BRITISH COLUMBIA
The Best Place on Earth

Ministry of Environment

BIOASSESSMENT OF STREAMS IN NORTH-CENTRAL BRITISH COLUMBIA USING THE REFERENCE CONDITION APPROACH

FINAL REPORT v.2a

March 31, 2007

Submitted to

B.C. Forest Science Program

LIMNOTEK

BIOASSESSMENT OF STREAMS
IN NORTH-CENTRAL BRITISH COLUMBIA
USING THE REFERENCE CONDITION APPROACH

Submitted to
B.C. Forest Science Program, PricewaterhouseCoopers LLP

Prepared by

C. J. Perrin
Limnotek Research and Development Inc., Vancouver, B.C.

S.A. Bennett
Bio Logic Consulting Ltd., Terrace, B.C.

S. Linke
University of Queensland, St. Lucia, Australia

A.J. Downie and G. Tamblyn,
B.C. Ministry of Environment, Smithers, B.C.

B. Ells
B.C. Ministry of Forests and Range, Smithers, B.C.

I. Sharpe
B.C. Ministry of Environment, Smithers, B.C.

and

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

March 31, 2007

Library and Archives Canada Cataloguing in Publication Data
Main entry under title:
Bioassessment of streams in north central British Columbia
using the reference condition approach : final report.

"Submitted to B.C. Forest Science Program"
Also available on the Internet.
ISBN 978- 0-7726-5790-9

1. Water quality biological assessment - British Columbia, Northern. 2. Forest management – British Columbia, Northern. 3. Benthic animals – Effect of logging on - British Columbia, Northern. I. British Columbia. Forest Science Program. II. British Columbia. Ministry of Environment.

QH106.2.B74B56 2007 363.739'4630971182 C2007-960135-9

Citation: Perrin, C.J., S. Bennett, S. Linke, A.J. Downie, G. Tamblyn, B. Ells, I. Sharpe, and R.C. Bailey. 2007. Bioassessment of streams in north-central British Columbia using the reference condition approach. Report prepared by Limnotek Research and Development Inc. and B.C. Ministry of Environment for the B.C. Forest Science Program. 121p.

EXECUTIVE SUMMARY

Forest harvesting in B.C. is moving from prescriptive to outcome based management as development proceeds on various regulatory and product certification systems. To reach this goal, new tools are needed to measure the effectiveness of forest land management practices in sustaining the health of forest ecosystems, including the quality of water that drains those ecosystems. In this project, a bioassessment tool known as the Reference Condition Approach (RCA) was developed for north-central British Columbia. In the RCA, benthic invertebrates and habitat descriptors from a large number of reference sites are used to build a predictive model that allows comparison of a test site with an appropriate reference condition. If the test site falls within the range of natural variability found at reference sites, the site is considered to be not stressed. If the site falls outside of the range natural variability found at the reference sites, the site is considered to be stressed in some way. The greatest value of the RCA is in providing an effective environmental screening tool for land and water management activities.

Using benthic invertebrate and habitat descriptions from three years of sampling, two different RCA models were built and compared for potential use in bioassessment in northern British Columbia. The study area extended from the west coast, mainly between Stewart and Kemano, across the Coast Mountains and the Interior Plateau to sites east of Prince George. Sites that were not affected by anthropogenic disturbance were selected for sampling and were called ***reference sites***. Habitat variables that were not affected by site disturbance were measured at each site (called ***natural gradient variables***) along with collections of invertebrates using kick net methods. Additional sites were sampled for testing and were called ***test sites***. Variables that may be influenced by anthropogenic disturbance were also measured and were called ***stressor gradient variables***. A standardized method including quality assurance and quality control for enumeration of the invertebrates was followed. A total of 256 complete observations were compiled from all years of sampling. Of this total, 86 observations were from reference sites and 170 were from test sites.

The two models were:

1. The Benthic Assessment of Sediment (BEAST) that followed protocols used in earlier development of Fraser Basin and Georgia Basin RCA models, which followed from its original application in the Great Lakes.
2. The Skeena River Assessment System (SkeenRIVAS), which followed Australian and UK protocols for RCA model development and site testing.

Development of the models and site testing followed four common steps after sampling in the field:

1. Groups of reference sites based on similarity of biological composition between samples from these sites were defined using clustering and ordination techniques,
2. Natural gradient habitat variables that best explained dissimilarities between the biological groups were selected using discriminant function analysis (DFA) and called predictor variables,

3. The predictor variables were used in another DFA to predict the probability that a test site belonged to each sample group.
4. The deviation between the composition of biota observed at a test site and the composition of biota expected at the test site if it was similar to that in a reference condition found in one (BEAST) or more (SkeenRIVAS) sample groups was measured to determine site status. The assignment of a test site to sample groups was based on probabilities of group membership defined in step 3.

Differences between the models mainly lay in different procedures used for sample grouping (step 1 above) and different procedures for testing a site against the reference condition (step 4 above).

The SkeenRIVAS model had 11 predictor variables and the BEAST model had 5. Both models were highly significant and passed a series of statistical tests to show they were acceptable for site testing.

Several criteria were examined to compare accuracy and precision of model performance. Both models had acceptable error in classifying sites to sample groups. The number of predictor variables and ease of measurement of those variables for routine use was acceptable. The distribution of test results among stress categories was similar for both models, which indicated adequate precision. There was a good rate of agreement between models in defining site status, again providing confidence in precision of both models. The models disagreed on 10% of all site tests. This difference was potentially related to small sample size in some sample groups in both models. In lieu of an evaluation of model accuracy with a known stressor gradient, agreement of test results between the models and known condition of selected sites was used to indicate model accuracy. Five out of eight stressed site assessments were accurate, and three were not accurate, one of which was related to sampling error. The models were completely accurate in predicting the presence of a reference condition at known reference sites that were not used in building the models. Both BEAST and SkeenRIVAS showed high precision by predicting the same condition at sites from which replicate samples were collected.

These comparisons between the two models suggested that either model would be a good choice for routine use. Since the website that will host the selected model (called CABIN (*http://cabin.cciw.ca/application/welcome.asp?Lang=en-ca*) and is managed by Environment Canada) is not set up to accept SkeenRIVAS, BEAST is recommended for immediate use in the Skeena Region. Results of the comparisons highlight the advantages of using both models. We therefore recommend that CABIN be modified to include SkeenRIVAS as well as BEAST for future use.

