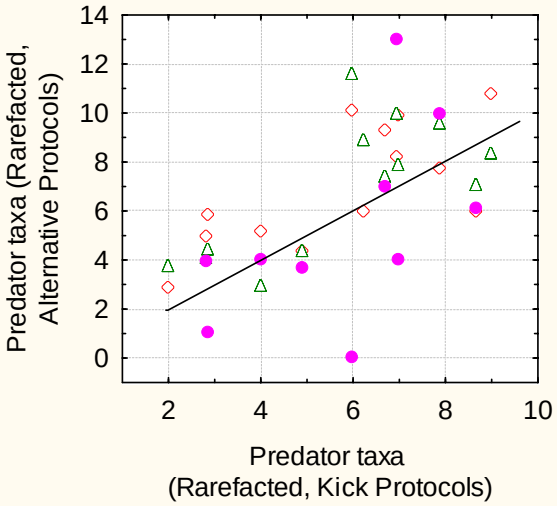

Beck's Biotic Index (Rarefacted, Alternative Protocols)
Beck's Biotic Index (Rarefacted, Kick Protocols)
HBI (Alternative Protocols)
HBI (Kick Protocols)
Hydropsychidae / Trichoptera (Alternative Protocols)
Hydropsychidae / Trichoptera (Kick Protocols)
Hydropsychidae / EPT (Alternative Protocols)
Hydropsychidae / EPT (Kick Protocols)
Baetidae / Ephemeroptera (Alternative Protocols)
Baetidae / Ephemeroptera (Kick Protocols)
Protocols:
Reachwide
Targeted
Surber

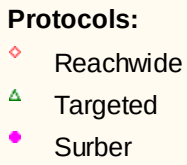

Scraper taxa (Rarefacted, Alternative Protocols)
Scraper taxa (Rarefacted, Kick Protocols)

