# BENTHIC MACROINVERTEBRATE SUSTAINABILITY INDICATOR DEVELOPMENT PROJECT: SUMMARY OF PROGRESS IN YEAR 2

Submitted to

BC Forest Science Program
PricewaterhouseCoopers LLP

Prepared by

Ian Sharpe
B.C. Ministry of Environment
Smithers BC

S.A. Bennett
Bio Logic Consulting Ltd.
Terrace, BC

S. Linke
University of Canberra
Australia

R.C. Bailey
University of Western Ontario
London, Ontario

C. J. Perrin
Limnotek Research and Development Inc.
Vancouver BC

April 30, 2006

*Citation*: Sharpe, I., S.A. Bennett, S. Linke, R.C. Bailey, and C.J. Perrin. 2006. Benthic macroinvertebrate sustainability indicator development project: Summary of progress in year 2. Report prepared by B.C. Ministry of Environment for the B.C. Forest Science Program. 119 p.

*Cover image*: a stylized ordination plot of groups of sampling sites based on benthic invertebrate composition and abundance.

## ACKNOWLEDGEMENTS

Financial support for this project was provided by the Forest Investment Account Forest Sciences Program, and by the B.C. Ministry of Environment, Houston Forest Products, B.C. Timber Sales, Babine Watershed Trust, and the Premier Mine.

The work was completed with the assistance of numerous people both in the field for the collection of samples, in various laboratories, and in the office. A.J. Downie and Angie Gagnon of B.C. Ministry of Environment assisted the senior author with project management throughout the year. GIS support and mapping services were provided by Matt Sakals, Blair Ells and Candice Wilson. The senior author led a "Best Professional Judgment" approach for site selection to augment the GIS-based methods. Individuals participating in selecting coastal streams included Joe Trainer, Dave Rolston and Jason Bennett. Staff of BC Parks and researchers of the University of Montana assisted with the site selections in the Kitlope watershed. Pierre Beaudry and individuals from the Babine Watershed Monitoring Trust assisted with site selection in the Babine watershed. Personnel from Canfor, Houston Forest Products, Babine Forest Products, Pacific Inland Resources, and Fraser Lake Sawmills were most helpful in suggesting sites and providing maps and access information for streams affected by the mountain pine beetle epidemic. Stephanie Sylvestre of Environment Canada provided site information and contributed funding for sampling reference sites in North Tweedsmuir Park. The authors were assisted in the field by Peter Christensen, A.J. Downie, Blair Ells, Angie Gagnon, Aaron Hill, Kevin Koch, Cory Koenig, Bryan Last, Jack Love, Kerrith McKay, Kevin Rieberger, Greg Tamblyn and Candice Wilson. The assistance of volunteers including Ken Rabnett and Regina Saimoto was greatly appreciated. Oona River Hatchery staff including Jan Lemon, Dave Rolston and Charmaine Carr-Harris are thanked for their hospitality and field assistance. Water quality samples were analyzed by Maxxam Analytics Inc. in Burnaby, and benthic invertebrate enumerations were completed by Linde Looy and Sue Salter of Fraser Environmental Services and Danusia Dolecki of Limnotek Research and Development Inc. Environment Canada's CABIN database was used for data storage, and data entry was performed by Blair Ells, Angie Gagnon, Lindsay Graff and Jeannette Lough. Lorne McIntosh helped contact participants and develop materials for implementation of the extension plan.

## TABLE OF CONTENTS

Page

## LIST OF FIGURES

Page

## LIST OF TABLES

Page

# 1 INTRODUCTION

Forest harvesting in B.C. is moving towards "outcome" based management. New tools are needed to measure the effectiveness of forest practices and other land uses in meeting a range of goals, including sustaining aquatic ecosystems. A benthic macroinvertebrate (BI) sustainability indicator system that is part of a performance based toolbox to assess impacts on aquatic ecosystems from forest harvesting activities has been under development for 5 years in Skeena Region. Major funding organizations include the Forest Science Projects envelope of the B.C. Forest Investment Account, Houston Forest Products, Timber Sales B.C., and the B.C. Ministry of Environment (MOE).

Other such BI monitoring and assessment systems that are based on multivariate modeling are being developed around the world (Bailey et al. 2004, CCME 2005) and it is intended that the present research project will rely on advances from this work, as well as local relevant research (e.g. Sylvestre et al. 2005, Rosenberg et al. 1999, Perrin 2006, Perrin and Sylvestre 2006). Some of this work is coupled with development of biotic indices of biological integrity (B-IBI) that was introduced by Karr (1981) and further developed by Karr and Chu (1999) and Kearns and Karr (1994) and applied in British Columbia in several studies (e.g. Bennett and Rysavy, 2003a, Croft 2004). This two pronged approach of applying multivariate modeling and IBI to bioassessment in British Columbia is made possible through a research team with relevant expertise from the United States, Australia, Ontario, and British Columbia.

To improve understanding and resolution of the BI monitoring and assessment system, 3 elements were identified for research: sampling protocol optimization, statistical design for impact classification/assessment and the adaptive management framework using bio-criteria. The 3 year project will lead to a proven sustainability indicator of aquatic ecosystem health, intended for use in Forest Stewardship Plans, Sustainable Forest Management Plans, Land and Resource Management Plans, as well as in Forest Product Certification systems. This tool, once set as a RISC Standard and combined with other indicators of aquatic ecosystem sustainability (fish and fish habitat) will serve as a monitoring and assessment feedback mechanism to determine the effectiveness of forest practices in protecting valued aquatic resources. It is intended that progress made in this project will lead to Province-wide application of the methods.

An adaptive approach using the existing forest harvesting landscape, and streams within it are providing the basis for developing models of how stressor gradients caused by anthropogenic disturbance in watersheds affect BI assemblages found in streams. Models that best explain these relationships have been adopted for use in generating a sustainability indicator scoring system. Analyses of resolution of effect, cost, operational limitations, and applicability to adaptive management of forest harvesting will be made in year 3 of the project.

The product will be embedded in an existing Environment Canada database management system and will include protocols for generating and using the statistically derived sustainability indicator models. Details will be provided on how stream condition ratings can be validated through statistical analyses on an ongoing basis. Steps necessary for quality control and assurance will be provided.

Year 1 deliverables (Sharpe et al. 2005) included the development, testing and reporting of a new protocol for laboratory procedures and sampling, watershed/stream characterization, stressor gradient analysis, and sustainability indicator model development. The intent of the protocol is to strike a balance among the following attributes: measuring appropriate habitat(s) that may be affected by human influence, contribution to impact identification success, contribution to meaningful “impact cause” hypothesis testing, cost, and ease of use. This year 2 report includes advances in the watershed/stream characterization, stressor gradient analysis, and sustainability indicator model development. It also includes recommendations for year 3 work in terms of optimizing the protocols, and the development of extension case study material for the 3-year project. Components of the plan for this year will include the dissemination of draft and final report(s) for wide review among interested scientists, existing and potential users of aquatic resource sustainability indicators, including industry, government and NGO stakeholders. An international workshop, to discuss this and other similar work, will be held in the fall of 2006 at the University of British Columbia (UBC). This workshop will be used to further publicize year 2 results, and seek feedback on fine tuning the workplan for year 3.

Having dealt with field and watershed/habitat characterization components in year 1, and optimizing statistical design for impact classification and sustainability indicator development in year 2 (as outlined in this report), the final year of the project will emphasize “putting the sustainability indicator into the hands of the intended users”. Given this emphasis, approximately 50 new sites will be added to the 180 already sampled in the first 2 years. These sites will be chosen to represent the expansion of the mountain pine beetle epidemic area to the NE, and to provide an opportunity to blend the dataset with one being developed concurrently in the Yukon. Additional stressor gradient analyses will be advanced to better determine relationships between forest harvesting disturbances and patterns in BI assemblages. These activities will further develop the sustainability indicator.

Results have, and will be reported and disseminated as dictated by the extension plan developed in year 1, and carried out in part in year 2. Further efforts in year 3 will include the preparation of one or more peer reviewed journal articles describing the results of stressor gradient effects modeling. Workshop venues will be used to bring the study team and other scientists together with forest management decision-makers to review the work to date, and fine-tune the workplan for year 3. The focus of this work will be on advancing the use of cause and effect interpretations, and developing the decision-making framework for use by forest managers.

As extension efforts expand, broader exposure of the research results to forest managers will occur. This exposure, along with the creation of a case study of how best to integrate sustainability indicator information into forest harvesting adaptive management will be used to put the sustainability indicator into practice. The focus will be on how best to create a sliding scale of management consequence based on the severity, extent (geographic) and duration of impacts demonstrated through the use of the sustainability indicator. Numeric stream condition scoring associated with this approach will serve as bio-criteria.

In addition to the ongoing extension efforts from the first 2 years, a demonstration package, complete with database management system will be made available, including training venues. Environment Canada's CABIN system is already set up for this, and training in its use occurs on an annual basis.

After year 3, it is intended that the BI sustainability indicator will become a "mainstream tool" with the backing of federal and provincial governments, and those involved in forest management decision making. This backing will include long term funding of the use of the system through partnership arrangements much like the one that is helping to fund this project. This outcome will be the result of the extension plan. As part of this process, it is essential that funding be committed to improving the statistical tools through ongoing review of analytical approaches and updating of the statistical modeling that will be the basis of decision making in the (BI) sustainability indicator system. Advancing this management tool will improve predictive capability and build confidence in the outcome of its applications among many user groups in British Columbia.

# 2 OUTLINE OF ANNUAL ACTIVITIES

## 2.1 Year 1 (2004 – 2005)

Research in the first year focused on the selection of field and laboratory methods and the beginning of model development for the BI) sustainability indicator system.

An initial task was the selection of a benthic invertebrate sampling design. The consensus reached by the project team was that the benthic macroinvertebrate sampling method was not a critical factor in terms of the resulting sustainability indicator resolution or achieving data quality objectives (each of the methods considered in this study have similar characteristics in this regard). Of two sampling devices that are typically used for benthic invertebrate collections in bioassessment programs, it was decided that a traveling kicknet rather than a Surber sampler be used as standard equipment. Use of the kick net provided two benefits:

1. Cost savings allowing more sites to be sampled compared to use of a Surber sampler,
2. The kick net conformed to standardized sampling methods that have been used for development of multivariate bio-assessment models that already exist over a large area of B.C., including the Fraser and Georgia basins as part of initiatives by Environment Canada (Reynoldson et al. 1997, Reynoldson et al. 2001, Sylvestre et al. 2005).

A total of 74 sites were sampled, including approximately 3 reference sites for every influenced site sampled. The sites were spread over a broad area, characterizing the Nadina Forest District and southern portions of the North Coast, Kalum, Skeena and Stikine Forest Districts. Reference site definition criteria were developed based on prior B-IBI work in the area (Rysavy 2000), and protocols used in Environment Canada's Fraser and Georgia Basins studies (Sylvestre et al 2005).

With the benthic macroinvertebrate sampling protocol resolved, the year 1 project moved to solving landscape and aquatic habitat data collection issues. These hydro-geomorphic data are needed to define how both nature and land use may determine what benthic macroinvertebrate assemblages are found in streams. First, a very broad set of variables were considered. Then the list was trimmed based on past Environment Canada experience with the Fraser and Georgia basins, and based on the need for rapid collection, including cutting out potential redundancies between variables. This process involved two separate, but related efforts. First, GIS analyses were used to extract as much relevant "macro" or watershed level hydro-geomorphic and land use information as possible, thus reducing the need for time consuming field work. Second, field testing a large list of site measured variables was used to determine which variables may provide the most accurate and useful characterization of the landscape attributes at the "micro" level.

Once all relevant data was available for sustainability indicator development, the following analyses began:

1. Natural and land use related stressor gradients were determined using GIS analyses and statistical modeling. Data from sites within the region that were sampled using a Surber sampler before the sustainability indicator project began (1999 – 2003) and preliminary sites that were sampled with a kicknet in 2003 and 2004 were used for this purpose. The outcome of this work was a set of GIS derived watershed and land use related attributes, which were found to occur in a variety of states in the watersheds, and had a high probability of being correlated to variation in benthic macroinvertebrate composition and abundance.
2. For both multivariate analytical methods (e.g. the reference condition approach to bioassessment (RCA; Bailey et al. 2004) and B-IBI work, reference site definitions were refined using GIS analyses of the watersheds under study. The outcome of this effort was the selection of a set of reference sites that were either pristine or minimally disturbed.

3. The first step in the RCA work for 2004 was the statistical classification of reference sites based on benthic macroinvertebrate assemblages found at each site. The second step was a comparison of hydro-geomorphic attributes at each test (influenced) site to those characterizing groups of reference sites with the use of statistical modeling. Hydro-geomorphic attributes chosen for modeling were those used in analysis #1 above. This was followed by the derivation of scores of an observed assemblage of taxa over an expected assemblage of taxa (O/E) for each of the influenced sites. These scores represented the degree to which test sites differed from a reference condition in terms of benthic assemblage from the statistically derived reference condition. A score of 1.0 indicated that the test site had the same benthic assemblage as reference sites that were hydro-geomorphically alike. A score of <1.0 indicated that the test site was stressed in some way and deviated from the reference condition. A score of >1.0, indicated a biodiversity "hot spot" where the benthic assemblage was richer and more diverse than the typical reference condition.
4. B-IBI scoring using previously published sets of metrics from Kispiox/Kalum, mid Bulkley, Upper Bulkley and Morice/Lakes studies was completed for the 2004 sites (Sharpe et al. 2005). These scores represented how each site differed from the graphically derived reference condition.
5. Comparisons of the outcomes of the 2 methods of sustainability indicator derivation methods were made by examining the rank order of each site sampled in 2004 in terms of stream condition represented by either O/E score or B-IBI score (Sharpe et al. 2005).

Outcomes of these analyses then formed the basis for year 2 study design.

Three other components of the project were also completed:
1. Recommended protocols for field characterization of hydro-geomorphic site variables were chosen and documented.
2. Laboratory protocols for benthic invertebrate sorting and identification were recommended
3. An extension plan to ensure the further development and use of the sustainability indicator in adaptive management of land and water uses was prepared.

### 2.2 Year 2 (2005 – 2006)

Year 2 focused on four objectives:

1. Expanding the number of sites sampled from 74 to 180. This included the addition of more sites in the mountain pine beetle affected area (New Nadina, Vanderhoof, Prince George, and Caribou Forest Districts) and to a lesser extent in the North Coast Forest District.
2. Optimizing the stressor gradient analysis method using more GIS layers and more complex data reduction and analysis techniques.
3. Optimizing the statistical design for impact classification and assessment: This included comparison of three multivariate modeling approaches (Canadian based Reference Condition Approach (RCA), the Australia based AUSRIVAS and ANNA approaches, and the multimetric B-IBI approach). Comparisons were

made among these approaches to assist in the selection of a sustainability indicator system that will occur in Year 3.

4. Moving forward with the extension plan, including hosting a major international workshop on aquatic biomonitoring and carrying out focus group sessions to provide feedback from land and water use practitioners on the usefulness of the products to date.

Results from each of these tasks are outlined in several sections of this report. An updated stressor gradient analysis was reported in Section 3. It included:

- Delineation of the watershed at each site
- Characterization of the natural and stressor environment of each site
- Definition of the position of each site on a stressor gradient(s)
- Definition of a boundary on the stressor gradient between Reference (low human activity) and Test (high human activity) sites
- Updated predictive models relating a site's benthic invertebrate community to its natural environment was built
- The models were to test sites to determine the deviation between the predicted biota (if the Test site was in Reference condition) and the observed biota.
- Next steps in the stressor gradient analysis was recommended.

In Section 4 a biological assessment using the Australian based AUSRIVAS and ANNA approaches is reported and compared with a multimetric IBI that is described in Section 5. A stream crossing quality index (SCQI) and associated water quality concern rating (WQCR) is compared with the results of the invertebrate based bioassessment approaches that are described in Sections 3 through 5. The results are provided in Section 6. Part of year 1 deliverables was an extension plan to assist in fine tuning the research project to meet the decision support needs of natural resource managers, and to put the aquatic sustainability indicator system in their hands for use. The plan was implemented in year 2 and is summarized in Section 7.

### 2.3 Year 3 (Proposed for 2006 – 2007)

A case study of how best to integrate sustainability indicator information gained from the project as a whole into forest harvesting adaptive management will be developed. It will include a description of how best to create a sliding scale of management consequence based on the severity, extent (geographic) and duration of impacts demonstrated through the use of the sustainability indicator. Both real and synthesized sustainability indicator data will be used. Workshops with forest managers from industry and governments will focus on how best to integrate the data and interpretations into operational and strategic level plans (FSPs, SFMPs and LRMPs).

The outcome will be a consensus on how best to structure adaptive management decision making processes with support from the sustainability indicator system that is developed.

