# Cramer Fish Sciences

## Diatom Biocriteria for Montana Streams

December 2010

**Lead Investigator:**

**Mark Teply**
Cramer Fish Sciences, Lacey Office
677 Woodland Sq. Lp SE, Suite D6
Lacey, WA  98503
(360) 456-4621
markt@fishsciences.net

## Introduction

The State of Montana (State) has been using benthic diatoms to assess water quality in wadeable streams since the 1970s. Most recently, the composition and structure of periphyton communities have been used to address two key questions that are directly relevant to the State's assessment of naturally flowing streams:

1) Do diatom sample results indicate impairment under 303(d) guidelines?
2) If so, do diatom sample results indicate the cause of impairment?

Of specific interest to the State is discriminating impairment to aquatic life use support in instances where diatom community response was most likely; that is, impairment due to sediment, nutrients, and/or metals.

Bahls et al. (2008) provides a brief history and status report of diatom biocriteria development and water quality assessment in Montana; a recap is presented in this introduction as background, as is a short description of recent events which enable further development of diatom biocriteria. In subsequent sections, we then describe our latest efforts in diatom biocriteria development since those reported by Bahls et al. (2008). Overall, this Technical Report provides scientific background for guidance being developed concurrently for using benthic diatoms in water quality assessments.

### *Brief Background*

Historically, *Periphyton Bioassessment Methods for Montana Streams* (Bahls 1993) has been used to assess biological integrity and impairment of aquatic life in Montana streams. Procedural guidelines and numeric biocriteria are provided for interpretation of the composition and structure of periphyton communities. This manual was based on the findings of a 1990 Montana reference stream study and follow-up surveys in 1991 (Bahls 1993). Three metrics - Sediment Index, Pollution Index, and Diversity Index - and two sets of biocriteria are provided, one for mountain and foothill streams and one

for plains streams. These original metrics have been supplemented by metrics adopted by the Montana Department of Environmental Quality (MT DEQ).

Teply and Bahls (2005) employed advances in biological data analysis and a large amount of periphyton data collected since 1993 to review existing metrics, test new ones, and develop revised biocriteria. Generally, existing metrics at the time were found to be descriptive and diagnostic and were presumptive and based on universal ecological attributes of diatom species and structural characteristics of benthic diatom associations. In all but a few instances, they were found to perform poorly in discriminating between non-impaired streams and those impaired for sediment, nutrients, and metals. Because they were presumptive, taxa lists used in these metrics were often too inclusive and thus diminished the discriminating ability of the associated metric. This led to investigation of empirically-derived metrics.

Methods for developing empirically-derived metrics are described in detail in Teply and Bahls (2005). These methods employed “Increaser Taxa” - taxa that, as a group, exist in detectable amounts in all stream classes and demonstrate a meaningful, measurable, and significant response to sediment, nutrients, or metals. Increaser Taxa were founded in much the same ecological basis as taxa used for “traditional” diatom metrics (e.g., Bahls 1993). But, they were based on empirical observations; that is, they were not presumptive. As in Bahls (1993), two sets of biocriteria were provided by Teply and Bahls (2005), one for mountain and foothill streams and one for plains streams. Increaser Taxa lists developed by Teply and Bahls (2005) showed demonstrable improvement over those that were in use by the State at the time.

Whereas the findings of Teply and Bahls (2005) provided demonstrable improvements over the guidance of Bahls (1993), they also recognized that refinement of Increaser Taxa at smaller ecoregion scales could improve biocriteria for local application (e.g., TMDL assessments). Teply and Bahls (2005) acknowledged that stream grouping at such a coarse scale – i.e., mountains v. plains – presumed a high degree of similarity among taxa assemblages from samples within each group. This is likely not the case.

Work by Tison et al. (2005) described how geology and relief were important determinants in the natural variability of diatom communities. Tison et al. (2005) cited several studies in which ecoregion factors such as water mineral content and current velocity were key determinants of diatom distribution (Biggs 1995, Pan et al. 1999, Potapova and Charles 2003, Soininen 2004, Gevery et al. 2004).

Therefore, Teply and Bahls (2006) employed methods similar to those used in Teply and Bahls (2005) to develop biocriteria for a selected Level III ecoregion per US EPA (2000) - Ecoregion 17 (the Middle Rockies). At the time, the Middle Rockies was the only ecoregion represented by enough diatom samples from Montana to support selection of modeling *and* validation data sets, validation being a requirement critical to the defensibility of the biocriteria. Even then, sufficient samples only existed to evaluate taxa potentially discriminating impairment due to sedimentation. Based on this analysis, a meaningful, significant, and robust model was developed, with demonstrable improvement over both the mountains Siltation Index presented by Bahls (1993) and the mountains Increaser Taxa lists presented by Teply and Bahls (2005).