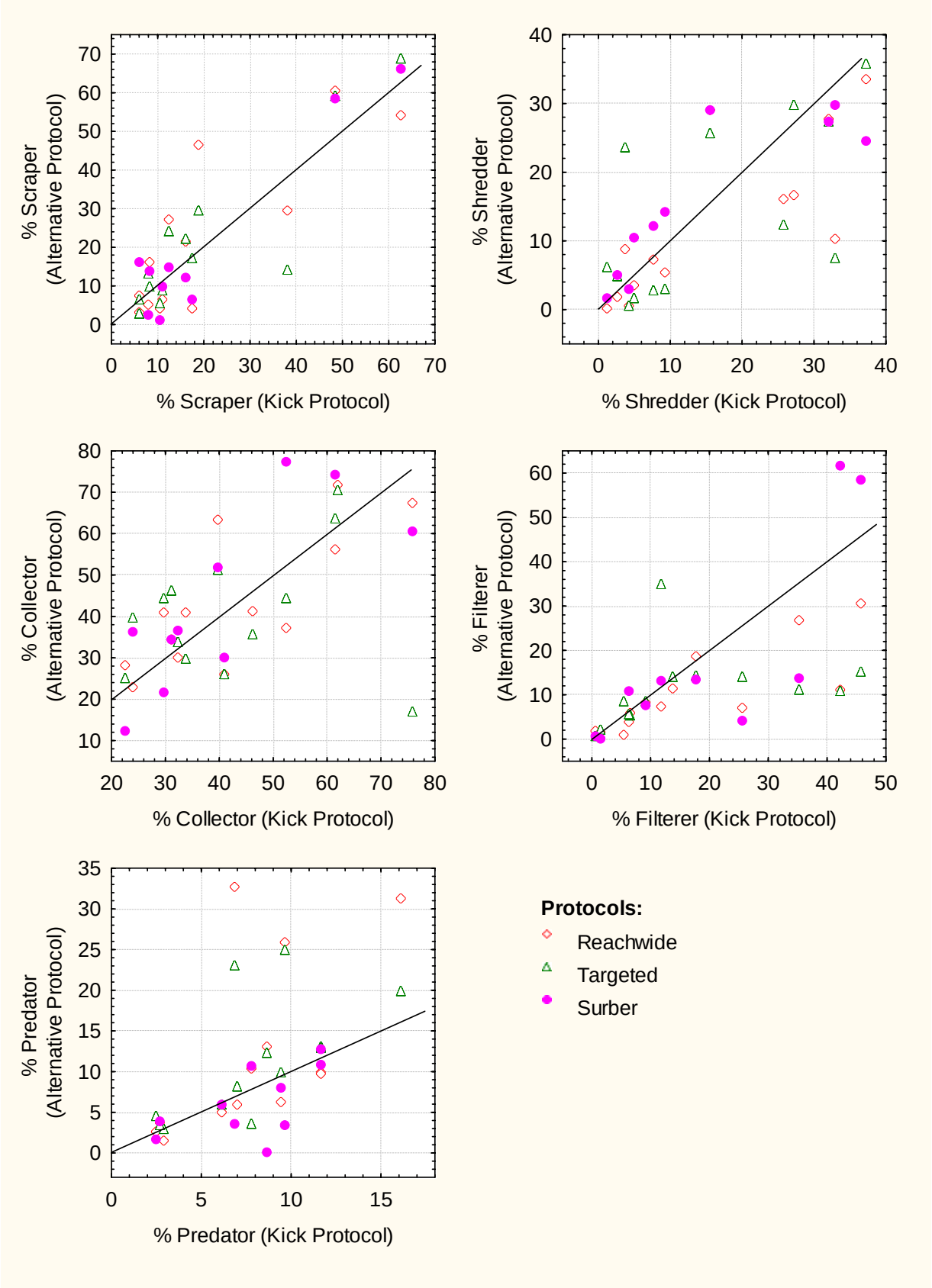

Shredder taxa (Rarefacted, Alternative Protocols)
Shredder taxa (Rarefacted, Kick Protocols)

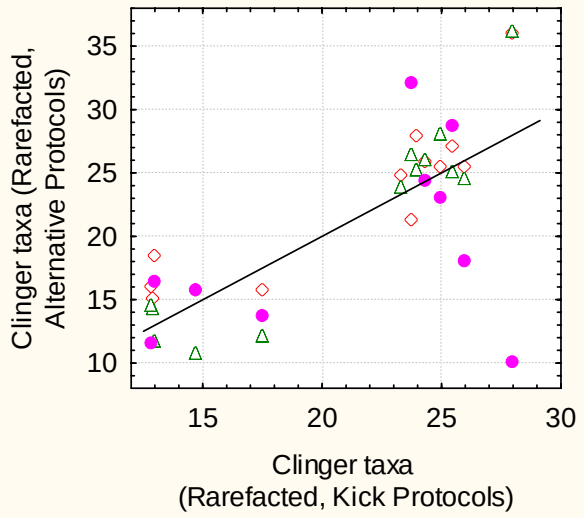

Collecter taxa (Rarefacted, Alternative Protocols)
Collecter taxa (Rarefacted, Kick Protocols)

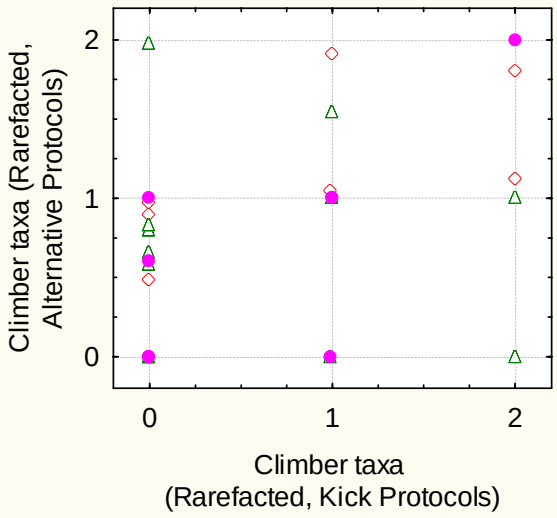

Filterer taxa (Rarefacted, Alternative Protocols)
Filterer taxa (Rarefacted, Kick Protocols)