# 3 STRESSOR GRADIENT DEFINITION AND BIOASSESSMENT USING THE REFERENCE CONDITION APPROACH (R.C. BAILEY)

## 3.1 Introduction

In fall 2005, landscape scale environmental and site scale benthic invertebrate data collected from Skeena Region streams in 2004 and 2005 were assembled. Field and laboratory methods were consistent with those described by Perrin et al. (2005). The tasks of this component of the study were to:

1. Delineate the watershed of each site.
2. Characterize the natural (e.g. climate) and stressor (e.g. road density) environment of each site.
3. Define the position of each site on stressor gradient(s). This is a multivariate characterization of human activity in the site's catchment area.
4. Define a boundary on the stressor gradient between Reference (low human activity) and Test (high human activity) sites.
5. Using Reference sites, build a predictive model relating a site's benthic invertebrate community to its natural environment.
6. Apply the predictive model to Test sites and determine the deviation between the predicted biota (if the Test site was in Reference condition) and the observed biota.
7. Suggest the next steps in the development of stressor gradients as needed for the sustainability indicator system.

A manuscript that was recently submitted to an international journal called "Freshwater Biology" is provided in Appendix A to show references and background information that supports the analyses that were used in developing the stressor gradients.

## 3.2 Watershed delineation at each site

The field crew determined the latitude and longitude of each site sampled using a Global Positioning System (GPS). These coordinates were imported as a point shapefile into ArcGIS and projected as a planar layer (from decimal degrees to meter units). Readings referenced to NAD1984 datum were made right at the sampling site to ensure accuracy of site positions that were recorded with GPS.

ArcHydro (Maidment 2002) was used to delineate watersheds, followed by inspection and screen editing of the polygons that defined catchment areas for each site. Accurate determination of the watersheds was dependent on accuracy of coordinates for the site locations, the resolution of the Digital Elevation Model (DEM; usually a polynomial fitting of point elevation model to the earth's surface, sometimes derived from a RADARSAT image), and an accurate line shapefile of the known stream network. The steps involved in delineating a watershed were:

1. The shapefile of site points and known streams was assembled, producing a raster file of the DEM that was called RawDEM (usually at 30x30m resolution, so each pixel may be thought of as a 30x30m tile at a given elevation).
2. The known stream network was burned into RawDEM by creating simulated ditches of 10-100m depth on the DEM corresponding to known streams. Output was called the BurnedDEM.
3. Pits (pixels where mistakes in the DEM left a low elevation pixel surrounded by higher elevation pixels) were filled in BurnedDEM. The result was called FilledDEM.
4. For each pixel on FilledDEM, the direction in which water would flow from the pixel was determined. If all surrounding pixels were at the same elevation, water was assumed to flow from a pixel in the same direction it entered the pixel. The resulting file was called the FlowDir raster file.
5. Using FlowDir, we determined how many pixels would feed water to each pixel. We call this the FlowAcc raster file.
6. We arbitrarily determined where streams would form based on either an area or pixel number criteria from the FlowAcc file. Output was called the StreamGrid raster file.
7. The "catchment" for each segment of the StreamGrid file was defined by where the pixels drained into each segment. Output was called the CatchGrid raster file.
8. Both the StreamGrid and the CatchGrid raster files were transformed into line and polygon shapefiles respectively and they were called DrainageLine and Catchment shapefiles.
9. The point shapefile of sites was then used with the DrainageLine and Catchment shapefiles to determine the catchment area of each sampled site. We called this polygon shapefile a mnemonic such as SkeenaBasins3.

Where necessary, the polygon shapefile of site catchment areas was edited in ArcMap. This was often necessary for very small catchment areas, where the StreamGrid might not include small streams that were sampled. It also happened where there were lakes or ponds in or near the site's catchment. These attributes were represented as flat areas in the DEM. If a polygon shapefile of these water bodies was available, we experimented with burning lakes in (as we do with the known stream network), but this technique is still under development.

For some of the sites, we were not able to delineate a watershed because they were small and there was no known stream network provided for their geographic area.

After delineation and editing of the watersheds, area (in hectares) and perimeter (in kilometers) were added to the attribute table of the SkeenaBasin polygon shapefile using Hawth's Tools (*http://www.spatialecology.com/index.php* ). The frequency distribution of basin areas was close to log-normal with mean basin size (23,000 ha) being much higher than the median (240 ha) (Figure 1). Supporting statistics are shown in Table 1.

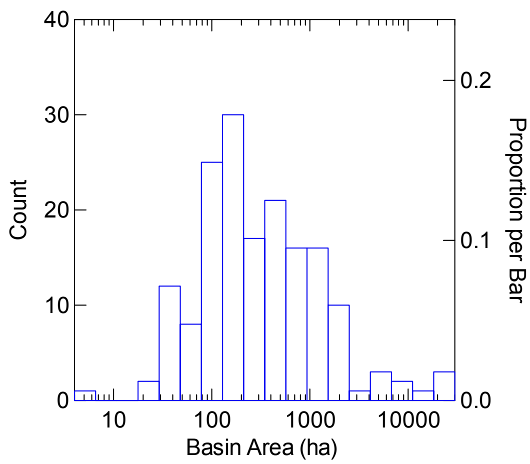


Figure 1. Frequency distribution of basin areas that were calculated among all sampled sites.

Table 1. Descriptive statistics of basin areas that were computed for all sites.

| **Statistic** | **Value** |
|---|---|
| n | 168 |
| Minimum | 4.2 ha |
| Maximum | 27,480 ha |
| Median | 240 ha |
| Mean | 1,115 ha |
| Standard deviation | 3,340 ha |

## 3.3 Characterization of the natural (e.g. climate) and stressor (e.g. road density) environment of each site

ESRI's ArcTools (*http://www.esri.com/products.html* ), followed by data management and manipulation with Microsoft Access and PC-ORD (*http://home.centurytel.net/~mjm/pcordwin.htm* ), were used to collect both natural and stressor environmental information about each site. An example is shown in Figure 2. In this case, a shapefile of areas of high agricultural activity in the Skeena Region is laid on catchment areas using the ArcTools *Intersect* command. The total area of high agricultural activity in each site's basin was summed by a query in Microsoft Access and this value, either as an area in hectares or a percentage of the basin size, became a column in the spreadsheet that described each site's environment. In the example in Figure 2, Site E242648 (2468ha basin) has 579ha (23%) of high agricultural activity. Contrast the several basins north of E242648 (e.g. E245179), which have no agricultural activity.

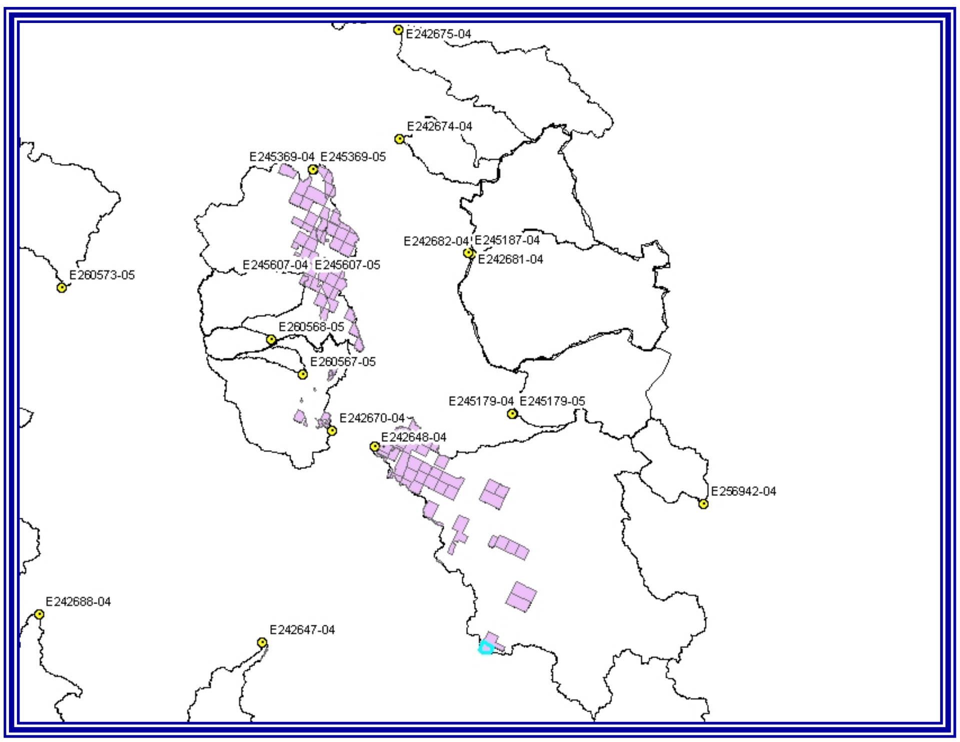


Figure 2. Example of the intersection of a shapefile showing areas of high agricultural activity (purple squares) and basin areas. In this case, Site E242648 (2468ha basin) has 579ha (23%) of high agricultural activity. Contrast the several basins north of E242648 (e.g. E245179), which have no agricultural activity.

A similar procedure was carried out to quantify the following properties of each basin. Data sources are shown in brackets following each attribute.

- Natural environment
  - Basin size (area and perimeter)

  - Climate (Agriculture & Agri-Food Canada)
  - Bedrock geology (British Columbia Ministry of Energy, Mines, and Petroleum Resources)
  - Surficial geology (Natural Resources Canada)
  - Coarse Land Cover (natural elements; Natural Resources Canada)
- Stressor environment
  - Road network (Geobase)
  - Agricultural activity (B.C. Ministry of the Environment)
    - High
    - Medium
    - Low
    - Vacant
    - Rangeland
  - Urban (human activity elements; Natural Resources Canada)
  - Protected areas (human activity elements; Natural Resources Canada)
  - Forest harvest (B.C. Ministry of the Environment)
  - Cutblocks
  - Mining (B.C. Ministry of the Environment)
    - Number and type of mines
    - Area of coal mines

We used Principal Component Analysis (PCA) run in SYSTAT v11 (Systat 2004) to synthesize the various gradients of natural and stressor environments. PCA is a method of summarizing multivariate data as accurately as possible using a few components. The intent is to seek fewer components than variables so that variation left over is negligible. By decomposing many variables down to a few components, the variables within each component are more highly correlated with variables in that component than they are with variables in other components. A component can thereby be interpreted with respect to the characteristics of variables that form that component. Principle components can be given new names that relate to each group of component variables (e.g. PC1, PC2, etc.).

The PCA revealed components corresponding to climate, bedrock geology, land cover, and agricultural activity. In the following output (Boxes 1 to 4), the component loadings for a given set of variables, the % variance explained of each component, and a brief explanation of each interpretable component is provided. The numbers at the top of each column (e.g. 1, 2, 3) indicate the number of each component. Each component can be considered a separate stressor gradient. The numbers under each component label are the component loadings, which are the correlations of the original variables with the components. For example, within component 1 (the first stressor gradient), there are four variables that are associated with climate (Box 1). Of these variables, the mean temperature in January (cold time of the year) and total precipitation is highly correlated with the first stressor gradient. The first gradient was from relatively cold and dry sites to warm and wet sites.

Box 1. Stressor gradients associated with climate.

| Climate | 1 | 2 | 3 |
|---|---|---|---|
| TMEANJAN | 0.965 | 0.210 | 0.010 |
| TMEANJUN | 0.427 | -0.635 | -0.643 |
| PRECTOTALMM | 0.940 | 0.275 | 0.135 |
| SNOWTOTALMM | -0.242 | 0.785 | -0.571 |
| % Variance | 51.391 | 28.466 | 18.952 |

Among geological variables (Box 2) the first gradient was from sites with a lot of Lower Jurassic and not much Middle Jurassic bedrock to those with a lot of Lower Jurassic and not much Middle Jurassic bedrock.

Among land cover variables (Box 3) the first gradient was from sites with mainly older growth forest to sites with younger growth forest. The second gradient was from sites with abundant alpine to sites having young forest cover.

Among agricultural variables (Box 4) the first gradient was from sites with low to intense agricultural activity.

Box 2. Stressor gradients associated with bedrock geology.

Bedrock Geology

| | GeoPC1 | 2 | 3 | 4 |
|---|---|---|---|---|
| CRETAC | 0.003 | 0.044 | 0.000 | 0.047 |
| DEVPERM | -0.003 | -0.001 | 0.000 | -0.000 |
| EARLYCR | -0.000 | 0.000 | 0.000 | 0.000 |
| EARLYJU | -0.006 | -0.003 | 0.001 | -0.000 |
| EARLYTE | 0.008 | 0.021 | 0.017 | 0.004 |
| EOCENE | 0.006 | 0.061 | 0.065 | 0.129 |
| EOCLMIO | 0.003 | 0.012 | 0.009 | 0.011 |
| EOCOLI | -0.000 | 0.000 | 0.000 | 0.000 |
| JURASSIC | -0.002 | -0.001 | 0.000 | -0.000 |
| LATECR | 0.004 | 0.015 | -0.016 | 0.005 |
| LCRETE | 0.000 | 0.000 | 0.000 | 0.000 |
| LCRPAL | 0.002 | 0.005 | 0.007 | 0.002 |
| LEOOLI | 0.016 | 0.054 | 0.095 | -0.174 |
| LATEJU | -0.000 | -0.000 | 0.000 | -0.000 |
| LTRECR | 0.000 | 0.001 | 0.001 | 0.001 |
| LTREJU | 0.001 | 0.002 | 0.001 | 0.000 |
| LOWCR | 0.013 | 0.079 | -0.216 | -0.036 |
| LOWERCR | 0.000 | 0.000 | 0.000 | 0.001 |
| LOWERJU | -0.299 | -0.164 | -0.016 | -0.008 |
| MIDCRET | 0.002 | 0.008 | 0.004 | 0.001 |
| MIDDLEJ | 0.232 | -0.225 | -0.018 | -0.001 |
| MIDDLET | 0.001 | 0.004 | 0.002 | 0.000 |
| MIOCENE | 0.001 | 0.008 | 0.006 | 0.006 |
| ORDOVICI | 0.006 | 0.026 | 0.017 | 0.003 |
| PALEOCEN | 0.000 | 0.002 | 0.001 | 0.002 |
| PAELOZ | 0.000 | 0.002 | 0.001 | 0.000 |
| PALEMESO | 0.000 | 0.000 | 0.000 | 0.000 |
| PALETERT | 0.005 | 0.019 | 0.012 | 0.003 |
| PENNJUR | 0.005 | 0.010 | 0.007 | 0.001 |
| PERMJURA | 0.003 | 0.012 | 0.007 | 0.001 |
| PLEIHOLO | -0.000 | 0.002 | 0.001 | 0.002 |
| PLEIREC | -0.001 | 0.002 | 0.002 | 0.000 |
| PROTPALE | 0.000 | 0.000 | 0.000 | 0.000 |
| TERTIARY | 0.001 | 0.003 | 0.005 | 0.001 |
| TRIAJURA | 0.000 | 0.000 | 0.000 | -0.000 |
| UPCREOC | -0.000 | 0.000 | -0.000 | 0.000 |
| UPJURA | -0.001 | -0.002 | -0.001 | -0.000 |
| UPPAMIJU | 0.000 | 0.001 | 0.000 | 0.001 |
| UPTRI | 0.003 | 0.001 | 0.003 | -0.002 |
| % Variance | 24.255 | 15.976 | 10.418 | 8.546 |

Box 3. Stressor gradients associated with land cover.

| Land Cover | LandCovPC1 | LandCov2 | 3 |
|---|---|---|---|
| ALPINE | 0.034 | -0.171 | -0.055 |
| BARREN | -0.001 | -0.000 | -0.000 |
| WATER | 0.003 | 0.003 | 0.002 |
| GLACIERS | 0.020 | -0.053 | 0.072 |
| OGFOREST | -0.252 | 0.072 | -0.026 |
| RECBURN | 0.001 | 0.001 | 0.000 |
| RECACT | 0.000 | -0.000 | -0.000 |
| RESAG | 0.000 | 0.000 | -0.000 |
| SUBALPCH | 0.003 | -0.041 | -0.002 |
| URBAN | 0.003 | 0.003 | -0.000 |
| WETLANDS | -0.000 | 0.012 | 0.004 |
| YGFOREST | 0.180 | 0.139 | -0.036 |
| % Variance | 52.918 | 32.412 | 6.250 |

Box 4. Stressor gradients associated with agricultural activity.

| Agricultural Activity | AgPC1 | 2 | 3 | 4 |
|---|---|---|---|---|
| PCAGPPRIV | 0.599 | -0.784 | 0.164 | 0.018 |
| PCAGHIGH | 0.959 | 0.198 | 0.179 | -0.028 |
| PCAGMED | 0.795 | -0.070 | -0.599 | 0.060 |
| PCAGLOW | 0.924 | 0.242 | 0.204 | 0.208 |
| PCAGVAC | 0.964 | 0.116 | 0.018 | -0.233 |
| % Variance | 73.883 | 14.596 | 9.207 | 2.042 |

### 3.4 Definition of boundaries on stressor gradients between Reference and Test sites

Both univariate descriptors of the stressor environment (e.g. road density) and multivariate descriptors of the stressor environment (e.g. AgPC1, the agricultural activity within a site's catchment area) were used to define multivariate stressor gradients (Box 5). The first gradient was from sites with low to high agricultural and mining activity, with associated road density. The second PC axis was an inverse stressor gradient. Sites low on the gradient had a relatively large area of cutblocks (high forestry) with associated road density and low amount of park land; those high on StressPC2 gradient had more park land and less cut block area and lower road density.