Extension activities from this project involved hosting of workshops and conferences, preparation of progress reports, training sessions, and presentations to industry, government, and public interest groups. These activities were invaluable in promoting communications among RCA system developers, researchers, and prospective users. Publication of findings from this project is ongoing and is expected to further extend communication, evaluation at senior scientific levels, and practical application of the RCA approach in British Columbia.

Four examples, using data from this project, were outlined to show how RCA can be applied in British Columbia.

The first was site testing as part of land and resource management plans (LRMP). The example focussed on the Morice Timber Supply Area (TSA), located in northwestern British Columbia. RCA test results were highly correlated with land use and were consistent with present knowledge. This example showed the RCA can provide a quick and simple means to assess and report the status of aquatic ecosystems at a landscape or watershed scale, which is fundamental to monitoring the effectiveness of LRMPs.

The second example was from the Toboggan Creek watershed where site specific water quality objectives are being developed because of highly valued aquatic resources in that area. Test results using BEAST, SkeenRIVAS, and an Index of Biotic Integrity (IBI) cumulatively showed variation in site status in the watershed, but most sites were in reference condition. In cases where site specific water quality objectives are being developed, this example showed that RCA can be a powerful line of evidence in supporting decisions on managing sensitive watersheds.

A third example was application of RCA to meet the needs of assessment for forest licencee certification. Certification is a tool that can be used by companies to support forest stewardship plans (FSP), to illustrate compliance with FSP strategies, and to demonstrate to product buyers and the public that their products are the result of sustainable practices. Operations by Pacific Inland Resources (PIR) were selected. Combined site assessments from BEAST, SkeenRIVAS, and an IBI showed that 84% of all sites where PIR operated were in reference condition and that 3 of 19 sites required attention pending further confirmation of site status. PIR has expressed interest in the continued use of the RCA as a tool to monitor its harvesting activities. The method can be attractive in this application because of relatively low cost, ease of testing, and because it is scalable to small and large forest management areas.

The final example was application of RCA to point source discharges. The closed Equity Mine was selected where there is a history of acid rock drainage. The mine could replace an existing monitoring design that is based on comparison of control and treatment sites over time with a layout using RCA testing and achieve objectives of its environmental effects monitoring (EEM) requirements. A challenge of relatively few control sites at the mine is presently overcome with reference to site-specific experiments that assist with the interpretation of monitoring data and the use of multiple lines of evidence to support conclusions. Results have shown that the mine is doing a good job in protecting downstream condition by collecting acid drainage and treating it before discharge. The RCA has built-in control data in the model, thus reducing the concern about the lack of reference information in routine monitoring. Another advantage would be a reduction in the ongoing monitoring cost. A rough estimate indicated that the cost of benthic invertebrate monitoring that occurs once every 4 years as part of EEM protocols, may be reduced by approximately 40% if the mine changed to RCA procedures. A disadvantage, however, would be loss of historical data in ongoing analyses because of the change in methods. Regardless of whether this change happens at Equity, where RCA models are in place, the RCA is an option that mines can consider in EEM plans, as defined in guidance documents prepared by Environment Canada.

There are five recommendations in addition to the main recommendation that the BEAST model be uploaded to CABIN for immediate use:

1. The BEAST model that was developed in this project should be frequently updated to ensure it remains valid and to maintain its accuracy and precision.
2. Present weak coverage of reference sites between Burns Lake and Prince George should be corrected with additional effort going into adding reference sites in that area. This area is important with respect to range and forest management, particularly with regards to tracking environmental and land management implications of the Mountain Pine Beetle infestation.
3. It is recommended that new reference sites be added to provide a validation data set for testing model accuracy and precision. This testing was completed on a limited scale in this project but it should be expanded as models are updated in future years.
4. We recommend that new reference sites be sampled that correspond to the sample groups that presently have a small sample size.
5. Finally, we recommend that researchers continue to be engaged in future development of the RCA in British Columbia. This involvement will ensure that the models remain “cutting edge” and reflect the current state of knowledge internationally. This approach will place British Columbia at the forefront in the rapidly changing science of bioassessment and its application to sustainable land management.

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, BC Timber Sales, Babine Watershed Monitoring 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. We would like to thank Angie Gagnon (Ministry of Environment) for her help with administrative aspects of project management. GIS support was provided by Matt Sakals and Candice Wilson in the first two project years before the sixth author of this report assumed those activities in the third year. Many people contributed to the selection of sampling sites throughout the large sampling region. They included Joe Trainer, Dave Rolston, Jason Bennett, John Rex, Scott Scholefield, John DeGagne, Kathleen Hebb, Leslie McKinley, Jim Hesse, Rob MacDonald, Andy MacDonald, Mark Connor, Brian Fuhr, Chris Broster, Mark Connor, and Anne Hetherington. Personnel from Canfor, Houston Forest Products, Babine Forest Products, Pacific Inland Resources, and Fraser Lake Sawmills were most helpful in suggesting sites and providing essential maps and access information. 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.

The authors were assisted in the field by Evangaline Alfred, Diane Barbetti, Emily Bulmer, George Butcher, Camille Callison, Peter Christensen ,Catherine Ducheminski, Deb Epps, Angie Gagnon, Jennifer Guay, Jim Hesse, Aaron Hill, James Jacklin, Kevin Koch, Cory Koenig, Julia Kokelj, Bryan Last, Derek Louie, Jeannette Lough, Jack Love, Kerrith McKay, Dave McEwan, Carrie Morita, Dean Peard, Kevin Rieberger, Marge Sidney, Sean Sharpe, Mark Connor, 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 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 Danielle Francis, Angie Gagnon, Lindsay Graff and Jeannette Lough. Lorne McIntosh and Jessica Chaplin helped contact participants and develop materials for implementation of the extension plan.

We especially extend our thanks to Stephanie Sylvestre of Environment Canada for ongoing scientific dialogue regarding the reference condition approach to biomonitoring and related modeling. In addition, we thank Stephanie for providing site information and funding for reference sites sampled in North Tweedsmuir Park.

TABLE OF CONTENTS

Page

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

Forest harvesting in B.C. is moving from prescriptive to outcome based management as development proceeds on various regulatory and product certification systems. To reach this goal, new tools are needed to measure the effectiveness of forest land management practices in sustaining the health of forest ecosystems, including the quality of water that drains those ecosystems. Instead of checking whether prescribed methods of stream protection are being properly implemented, the state of the biota itself can be used as an indicator of whether protective measures have been adequate.