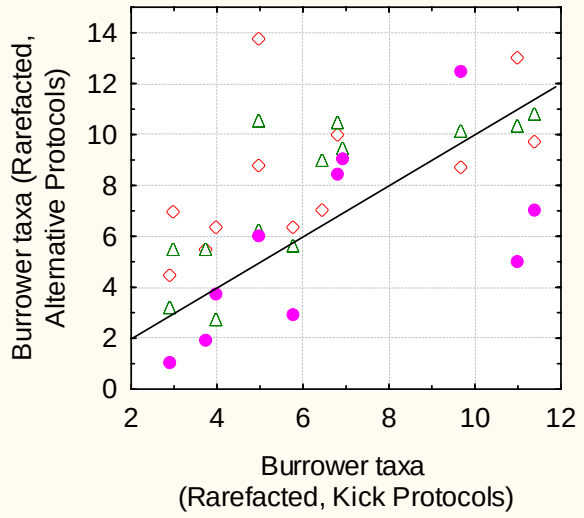

Predator taxa (Rarefacted, Alternative Protocols)
Predator taxa (Rarefacted, Kick Protocols)

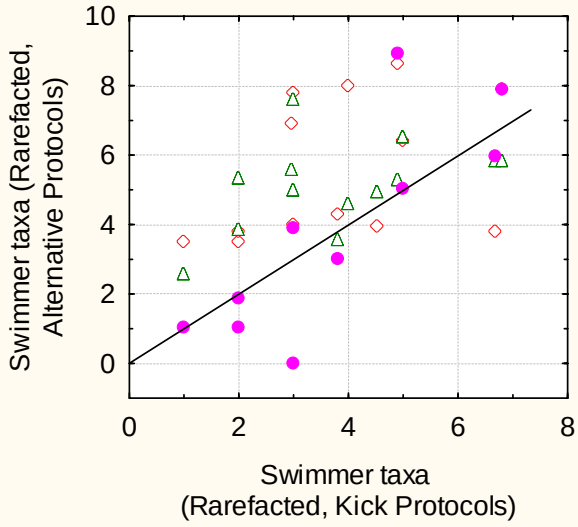

Protocols:
Reachwide
Targeted
Surber

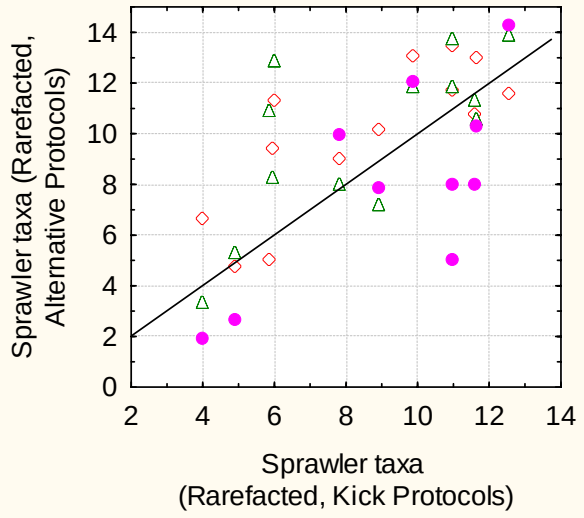

% Scraper (Alternative Protocol)
% Scraper (Kick Protocol)
% Shredder (Alternative Protocol)
% Shredder (Kick Protocol)
% Collector (Alternative Protocol)
% Collector (Kick Protocol)
% Filterer (Alternative Protocol)
% Filterer (Kick Protocol)
% Predator (Alternative Protocol)
% Predator (Kick Protocol)
Protocols:
Reachwide
Targeted
Surber

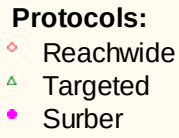

Clinger taxa (Rarefacted, Alternative Protocols)
35
30
25
20
15
10
15
20
25
30
Clinger taxa
(Rarefacted, Kick Protocols)