*Recent Events*

Since 2006, several events have led to better information to support further development of diatom biocriteria. First and foremost, sampling and analysis of impaired and non-impaired streams has been conducted to fill data gaps specifically identified in previous analyses. This enables screening of Increaser Taxa lists across more stream groups. Additional sampling of reference assemblages has also been completed, providing evidence for meaningful and significant stream groups in Montana (Teply and Bahls 2007). This effort yields strata within which there is a high degree of similarity among taxa assemblages. Finally, MDEQ has published the 2008 revised 303(d) list of impaired streams and an expanded list of reference sites. This furthers confidence in stream classifications. In combination, these events lead to the effort reported herein.

The following sections describe methods and findings from our latest efforts in diatom biocriteria development since those briefly described above and reported by Bahls et al. (2008). Improved information has led to improved Increaser Taxa lists for most stream groups and impairment causes. In turn, this improves the State's ability to assess impairment for various causes in naturally flowing streams.

## Methods

Diatom samples used in our evaluation were drawn from the Montana Diatom Database described in Teply and Bahls (2005) and from over 800 diatom samples collected and analyzed by the State since 2004. Recent sampling was conducted to fill data gaps identified by previous diatom biocriteria development efforts and to support TMDL water quality assessments. In filling data gaps, sample streams were randomly selected such that they were more or less geographically uniformly distributed within stream groups identified by Teply and Bahls (2007). The basis for the sample frame at the time was the Montana 2004 303(d) list described by Teply and Bahls (2006). Upon their own initiative, the State often sampled multiple reaches per stream and, sometimes, on multiple occasions. Samples collected by the State followed the PERI-1 or PER-1mod protocol (MT DEQ 2005; MT DEQ 2006) and were analyzed by two laboratories (*Hannaea* and the Academy of Natural Sciences of Philadelphia) following the 400 cell count (800 valve) protocol described in DEQ contract (MT DEQ 2004).

Diatom samples were filtered using information in the Montana 303(d) list in the 2008 Water Quality Assessment Database (http://cwaic.mt.gov/) to exclude those occurring streams impaired for unusual causes which would confound analyses (e.g., pesticides). Only samples occurring along the streams filtered from this database were carried forward. Diatom samples were classified according to impairment classifications. Separate classifications were made for each impairment cause covered in our analysis – sediment, nutrients, or metals. **Impaired** streams were classified by the subject cause of impairment where aquatic life use support was listed as none. Although metals impairment was found to be naturally occurring along plains streams, it was not

considered in our analysis. **Non-impaired** streams were classified as those where the cause of impairment was other than that being considered. Non-impaired streams also included those listed as impaired but for causes other than sediment, nutrients, and metals, such as habitat degradation, flow alteration, or thermal loading; i.e., stressors not considered to have an affect on the diatom community.

Each list of impaired and non-impaired samples was then classified according to stream groupings. We chose to base stream groups on the analysis of reference taxa assemblages across Montana (Teply and Bahls 2007). Of the two alternatives were presented by Teply and Bahls (2007), we chose the one which used rules based on MT DEQ fisheries classification and Level III ecoregions (see **Figure 1**). These rules make effective use of a streams fisheries classification (cold or warm) and Level III ecoregion (US EPA 2000). To classify streams into cold or warm water fisheries, we used rules implemented by the State based on Level III and IV ecoregions (see **Appendix A**). Classification accuracy using these rules was found to be about 71%. Another classification system was presented by Teply and Bahls (2007) which had slightly better accuracy, but also considered stream gradient and river basin. Whereas information on river basin is readily available, information on stream gradient is not without modeling and/or field observation. Thus, for simplicity and expedience, we chose to carry forward the classification system outlined below.

**Figure 1. Dichotomous key to stream groupings using MDEQ Fisheries Classification and Predominant Level III Ecoregion (Alternative I).**

1. Site in a warm water fishery?
    - a. Yes..... Warm-Water Stream Group (I.1a)
    - b. No ..... Go To 2
2. Site within the Middle Rockies Ecoregion?
    - a. Yes..... Middle Rockies Stream Group (I.2a)
    - b. No ..... Northern/Canadian Rockies Stream Group (I.2b)

Increaser Taxa and associated thresholds were identfied from these samples using a three step process described below. These three steps were applied to each combination of stream group and impairment cause described above. The objective of Step I was to identify three separate candidate Increaser Taxa lists using three independent random draws of filtered samples. Step II was used to identify selected taxa that occurred on two more more candidate lists. This step took advantage of the greater number of samples now available to us. Step III used the resulting Increaser Taxa to identify impairment thresholds discriminating between impaired and non-impaired streams. This process yielded a single model for each combination of stream group and impairment cause described above.

*Step I – Preliminary Increaser Taxa Screening*

The first step was, in itself, a three-step process to identify multiple candidate Increaser Taxa lists. For each combination of stream group and impairment cause, three independently derived Increaser Taxa lists were generated. We did this to take advantage of the large, robust data set compiled by the State since our last analysis. We now found ourselves with the opportunity to draw multiple independent modeling and validation data sets – not just one as in previous analyses. In doing so, we were now able to consider "repeatability" of Increaser Taxa. This is was an important improvement in our methods compared to previous efforts and it was yet one more benefit of the commitment the State had made to increased sampling.