Box 5. Stressor gradients associated with components of the stressor environment.

Stressor Gradient

| | StressPC1 | StressPC2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| MINEPC1 | 0.862 | 0.320 | 0.087 | 0.108 | 0.369 |
| AGPC1 | 0.867 | 0.275 | 0.165 | 0.110 | -0.366 |
| ROADDEN | 0.540 | -0.505 | -0.206 | -0.641 | -0.001 |
| PCPARK | -0.311 | 0.665 | 0.470 | -0.491 | 0.007 |
| PCCUTBLOCK | 0.053 | -0.655 | 0.747 | 0.095 | 0.033 |
| % Variance | 37.704 | 26.085 | 17.107 | 13.688 | 5.416 |

Basic descriptive statistics of the stressor gradients suggested break points between Reference and Test sites on each axis (Box 6). Sites were defined as Reference on StressPC1 if they had values below the median (-0.300). Recall that this reflected mining and agricultural activity with associated road networks. Sites were defined as Reference on StressPC2 if they had values greater than the median (0.053), which reflected a relative lack of forest harvest but abundant park land. It is worth noting that a site that was deemed Reference in one dimension was not necessarily Reference in another.

Box 6. Descriptive statistics of the stressor gradients.

| | STRESSPC1 | STRESSPC2 | STRESSPC3 | STRESSPC4 | STRESSPC5 |
|---|---|---|---|---|---|
| N of cases | 173 | 173 | 173 | 173 | 173 |
| Minimum | -1.301 | -5.177 | -1.506 | -3.113 | -2.734 |
| Maximum | 9.020 | 2.849 | 5.665 | 0.962 | 3.798 |
| Range | 10.321 | 8.026 | 7.171 | 4.075 | 6.532 |
| Sum | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| Median | -0.300 | 0.053 | -0.399 | 0.248 | -0.084 |
| Mean | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| Standard Dev | 1.373 | 1.142 | 0.925 | 0.827 | 0.520 |

StressPC1 categories are shown as rows in Box 7, while StressPC2 categories are columns.

Box 7. Categories, by Reference and Test sites, associated with stressor gradients 1 (StressPC1 shown in rows) and 2 (StressPC2 shown in columns).

| | Ref | Test | Total |
|---|---|---|---|
| Ref | 68 | 19 | 87 |
| Test | 19 | 67 | 86 |
| Total | 87 | 86 | 173 |

### 3.5 Use of reference sites to build a predictive model relating a site's benthic invertebrate community to its natural environment.

Several models were explored to examine relationships between selected attributes of the invertebrate communities and combinations of stressor gradients and selected habitat variables.

If we try to predict the taxonomic richness at sites defined as Reference along StressPC1, we see that latitude, longitude, and climate explained over 25% of the variation based on multiple regression output in Box 8.

Box 8. Output of a regression model predicting richness from location coordinates and climate principle component scores.

Cases are weighted by the value of variable REF1.

Dep Var: RICHNESS N: 85 Multiple R: 0.502 Squared multiple R: 0.252

Adjusted squared multiple R: 0.224 Standard error of estimate: 3.519

| Effect | Coefficient | Std Error | Std Coef | Tolerance | t | P(2 Tail) |
|---|---|---|---|---|---|---|
| CONSTANT | -125.834 | 62.207 | 0.000 | . | -2.023 | 0.046 |
| LATITUDE | -1.604 | 0.548 | -0.339 | 0.688 | -2.928 | 0.004 |
| LONGITUDE | -1.807 | 0.559 | -0.555 | 0.313 | -3.232 | 0.002 |
| CLIMATEPC1 | -2.295 | 0.462 | -0.894 | 0.286 | -4.971 | 0.000 |

In predicting the taxonomic composition (defined as the Bray-Curtis distance of each community to its group's median community) at sites defined as Reference along StressPC1, we see that latitude, altitude, basin principle component scores, and geology principle component scores explained over 30% of the variation of the first ordination axis (Box 9).

Box 9. Output of a regression model predicting taxonomic composition defined as ordination scores on a first ordination axis from latitude, altitude, basin principle component scores, and geology principle component scores.

```
 Cases are weighted by the value of variable REF1.

Dep Var: BUGNMS1   N: 85   Multiple R: 0.555   Squared multiple R: 0.308

Adjusted squared multiple R: 0.273   Standard error of estimate: 0.492
```

| Effect | Coefficient | Std Error | Std Coef | Tolerance | t | P(2 Tail) |
|---|---|---|---|---|---|---|
| CONSTANT | -22.168 | 3.866 | 0.000 | . | -5.734 | 0.000 |
| LATITUDE | 0.410 | 0.071 | 0.601 | 0.794 | 5.754 | 0.000 |
| ALTITUDE | -0.000 | 0.000 | -0.256 | 0.922 | -2.640 | 0.010 |
| GEOPC1 | -0.264 | 0.145 | -0.184 | 0.839 | -1.812 | 0.074 |
| BASINPC1 | 0.058 | 0.035 | 0.157 | 0.952 | 1.646 | 0.104 |

In predicting the taxonomic composition (ordination scores for benthic invertebrates) at sites defined as Reference along StressPC1, we see that the basin principle component scores alone explained just 14% of the variation of the second ordination axis (Box 10).

Box 10. Output of a regression model predicting taxonomic composition defined as ordination scores on the second ordination axis from basin principle component scores.

```
 Cases are weighted by the value of variable REF1.

Dep Var: BUGNMS2   N: 85   Multiple R: 0.375   Squared multiple R: 0.141

Adjusted squared multiple R: 0.130   Standard error of estimate: 0.768

Effect         Coefficient    Std Error     Std Coef Tolerance     t   P(2 Tail)

CONSTANT             -0.080        0.086        0.000      .      -0.932    0.354
BASINPC1             -0.197        0.053       -0.375     1.000   -3.689    0.000
```

## 3.6 Application of the predictive model to Test sites and determination of the deviation between the predicted and observed biota.

We applied the richness predictive model, failing all sites that were in the outside 25% (either too high or too low) of values predicted by our model (Box 11).

Box 11. Pass and fail results from application of the richness predictive model to reference and test sites (Box 8).

```
          FAIL  PASS    Total
        +-------------+
   Ref  |   21    64  |    85
   Tes  |   33    50  |    83
        +-------------+
Total      54   114       168
```

We applied the composition predictive models, failing all sites that were in the outside 25% (either too high or too low) of values predicted by our model (Box 12).

Box 12. Pass and fail results from application of the two taxonomic composition models (Boxes 9 and 10) to reference and test sites (Box 8).

Composition 1 vs 2: Reference Sites

| | FAIL | PASS | Total |
|---|---|---|---|
| FAIL | 7 | 16 | 23 |
| PASS | 16 | 46 | 62 |
| Total | 23 | 62 | 85 |

Composition 1 vs 2: Test Sites

| | FAIL | PASS | Total |
|---|---|---|---|
| FAIL | 12 | 12 | 24 |
| PASS | 15 | 44 | 59 |
| Total | 27 | 56 | 83 |

The composition models did not appear to be as sensitive as the richness model in detecting departures from Reference on the first stressor gradient (primarily agricultural and mining activity).

### 3.7 Next steps to improve the stressor gradient analysis.

The following analyses will help to improve the stressor gradient analysis and should be implemented in the final year of the project:

- *lakes and ponds*: incorporation of lake and pond features into determination of the site catchment areas.
- *known stream network*: subdivision of known stream network to enable drainage density and other related calculations to help describe natural environment.
- *proximity of natural and stressor features*: calculate more comprehensive description of natural and stressor environment with proximity weighting and basin bathymetry.
- *better characterization of stressors*: particularly with regard to residential and industrial water use, as well as mining, forestry, and agriculture.
- *sampling the full stressor gradients*: choose sites to sample that fill gaps in the stressor gradients, especially correlations among the gradients.

# 4 BIOASSESSMENT USING *AUSRIVAS* AND *ANNA* MODELING APPROACHES (S. LINKE)

## 4.1 Approach

Model development was advanced in 2005 using procedures of the Australian River Assessment System (AUSRIVAS), which is a rapid prediction system used to assess the biological health of running waters (*http://ausrivas.canberra.edu.au/* ). The first step in creating an AUSRIVAS model is to classify the reference sites into groups, based on the faunal composition using UPGMA (Unweighted Pair-Group arithMetic Averaging) as the classification algorithm (Simpson & Norris 2000). A Stepwise Multiple Discriminant Function Analysis (MDFA) is then carried out to determine which environmental variables discriminate best between the groups are most closely related to the structure of the faunal data. To predict the expected community from a certain combination of environmental variables at a test site, the discriminant functions are used to determine the standardized, multivariate distance of the site from the groups (Figure 3).

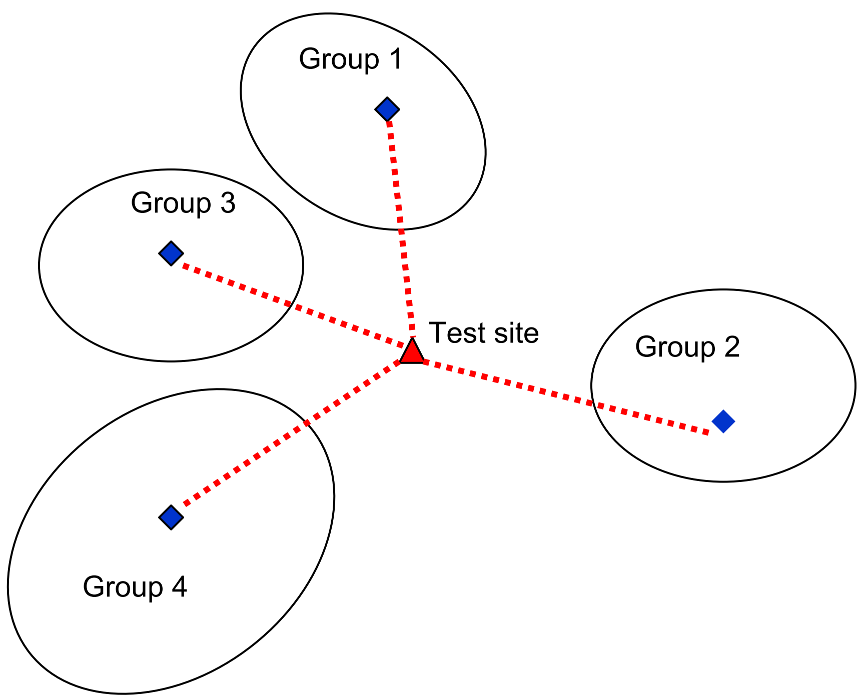


Figure 3. Graphical representation of the probability weighting applied in AUSRIVAS.

Based on this distance, a weighted average of the probability of the taxon occurring at the test site is calculated as described in Clarke et al. (1996) and Moss et al. (1999) and demonstrated in Table 2.

Table 2. Calculation of the probability of a taxon occuring at a site. The final probability is the sum of the contributions of each group, calculated by the probability of the site belonging to that group and the frequency of the taxon being found in that group.

| Classification Group | Probability that test site Y belongs to group | Frequency of taxon X in group (%) | Contribution to probability that taxon X will occur at site Y (%) |
|---|---|---|---|
| 1 | 0.1 | 60 | 6 |
| 2 | 0.6 | 50 | 30 |
| 3 | 0.2 | 60 | 12 |
| 4 | 0.1 | 90 | 9 |
| | | | Σ=Total Probability = 57% |

Unlike the British RIVPACS (Moss et al. 1999) system for bio-assessment, only taxa that have probability of >50% are considered. The rationale behind this is to exclude taxa with a low chance of occurrence from the prediction, so that sampling variability will have a low impact on the sensitivity of the model. On the other hand, enough taxa have to be included to be able to measure the reaction of a community to damage caused by humans.

Simpson and Norris (2000) showed that the 50% cut-off seems to be appropriate for achieving both robustness and sensitivity, as demonstrated in Figure 4, where taxa with a probability >50% provide most of the information.

The observed number (O) of taxa is the number of taxa with >50% chance of occurrence that were found at a site, while the expected number of taxa (E) is the sum of the probabilities of those taxa predicted to occur at the test site. When all of the expected taxa occur, the ratio of observed/expected (O/E) will be close to one. In case of an unnatural change in the community, the number of observed taxa will usually drop and the O/E will decrease. The acceptable range of O/E scores in AUSRIVAS has been defined as the range between the 10th and the 90th percentile of the reference sites Simpson and Norris (2000). An O/E below the 10th percentile indicates an unnatural loss of taxa, an O/E higher than the 90th percentile is judged to be richer than expected and the site is reviewed.

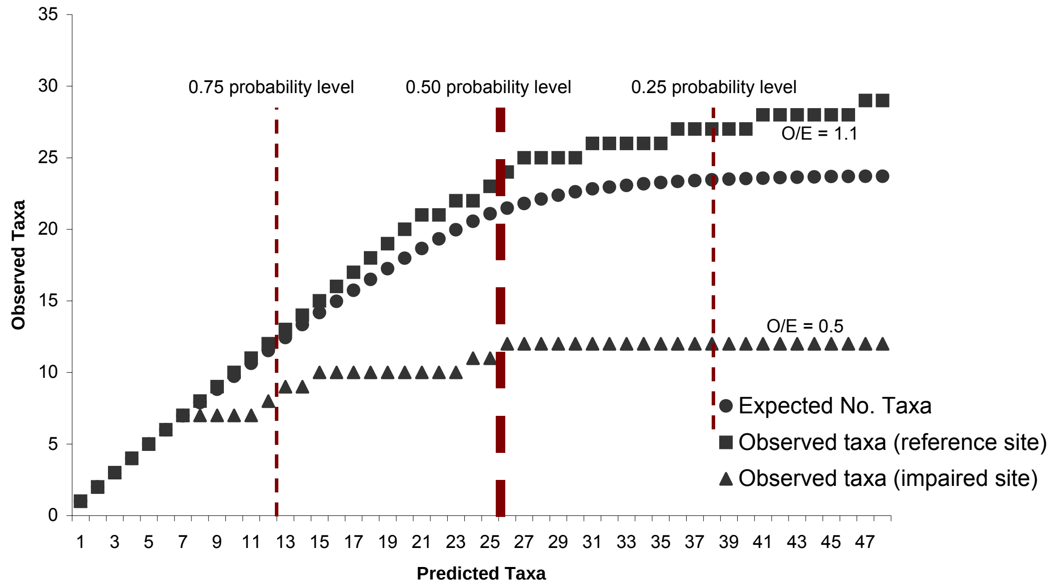


Figure 4 Figure reprinted from Simpson and Norris (2000). Taxa accretion curves for a predicted reference site (circles), an actual reference site (squares) and an impaired site. Taxa below the 0.5 probability of occurrence cut-off contribute little information that is useful for making an assessment of biological condition at a site.

To summarize output in AUSRIVAS, a banding scheme has been developed (Table 3). Hereby the band width is determined by the width of band A, i.e. the 10$^{th}$ and 90$^{th}$ percentiles. Band B starts at the 10$^{th}$ percentile (typically about O/E=0.85) and has the same bandwidth as band A. Band C will have the same bandwidth again, whereas the width of band D will be determined by the difference of its start and an O/E of 0. Sites richer than reference will be assigned band X, which usually characterizes mild organic enrichment and a potential biodiversity "hot spot".

AUSRIVAS models are validated by running the reference sites back through their own model and checking whether they ended up placed in the correct group, based on their environmental variables. The measure of model quality is the percentage of sites that were predicted into the correct groups. As a rule-of-thumb, a model is accepted if the error rate is ≤30%.

Table 3. AUSRIVAS/E-Ball banding schemes

| Band Label | Band name | Comments |
|---|---|---|
| X | Richer than reference | More taxa found than expected.<br>Potential biodiversity "hot-spot"<br>Mild organic enrichment<br>Continuous irrigation flow in a normally intermittent stream |
| A | Reference | Index value within range of central 80% of reference sites |
| B | Below reference | Fewer taxa than expected<br>Potential impact either on water quality or habitat quality or both resulting in a loss of taxa |
| C | Well below reference | Many fewer taxa than expected<br>Loss of taxa due to substantial impacts on water and/or habitat quality |
| D | Impoverished | Few of the expected taxa remain<br>Severe impairment |

### 4.2 Sites and Variables

Sampling sites that were thought to be a reference condition were selected using a combination of local knowledge (“best professional judgement”) and a stressor gradient approach. Our rule was to mark sites as reference when they were labeled reference on the first two PC axes as defined in the stressor gradient analysis that was outlined in Section 3. These sites were accepted as reference if they passed acceptance by local expert opinion (e.g. if a site that was determined to be reference on a PC axis was known to be disturbed to some extent, it was rejected as a reference site). The first PC was related to mining, agricultural activity and associated road networks in the catchment. The second PC was related to old growth forest and cut areas.

A total of 68 sites were selected as reference sites (Appendix D). Invertebrates that were present at only a few sites were removed because they would mathematically not have a chance of being predicted at the $p>0.5$ level.