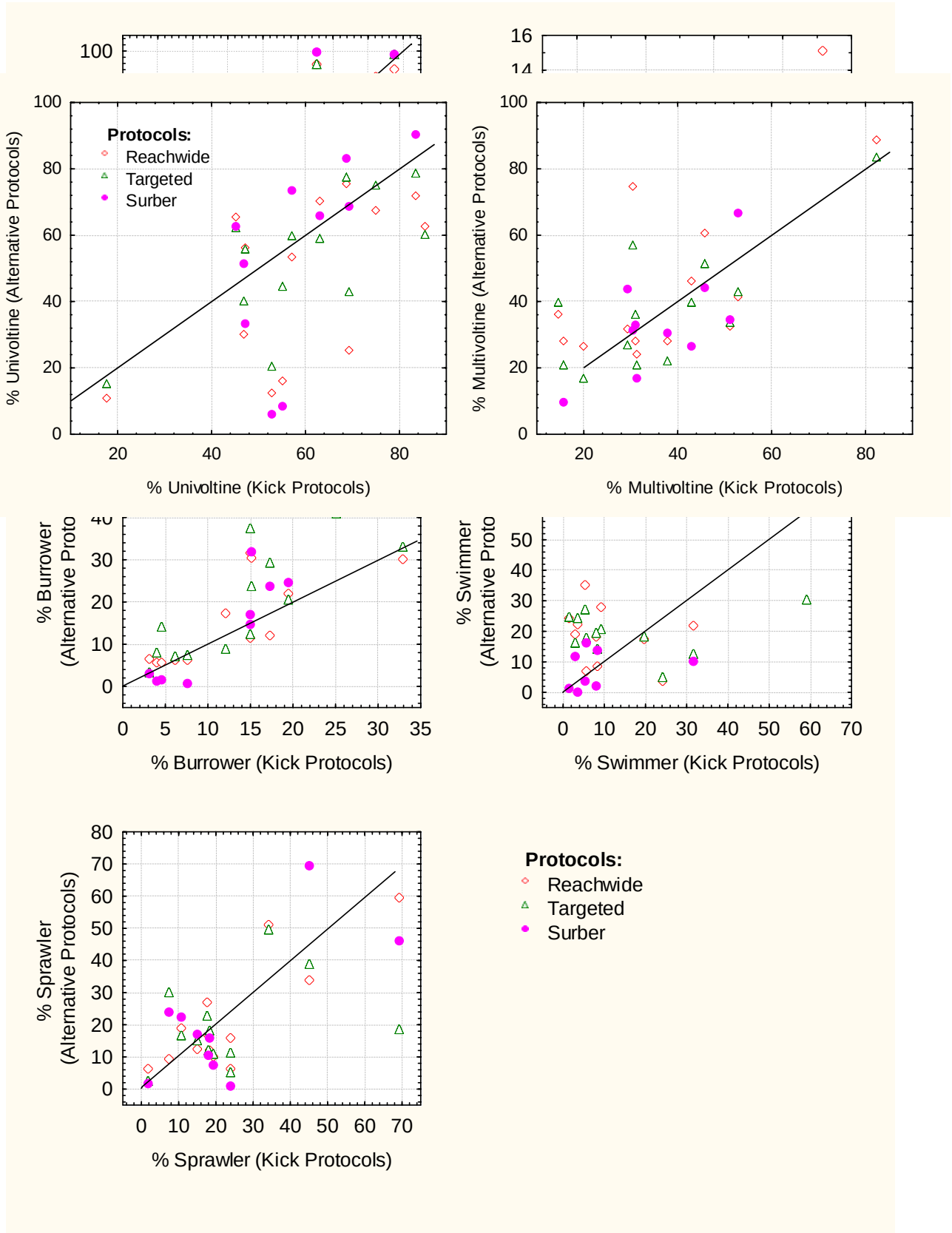

Climber taxa (Rarefacted, Alternative Protocols)
2
1
0
0
1
2
Climber taxa
(Rarefacted, Kick Protocols)


Burrower taxa (Rarefacted, Alternative Protocols)
14
12
10
8
6
4
2
0
2
4
6
8
10
12
Burrower taxa
(Rarefacted, Kick Protocols)


Swimmer taxa (Rarefacted, Alternative Protocols)
10
8
6
4
2
0
0
2
4
6
8
Swimmer taxa
(Rarefacted, Kick Protocols)


Sprawler taxa (Rarefacted, Alternative Protocols)
14
12
10
8
6
4
2
0
2
4
6
8
10
12
14
Sprawler taxa
(Rarefacted, Kick Protocols)


Protocols:
Reachwide
Targeted
Surber


% Univoltine (Alternative Protocols)
% Univoltine (Kick Protocols)
Protocols:
Reachwide
Targeted
Surber
% Multivoltine (Alternative Protocols)
% Multivoltine (Kick Protocols)
% Burrower (Alternative Proto
% Burrower (Kick Protocols)
% Swimmer (Alternative Proto
% Swimmer (Kick Protocols)
% Sprawler (Alternative Protocols)
% Sprawler (Kick Protocols)
Protocols:
Reachwide
Targeted
Surber