Benthic invertebrates are good indicators of water quality (Rosenberg and Resh 1993) and ecosystem health (Reice and Wohlenberg 1993, Norris and Hawkins 2000). In forested drainages, aquatic ecosystem health can be defined in terms of attributes of benthic communities (e.g. abundance and composition) that are part of the river food web. Because of continuous exposure to water flow, benthic biota can provide an integrated record of physical and chemical environmental quality. They are ubiquitous, largely sedentary, and there are large numbers of species that can provide an integrated measure of response to stress. Their characteristics allow effective spatial and temporal analyses of disturbance among streams, within reaches of streams, and between streams over wide geographic areas (Bailey et al. 2004). The invertebrates are a major food supply for fish, particularly salmonids in northern British Columbia streams, and thus provide an indication of food availability for fish populations through time and space. Benthic invertebrates respond rapidly to change in environmental conditions. They, along with epilithic algae, are often the first organisms of an aquatic community to respond to environmental stress and they are usually the first to recover from it. The result is that monitoring of benthic invertebrates can provide a clear indication of ecosystem health and change in the quality of the water they inhabit.

Over large regional scales, bioassessment procedures based on multimetric indices of benthic invertebrate composition and abundance have been used to monitor water quality, particularly in the United States (Karr 1981, Karr and Chu 1999, Barbour et al. 1999). A multimetric index is the combination of a number of individual metrics (e.g. number of mayflies, stoneflies, caddisflies (EPT), percent chironomids) to form a single score. It is developed from the biota found at a set of sites thought to be on a gradient from no disturbance to highly disturbed, and then applied to sites with an unknown degree of disturbance (e.g. Kearns and Karr 1994). The Index of Biotic Integrity (IBI) that was developed by Karr (1981) and Karr and Chu (1999) is perhaps the best known and most widely used of the many multimetric bioassessment methods. Because the IBI requires development of a score from observations along a gradient from undisturbed to very disturbed sites, it requires calibration throughout the region to which the IBI assessment is being applied. While best known in the United States, a multimetric IBI based on the methods developed by Karr (1981) was successfully developed for the

Skeena region of British Columbia (Rysavy 2000, Bennett and Rysavy 2003, Croft 2004). This work provided an initial step in the process of developing a forest ecosystem sustainability indicator system that is now part of a performance based toolbox to assess impacts on aquatic ecosystems from forest harvesting activities in the Skeena region.

Another biological assessment approach known as the Reference Condition Approach (RCA) is based on characterization of undisturbed reference sites in a wide variety of environments, relating the natural environment of these sites to their biota, and then predicting the biota that would be found at a new, “test” site if it was in reference condition. The deviation between what is observed at the test site and this prediction is a measure of how disturbed the site is, and it is a measure of the nature of the disturbance (Bailey et al. 2004). It is used as a standard procedure for testing site quality in many countries, particularly in the UK (Wright et al. 2000), Australia (Parsons and Norris 1996), and Canada (Bailey et al. 2004, Sylvestre et al. 2005, Reynoldson et al. 1997, Reynoldson et al. 2001).

Both the IBI and the RCA can be considered screening tools for water quality assessment within a large region. Both approaches are based on the concept of comparison to a reference condition and can be considered complimentary (Reynoldson et al. 1997). The IBI is based on the sum of a selected number of biological metrics that are found to be sensitive to a known gradient of water quality or ecosystem health within a region. The RCA combines the ideas of multivariate modeling of entire biological communities (Wright et al. 2000) with the concept of comparison to a reference condition. RCA is more comprehensive because it includes complete communities rather than parts of communities in a final predictive model. While the RCA is more computationally complex than IBI, the computations can easily be run on a web site wherein calculations run behind the scenes, making site testing a very rapid and simple process. The website called CABIN (Canadian Aquatic Biomonitoring Network; *http://cabin.cciw.ca/cabin/asp/english/welcome.asp*) is the portal where testing of sites in Canada using the RCA can be run. CABIN is a database management system capable of archiving biological, GIS derived basin characterization information and habitat data for all reference and test sites, It houses and enables use of both RCA and B-IBI models to calculate stream condition scores. It includes standard sets of protocols and methods for all phases of data collection and processing, including standard field sheets and laboratory forms, and will soon contain on-line training tools.

Over the past three years, various RCA models have been under development for the Skeena Region of British Columbia. In this final year of the project, a model has been selected for routine site quality testing in forested ecosystems of northern British Columbia and it will be uploaded to the CABIN website. This bioassessment tool is particularly intended for use in Forest Stewardship Plans, Sustainable Forest Management Plans, Land and Resource Management Plans, state of forest reporting, as well as in Forest Product Certification systems. It is intended for environmental screening of site quality. In the future it can be set as a RISC Standard Method (*http://ilmbwww.gov.bc.ca/risc/*) and combined with other indicators of aquatic ecosystem sustainability (fish and fish habitat), it can serve as a major part of monitoring and assessment procedures to determine the effectiveness of forest and other land management practices in protecting valued water resources. We anticipate that progress made in this project will lead to Province-wide application of the methods.

This report outlines results from three years of RCA model development for the Skeena Region. It includes:

- Development of standard methods for collecting and analysing data for use in the RCA for biomonitoring in the Skeena region,
- Results of statistical analyses used in the development of alternative RCA models for the Skeena Region,
- A comparison of modeling outcomes and a recommendation of a final model for uploading to the CABIN website for routine use,
- Case examples to show applicability of the RCA approach to assessing site quality in northern British Columbia.
- Recommendations for improvements to field data collection forms and application and use of the RCA model on the CABIN website.

Major funders contributing RCA model development included the Forest Science Project's envelope of the B.C. Forest Investment Account, West Fraser Timber Company Ltd., the Morice and Lakes Innovative Forest Practices Agreement, BC Timber Sales, and the B.C. Ministry of Environment (MOE).