The first step in screening was to make three independent, random sample draws. For each draw, modeling and validation data sets where selected, each consisting of 20 impaired samples and 20 non-impaired samples. This target of 20 samples is generally considered a "rule-of-thumb" for discriminant analyses, described later in our methods. For each draw, one sample was randomly selected for each stream segment filtered from the 303(d) database where sampling had occurred; successive draws were conducted without replacement. Then, samples in each draw were randomly assigned to the modeling and validation data sets. This was accomplished by sorting samples by

impairment classification, watershed, and stream and then systematically selecting samples using a random start and a selection frequency that achieved the target sample size. This procedure had the effect of uniformly distributing samples geographically and temporally such that there was little to no bias or dependence.

The second screening step was to identify Increaser Taxa in the modeling data sets. Generally, we employed the screening methods described in detail in Teply and Bahls (2005 and 2006), but we used refined screening criteria, as follows:

1. Taxa were selected if they occurred in detectable amounts in at least one-quarter of the impaired samples (i.e., 5) in the modeling data set;

2. Taxa were selected if they exhibited at least an 25% increase in average relative abundance in impaired samples (vs. non-impaired samples); and,

3. If by this point more than twenty taxa remained, they were sorted by their percent increase in average relative abundance and the top twenty were selected.

The first two criteria sought to identify common taxa that existed in detectable amounts in all impairment classes and to identify taxa that demonstrate a meaningful, measurable response to sediment, nutrients, or metals impairment. The third criterion prioritized taxa where the first two criteria did not provide adequate screening.

The final screening step was to evaluate the signficance and reliability of candidate Increaser Taxa lists in predicting impaired and non-impaired streams. Generally, we employed the methods described in detail in Teply and Bahls (2005 and 2006). We used discriminant analysis where the dependent variable was impairment class and the independent variable represented the total percent relative abundance of taxa found on the Increaser Taxa list. Discriminant models were evaluated based on their statistical significance (preferably $p<0.05$), and overall classification accuracy (preferably greater than 65%) and false positive rates (preferably less than 30%) in the validation data set.

Some discretion was used to consider the influence of outliers and unusual sample circumstances. Furthermore, we also considered diatom autoecology through published sources (e.g., van Dam et al. 1994) and professional judgment.

*Step II – Identify Repeatable Taxa*

In this step, we identified taxa that occurred in two or more of the three independently derived Increaser Taxa lists for each combination of stream group and impairment cause. If a particular taxon was screened in each of the three lists, we felt it had a consistent ability to indicate the cause of impairment being evaluated. If a taxon appeared on only one of the lists, is was clearly less consistent and and there was less evidence to suggest it is a reliable indicator of the impairement cause. Furthermore, it could have simply been a chance result. Recognizing that Increaser Taxa lists would be used in regulatory water quality assessments, we felt is was important to minimize the possibility that we were including a taxon on the basis of a chance occurrence.

In identifying repeatable taxa, we preferred taxa that were screened from all three sample draws. If we could identify a set of 10 to 20 taxa that were screened in all three sample draws, we carried that select list forward to Step III. If less than 10 taxa occurred on all three lists, we supplemented the list with taxa screened in two sample draws. If we couldn't identify at least 10 in this manner, we then felt we had insufficient evidence to forward an Increaser Taxa list for the stream group and impairment cause being evaluated; however, this never happened. Conversely, if this process resulted in more than 20 taxa, we then felt we had too much noise to carry an Increaser Taxa list forward; this also never happened. We were always able to identify 10 to 20 repeatable taxa for each combination of stream group and impairement cause.

Using repeatable taxa, discriminant models were generated and evaluated for their ability to in predict impaired and non-impaired streams. This was accomplished in the same manner as described in Step II and using the same three independently drawn modeling and validation data sets. The same performance criteria were used to gage

statistical signficance and classification accuracy. As above, some discretion was used to consider the influence of outliers and unusual sample circumstances. Generally, we found little change in model performance using the repeatable taxa; however, that we found that a select list of taxa could discriminate impairment in three independently drawn data sets increased our confidence in their utility.

*Step III – Develop a Unified Model*

The final step was to develop one model discriminating impairment for each combination of stream group and impairment cause. Steps I and II yield multiple models each with a slightly different probability model. Multiple models would be problematic in water quality assessments, because they would have slightly different thresholds for discriminating impairment. To address this, we simply calculated an average threshold value (in terms of percent relative abundance of taxa) and covariance from the models generated in Step II. Generally, these statistics varied only a small amount among each sample draw for each stream group and impairment cause. Average thresholds did not appreciably degrade the classification accuracies in validation data sets. Final model statistics are reported and guidance for their use and interpretation is provided in a separate document being prepared concurrently with this Technical Report.