Over the course of developing linear models and examining stressor gradients, many landscape variables were calculated and accessed using GIS techniques. Landscape variables were derived using the procedures outlined in Section 3. Geological variables were summarised to three main axes using PCA. These variables were considered potential predictors for calculation of O/E scores and for development of linear models describing relationships between the landscape attributes and various measures of biological community diversity, richness, and other metrics. Additional variables were taken from the CABIN description of the sites. The final list of variables is shown in Table 4.

Table 4. Final list of landscape variables used in analysis of stressor gradients and calculation of O/E scores with the Skeena 2005 bioassessment data.

| Variable coding | Description |
|---|---|
| Channel_Depth__Avg___cm_ | Average Channel Depth (cm) |
| latitude | Latitude |
| longitude | Longitude |
| altitude | Altitude (m) |
| Bnkfl_width__m_ | Bankfull width in m |
| Canopy____ | Canopy cover (%) |
| Channel_Depth__Max___cm_ | Maximum Channel depth (cm) |
| Channel_Width__m_ | Channel width (m) |
| Contrees__None_ | Presence of Coniferous trees |
| Dectrees__None_ | Presence of Decideous trees |
| Grasses__None_ | Presence of Grasses |
| Macrophyte__None_ | Presence of Macrophytes |
| Shrubs__None_ | Presence of Shrubs |
| Pool____ | Presence of Pools in the reach |
| Rapid____ | Presence of Rapids in the reach |
| Riffle____ | Presence of Riffle in the reach |
| Run____ | Presence of Run in the reach |
| Slope__m_m_ | Slope (m/m) |
| AREAHA | Area in ha |
| PERIMKM | Total perimeter |
| TMEANJAN | Mean temperature in January |
| TMEANJUN | Mean temperature in June |
| PRECTOTALMM | Total precipitation in mm |
| SNOWTOTALMM | Total snow in mm |
| GLACIERM2 | Area of glacier present ($m^2$) |
| GEOPC1 | Geology PC1 |
| GEOPC2 | Geology PC2 |
| GEOPC3 | Geology PC3 |

### 4.3 AUSRIVAS results

Classification of the reference sites in PCORD based on macroinvertebrate communities showed a strong pattern with relatively little chaining. Seven groups were identified.

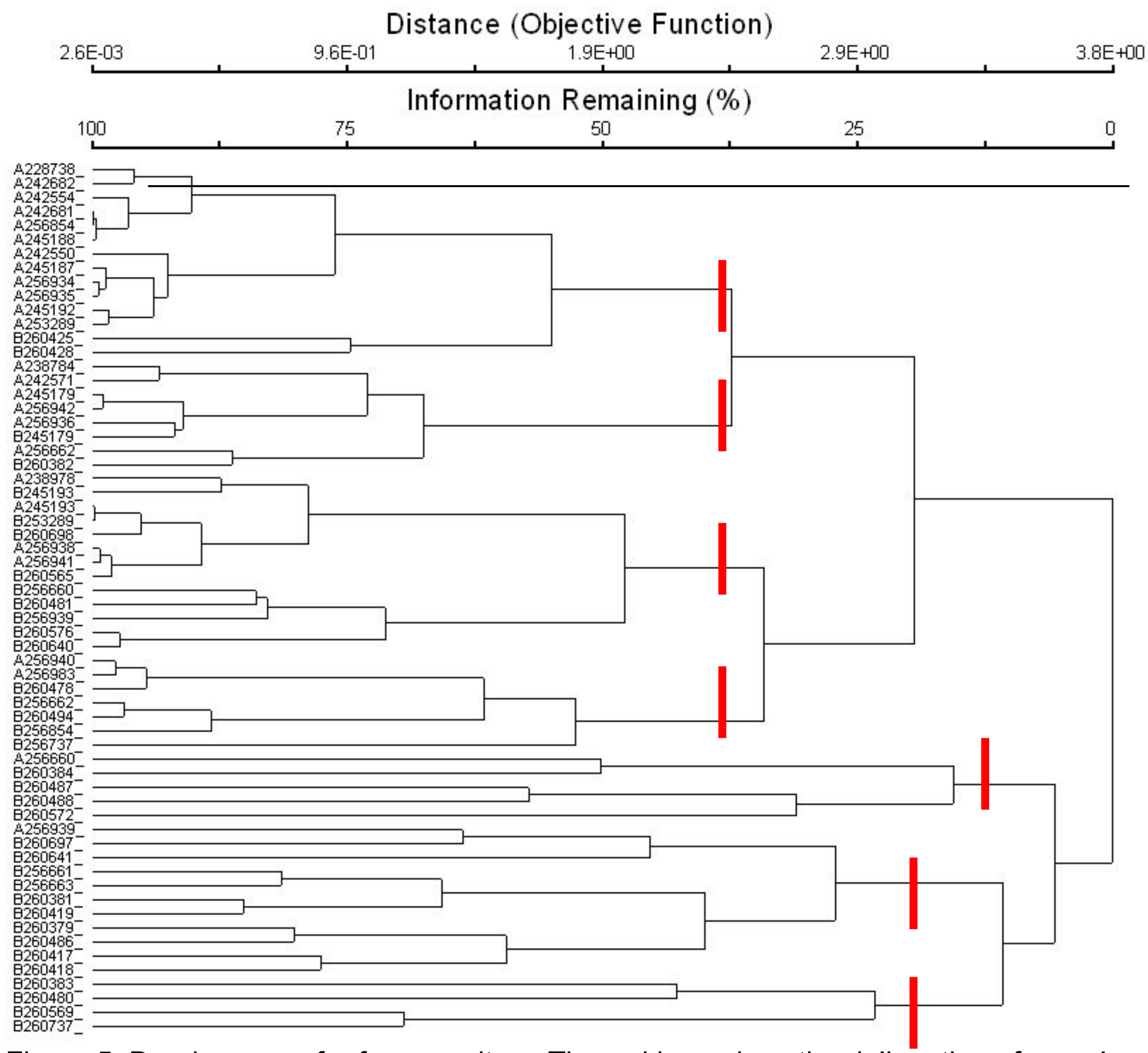


Figure 5. Dendrogram of reference sites. The red bars show the delineation of sample groups.

A list of 8 habitat variables were found by the DFA to best discriminate between the 7 groups of reference sites. Those variables were:

- Average elevation,
- Geo PC 2 and 3,
- Canopy cover,
- Average depth,
- Presence of grasses and macropytes,

- Ecoregion.

Note that the DFA does not infer a causal link between these habitat variables and the invertebrate communities. They were only found to best discriminate between the sample groups and may be surrogates for the real drivers of community composition.

The macroinvertebrate predictions as measured by correlating O/E scores among reference sites indicated very good models. The r-square of the correlation was 0.41, which is considered high in these data where spread around perfect agreement of observed and expected numbers of taxa is anticipated. Remember that the acceptable range of O/E scores in AUSRIVAS has been defined as the range between the 10th and the 90th percentile of the reference sites (Simpson and Norris 2000). The slope of a regression line fit to the observed and expected data was greater than one, which indicated some bias (Figure 6). Banding was tight, with the central 80 percent of reference sites between 0.88-1.21 and a mean very close to 1, which was another indicator of an acceptable model (Figure 7).

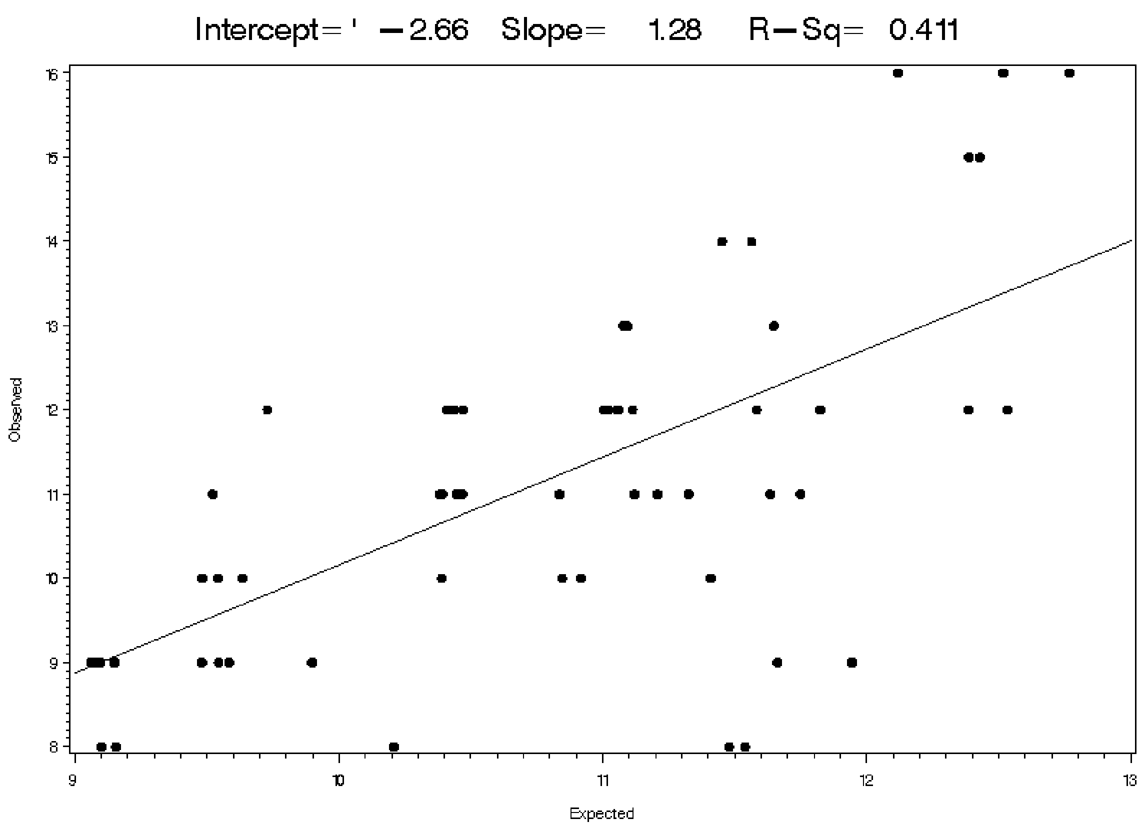

Figure 6. Model validation based on regression analysis of the reference O/E scores.

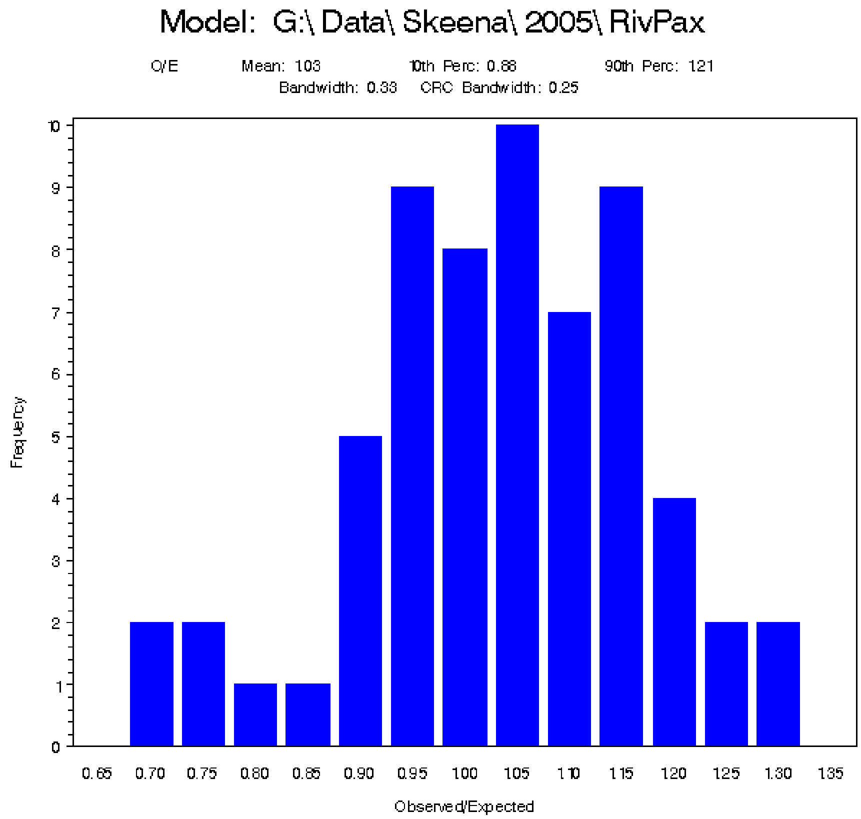


Figure 7. Frequency distribution of reference O/E scores among Skeena sites.

## 4.4 Application of ANNA modeling

### 4.4.1 ANNA method

A modeling approach called ANNA was introduced in Year 1 of the Skeena project and it was repeated with the addition of the 2005 data. The fundamental difference between RIVPACS/AUSRIVAS predictive models and ANNA models is that ANNA models avoid the classification and the discriminant function analyses when matching test sites with reference sites. ANNA finds the reference sites that most resemble the test sites in their values for environmental predictor variables. It then predicts community composition of these sites based on the community composition of those nearest neighbours (Figure 3), thus treating the macroinvertebrate assemblage as a continuum instead of discrete groups.

Once the probability of a taxon occurring has been estimated, O/E scores are computed in the same way as in AUSRIVAS or RIVPACS. To reduce noise created by random occurrence of taxa in reference sites and to be consistent with AUSRIVAS (Simpson and Norris 2000), only taxa with $p > 0.5$ are considered.

There are four steps to constructing an ANNA model:

1. Influential environmental variables are weighted and predictors that are not correlated with the biota are discarded (like the stepwise DFA in the AUSRIVAS or RIVPACS methods) using an approach similar to Principal Axis Correlation (PCC). The Bray–Curtis distance matrix (Bray and Curtis 1957) is calculated for the sites, based on the presence/absence of taxa, and then a non-metric multidimensional scaling (NMDS) is used to ordinate the sites in three dimensions (Clarke 1993). In all our trial datasets, we found three dimensions to be sufficient, because stress levels have never exceeded the recommended 25%.

2. To choose meaningful predictors, a stepwise multiple regression with an entry- and removal criterion of $r^2$=0.1 of environmental predictors is carried out on each of the three ordination axes. Only environmental predictors that correlate with the structure of the biotic data are selected. This approach is used instead of the normal PCC to ensure that cross-correlated variables are not included in the predictions.

3. The final distance of the reference sites to the test sites in multivariate space is calculated by combining the weights on the three axes. This is done by substituting the environmental predictors into the regression equation. The intercepts can be left out, because they would be subtracted from each other in the distance equation. The modified Euclidean distance (*d*) is calculated by:

$$d = \sqrt{(\sum \left| a_n chem_{in} - a_n chem_{jn} \right|)^2 + (\sum \left| b_n chem_{in} - b_n chem_{jn} \right|)^2 + (\sum \left| c_n chem_{in} - c_n chem_{jn} \right|)^2} \quad (1)$$

where $chem_i$ is the $i$ th environmental variable, *a,b,c* are regression coefficients, $i,j$ are observations (sites), and $n$ is the number of variables.

4. Initial trials have shown that weighting the contribution of reference sites to taxon composition at the test site by the reciprocal ecological distance (as calculated in equation (1)) sometimes puts too much emphasis on the very close sites. Therefore, we use the square root of the reciprocal ecological distance to weight the contribution, as demonstrated in equation (2).

The probability of a taxon occurring (*p*) is estimated as

$$p = \frac{\sum_{i=1}^{n} x_i \frac{1}{\sqrt{d_i}}}{\sum_{i=1}^{n} \frac{1}{\sqrt{d_i}}} \quad (2)$$

where $n$ is the number of reference sites, $x_i = 0$ for absence at reference site $i$, $x_i = 1$ for presence at reference site $i$, and $d_i$ is the distance to reference site $i$.

### 4.4.2 ANNA Results

The ANNA model improved greatly, compared to findings from 2004 – 2005 (Sharpe et al. 2005). Despite the mean of the frequency distribution being less than desired (0.93, Figure 8), banding was tighter with inclusion of the 2005 data, with the cut-off from Band A being 0.76 instead of 0.7 (Figures 8 and 9).