2 OVERVIEW OF RCA MODELING

Layouts and approaches for testing an effect or degree of disturbance in surface waters can involve multiple lines of evidence using a suite of univariate statistical tests. A common layout involves sampling at reference and potentially impacted sites before and after start-up of a disturbance or discharge, thus facilitating a layout known as a before-after/control-impact (BACI) design (Stewart-Oaten et al. 1986). It can be used to test for effects of a known point or non-point source discharge on ecological endpoints of interest (Bowman and Somers 2005) using analysis of variance (ANOVA). There are several variations of the BACI design ranging from single control and treatment sites from which replicates are considered to be samples collected through time, to stronger layouts involving multiple control sites that are analysed by asymmetric analysis of variance (Underwood 1994). All of these designs involve univariate analytical approaches. They require the "correct" selection of an endpoint among many taxa. To avoid missing taxa of potential importance, total abundance or biomass may be analysed or a series of metrics (e.g. abundance of the combination of mayflies, stoneflies, caddisflies; abundance of chironomids; etc.) may be selected for independent analysis. Evidence from these multiple tests and other observations can then be combined to determine the effect of specific disturbances and to examine cause – effect pathways that are critical for supporting water management decisions. Although the time required for field testing can be shortened and additional control can be applied to the tests by running experiments, for example at mesocosm scales (e.g. Perrin and Richardson 1997), most of these approaches require long time periods before a definitive description of water quality and cause – effect pathways may be found. There must be enough foresight of the putative impact to collect the 'before' data at both the test and suitable control sites. In this regard, the stressor or impact must be known before it actually occurs. This requirement eliminates BACI designs from assessments that are required to test for effects of "accidents" or surprising events (Bowman and Somers 2005). The designs can be expensive and impractical to complete on a large regional scale. In addition, basic assumptions of the statistical analyses may be violated, perhaps due to insufficient funding to collect enough samples at any site, difficult logistics that prevent repeated sampling, or other factors that constrain an ideal layout of sample collection (Bailey et al 2004). Hence, these approaches are most powerful at the site specific level and are best suited to definitive experimentation rather than providing evidence of water quality condition at a regional level.

An alternative is the reference condition approach, commonly known as RCA (Bailey et al. 2004). Reference condition describes a suite of attributes found at sites having little or no exposure to stressors caused by land use and other human activities. The premise behind the RCA is to sample a large number of sites in reference condition and use the covariance between biological and environmental descriptors to build a predictive model that allows comparison of a test site with an appropriate reference condition. Invertebrate communities are naturally variable and continually changing. The RCA is a method for ensuring an appropriate description of this variability and the power to detect a change from a reference condition over the 'noise' of natural variation (Green 1999). Similar to an ANOVA, a test site is compared to an appropriate group of reference sites and the test site is determined to be in reference condition if the biological community is similar to that of the reference sites (falls within the range of error variance, Figure 1). If the test site falls outside the range of natural variability found at reference sites, the null hypothesis that the test site is the same as the reference group is rejected (Bailey et al 2004). Hereby the pool of all sites, as opposed to repeat samples of the same site that can be the basis of a BACI design, serve as replicates (Reynoldson et al. 1997).

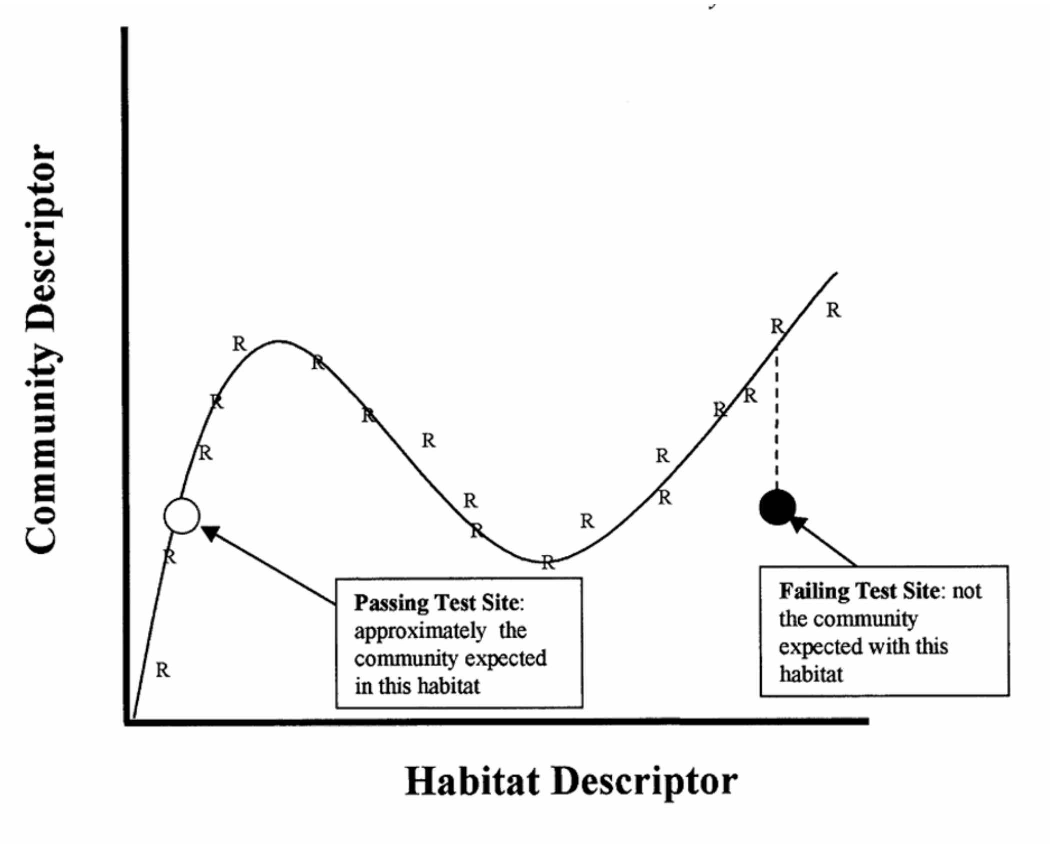

Figure 1. Conceptual view of site testing using the Reference Condition Approach (from Bailey et al. (1998)). A site passes if its biological community is similar to that expected along a continuum described by community and habitat descriptors. A site fails if its biological community is dissimilar to that expected.