## Results

**Table 1** summarizes key modeling output for each combination of stream group and impairment cause. Key results from each Step are tabulated. From Step I, the number of taxa screened from each sample draw is reported along with mean percent relative abundances for impaired and non-impaired samples in the modeling data set; statistically significant differences ($p<0.05$) are in boldface type. Overall classification accuracies determined from independently drawn validation data sets are also reported, as are false positive rates for impaired and non-impaired predictions. Similar output is reported for Step II, but the number of taxa represents the number of taxa that were screened from two or three of the independent sample draws. Results from Step III

include the average discrimination threshold and the within group covariance needed for an associated probability model. Classification accuracies using the average threshold are also reported. Final Increaser Taxa lists are in **Table 2**.

Stream groups and impairment categories for which statistical and validation modeling criteria were met are highlighted. Most discriminant models met our statistical screening criteria. Consistently strong models were developed for discriminating sediment impairment in all three stream groups. In all but one test, models were statistically significant in multiple independent sample draws. In all but one validation data sets, overall classification accuracy exceeded 65%. And, in all but a few instances, false positive rates were below 30%. We also found a high level of repeatability among independently derived Increaser Taxa and threshold averaging resulted in only very slight shifts in classification accuracy. And, whereas there was repeatability in taxa lists within stream groups, there was little overlap in Sediment Increaser Taxa among stream groups. This is consistent with our expectations given meaningful differences among the reference taxa assemblages (in Teply and Bahls 2007). Finally, Increaser Taxa were also considered to be ecologically meaningful (Bahls, personal communication, 2010).

Results for other impairment categories were not as robust. Moderately strong models were developed for discriminating nutrient impairment in the Warm Water Fisheries and in the Northern/Canadian Rockies stream groups. Significant differences were evident in all instances; however, overall classification accuracy was rarely over the preferred 65% threshold. Yet, because the overall classification accuracy was consistently above 50% and impairment false positive rate mostly below 30%, we used discretion in determining that they met the statistical screening criteria. Even though marginal, these results suggest some measure of reliability. In comparison, we found that two models – Middle Rockies nutrients and Middle Rockies metals –were not sufficiently reliable. Significant models could be developed in most instances, but overall classification accuracy for these models was often below 50% and false positive rates mostly well over 30%. These results suggest that use of Increaser Taxa in these instances is not even marginally reliable for regulatory use in stream impairment determinations.

**Table 1. Summary of modeling and validation results for development of diatom Increaser Taxa.**

### a) Step I Results

| Stream Group - Impairment Cause | Step I Results | | | | | |
|---|---|---|---|---|---|---|
| | No. Taxa | Percent Relative Abundance Non-Impaired | Percent Relative Abundance Impaired | Overall Classification Accuracy | False Positive Rate Impaired | False Positive Rate Non-Impaired |
| Warm Water Fisheries - Sediment | 20 | **17.16** | **42.30** | 70.4% | 30.0% | 28.6% |
| | 19 | **12.60** | **28.49** | 63.0% | 35.0% | 42.9% |
| | 19 | **8.04** | **29.89** | 73.1% | 26.3% | 28.6% |
| Warm Water Fisheries - Nutrients | 13 | **7.71** | **19.45** | 50.0% | 25.0% | 75.0% |
| | 20 | **6.57** | **23.42** | 55.0% | 25.0% | 65.0% |
| | 20 | **8.81** | **30.65** | 60.0% | 15.0% | 65.0% |
| Middle Rockies - Sediment | 20 | **13.94** | **31.97** | 67.5% | 30.0% | 35.0% |
| | 20 | **10.08** | **27.35** | 70.0% | 30.0% | 30.0% |
| | 20 | **13.49** | **26.90** | 65.0% | 25.0% | 45.0% |
| Middle Rockies - Nutrients | 20 | **13.86** | **32.29** | 50.0% | 50.0% | 50.0% |
| | 20 | **11.76** | **26.79** | 45.0% | 35.0% | 75.0% |
| | 20 | **14.78** | **32.61** | 52.5% | 40.0% | 55.0% |
| Middle Rockies - Metals | 20 | **8.31** | **26.92** | 40.0% | 45.0% | 75.0% |
| | 20 | **15.78** | **40.98** | 55.0% | 25.0% | 65.0% |
| | 19 | **7.80** | **28.68** | 55.0% | 30.0% | 60.0% |
| Northern/Canadian Rockies - Sediment | 20 | **9.15** | **27.50** | 60.0% | 40.0% | 40.0% |
| | 20 | **8.68** | **31.12** | 52.5% | 40.0% | 55.0% |
| | 20 | **8.09** | **30.23** | 55.0% | 40.0% | 50.0% |
| Northern/Canadian Rockies - Nutrients | 20 | **8.85** | **24.66** | 60.6% | 25.0% | 61.5% |
| | 20 | **19.42** | **36.84** | 57.6% | 25.0% | 69.2% |
| | 20 | **6.53** | **22.90** | 42.4% | 50.0% | 69.2% |