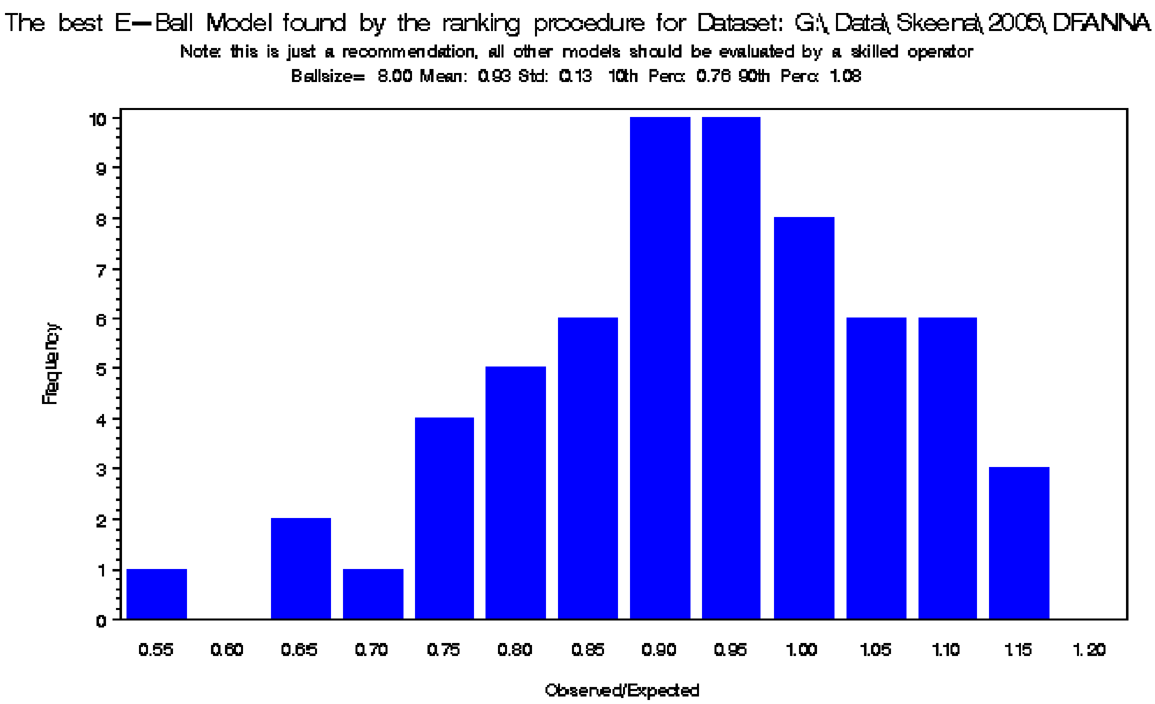


Figure 8. Frequency distribution of reference O/E scores using ANNA among Skeena sites.

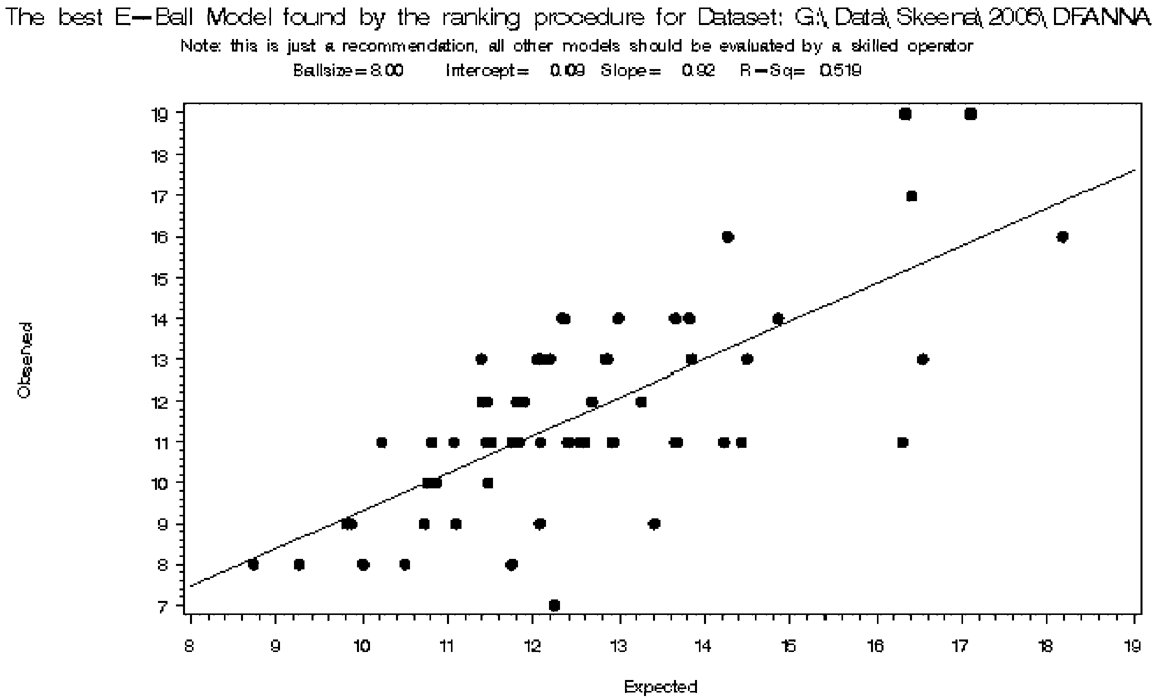


Figure 6. Model validation based on regression analysis of the reference O/E scores.

Application of ANNA in analyses of the 2004 data (Sharpe et al. 2005) showed that taxa were predicted reasonably well, but the associated probabilities of occurrence (e.g. Table 2) were too high, resulting in an increased aggregated expected value (E). With inclusion of the 2005 reference sites, bias was lower (slope of the curve in Figure 6 was closer to 1), and the fit of a line to the observed and expected scores was relatively high ($r^2$ = 0.52).

## 4.5 Comparing AUSRIVAS, ANNA and IBI

Correlation between AUSRIVAS and ANNA scores was very high (Figure 9). This indicated consistency of output between the two methods. Given the general acceptance of AUSRIVAS, the high correlation indicated accuracy of the ANNA output.

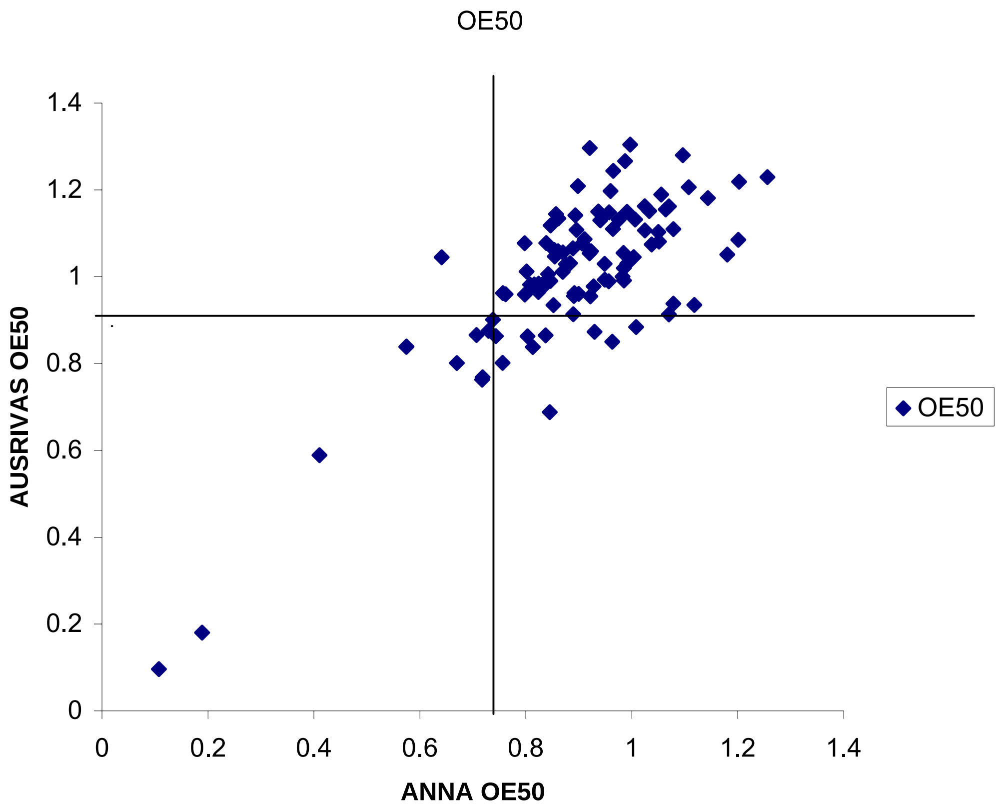


Figure 9. Comparison of AUSRIVAS and ANNA scores.

AUSRIVAS and IBI scores were also highly correlated (Figure 10). The lack of outliers compared to findings from 2004 (Sharpe et al. 2005) indicated that the stressor gradients can be effective in assisting with the selection of reference sites as new sites are added to the database.

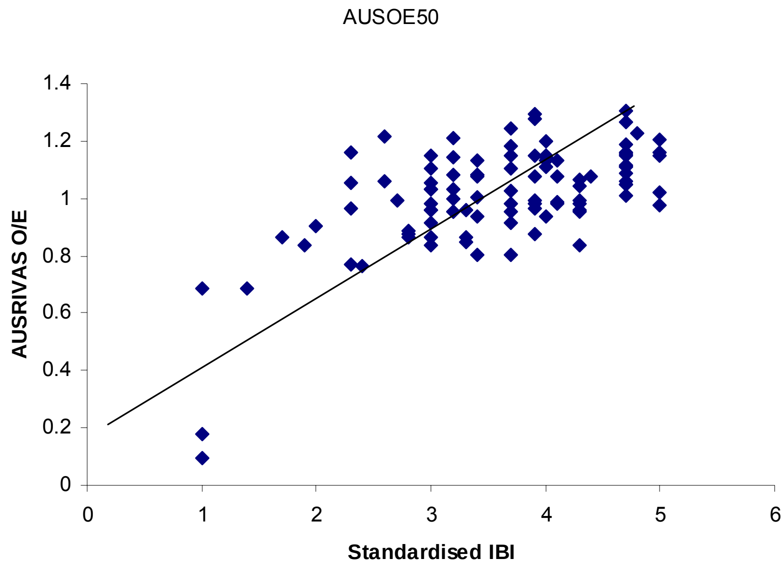


Figure 10. Comparison of AUSRIVAS and IBI scores.

Figure 11 illustrates that ANNA performed better this year with the addition of more samples. A strong correlation with the IBI showed that ANNA was actually usable. However, a group of degraded sites was only detected as mildly, yet not significantly impacted. We conclude that AUSRIVAS is still more sensitive than ANNA.

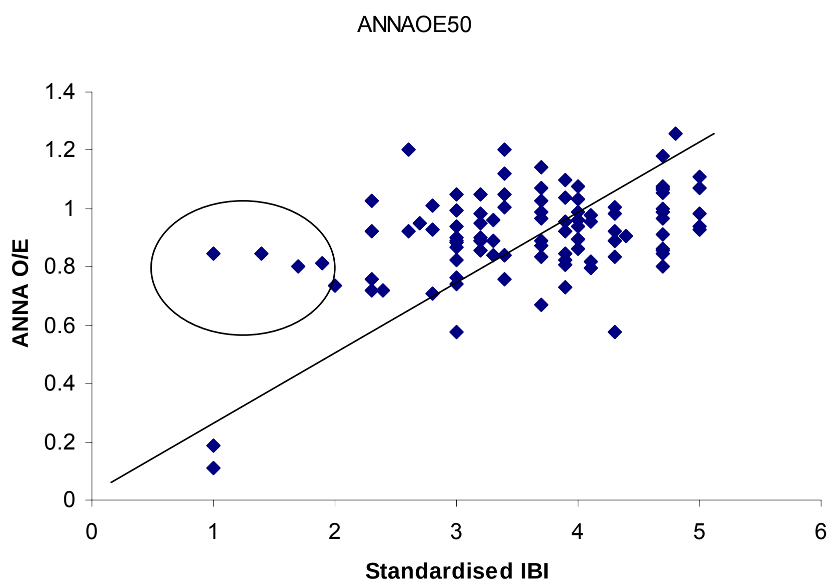


Figure 11. Comparison of ANNA and IBI scores.

### 4.6 Conclusions

Compared to findings from 2004 (Sharpe et al. 2005), several trends are apparent. The $r^2$ of the AUSRIVAS model was slightly lower than that found last year. This can be explained with slight overfitting last year, when there were fewer sites available. The frequency distribution of O/E bands were tighter than last year, giving the model increased sensitivity.

The ANNA model improved greatly compared to the first try last year and can be described as usable now. Bias has been reduced compared to last year, which enabled us to report on scores and make comparisons with the AUSRIVAS scores.

The comparisons on test sites were very encouraging. AUSRIVAS and ANNA showed almost identical outputs. When comparing both with the IBI, AUSRIVAS seems to fail borderline sites, while ANNA passes them, indicating AUSRIVAS is still more sensitive.

For next year, we definitely recommend calculating AUSRIVAS and the IBI. With the use of more sites and improved stressor gradients in 2006, the sensitivity of ANNA may increase. If this does not occur we recommend that ANNA be dropped from further investigation.

## 5 COMPARISON OF B-IBI METRIC SCORES WITH MULTIVARIATE MODELING APPROACHES (S. BENNETT)

### 5.1 Introduction

The benthic invertebrate index of biological integrity (B-IBI) is a multimetric approach to interpreting biological data to assess the condition of a stream (Karr and Chu 1999). A metric is a descriptive statistic of the benthic invertebrate community. Metrics chosen for inclusion in the index must have a consistent and measurable response to increasing human influence at a stream station.

Development of a benthic invertebrate index of biological integrity (B-IBI) in the Kispiox Forest District and Upper Bulkley watershed began in 1999 (Bennett and Rysavy, 2003a, Bennett and Ohland 2002). In the following years, similar projects began in the Bulkley and Kalum Forest Districts Bennett and Rysavy 2003b) and the Lakes and Morice IFPA (Croft 2004). From the beginning, the objective was to develop a ‘results-based’ water quality assessment system to promote biological assessment in streams and use the results to support forest management decisions.

In this section, the metric results and B-IBI scores for the sites sampled in 2004 and 2005 are compared with and evaluated against the stressor gradient modeling results that are reported in Section 3. \

## 5.2 Methods

Field and laboratory methods were consistent with those reported by Perrin et al. (2005).

Metrics were calculated and defined as described by Bennett (2004) and Croft (2004). The results were compiled into a master data sheet list that included a list of taxa, with assigned functional feeding groups, life history and tolerance designations. These data were used for all subsequent calculations.

Before development of a multimetric index, sites were assigned *a priori* to either a reference condition or a test group. For B-IBI development, sites were defined as reference condition group if they met the following criteria as defined by Bennett (2004).

- Less than 5% harvesting or cleared land in catchment,
- No mining in watershed,
- No channelization,
- No upstream impoundments,
- No known point or non-point source discharges (did not include natural slides),
- No urban land use in catchment, and
- An extensive riparian buffer on both river edges separating the stream from the adjacent land use.

Using the methods described in Bennett and Rysavy (2003a, 2003b), metrics were selected for inclusion in the Kispiox and Bulkley B-IBIs if they:

1. Separated uninfluenced from heavily influenced stations using scatter plots and box plots,
2. Had a coefficient of variation less than 1 between reference stations and between multiple replicates collected at a single reference station,
3. Had a proportion of total variance between human influence groups that was greater than the proportion of total variance within human influence groups, and
4. Were shown to contribute unique and biologically relevant information to the index through correlation analyses.

Using these criteria, the same six metrics were chosen for inclusion in B-IBI's for the Bulkley and Kispiox as shown in Tables 5 and 6. In the Upper Bulkley (Bennett and

Ohland 2002) and Lakes and Morice IFPA (Croft 2004), a similar process was used to select metrics. The Upper Bulkley B-IBI was based on 10 metrics as shown in Table 7, while the Lakes and Morice IFPA B-IBI was based on 9 metrics as shown in Table 8.

Table 5: Summary of scoring cutoff points for the six metrics included in the Bulkley TSA benthic invertebrate index of biological integrity.

| **Metric** | **Metric Score** | | |
|---|---|---|---|
| | **1** | **3** | **5** |
| # Plecoptera Taxa | ≤ 3 | 3.1 - 5 | ≥ 5 |
| # Trichoptera Taxa | ≤ 1.5 | 1.6 – 3.4 | ≥ 3.5 |
| # Intolerant Taxa | ≤ 2 | 3 - 4 | ≥ 5 |
| Hilsenhoff Biotic Index | ≥ 4.75 | 3.76 – 4.74 | ≤ 3.75 |
| # of Clingers | ≤ 8 | 8.1 - 10.9 | ≥ 11 |
| % Dominance | > 70 | 60-70 | < 60 |

Table 6: Summary of scoring cutoff points for the six metrics included in the Kispiox TSA benthic invertebrate index of biological integrity.

| **Metric** | **Metric Score** | | |
|---|---|---|---|
| | **1** | **3** | **5** |
| # Plecoptera Taxa | ≤ 3 | 3.1 - 4.69 | ≥ 4.7 |
| # Trichoptera Taxa | < 2 | 2 – 2.99 | ≥ 3 |
| # Intolerant Taxa | ≤ 1 | 2 - 3 | ≥ 4 |
| Hilsenhoff Biotic Index | ≥ 5 | 3.76 – 4.99 | ≤ 3.75 |
| # of Clingers | ≤ 6 | 6.1 – 8.4 | ≥ 8.5 |
| % Dominance | > 75 | 60-75 | < 60 |

Table 7: Summary of scoring cutoff points for the ten metrics included in the Upper Bulkley benthic invertebrate index of biological integrity (Bennett and Ohland 2002).

| Metric | Metric Score | | |
|---|---|---|---|
| | 1 | 3 | 5 |
| # Plecoptera Taxa | ≤ 3.5 | 3.6 - 4.5 | ≥ 4.6 |
| # Trichoptera Taxa | < 1.8 | 1.8 – 2.3 | ≥ 2.4 |
| % Diptera & Non-insects | > 50 | 30 – 50 | < 30 |
| % Ephemeroptera | < 22 | 22 - 34 | > 34 |
| # Intolerant Taxa | ≤ 1 | 2 - 3 | ≥ 4 |
| % Predators | < 4.5 | 4.5 - 10 | > 10 |
| % Dominance | > 75 | 55-75 | < 55 |
| % Sediment Tolerants | > 10 | 2.1 - 10 | ≤ 2 |
| % Clingers | < 20 | 20 - 40 | > 40 |
| Hilsenhoff Biotic Index | > 4.75 | 3.75 – 4.75 | < 3.75 |

Table 8: Summary of scoring cutoff points for the nine metrics included in the IFPA benthic invertebrate index of biological integrity (Croft 2004).

| Metric | Metric Score | | |
|---|---|---|---|
| | 1 | 3 | 5 |
| # Ephemeroptera Taxa | < 5 | 5 - 7 | >7 |
| # Plecoptera Taxa | < 5 | 5 – 7.5 | > 7.5 |
| % Non-insects | > 3 | 1.5 - 3 | < 1.5 |
| # Taxa | < 16 | 16 - 22 | > 22 |
| % Diperta individuals | > 4 | 1.5 - 4 | < 1.5 |
| # Intolerant Taxa | < 2 | 2 - 4 | > 4 |
| % Sediment Intolerant | < 0.5 | 0.5 – 1.5 | > 1.5 |
| % Predators | < 2.5 | 2.5 - 6 | > 6 |
| # of Clingers | < 7 | 7 – 11.5 | > 11.5 |

For each of the selected metrics, scoring cutoffs were chosen from scatterplots, using natural slope breaks where possible, and metrics were scored 5 points if values were similar to uninfluenced streams, 3 points if values were similar to moderately influenced streams, and 1 point if values were similar to heavily influenced streams (Karr and Chu 1999).