While the RCA might be used as part of an impact assessment study, its greatest value is in providing an effective environmental screening tool for routine use. Once an RCA model is established for a region of interest, site testing can be completed quickly and inexpensively, thus meeting schedules for making environmental decisions. An example of a decision process is shown in Figure 2. In a region where environmental values of surface waters have been identified, site testing using the RCA can be run on a defined schedule. If a site passes the RCA test, it can be queued for follow-up testing at a later date. If the site fails the test, further investigation of available data or experimentation can be conducted to identify the specific cause of the failure, if not obvious, thereby providing technical criteria to support management actions to improve site quality. Principal component analysis followed by development of regression models using existing data is one approach that can be used for follow-up interpretation. Other approaches may include site-specific experimentation to test for cause and effect (e.g. Perrin et al. 1992, Richardson and Perrin 1994, Perrin and Richardson 1997) or detailed monitoring that supports modeling for management decision-making (Perrin 2006). The RCA model can then be used to repeatedly and inexpensively test the site to determine if improvements have been realized or to track change in site status over time. Used in this way, the RCA is an effective environmental screening tool that is scientifically defensible and can be used to quickly assess a site at low cost.

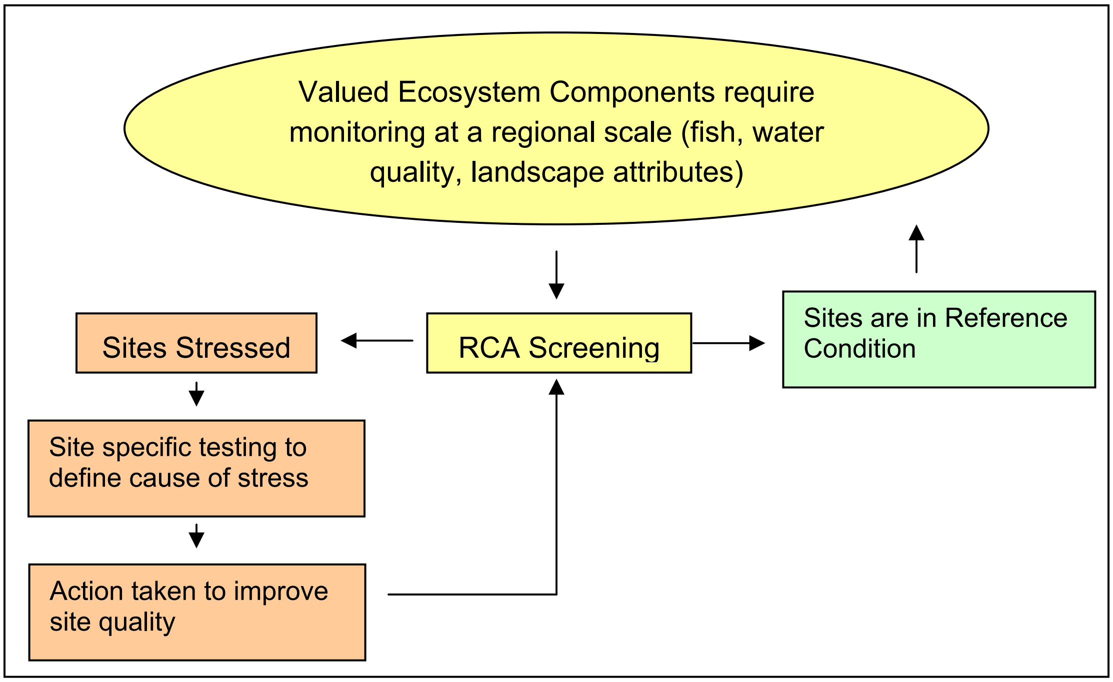

Figure 2. Schematic illustration of where RCA fits in environmental decision-making.

3 STUDY AREA DESCRIPTION

The study included sampling of biological and habitat attributes of streams in north central British Columbia over three years (Figure 3). The sites were selected according to methods described in Section 4.1. They extended from 121.52ºW in drainages east of Prince George to the west coast over a range from Stewart, south to exposed fjords west of Kitimat and Kemano (near 130ºW). The latitudinal range (58.82ºN to 52.97ºN) extended from Atlin in the north to remote drainages of the Kitlope Heritage Conservancy and Tweedsmuir Park in the south. Ecoprovinces (Perrin and Blyth 1998) included in the study area were mainly the Coast and Mountains to the west, the Sub-Boreal Interior, and northern ecoregions of the Central Interior (Bulkley Basin, Lower Nechako, Nechako Plateau). A few sampling sites were in the Northern Boreal Mountains and the northern tip of the Southern Interior Mountains.