### b) Step II Results

| Stream Group - Impairment Cause | Step II Results | | | | | |
|---|---|---|---|---|---|---|
| | No. Taxa | Percent Relative Abundance Non-Impaired | Percent Relative Abundance Impaired | Overall Classification Accuracy | False Positive Rate Impaired | False Positive Rate Non-Impaired |
| Warm Water Fisheries - Sediment | 15 | **9.17** | **24.17** | 70.4% | 30.0% | 28.6% |
| | | 15.84 | 22.36 | 70.4% | 20.0% | 57.1% |
| | | **9.12** | **26.87** | 74.1% | 25.0% | 28.6% |
| Warm Water Fisheries - Nutrients | 12 | **7.86** | **16.31** | 55.0% | 20.0% | 70.0% |
| | | **6.77** | **10.96** | 57.5% | 30.0% | 55.0% |
| | | **5.74** | **19.63** | 55.0% | 20.0% | 70.0% |
| Middle Rockies - Sediment | 16 | **10.33** | **21.88** | 70.0% | 20.0% | 40.0% |
| | | **9.65** | **17.42** | 57.5% | 40.0% | 45.0% |
| | | **12.38** | **20.38** | 62.5% | 20.0% | 55.0% |
| Middle Rockies - Nutrients | | | | | | |
| Middle Rockies - Metals | | | | | | |
| Northern/Canadian Rockies - Sediment | 19 | **12.55** | **24.01** | 80.0% | 20.0% | 20.0% |
| | | **10.49** | **26.92** | 67.5% | 25.0% | 40.0% |
| | | **8.57** | **28.33** | 65.0% | 35.0% | 35.0% |
| Northern/Canadian Rockies - Nutrients | 15 | **8.19** | **19.70** | 75.8% | 20.0% | 30.8% |
| | | **10.31** | **22.00** | 60.6% | 30.0% | 53.8% |
| | | **8.63** | **22.28** | 57.6% | 35.0% | 53.8% |

**c) Step III Results**

| Stream Group - Impairment Cause | Step III Results | | | | |
|---|---|---|---|---|---|
| | Percent Relative Abundance Impairment Threshold | Within Group Covariance | Overall Accuracy | False Positive Rate Impaired | False Positive Rate Non-Impaired |
| Warm Water Fisheries - Sediment | 17.92 | 351.40 | 74.1% | 25.0% | 28.6% |
| | | | 70.4% | 20.0% | 57.1% |
| | | | 74.1% | 25.0% | 28.6% |
| Warm Water Fisheries - Nutrients | 11.21 | 163.84 | 57.5% | 25.0% | 60.0% |
| | | | 57.5% | 20.0% | 65.0% |
| | | | 55.0% | 30.0% | 60.0% |
| Middle Rockies - Sediment | 15.34 | 160.59 | 70.0% | 25.0% | 35.0% |
| | | | 57.5% | 35.0% | 50.0% |
| | | | 67.5% | 20.0% | 45.0% |
| Middle Rockies - Nutrients | | | | | |
| Middle Rockies - Metals | | | | | |
| Northern/Canadian Rockies - Sediment | 18.48 | 293.25 | 80.0% | 20.0% | 20.0% |
| | | | 67.5% | 25.0% | 40.0% |
| | | | 65.0% | 35.0% | 35.0% |
| Northern/Canadian Rockies - Nutrients | 15.18 | 244.69 | 75.8% | 10.0% | 46.2% |
| | | | 63.6% | 30.0% | 46.2% |
| | | | 57.6% | 35.0% | 53.8% |

**Table 2. Final Increaser Taxa lists (pink – taxa area repeatable on all three candidate Increaser Taxa lists; tan – taxa are repeatable on two candidate Increaser Taxa lists).**

**a) Sediment Increaser Taxa**

| Warm Water Fisheries - Sediment |
|---|
| *Amphora pediculus* |
| *Caloneis bacillum* |
| *Cocconeis placentula* |
| *Diatoma moniliformis* |
| *Epithemia sorex* |
| *Gomphonema minutum* |
| *Navicula capitatoradiata* |
| *Navicula cryptotenella* |
| *Navicula gregaria* |
| *Navicula reichardtiana* |
| *Nitzschia inconspicua* |
| *Nitzschia liebethruthii* |
| *Nitzschia linearis* |
| *Reimeria sinuata* |
| *Surirella brebissonii kuetzingii* |

| Middle Rockies - Sediment |
|---|
| *Amphora inariensis* |
| *Cocconeis pediculus* |
| *Cocconeis pseudolineata* |
| *Eolimna minima* |
| *Geissleria acceptata* |
| *Gomphonema drutelingense* |
| *Meridion circulare* |
| *Navicula gregaria* |
| *Navicula lanceolata* |
| *Navicula tripunctata* |
| *Nitzschia recta* |
| *Planothidium frequentissimum* |
| *Planothidium lanceolatum* |
| *Reimeria sinuata* |
| *Sellaphora pupula* |
| *Staurosirella leptostauron* |

| Northern Rockies - Sediment |
|---|
| *Achnanthidium deflexum* |
| *Aulacoseira italica* |
| *Eolimna minima* |
| *Gomphonema minutum* |
| *Gomphonema pumilum* |
| *Gomphonema rhombicum* |
| *Gomphosphenia sp.* |
| *Melosira varians* |
| *Meridion circulare* |
| *Navicula cryptocephala* |
| *Navicula radiosa* |
| *Nitzschia archibaldii* |
| *Nitzschia palea* |
| *Nitzschia perminuta* |
| *Planothidium frequentissimum* |
| *Pseudostaurosira brevistriata* |
| *Reimeria sinuata* |
| *Rhopalodia gibba* |
| *Staurosira construens* |