The final metric scores were standardized by dividing through by the number of contributing metrics to give a score with a maximum of 5 and a minimum of 1. This process was done to aid interpretation of the results and it provided a consistent approach for comparing stream conditions in areas with different B-IBIs.

Metric results and standardized IBI scores for each site are included in Appendix C.

## 5.3 Results and Discussion

### 5.3.1 Evaluation of Multimetric Results Using the RCA Stressor Gradients

Using GIS-derived landscape variables and field collected habitat variables, the natural (e.g. climate) and stressor (e.g. road density) environment was characterized for each watershed above 165 sampled stream sites in the Skeena Region (Section 3). Two multivariate stressor gradients were identified and each site was scored along each of the stressor gradients (defined in Section 3). The first stressor gradient (STRESS PC1) reflected mining, agricultural activity and associated road networks. The second stressor gradient (STRESS PC2) reflected old growth forest and cut areas. In Section 3 a boundary along each multivariate stressor gradient was used to separate reference condition sites from human-influenced sites. Many of the sites defined as reference were common to both stressor gradients although there were some sites that were defined as reference along one gradient but not the other (see Section 3).

Since a similar, although less sophisticated, approach had been used to define reference sites for building the multimetric models, it makes sense to expect the IBI scores to be higher for stressor defined reference sites than the test sites. As shown in Figure 12, this was not the case. While the majority of the stressor defined reference sites had a score of 3.5 or greater, there were some sites (shown circled) that had IBI scores less than 3, suggesting poor stream condition.

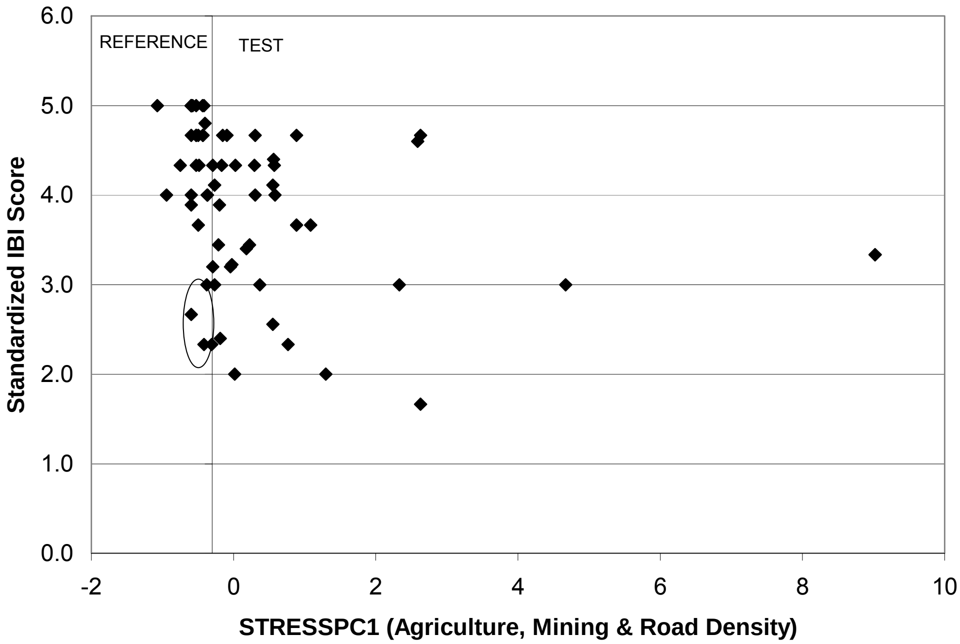


Figure 12 Agriculture, Mining and Road Density stressor gradient (STRESSPC2) plotted against the standardized IBI scores. Reference sites were defined along the stressor gradient as any sites with a value less than –0.3

As shown in Section 3 there were some sites defined as reference using the stressor gradients that did not meet predicted richness or community composition targets. Reference and test sites that failed to meet predicted richness targets (RICH1) are shown in red in Figure 13. As expected, some of the sites with low IBI were also failed using the stressor gradient approach when predicted richness targets were not met. The comparison with IBI results was limited to the richness predictive model (RICH1) rather than the community composition models. As described in Section 3, we found that the richness models were more sensitive in detecting departures from Reference on the first stressor gradient (STRESSPC1).

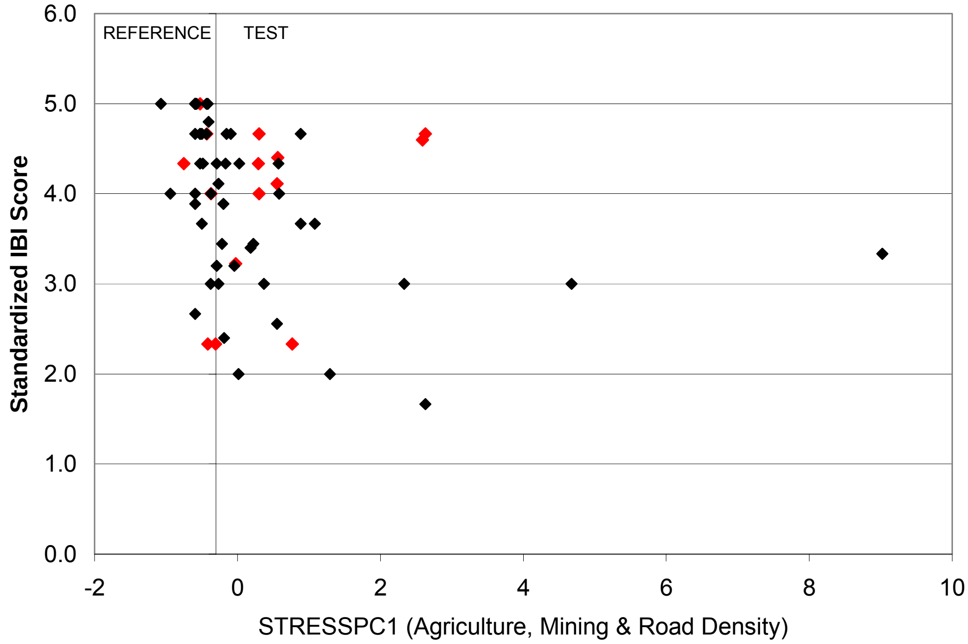


Figure 13 Agriculture, Mining and Road Density stressor gradient (STRESSPC2) plotted against the standardized IBI scores. Reference sites were defined along the stressor gradient as any sites with a value less than –0.3. With respect to the RICH1 RCA model (Section 3), sites in black were passed and sites in red were failed.

A similar approach was taken to investigate the pattern of IBI scores along the forest harvest stressor gradient as shown in Figure 14. Again, while the majority of the stressor defined reference sites had a score of 3.5 or greater, there were some sites (shown circled) that had IBI scores less than 3, suggesting poor stream condition. Sites defined as reference using the stressor gradients that did not meet predicted richness targets (RICH1) are shown in red in Figure 15. Only one of the sites with low IBI was also failed using the stressor gradient approach when predicted richness targets were not met. Most other failed reference sites were sites that had very high IBI scores.

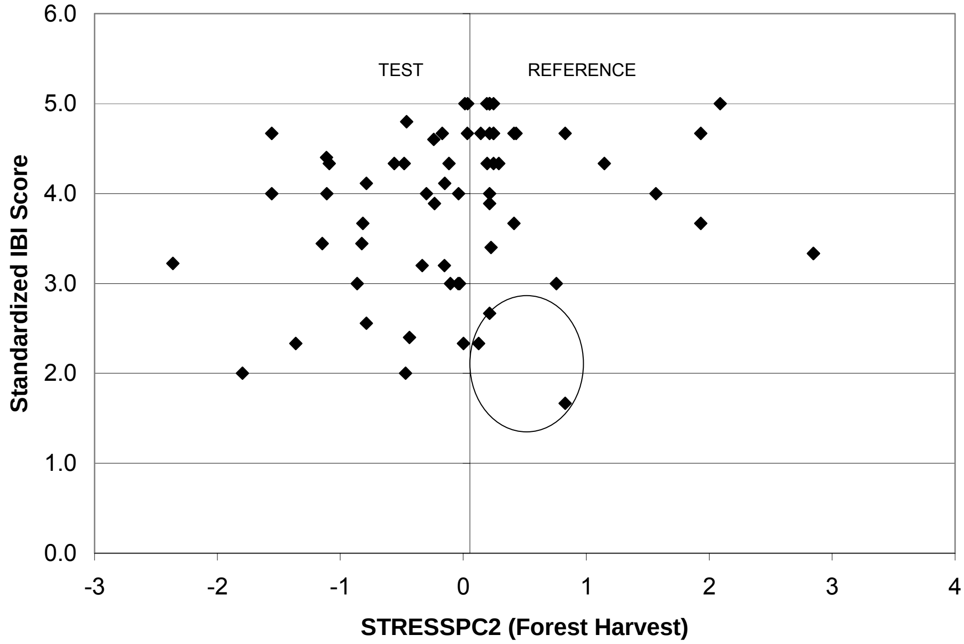


Figure 14 Forest harvest related stressor gradient (STRESSPC2) plotted against the standardized IBI scores. Reference sites were defined along the stressor gradient as any sites with a value greater than 0.053.

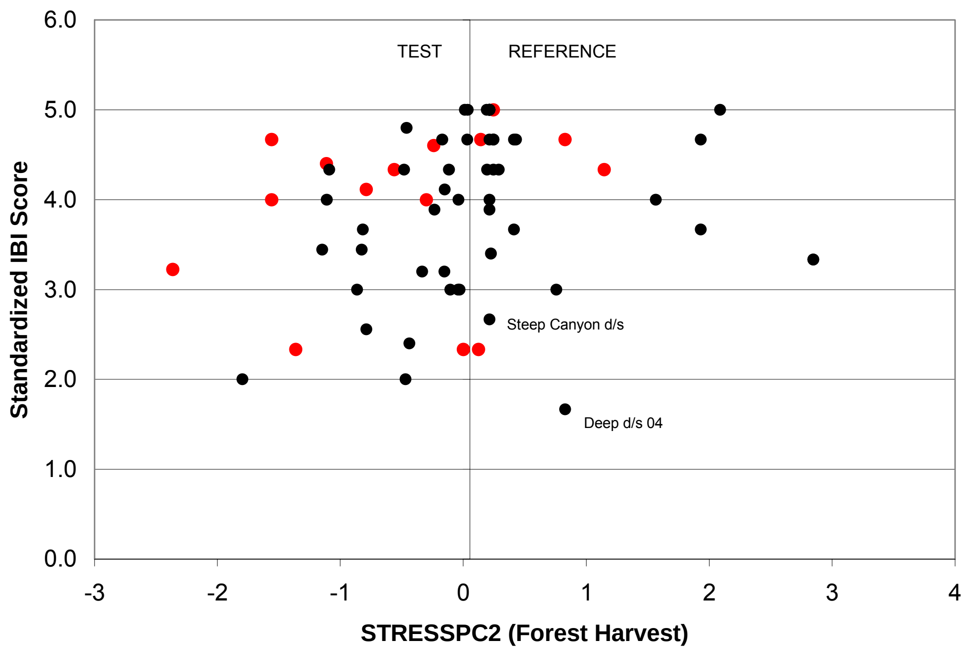


Figure 15 Forest harvest related stressor gradient (STRESSPC2) plotted against the standardized IBI scores. Reference sites were defined along the stressor gradient as any sites with a value greater than 0.053. With respect to the RICH1 RCA model (Section 3), sites in black were passed and sites in red were failed.

As shown in Figure 15, the stressor gradient defined reference sites with lower than expected IBI scores were Deep Creek d/s and Steep Canyon d/s. Both sites were sampled in 2004. Steep Canyon downstream had an IBI of 2.7, a Hilsenhoff biotic index of 3.64, a Simpson's index of diversity of 0.8 and an evenness of 0.26. Simpson's index of diversity[1] approaches 1 as the taxa richness increases, while evenness reflects the abundance and distribution of the individuals in a sample among the taxonomic groups (Environment Canada 2002). Deep Creek had an IBI score of 1.7, a Simpson's index of diversity of 0.34 and evenness of 0.08. The lower Simpson's index of diversity and evenness scores suggest that taxa richness is low (18) and that there were a few dominant taxa at the site (% dominance of three taxa was 90%). In addition, the Deep Creek site was failed using the community composition stressor gradient models (NMS11 and NMS21).

When the results of specific sites are in question, it may be helpful to have more than one model for bioassessment, and a weight of evidence approach can be used. It has also been useful to review individual metric results (*e.g.* HBI or evenness) when specific site interpretation is needed.

### 5.3.2 Comparison of the IBI Scores and Stressor Gradient Approach Results

The IBI scores were plotted against SkeenRIVAS results (see Section 4) and there was a strong positive correlation present. Comparison of the IBI results to the stressor gradient models is a bit more of a challenge since there are numerous richness and community composition models at this point in the development process. However, to compare the two bioassessment approaches, the standardized IBI scores were plotted against the residual richness values as shown in Figure 16.

[1] Simpson's Index of Diversity (D) was calculated as $1-\sum(p_i^2)$

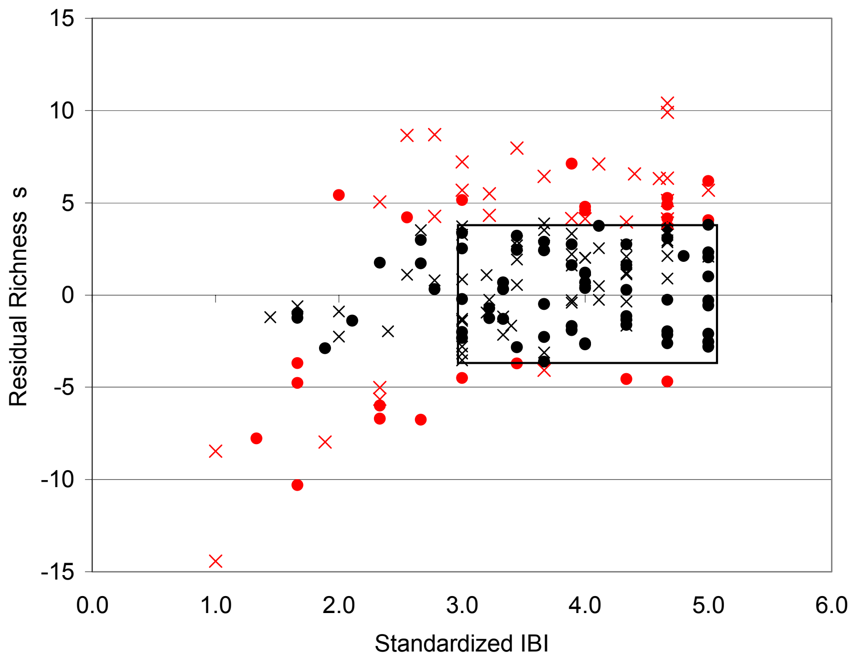


Figure 16 Multivariate stressor gradient model residual richness plotted against a standardized multimetric IBI score for the 2004 and 2005 Skeena region sites. As defined by the stressor gradient PC1, reference sites are shown as circles and test sites are shown as x's. Sites with measured richness much less or much greater than the predicted richness were failed and are shown in red.