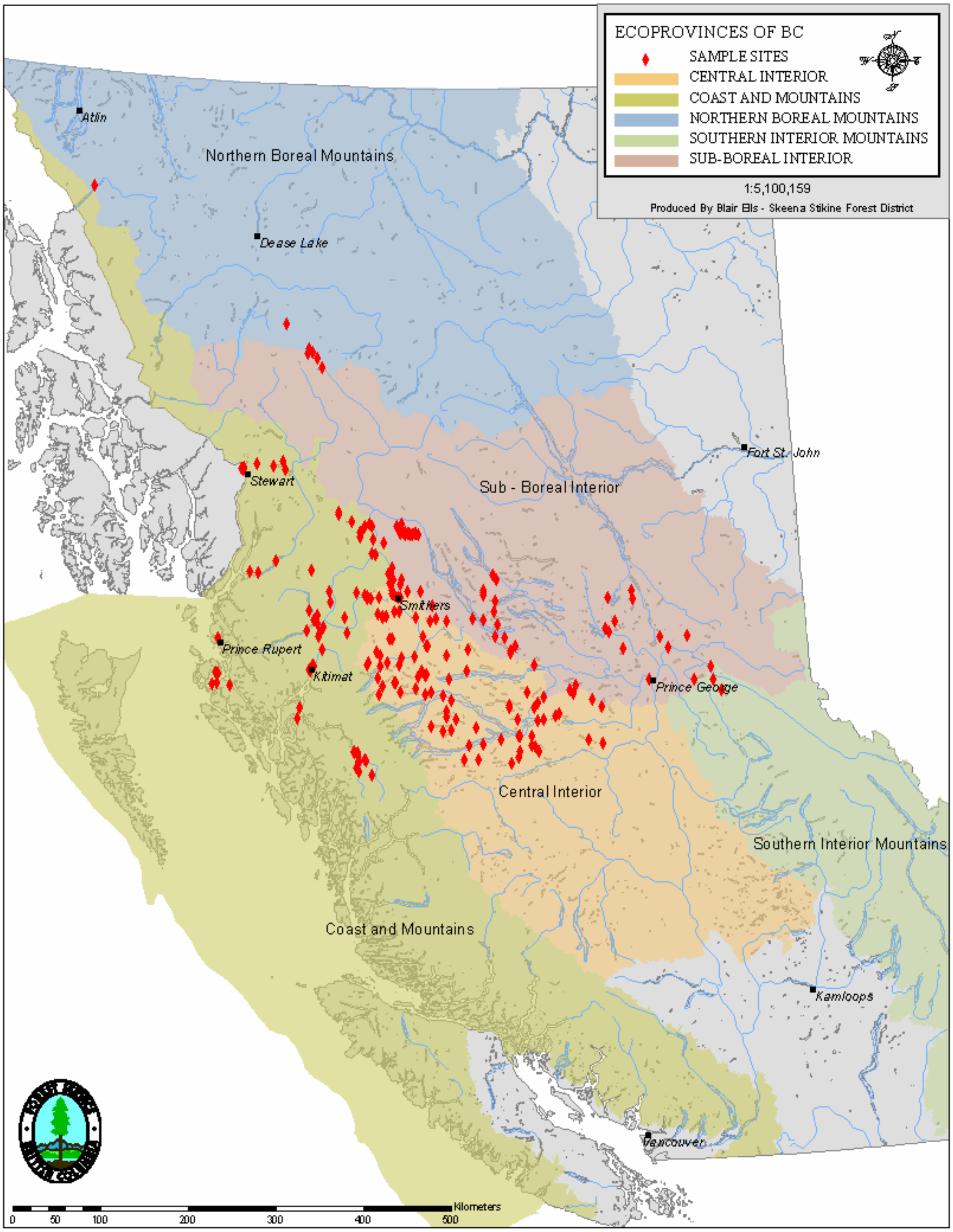

Figure 3. Distribution of sampling sites among the Ecoprovinces of British Columbia.

Characteristics of the main ecoprovinces in which sampling sites were established are as follows. The information is mostly from Perrin and Blyth (1998) and Demarchi et al. (1990).

The Coast and Mountains Ecoprovince extends the full length of the British Columbia coastline and is the largest and most diverse of all ecoprovinces. Sampling sites in this ecoprovince were located on the windward side of the Coast Mountains (Figure 3). The main feature of the ecoprovince is a north-south continuum of large rugged mountain ranges, high amounts of precipitation, and large elevational variation in aquatic ecosystems. Glacial scouring modified massive granitic intrusions that formed from heating with the docking of superterranes. As glaciation receded, massive moraines were left in valleys and outwash areas at the ocean interface. Subsequent drainage formed high densities of small streams, and small to large sized lakes which can have steep littoral zones in fjord-like basins. Glaciation remains typical among highest peaks. Major northern rivers cut through Coast Mountains to the west emptying into deep fjords that cut several hundred km into the exposed coastline. Small glacial pothole lakes are found at moderate elevations, and at the lowest elevations, small to moderate sized lakes are abundant as remnants of glacially-formed depressions. The mild coastal climate which dominates the Coast and Mountains Ecoprovince favours warm monomictic lakes and associated streams. Ice cover generally does not occur or it is transient. Streams along the coast have well developed riparian communities. Hemlock and amabilis fir forests are prevalent in the ecoprovince with sitka spruce and western hemlock common in the north. Floodplains have forests of spruce, black cottonwood, and red alder.

In the Northern Boreal Mountains only Takwahona Creek, located south of Atlin and a group of sites over a latitudinal range of 57.12ºN to 57.56ºN along the southern boundary of the ecoprovince were sampled (Figure 3). This ecoprovince was largely north of the main area of interest. The ecoprovince is characterized by mountain ranges separated by wide valleys. Major physiographic features from west to east are the Alsek Ranges, Cassiar Mountains, Liard Ranges, the Northern Rocky Mountain Trench, the Muskwa Ranges, and the Liard Gorge through which the Liard River flows between the Rocky and MacKenzie Mountains. In many broad valleys of the ecoprovince, particularly in the Stikine Plateau, there are two treelines. One separates extensive alpine tundra from subalpine fir and white spruce and a second occurs in lower elevation valleys where cold air drainage tends to keep river valleys cool enough to limit forest communities and favour wetlands with willows and moss cover. Small lakes and slow meandering streams are typical in these valleys.

The Sub-Boreal Interior Ecoprovince is located in the north-central part of B.C., east of the coast mountains, west of the Alberta plains, south of northern boreal plateaus, and north of the Central Interior plateau. This ecoprovince contained the eastern extent of sampling sites. It is characterized by large lakes, reservoirs, and rivers that are important for power production, industrial water supplies, fish production, and transportation corridors. Mountains and flat plateau are found in this ecoprovince. In the northwest are the Omineca and Skeena Mountains, which originated as massive granitic intrusions. Drainage is to the east in the Omineka Mountains but it is to the south and west in the Skeena Mountains. Further to the south in the Takla/Manson Plateau, Babine Uplands and Upper Fraser Ecoregions, the bedrock is sedimentary with some volcanic intrusions making the parent materials highly erodable. These areas are flat or gently rolling with abundant small lakes and wetlands that have formed in surface depressions where drainage is generally poor. Deep incisions are formed by the lower reaches of the Nechako River near Prince George due to fluvial erosion which has created long ridges of low relief that follow the river channel. Many slopes from these ridges consist of loose gravel and sand that is constantly being eroded by precipitation and freezing and thawing, thus contributing to a sand and small gravel substratum in many reaches of the Nechako River. Lodgepole pine forests are extensive in this ecoprovince with much of the area affected by the Mountain pine beetle infestation (*http://www.for.gov.bc.ca/hfp/mountain_pine_beetle/*). Engelmann spruce and subalpine fir are dominant in the subalpine. In areas where climax pine forests are not established, trembling aspen and birch form extensive deciduous cover. Wetlands are extensive in lower relief areas. Sphagnum bogs are common at low elevations with black spruce, Labrador tea, and sedges being the main vegetation.