**b) Nutrient Increaser Taxa**

| Warm Water Fisheries - Nutrients |
|---|
| *Amphora pediculus* |
| *Gomphonema parvulum* |
| *Navicula cryptotenella* |
| *Navicula libonensis* |
| *Navicula tripunctata* |
| *Nitzschia acicularis* |
| *Nitzschia amphibia* |
| *Nitzschia archibaldii* |
| *Nitzschia fonticola* |
| *Nitzschia gracilis* |
| *Nitzschia inconspicua* |
| *Nitzschia linearis* |

| Northern Rockies - Nutrients |
|---|
| *Achnanthes nodosa* |
| *Achnanthidium deflexum* |
| *Adlafia minuscula* |
| *Eolimna minima* |
| *Geissleria acceptata* |
| *Gomphonema minutum* |
| *Gomphonema pumilum* |
| *Gomphosphenia sp.* |
| *Meridion circulare* |
| *Navicula cryptocephala* |
| *Nitzschia fonticola* |
| *Nitzschia inconspicua* |
| *Nitzschia perminuta* |
| *Planothidium frequentissimum* |
| *Synedra rumpens* |

## Discussion

Two factors have led to a significant advancement in the development of diatom biocriteria since Teply and Bahls (2006). First and foremost is the investment the State of Montana has made in sample data collection – in streams meeting reference criteria developed by Suplee et al. (2005) and in streams where support for aquatic life use was partial or none. Reference stream sampling enabled the discrimination of meaningful and significant differences among taxa assemblages which, in turn, provided the basis for stream groups defined by Teply and Bahls (2007). Heretofore, analyses were conducted at a coarse scale of resolution within which there was likely a great amount of variability in baseline taxa assemblages. In previous analyses, this likely obscured our ability to detect taxa assemblage responses to impairment.

Sampling in other streams further increased our ability to evaluate potential differences among impaired and non-impaired streams within each of these stream groups. It was usually only possible to select one independent data at a coarser scale – i.e., plains streams versus mountain streams – where differences in taxa assemblages within these regions likely obscured the ability to detect trends. And, where there were sufficient

samples at a finer scale, it only occurred within the Middle Rockies ecoregion (Teply and Bahls 2006) which had been traditionally heavily sampled. The State's commitment to sampling enabled investigation at a finer scale throughout Montana. And, as an unexpected benefit, the quantity of samples enabled repeated independent sample draws which enabled repeated validation of statistical models – this is of great value.

The other significant advancement was the 2008 revised 303(d) list of impaired streams. This list is a result of diligent review of stream impairment determinations throughout Montana. Although it would likely be difficult and contentious to determine the classification accuracy of this list, our view was that the list appears to have resolved many false positives and false negatives through focused monitoring and interpretation. In previous analyses, misclassified streams likely obscured our ability to detect taxa assemblage responses to impairment. Although we were not tasked with quantifying this effect, we can qualitatively state that we found it much easier to identify patterns and trends. We feel this is a direct effect of improved stream grouping *and* stream impairment classifications. Again, the State's commitment, this time to stream impairment determinations, had direct benefit to our investigations.

Overall, we were able to develop several meaningful and significant Increaser Taxa lists (see **Table 2**) and thresholds (**Table 1**). Two sets of lists are discussed below – Sediment Increaser Taxa and Nutrients Increaser Taxa – along with a general discussion regarding interpretation of discriminant models. Specific interpretation of these models for stream impairment determinations is presented in a guidance document being prepared concurrently. The following provides background for their use.

*Sediment Increaser Taxa*

Sediment Increaser Taxa lists in **Table 2**, and associated thresholds in **Table 1**, are offered as a useful, meaningful, significant, and robust alternative to the Siltation Index presented by Bahls (1993) and lists presented by Teply and Bahls (2005) and Teply and Bahls (2006). Given the lack of numeric State standards for sediment impairment, this

model can offer an important piece of evidence in water quality assessments. About half of the impaired streams in Montana are impaired due to sediment – either solely or in combination with other causes. Therefore, this model is offered as a useful alternative to existing metrics (e.g., Siltation Index) and biocriteria used by the State. In application, it is specific to impairment due to sediment and it has a demonstrated reliability in discriminating sediment impairment when it is the sole impairment cause and in combination with other impairment causes – the latter being more likely the case.