Residual richness is the model predicted richness minus the actual richness for a given site. A value close to zero suggests that the actual value was similar to the model predicted value. When the actual and predicted richness were similar, the site was given a 'pass'. Large positive residual richness values show greater than predicted richness at a given site, while large negative residual richness values suggest an impoverished taxa richness compared with reference condition.

As Figure 16 illustrates, if the richness at a given site was within 4 units of the model predicted richness, the site passed (shown in black). One big advantage to this type of modelling is that positive and negative deviations from reference condition can be detected. Sites with a score greater than 3 standardized IBI units are similar to reference condition as defined by the IBI method (see methods). The box in Figure 16 outlines the area of overlap where both bioassessment methods agree that sites are similar to reference condition.

One known weakness of the multimetric approach is that the change from reference condition along a stressor gradient is either considered to be a one-way action or the multimetric model must be limited to one type of stream stressor (e.g. agriculture). For example, a change from reference condition to a stressed condition in a stream is assumed to result in decreased taxa richness. However, in cases of nutrient enrichment, often the stream is in a stressed state, but the taxa richness might be increased. The advantage of the reference condition multivariate modelling approaches over the IBI are clearly shown in Figure 16 where many of the sites failed by the richness model were enriched compared with the reference condition. The multivariate models can detect a positive or negative departure from the expected reference condition.

Similar results were found when the standardized IBI score was plotted against the residual community composition value (NMS12) as shown in Figure 17.

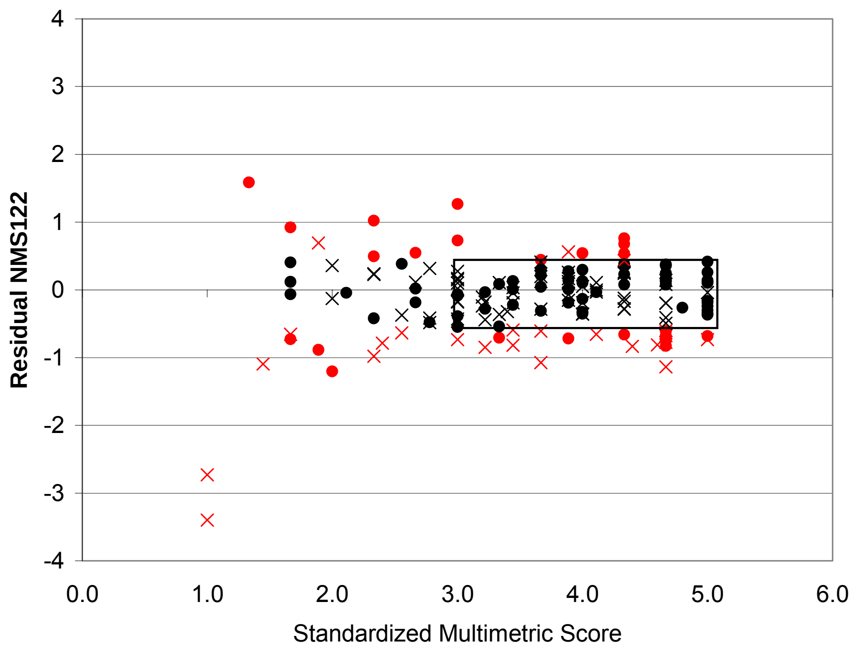


Figure 17 Multivariate stressor gradient model residual community composition plotted against a standardized multimetric IBI score for the 2004 and 2005 Skeena region sites. As defined by the stressor gradient PC1, reference sites are shown as circles and test sites are shown as x's. Sites that failed to match with the reference condition in the multivariate model are shown in red (FAIL), while sites that were comparable to reference condition are shown in black (PASS).

There are also a number of passing sites with very low IBI scores in both Figure 16 and Figure 17 and it's not immediately clear what might be the cause.

### 5.3.3 Recommendations

If we continue with multimetric work, we need to develop a method for detecting enrichment relative to reference condition. It may be prudent to abandon the multimetric indices that are geographically limited and just calculate some metrics to enhance discussion and interpretation of multivariate model results. Currently there are four separate IBI's that cover roughly 40% of the Skeena bioassessment area. If we proceed with the multimetric indices, the GIS variables and environmental gradients calculated and defined in Section 3 may be used to better define natural and geographic gradients for IBI's, rather than the arbitrary Forest District defined areas that are currently being used. In addition, the stressor gradients and individual contributing parameters could be used to refine and re-select appropriate metrics.

# 6 COMPARISON OF A STREAM CROSSING QUALITY INDEX (SCQI) AND INVERTEBRATE BASED BIOLOGICAL ASSESSMENT OF NICHYESKWA WATERSHED STREAMS (S. BENNETT)

## 6.1 Background

In the Skeena Stikine Forest District, maintenance of water quality is a sustainable forest management (SFM) objective within the Babine watershed. Using good road building and maintenance practices should minimize or even eliminate sediment delivery to streams. The Babine Watershed Monitoring Trust (BWMT) has the task of monitoring the effectiveness of the planned management strategies in maintaining water quality within the Babine watershed (Beaudry 2006). The BWMT commissioned a consultant to survey 60 stream crossings in 2005 and implement a stream crossing quality index (SCQI) to assess and score the risk of accelerated erosion and sediment delivery to streams (Beaudry 2006). The SCQI is an effective risk management tool in other areas of B.C. and western Alberta (Beaudry 2006).

The SCQI is a measurement of the size of the adjacent sediment sources modified to reflect the other variables that would increase or decrease the risk of erosion or delivery potential (Beaudry 2006). The SQCI is a predictor of increased turbidity caused by erosion near stream crossings. Stream crossing induced turbidity has been measured to validate the SCQI procedure in B.C. and western Alberta. Beaudry (2006) has shown a strong relationship between SCQI and induced turbidity. A water quality concern rating (WQCR) was developed to quantify the predicted risk to fish habitat due to increased turbidity (Beaudry 2006).

## 6.2 Objectives

The objective of this study was to compare the SQCI and associated water quality concern rating (WQCR) with the results of three invertebrate based bioassessment approaches (a stressor based multivariate reference condition approach (RCA), SkeenRIVAS and a multimetric index) used to evaluate biological integrity of streams using invertebrate communities.

## 6.3 Site selection

Sixty stream sites were surveyed using the SCQI methods in 2005. Of those streams surveyed, 83% had low WQCR (0.7 or less). Twenty-one of the sites were on class 2 and 3 streams, which are generally an adequate depth and width for invertebrate sampling. Invertebrate sampling was conducted on 13 of the 21 sites, across a range of SCQI scores. As shown in Table 9, 10 of the 13 sites had a WQCR of none to low, while two sites had a moderate WQCR and one site scored a high WQCR.

Using GIS-derived landscape variables and field collected habitat variables, the natural (e.g. climate) and stressor (e.g. road density) environment was characterized for each watershed above 165 sampled stream sites in the Skeena Region as described in Section 3. Two of the sampled sites were not included in the multivariate modelling exercises because of missing habitat data or unavailable GIS landscape level data at the time of analyses. The Nichyeskwa sites were used in part to develop the stressor gradients, and each site was scored along each of two multivariate stressor gradients (Section 3). The first stressor gradient (STRESS PC1) reflected mining, agricultural activity and associated road networks. The second stressor gradient (STRESS PC2) reflected old growth forest and cut areas. Along STRESS PC1, 6 of the 11 sites were identified as similar to reference condition for model building (shown in Table 9 as REF). Those six sites included 5 with a WQCR of none or low, and one site with a WQCR of moderate. Along STRESSPC2, 5 sites with a WQCR of none or low were identified as similar to reference condition. All three sites with a WQCR of moderate or high were identified as test sites along the forestry related gradient (STRESS PC2 ), shown in Table 9 as TEST.

Table 9 SCQI streams sites, scores and associated WQCR for 13 Nichyeskwa watershed sites included in the 2005 FSP invertebrate project.

| Stream Crossing Quality Index | | | Reference Sites for RCA and SkeenRIVAS Models | | |
|---|---|---|---|---|---|
| **SCQI Site Name** | **SCQI Total_xing_ score** | **WQCR** | **EMS Code** | **STRESS PC1** | **STRESS PC2** |
| RC09 | 0 | None | E260641 | REF | REF |
| RC11 | 0 | None | E260640 | REF | REF |
| RC16 | **1.18** | High | E260479 | TEST | TEST |
| RC19 | **0.51** | Moderate | E260577 | TEST | TEST |
| RC22 | 0.38 | Low | E260478 | REF | REF |
| RC25 | 0.14 | Low | E242566 | REF | TEST |
| RC37 | 0.39 | Low | E260576 | REF | REF |
| RC40 | **0.6** | Moderate | E260477 | REF | TEST |
| RE02 | 0.01 | Low | E260579 | TEST | TEST |
| RE04 | 0.19 | Low | E260578 | TEST | TEST |
| RE13 | 0.34 | Low | E260637 | n/a | n/a |
| RE14 | 0.12 | Low | E260639 | TEST | REF |
| RE15 | 0.06 | Low | E260638 | n/a | n/a |

## 6.4 Methods

Invertebrates were collected at each of the 13 stream sites in early September of 2005 (see Section 5 for field methods). All Nichyeskwa sites were sampled 100 meters downstream from the road crossing. Three approaches to invertebrate-based bioassessment included a multivariate reference condition approach (RCA) based on natural and human influenced gradients (see Section 3), a multivariate AUSRIVAS type approach (SkeenRIVAS) (see Section 4), and a multimetric index approach (see Section 5).

## 6.5 Results and Discussion

The results of the three invertebrate modelling approaches are summarized in Table 10. The RCA model failed 7 of the 11 sites tested. In comparison, SkeenRIVAS failed just 1 of the 7 test sites (reference sites identified in Table 9 are not scored using this method), and the multimetric index method scored 2 of the 13 sites lower than the remaining 11.

Table 10. SCQI scores and associated WQCR compared with invertebrate modelling results using two multivariate methods (RCA and SKEENRIVAS) and a multimetric method (Kispiox Multimetric model) for the 13 sites included in the 2005 FSP invertebrate project. Results shown in bold print indicate a failed site or non-reference stream condition.

| Stream Crossing Quality Index | | | Invertebrate Model Results | | |
|---|---|---|---|---|---|
| **SCQI Site Name** | **SCQI Total_xing_ score** | **WQCR** | **RCA Richness Model** | **SkeenRIVAS** | **Kispiox Multimetric Index** |
| RC09 | 0 | None | PASS | | **2.3** |
| RC11 | 0 | None | **FAIL** | | 4.7 |
| RC16 | **1.18** | High | **FAIL** | **1.266** | 5.0 |
| RC19 | **0.51** | Moderate | PASS | 1.189 | 4.7 |
| RC22 | 0.38 | Low | PASS | | 4.7 |
| RC25 | 0.14 | Low | **FAIL** | 1.206 | 5.0 |
| RC37 | 0.39 | Low | **FAIL** | | 5.0 |
| RC40 | **0.6** | Moderate | **FAIL** | 1.110 | 5.0 |
| RE02 | 0.01 | Low | **FAIL** | 1.155 | 5.0 |
| RE04 | 0.19 | Low | **FAIL** | 1.162 | **3.3** |
| RE13 | 0.34 | Low | n/a | n/a | 4.7 |
| RE14 | 0.12 | Low | PASS | 1.110 | 5.0 |
| RE15 | 0.06 | Low | n/a | n/a | 5.0 |

There was only moderate agreement between the three invertebrate bioassessment methods and the WQCR. Site RC16, which had the highest WQCR, was failed by both the RCA richness model and the SkeenRIVAS model, but given the best possible score with the multimetric index approach. Additionally, 5 sites with low SCQI scores were failed using the RCA bioassessment approach.

RCA bioassessment results were plotted against the SCQI score as shown in Figure 18. Residual richness is the model predicted richness minus the actual richness for a given site. A value close to zero suggests that the actual value was similar to the reference condition based predicted value. When the actual and predicted richness are similar, the site is given a ‘pass’. Large positive residual richness values show greater than predicted richness at a given site, while large negative residual richness values suggest an impoverished taxa richness compared with reference condition. Figure 18 illustrates that the sites that failed had higher than expected taxa richness compared to the reference condition. Only 2 of the 11 sites had lower than expected richness (a negative residual richness) and the differences were not great enough to fail the sites.

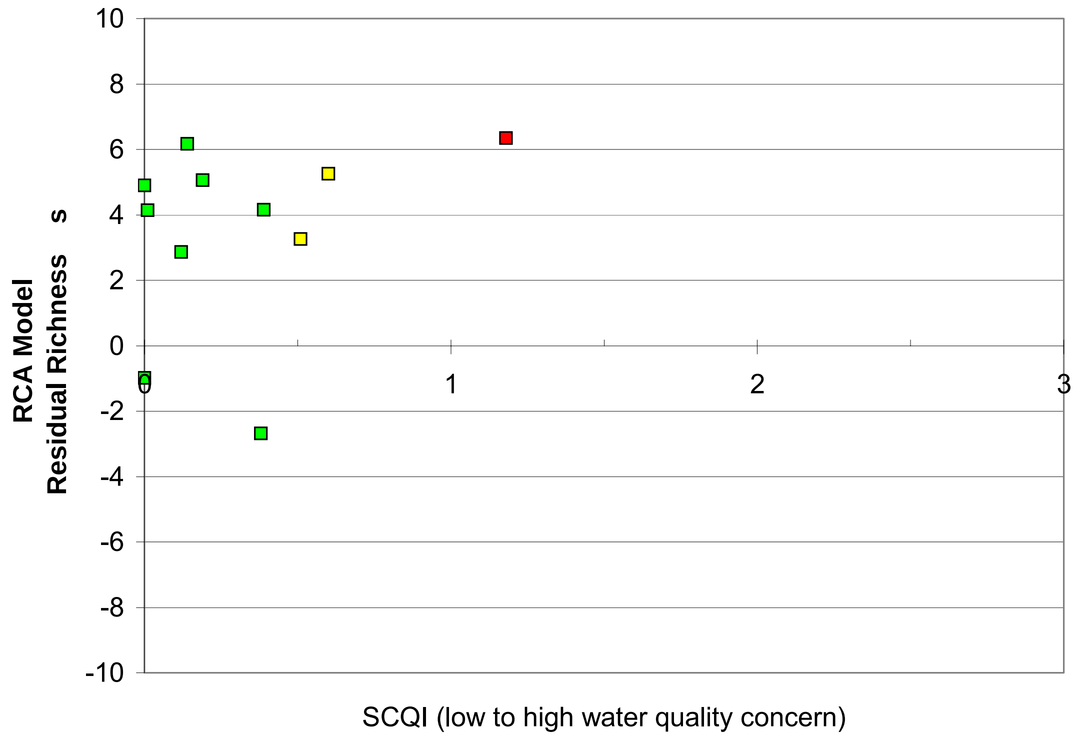


Figure 18 SCQI scores plotted against the residual richness of the RCA model approach. Residual richness is the model predicted richness minus the actual richness at a given site. Stream sites are colour coded by WQCR where red is high, yellow is moderate and green is low or none.

In Figure 19, the SQCI scores are plotted against the SkeenRIVAS model results. During the development of the SkeenRIVAS model, scoring bands are drawn around the reference condition group to aid interpretation of the results. The letters associated with 'bands' in the SkeenRIVAS scores indicate the position relative to reference condition in the multivariate model. Sites that fall within band A are similar to reference condition. Sites that fall in band B are slightly impaired. Sites that fall in band C are moderately impaired. Sites that fall in band D are highly impaired. Band X represents sites that are enriched relative to reference condition. All site scores fall within band A except RC16, which falls in band X suggesting an enrichment of the invertebrate community. These results agree quite well with the RCA results that found that many of the sites were enriched compared to the expected taxa richness. However, it appears that the two models differ in sensitivity, with SkeenRIVAS placing most sites within the reference condition group, while the RCA found those same sites to be enriched compared with the reference group.