The main feature of the Central Interior Ecoprovince is a wide plateau spread between the Coast Mountains and ranges of the Southern Interior Mountains Ecoprovince. Sampling sites were mainly in northern ecoregions including the Bulkley Basin, Lower Nechako, and Nechako Plateau. A major drainage is the the Bulkley River that flows north emptying into the Skeena River at Hazelton, at the extreme north end of the ecoprovince. The Nechako Plateau contains the Nechako Reservoir, Eutsuk Lake and the associated local inflows. While most of the ecoregion is characterized by rolling hills and moderate relief, extreme southwestern areas include higher relief of the east slope of the Coast Mountains. A large water storage reservoir system in the Central Interior is the Nechako Reservoir, which is 160 km long and includes 95,000 ha of lakes, rivers and submerged forest that was not logged prior to flooding in 1954. Lodgepole pine forests are in this ecoprovince with much of the area severely affected by the Mountain pine beetle infestation (*http://www.for.gov.bc.ca/hfp/mountain_pine_beetle/*). In northern areas, white spruce is the climax species, occurring with subalpine fir. Stands of trembling aspen and birch are common. Wetlands are common although they are mostly covered with shrubs and trees.

4 METHODS

4.1 Sample Site Selection

Site selection began as a theoretical mapping exercise each year that was fine-tuned in the field. Potential sample sites and areas of interest were identified in consultation with people who were familiar with the study area. GIS analysis was used to confirm land use activity and field visits were used to verify the GIS exercise and confirm suitability for sampling.

Sites covered a broad geographic area with many types and degrees of land use. Reference sites were expected to be naturally variable according to natural landscape attributes. This variability was captured by sampling streams of all sizes, with a wide range of altitudes and physical and chemical conditions, over a large geographic area.

Prior to the field season each year, a list of high priority areas and a short list of potential reference and test sites was created, largely using local knowledge from people who have worked in the region for many years. We met with staff from various provincial and First Nations government programs, including the Integrated Land Management Bureau, Ministry of Environment (Environmental Protection Division, Water Stewardship Division and Environmental Stewardship Division - Fish and Wildlife Branch, Ecosystems Branch and Parks and Protected Areas Branch), Ministry of Forests and Range (District Office field staff and Forest and Range Evaluation Program staff) and BC Timber Sales, as well as representatives from the forest industry including CANFOR, West Fraser Timber, and Hampton and Affiliates. In many cases, maps were provided with high priority streams and areas of interest highlighted. The exact location of sites within the areas of interest or along high priority streams were determined by field crews on the day of sampling based on access and local conditions.

In some areas we were able to conduct a reconnaissance visit of potential sites prior to the sampling season; however, in most cases, time constraints did not allow this to happen so multiple sites were identified in areas where the suitability of individual streams (or reaches) was uncertain. The field crews were then able to select the best candidate site when they were in the field.

4.2 Habitat Variables Selection

Two groups of variables describing habitat attributes were measured. One was a group of measurements made in the field (mainly physical and chemical variables) and the other was a group of variables that were compiled or calculated from GIS databases (mainly watershed characteristics and geomorphic variables).

Those variables that did not vary with anthropogenic disturbance (Reynoldson et al. 2001, Sloane and Norris 2003) were called ***natural gradient variables***. They mainly described geomorphological and other physical attributes including:

- Attributes of stream morphology, gradient, and the drainage basin at the sampling site (e.g. bankfull width, wetted width, channel depth, percent of different flow habitats (pools, glide, riffle, cascade), area of drainage basin upstream of the sampling site, elevation, relief, percent of the watershed area that is in the alpine, percent of avalanche chute area in the watershed, water temperature);
- Substrate characteristics including relative abundance of particle categories (e.g. sand, gravel, cobble);
- Water attributes including drainage density of streams, stream length, and percent of the drainage area comprised of wetlands, lakes, and ice;
- Characteristics of riparian vegetation development (e.g. grasses present or absent, over-stream cover, riparian species composition);
- Composition of riparian vegetation (e.g. barren, grass/herb, shrub, tree type);
- Parent material geology (e.g. presence/absence or proportion of intrusives, volcanics, sedimentary, metamorphic, and ultramafic rock); and
- Geographic location (e.g. latitude and longitude).

Some groups of well known variables were not included in this list. Nutrients were not included because anomalous discharges can modify growth of periphyton (Stockner and Shortreed 1978, Perrin et al. 1987, Bothwell 1989) and cause change in whole system production (Johnston et al. 1990, Deegan and Peterson 1992). Concentration of metals were not included because they can cause toxicity in stream biota (Campbell and Stokes 1985, Hickey and Clements 1998) while treatment of mine water discharge with lime (e.g. major cations) can reduce this toxicity (Perrin et al. 1992). Even basic electrochemical analytes including total dissolved solids/conductivity, alkalinity, pH, and dissolved oxygen were not included because they can be modified by anthropogenic disturbance or water treatment.

Stressor gradient variables were those that could be affected by human activity. A suite of these variables were measured at each of the sites but they were not directly used in developing the RCA models. These data were collected as a general protocol for possible future use in examining cause of site impairment at test sites, should one or more be stressed. This process of site assessment following RCA screening is described in Section 2.

Natural and stressor gradient variables included those used in development of the Environment Canada RCA (Sylvestre et al. 2005, Reynoldson et al. 2001), the Australian River Assessment System called AUSRIVAS (*http://ausrivas.canberra.edu.au/Geoassessment/Physchem/Man/Protocol/index.html*), and others considered potentially important in assessing sites in northern British Columbia.

In compiling the list, we considered logistics of being able to complete the measurements in reasonable time in the field and the benefits of mining information from GIS databases that is more cost effective than operation of field crews. We wanted a variable list and sampling protocol that would limit time for sampling and field measurements at any given site to an average of 1.5 to 2 hours. Including travel time by vehicle to stream access points, this goal would allow completion of approximately 3 sites per day per field crew working a standard 8 hour day.

Using a list of candidate natural gradient variables, a consensus-based exercise was used to identify redundancies and compile a final list for development of the RCA models (Table 1).

Table 1. List of candidate variables and those selected for RCA modeling.