When used in stream impairment determinations, Sediment Increaser Taxa will likely provide better discriminating ability than the Siltation Index presented by Bahls (1993). Teply and Bahls (2005) found the overall classification accuracy using Siltation Index to be much less than 50%; average overall classification accuracies of the Sediment Increaser Taxa is about 70%. The Siltation Index is based on the percent relative abundance of valves in the genera *Navicula, Nitzschia and Surirella*. Whereas many Sediment Increaser Taxa belong to these genera, there are many members of these genera that did not respond to sediment impairment. Furthermore, there are members of other genera that did respond. Being empirically derived, Sediment Increaser Taxa are not presumptive and have a demonstrated ability to discriminate impairment.

Sediment Increaser Taxa will likely provide better discriminating ability than lists presented by Teply and Bahls (2005) and Teply and Bahls (2006). Those identified in this latest effort are both locally relevant and repeatable. Increaser Taxa reported in Teply and Bahls (2005) only distinguished between mountains and plains and did not account for meaningful differences in taxa assemblages identified in Teply and Bahls (2007). Increaser Taxa reported in Teply and Bahls (2006) did account for regional differences – i.e., the Middle Rockies Ecoregion – however, there is some overlap with stream groups we now know to be distinct. Compared to previous Sediment Increaser Taxa Lists, those reported in **Table 2** represent a meaningful and significant refinement.

*Nutrient Increaser Taxa*

Nutrient Increaser Taxa lists in **Table 2**, and associated thresholds in **Table 1**, are offered as a useful, meaningful, significant, and robust alternative to the Pollution Index presented by Bahls (1993) and lists presented by Teply and Bahls (2005). About one-third of the impaired streams in Montana are impaired due to nutrients – either solely or in combination with other causes. Therefore, this model is offered as a useful alternative to existing metrics (e.g., Pollution Index) and biocriteria used by the State to help determine nutrient problems. In application, it is specific to impairment due to nutrients and it has a demonstrated reliability in discriminating nutrient impairment when it is the sole impairment cause and in combination with other impairment causes – the latter being more likely the case.

The performance of Nutrient Increaser Taxa is slightly less than that of Sediment Increaser Taxa, but it still offers a measure of reliability for regulatory use in stream impairment determinations. When used in stream impairment determinations, Nutrient Increaser Taxa will likely provide better discriminating ability than the Pollution Index presented by Bahls (1993) or lists presented by Teply and Bahls (2005). Teply and Bahls (2005) found the overall classification accuracy using Pollution Index to be much less than 50%; average overall classification accuracies of the Sediment Increaser Taxa is about 60%. Compared to the Pollution Index, Nutrient Increaser Taxa are not presumptive and have a demonstrated ability to discriminate nutrient impairment. And, Nutrient Increaser Taxa also account for meaningful differences in taxa assemblages identified in Teply and Bahls (2007), albeit for only two stream groups.

*Discriminant Models*

An advantage of discriminant models (e.g., versus ANOVA or percentile-based threshold setting) is that discriminant models can be used to directly calculate the probability of class membership (i.e., the probability that a sample represents impaired conditions). Specifically, for a given percent relative abundance of Increaser Taxa, discriminant models permit us to state the probability that the associated sample is from an impaired stream. This improves the State's ability to assess naturally flowing streams

and was an attraction to selecting the approach. Meaningful application of discriminant models assumes randomness and sample independence and normality and equal variance in metric values among impaired and non-imparied samples. In general, we found that these assumptions were not saliently violated and that any bias was minimal.

It is important to remember that the probability yielded by discriminant models is just that – a probability. When the percent relative abundance of Increaser Taxa is just above the threshold reported in **Table 1**, there is a slightly better than 50% chance that the associated sample is from an impaired stream. Model validation indicates slightly better overall classification accuracies. The false positive rates for predicting impairment averages about 25% (i.e., a one-in-four chance that the model indicates impairment when in fact the stream *is not* impaired) and for non-impairment predictions averages about 45% (i.e., almost a one-in-two chance that the model indicates non-impairment when in fact the stream *is* impaired). This error is not insubstantial, but it is a dramatic improvement compared to Bahls (1993). Further guidance for interpretation of these probabilities is provided in a companion paper being prepared concurrently.

## References

Bahls, L.L. 1993. Periphyton Bioassessment Protocols for Montana Streams. Montana Department of Health and Environmental Sciences; Water Quality Bureau; Helena, Montana. Revised January 1993.

Bahls, L., M. Teply, R. Sada de Suplee, and M. Suplee. 2008. Diatom biocriteria development and water quality assessment in Montana: a brief history and status report. Diatom Research 23(2):533-540.

Biggs, B.J.F. 1995. The contribution of flood disturbance, catchment geology, and land use to the habitat template of periphyton in stream ecosystems. Freshwater Biology 33:419-438.

Gevery, M., F. Rimet, Y.-S. Park, J.L. Giraudel, L. Ector, and S. Lek. 2004. Water quality assessment using diatom assemblages and advanced modeling techniques. Freshwater Biology 49:208-220.

MT DEQ, 2004. Biological Contract; DEQ Contract No. 205012. Water Quality Planning Bureau. December 6, 2004.