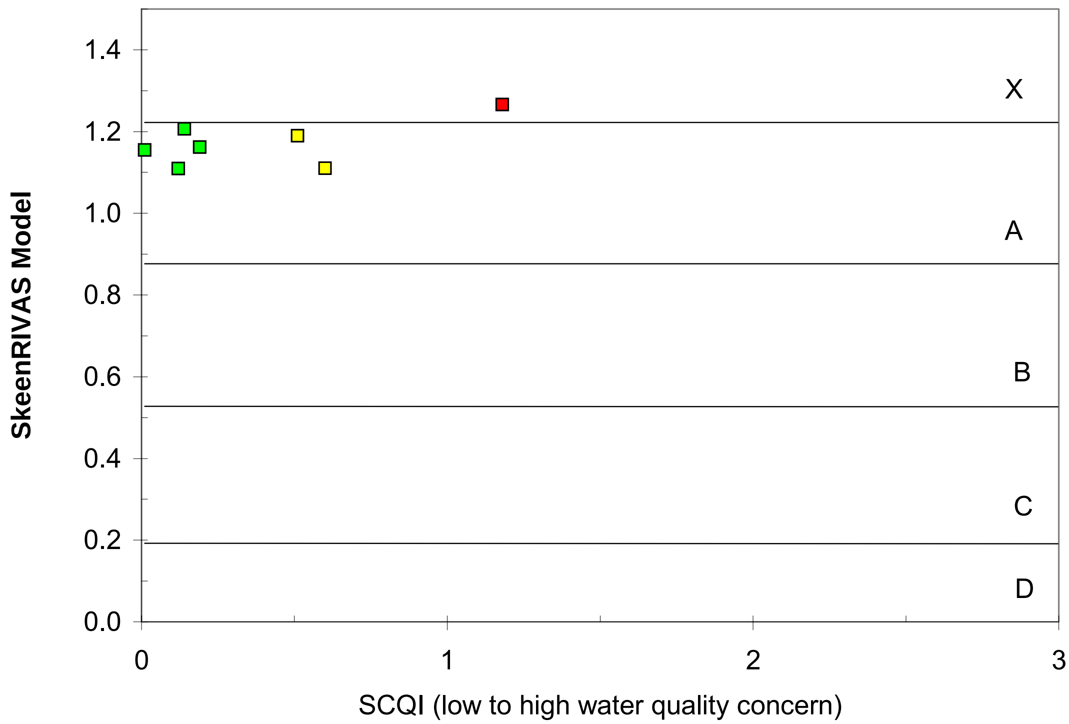


Figure 19 SCQI scores plotted against the SkeenRIVAS model scores. The letters associated with 'bands' in the SkeenRIVAS scores indicate the position relative to reference condition in the multivariate model. Sites that fall within band A are similar to reference condition. Sites that fall in band B are slightly impaired. Sites that fall in band C are moderately impaired. Sites that fall in band D are highly impaired. Band X represents sites that are enriched relative to reference condition. Stream sites are colour coded by WQCR where red is high, yellow is moderate and green is low or none.

Similar to the results of the two multivariate approaches, the multimetric index scores were mostly at the higher end as shown in Figure 20, suggesting taxa rich communities. One big disadvantage to using the multimetric approach is that the metrics are expected to have a one-way response to human influence, which is not always the case. Human influence in a watershed can often cause either impoverishment or enrichment of invertebrate community richness, depending on the type and intensity of the stressor(s) present. The Kispiox multimetric index was not built to identify unacceptable enrichment over reference condition. Both the RCA richness and the SkeenRIVAS models have the ability to detect streams that are impoverished or enriched when compared to the reference condition.

Stream site RC16 had a SQCI of 1.18 (a high WQCR) and was failed by both the RCA richness and SkeenRIVAS models due to enrichment. RC16 scored very high with the multimetric method indicating a stream in excellent condition when it may have unacceptable enrichment compared with the reference condition. If only the multimetric method had been used for bioassessment of these streams, several sites (*e.g.* RC16, RC25) would have been mistakenly assessed as uninfluenced.

Stream site RC09 had a SQCI of zero and an associated WQCR of none. In terms of invertebrate richness, the site passed using the RCA approach. It was determined to be a reference site using the stressor gradients and was not assessed using SkeenRIVAS because reference sites were used to build the model. Using the multimetric approach, the site had the lowest score, indicating the poorest stream integrity. A closer look at the metrics (see Appendix C) for the site found that the site had a low Simpson's Diversity index, low evenness, and high Hilsenhoff biotic index, all indicators of poor water quality.

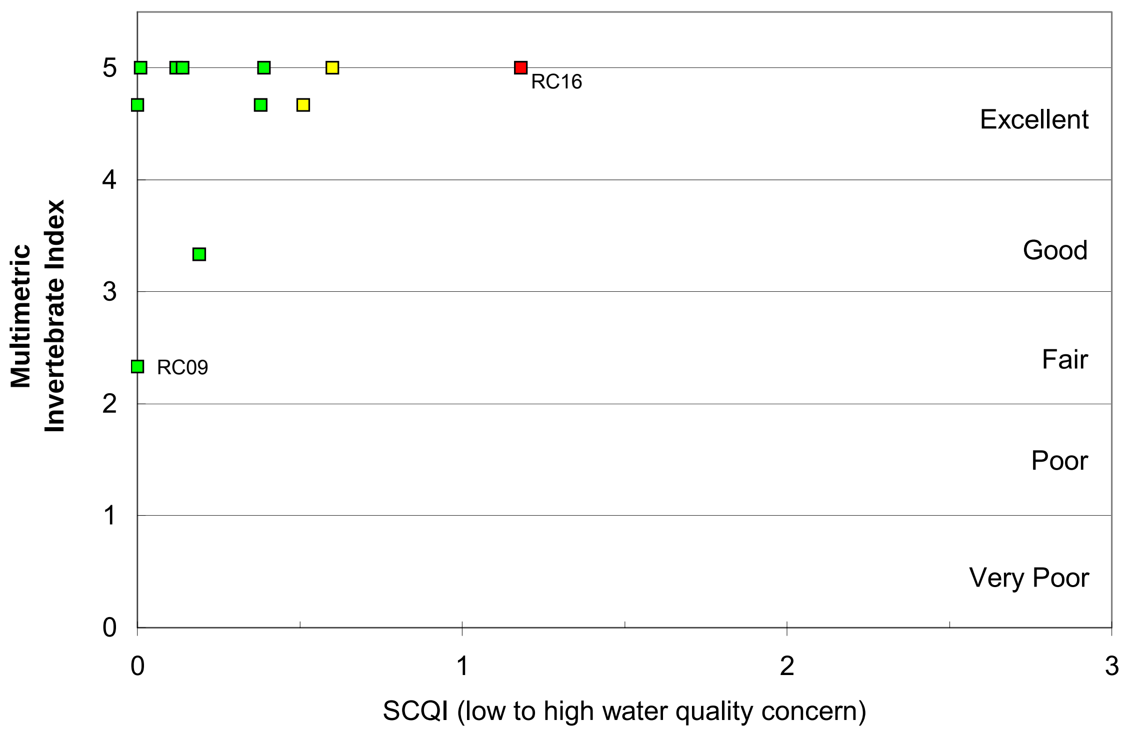


Figure 20 SCQI scores plotted against multimetric invertebrate index results. A multimetric score close to 5 suggest a stream in excellent condition, while a score close to 1 suggests a stream in very poor condition. Stream sites are colour coded by WQCR where red is high, yellow is moderate and green is low or none.

As shown in Table 10, scores for the high and moderate SCQI rated sites were all at the high end of the multimetric results (close to 5) suggesting a condition similar to reference condition. These results are important in that they highlight the need for the ability to detect enrichment with the multimetric approach. This would be especially important in areas where levels of land use for agriculture and range are high.

### 6.6 Conclusions

Overall, the SCQI suggest that there is a low WQCR for the majority of streams surveyed in 2005 (Beaudry 2006). The invertebrate bioassessment results for the 11 sites sampled show that many of the sites are mildly enriched compared to the reference condition. The results may be reflecting an increase in nutrients downstream from the road crossings due to nutrient-rich inputs along ditches and road cuts or an increase in UV and periphyton (food source for invertebrates) downstream of the road cuts. Increased richness and high multimetric scores would not be expected if there were severe sedimentation impacts to fish habitat. The bioassessment tools have shown that the sites sampled (except possibly RC09) are in relatively good condition. However, if nutrient enrichment is a concern in the Nichyeskwa and Babine watersheds, this may warrant further investigation.

The weak correspondence between the SCQI and the bioassessment methods was not unexpected. The SCQI is a measure of risk and the bioassessment methods provide an insight to the current condition based on invertebrate community and habitat attributes that are rated to a reference condition. Beaudry (2006) states: "*The SCQI is certainly not an assessment tool to evaluate the specific impacts of road crossing on the aquatic environment, but rather a tool to score the hazard level that forest roads have on increasing erosion and sediment delivery to the stream network*". Both the WQCI and the invertebrate bioassessment methods contribute to a successful effectiveness monitoring program designed to maintain water quality in the Babine River watershed.

## 7 BENTHIC MACROINVERTEBRATE SUSTAINABILITY INDICATOR PROJECT EXTENSION PLAN PROGRESS REPORT (I. SHARPE)

Part of year 1 deliverables was an extension plan to assist in fine tuning the research project to meet the decision support needs of natural resource managers, and to put the aquatic sustainability indicator system in their hands for use. The plan has now been implemented in year 2 as follows:

- May 2005 international workshop (2 days) on biomonitoring at UBC was held. Presentations can be found at: *http://faculty.forestry.ubc.ca/richardson/home.html* (see benthic biomonitoring 2005 link)
- 2 phone-in focus group sessions including presentation and Q&A session totalling 15 forest management practitioners (see presentation in Appendix B)
- 1 provincial watershed sensitivity workshop presentation to 20 scientists with Q&A session.

The phone-in focus groups and provincial watershed sensitivity classification workshop were held in April 2006. The feedback from these sessions has been valuable in providing advice which is helping to focus year 3 efforts to develop case study materials as a means of transferring the newly developed sustainability indicator system to the intended users. The following is a summary of comments and advice received in these sessions:

- System has potential for state of the environment reporting.
- Benthic macroinvertebrates are a useful tool in terms of being able to explain effects of land and water use on ecosystems, unlike water physical / chemical analyses.
- It is still not clear whether benthic assemblages are responding to reach level or basin wide disturbances. One person familiar with some of the Fraser Basin RCA work stated that one of the streams that had a score reflecting a “not stressed” condition is actually highly influenced upon observation. This needs more study and explanation in terms of the pervasiveness of contradictory results.
- Some people do not believe that enough watershed variability can be characterized to adequately define the reference condition or the range of land and water use related disturbances. How many sites over a given landscape would be needed to ensure adequate characterization of reference condition?
- The complexity of the mathematical models used in RCA is confusing to some. Being able to simplify the outputs is important. B-IBI system seems easier to understand in terms of biological response to disturbance.
- There is some scepticism as to how the BI sustainability indicator can be used to support watershed management decisions. How practical is it?
- How can cause and effect hypotheses be tested using the RCA approach? This question is important because the RCA was never intended to test cause and effect hypotheses. It is a screening tool from which detailed experimentation or other techniques are required for site specific testing of cause of site impairment. This point requires clarification to potential users of the RCA.
- The RCA approach book by Bailey et al is vague on the modelling. Explanations of how the stratifications of reference sites is done are needed.
- This kind of work should not proceed until there are well researched “process models” explaining how certain watershed disturbances affect benthic assemblages (dose / response relationships). Beyond this, there must be a means built into the system to show causation.

- The models generated in this work must have confidence intervals defined so that the validity of the results can be judged.
- Classical statistics would dictate that at least 25 samples would be needed to determine a statistical norm for one riffle. How can a single kicknet sample in a riffle be of value?
- Temporal variability does not seem to be addressed in this work. Has there been any attempt to address “lag effects”?
- It is unfortunate that there is no long term water quality monitoring across the landscape.
- It is surprising that there is little change in benthic assemblages from logging. Is there a threshold beyond which change becomes more pronounced? How will we choose thresholds of change (biocriteria) for logging?

These issues and others which are identified in future focus group sessions will be addressed in the year 3 case study development. The intention is to satisfy the questions about scientific rigor of the proposed sustainability indicator, and inform about the limits and advantages of it.

## 8 LIST OF REFERENCES


Bailey, R.C., R.H. Norris, and T.B. Reynoldson. 2004. *Bioassessment of freshwater ecosystems using the reference condition approach*. Kluwer Academic Publishers. New York.

Beaudry, P and Associates Ltd. Results of the Stream Crossing Quality Index (SCQI) Survey for the Nichyeskwa Watershed, Skeena Stikine Forest District. January 2006. Prepared for The Trustees of the Babine Watershed Monitoring Trust c/o BV Research Centre, Smithers, BC.

Bennett, S. 2004. Comparison of the Effects of Four Rapid Bioassessment Sample Collection and Processing Methods on Benthic Invertebrate Community Metrics. Unpublished Manuscript.

Bennett, S. and K. Ohland. 2002. Expansion and re-calibration of a multimetric benthic invertebrate index of biological integrity for the Upper Bulkley River. Prepared for Community Futures Development Corporation of Nadina.

Bennett, S. and K. Rysavy. 2003a. A benthic invertebrate index of biological integrity for streams in the Kispiox Forest District: field season 2002. Prepared for Pacific Inland Resources and B.C. Environment, Smithers, BC.

Bennett, S. and K. Rysavy. 2003b. A benthic invertebrate index of biological integrity for streams in the Bulkley TSA: field season 2002. Prepared for Pacific Inland Resources and B.C. Environment, Smithers, BC.

Bray, J.R. and J.T. Curtis. 1957. An ordination of the upland forest communities of southern Wisonsin. Ecological Monographs. 27: 325-349.

CCME 2005. Developing biocriteria as a water quality assessment tool in Canada: Scoping assessment project #344-2005. Canadian Council of Ministers of the Environment. National Guidelines and Standards Office. Gatineau, PQ. 37p.

Clarke, K.R. 1993. Non-parametric multivariate analyses of changes in community structure. Aust. J. Ecol. 18: 117-143.

Clarke, R. T., Furse, M. T., Wright, J. F. & Moss, D. 1996. Derivation of a biological quality index for river sites: Comparison of the observed with the expected fauna. *Journal of Applied Statistics*, 23, 311-332.

Croft, C. 2004. Expansion and re-calibration of a multimetric benthic invertebrate index of biological integrity for the Morice and Lakes IFPA. Prepared for the Morice and Lakes IFPA.

Environment Canada. 2002. Metal Mining Guidance Document for Aquatic Environmental Effects Monitoring. June 2002. *http://www.ec.bc.ca/eem*

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

Karr, J.R. and E.W. Chu. 1999. Restoring life in running waters: Better biological monitoring. Island Press, Washington, DC.

Kearns, B.L. and J.R. Karr. 1994. A benthic index of biotic integrity (B-IBI) for rivers of the Tennessee Valley. Ecological Applications 4: 768-785.

Maidment, D.R. 2002. Ed. ArcHydro: GIS for Water Resources. ESRI. Redlands, CA.

Moss, D., Wright, J. F., Furse, M. T. & Clarke, R. T. 1999. A comparison of alternative techniques for prediction of the fauna of running-water sites in Great Britain. Freshwater Biology, 41, 167-181.

Perrin, C.J. 2006. Application of multivariate techniques to examine quality of streams in the Okanagan region. Report prepared by Limnotek Research and Development Inc. for B.C. Environment. 39p.

Perrin, C.J., D. Dolecki, and S. Salter. 2005. A recommended approach for sorting and sub-sampling benthic invertebrate samples for use in the Skeena RCA analysis. Report prepared by Limnotek Research and Development Inc. for B.C. Ministry of Water Land and Air Protection. Smithers, B.C. 9pp.

Perrin, C.J. and S. Sylvestre. 2006. Assessment of deep water habitats of a large river estuary using the reference condition approach. Proceedings of 54th Annual Meeting of the North American Benthological Society, Anchorage, Alaska, June 4-9, 2006.

Reynoldson, T.B., R.H. Norris, V.H. Resh, K.E. Day, and D.M. Rosenberg. 1997. The reference condition approach: a comparison of multimetric and multivariate approaches to assess water quality impairment using benthic macroinvertebrates. Journal of the North American Benthological Society 16: 833-852.

Reynoldson, T.B., D.M. Rosenberg and V.H. Resh. 2001. Comparison of models predicting invertebrate assemblages for biomonitoring in the Fraser River catchment, British Columbia. Canadian Journal of Fisheries and Aquatic Sciences 58: 1395-1410.

Rosenberg, D.M., T.B. Reynoldson and V.H. Resh. 1999. Establishing reference conditions for benthic invertebrate monitoring in the Fraser River catchment, British Columbia, Canada. Environment Canada's Fraser River Action Plan Report DOE FRAP 1998-32. *http://www.rem.sfu.ca/FRAP/9832.pdf*.

Rysavy. S. 2000. Calibration of a multimetric benthic invertebrate index of biological integrity for the upper Bulkley River watershed: A tool for assessing and monitoring stream condition. Report prepared by Bio Logic Consulting Ltd. for Community Futures Development Corporation of Nadina. Houston, B.C. 20p plus appendices.

Sharpe, I., C.J. Perrin, S. Bennett, S. Linke, and B. Bailey. 2005. Benthic macroinvertebrate sustainability indicator development project: Summary of progress in year 1. Report prepared by B.C. Ministry of Water Land and Air Protection. Smithers, B.C. 40 pp.

Simpson, J. & Norris, R. H. (2000) Biological assessment of water quality: development of AUSRIVAS models and outputs. In *RIVPACS and similar techniques for assessing the biological quality of freshwaters* (ed. J. F. Wright, D. W. Sutcliffe & M. T. Furse), pp. 125-142. Freshwater Biological Association and Environment Agency, U.K., Ableside, Cumbria, U.K.

Sylvestre, S., M. Fluegel, and T. Tuominen. 2005. Benthic invertebrate assessment of streams in the Georgia Basin using the reference condition approach: expansion of the Fraser River invertebrate monitoring program 1998-2002. Environment Canada report EC/GB/04/81.

Systat. 2004. SYSTAT 11. Systat software Inc. Richmond CA.