MT DEQ, 2005. Field Procedures Manual for Water Quality Assessment Monitoring. Water Quality Planning Buruea, WQPBWQM-020, Revision 2, 4/21/2005. *Available at* http://deq.mt.gov/wqinfo/QAProgram/default.mcpx

MT DEQ, 2006. Sampling and Analysis Plan 2006 for the University of Montana Watershed Health Clinic, Wadeable Stream Reference Study. Water Quality Planning Bureau. June 8, 2006.

Pan, Y., R.J. Stevenson, B.H. Hill, P. Kauffmann, and A.T. Herlihy. 1999. Spatial patterns and ecological determinants of benthic algal assemblages in Mid-Atlantic streams, USA. Journal of Phycology 35:460-468.

Potapova, M. and D.F. Charles. 2003. Distribution of benthic diatoms in US rivers in relation to conductivity and ionic composition. Freshwater Biology 48:1311-1328.

Soininen, J. 2004. Determinants of benthic community structure in boreal streams: the role of environmental and spatial factors at different scales. International Review of Hydrobiology 89:139-150.

Suplee, M., R.S. de Suplee, D. Feldman, and T. Laidlaw. 2005. Identification and Assessment of Montana Reference Streams: A Follow-up and Expansion of the 1992 Benchmark Biology Study. Montana Dept. of Environmental Quality, Helena, MT. 23p.

Teply, M. and L. Bahls. 2005. Diatom Biocriteria for Montana Streams. Prepared for the Montana Department of Environmental Quality by Larix Systems, Inc. of Helena, Montana. September 2005.

Teply, M. and L. Bahls. 2006. Diatom Biocriteria for Montana Streams – Middle Rockies Ecoregion. Prepared for the Montana Department of Environmental Quality by Larix Systems, Inc. of Helena, Montana. February 2006.

Teply, M. and L. Bahls. 2007. Statistical Evaluation of Periphyton Samples from Montana Reference Streams. Prepared for the Montana Department of Environmental Quality by Larix Systems, Inc. of Helena, Montana. February 2007.

Tison, J., Y.-S. Park, M. Coste, J.G. Wasson, L. Ector, F. Rimet, and F. Delmas. 2005. Typology of diatom communities and the influence of hydro-ecoregions: A study on the French hydrosystem scale. Water Research 39:3177-3188.

US EPA. 2000. Level III Ecoregions of the Continental United States (Map). National Health and Environmental Effects Research Laboratory. Corvallis, Oregon.

van Dam, H., A. Mertens, and J. Sinkeldam. 1994. A coded checklist and ecological indicator values of freshwater diatoms from The Netherlands. Netherlands Journal of Aquatic Ecology 28(1):117-133.

# Appendix A: Stream Fisheries Classifications

## Cold Water Fisheries

Table 3.1. Ecoregions in which Nutrient-impairment Assessment Methodologies Described in this Section Best Apply.

| Ecoregion Scale | Ecoregion Name | Ecoregion Number |
|---|---|---|
| Level III | Northern Rockies | 15 |
| Level III | Idaho Batholith | 16 |
| Level III | Middle Rockies | 17 |
| Level III | Canadian Rockies | 41 |
| Level IV | Sweetgrass Uplands | 42l |
| Level IV | Milk River Pothole Upland | 42n |
| Level IV | Rocky Mountain Front Foothill Potholes | 42q |
| Level IV | Foothill Grassland | 42r |
| Level IV | Unglaciated Montana High Plains | 43o |
| Level IV | Non-calcareous Foothill Grassland | 43s |
| Level IV | Shields-Smith Valleys | 43t |
| Level IV | Limy Foothill Grassland | 43u |
| Level IV | Pryor-Bighorn Foothills | 43v |

## Warm Water Fisheries

Table 4.1. Level III Ecoregions in which Nutrient-impairment Assessment Methodologies in this Section Best Apply, along with Specified Level IV Ecoregions for which the Methodologys do not Apply.

| Ecoregion Scale | | | |
|---|---|---|---|
| Level III | Level III | Ecoregion Name | Ecoregion Number |
| **Level III** | | **Northwestern Glaciated Plains** | **42** |
| *Level IV ecoregions in the Northwestern Glaciated Plains for which the methods do not apply:* | | | |
| | Level IV | Sweetgrass Uplands | 42l |
| | Level IV | Milk River Pothole Upland | 42n |
| | Level IV | Rocky Mountain Front Foothill Potholes | 42q |
| | Level IV | Foothill Grassland | 42r |
| **Level III** | | **Northwestern Great Plains** | **43** |
| *Level IV ecoregions in the Northwestern Great Plains for which the methods do not apply:* | | | |
| | Level IV | Unglaciated Montana High Plains | 43o |
| | Level IV | Non-calcareous Foothill Grassland | 43s |
| | Level IV | Shields-Smith Valleys | 43t |
| | Level IV | Limy Foothill Grassland | 43u |
| | Level IV | Pryor-Bighorn Foothills | 43v |