Candidate variables	x indicates selected for modeling	Method of measurement (field, GIS, calculation from GIS)	Description
Site Name		n/a*	Site name
SITE CODE		n/a*	CABIN site code name
Year		n/a*	Year
SITE		n/a*	EMS code with year coding
EMS		n/a*	EMS site code
ref_test		n/a*	Status of site (reference or test)
Area (m^2)	x	GIS	Area of watershed - square meters
Pct_Wtlnd	x	GIS	Percent of wetland area in watershed
Pct_Lake	x	GIS	Percent of lake area in watershed
Pct_Rvr	x	GIS	Percent of river area in watershed
Rvr_Lngth	x	GIS	Length of rivers (approximation)
Pct_Ice	x	GIS	Percent of ice in watershed
Strdefin_Lgth	x	GIS	Definite (always flowing) stream length
SL_ratio	x	calculation	Ratio of definite stream length to total length
Tot_StrLgth		GIS	Total stream length
StrRiv_DrnDnsty	x	GIS	Drainage density of all rivers and streams
Fsh_DrnDnsty		GIS	Drainage density of fish bearing streams
Pct_Agecl67		GIS	Percent of forest age classes 6 and 7 (101-140 years old) in the watershed area
Pct_Agecl89		GIS	Percent of forest age classes 8 and 9 (>141 years) in the watershed area
Pct_old_growth		calculation	Percent of the watershed area comprised of old growth forest (sum of Pct_Agecl67 and Pct_Agecl89)
Geo_Class1		GIS	Largest geologic class
Pct_Geo1		GIS	Percent of largest area rock type (geology) within the watershed
Pct_sedimentary Geo1		calculation	Percent of watershed with sedimentary rock as the largest rock type
Pct_Intrusive Geo1		calculation	Percent of watershed with intrusive rock as the largest

Candidate variables	**x indicates selected for modeling**	**Method of measurement (field, GIS, calculation from GIS)**	**Description**
			rock type
Pct_Volcanic Geo1		calculation	Percent of watershed with volcanic rock as the largest rock type
Geo_Class2		GIS	Second largest geologic class
Pct_Geo2		GIS	Percent of second largest area rock type (geology) within the watershed
Pct_sedimentary Geo2		calculation	Percent of watershed with sedimentary rock as the second largest rock type
Pct_Intrusive Geo2		calculation	Percent of watershed with intrusive rock as the second largest rock type
Pct_Volcanic Geo2		calculation	Percent of watershed with volcanic rock as the second largest rock type
Geo_Class3		GIS	Third largest geologic class
Pct_Geo3		GIS	Percent of third largest area rock type (geology) within the watershed
Pct_sedimentary Geo3		calculation	Percent of watershed with sedimentary rock as the third largest rock type
Pct_Intrusive Geo3		calculation	Percent of watershed with intrusive rock as the thrid largest rock type
Pct_Volcanic Geo3		calculation	Percent of watershed with volcanic rock as the third largest rock type
Pct_tot_sedimentary	**x**	calculation	Percent of watershed with sedimentary rock
Pct_tot_intrusive	**x**	calculation	Percent of watershed with intrusive rock
Pct_tot_volcanic	**x**	calculation	Percent of watershed with volcanic rock
total rocks		calculation	Check on addition of percentages
Pct_other _rocks	**x**	calculation	Percent of watershed with rock other than sedimentary, intrusive, volcanic
Pct_Alpine	**x**	GIS	Percent of alpine area in watersheds
Pct_Ava	**x**	GIS	Percent of avalanche chute area in watersheds
Latitude	**x**	Field	Latitude
Longitude	**x**	Field	Longitude
Min_Elev	**x**	GIS	Minimum elevation of the watershed measured from the sampling site. This measure is the elevation of the sampling site
Relief	**x**	GIS	Maximum minus minimum elevation
Pct_SC4	**x**	GIS	Percent of watershed area with slope class 4
Pct_SC5	**x**	GIS	Percent of watershed area with slope class 5
WT	**x**	Field	Water temperature at time of sampling
%Gradient	**x**	Field	Gradient measured in percent
%Pools	**x**	Field	Percent pools
%glides	**x**	Field	Percent glides
%Riffles	**x**	Field	Percent riffle
%cascades	**x**	Field	Percent cascade
BWAve	**x**	Field	Average bankfull width (from 3 field measurements)
Substrate_%gravel		Field	Visual estimate of percent gravel on stream substrate
substrate_%pebble		Field	Visual estimate of percent pebble on stream substrate

Candidate variables	**x indicates selected for modeling**	**Method of measurement (field, GIS, calculation from GIS)**	**Description**
substrate_%cobble		Field	Visual estimate of percent cobble on stream substrate
substrate_%boulder		Field	Visual estimate of percent boulder on stream substrate
substrate_%bedrock		Field	Visual estimate of percent bedrock on stream substrate
dom_substrate	x	calculation	Dominant substrate particle size by percent composition
			1=gravel
			2=pebble
			3=cobble
			4=boulder
			5=bedrock
			If 2 substrate types had the same percentage composition, the larger size class was selected
D50		Field	Median pebble diameter (from pebble count, n=100)
%unveg		Field	Percent of riparian that is unvegetated
%grass_herb		Field	Percent of riparian that is grass/herb
%shrub		Field	Percent of riparian that is shrub
%dec		Field	Percent of riparian that is deciduous trees
%con		Field	Percent of riparian that is coniferous trees
dominant Rveg	x	Calculation	Dominant riparian vegetation
			1=barren dominant
			2=grass/herb dominant
			3=shrub dominant
			4=deciduous tree dominant
			5=conifer dominant
			If 2 vegetation types had the same percentage composition, the higher code was selected

*these metadata were assigned to code sample name, time and location of collection and to define whether the sample was a reference or test site.

4.3 Field Logistics

Fieldwork was performed during the late summer low flow period between mid-August and mid-September in each year. This time of year is standard for RCA sampling mainly because it is the easiest time to collect samples. Benthic invertebrates are abundant at that time. The low stream flows improve safety and wadeability, and leaf litter that can hinder sample processing in the lab has not yet accumulated in the stream.

Prior to sampling, a field team of 6-12 individuals (including Ministry staff, contractors, volunteers and representatives from partner agencies) completed a half day training session on a local stream. Experienced “leaders” were selected and sampling crews of two to four members were assigned. Up to three crews were required to sample the various locations in the region in any one year.

Schedules were developed and modified throughout the sampling season, with consideration for travel requirements and sample shipping logistics. Smithers served as the staging point for the sampling crews, but in many cases the crews departed for 3-4 days at a time, staying overnight in communities near their sampling area. While most sites were accessed by truck using logging and mining roads, some of the more remote sites such as those in North Tweedsmuir Park and Entiako Park were accessed by helicopter to improve cost efficiency. Many of the North Coast sites were accessed by boat, and required overnight stays in Oona River and Kitlope Park.

In many cases, time constraints did not allow for reconnaissance visits to potential sites, so multiple sites were identified in areas where the suitability of streams was uncertain. When visiting reference streams (defined in Section 4.8), samples were collected upstream of all resource developments (including roads and road crossings, cutblocks, etc.). Test sites were chosen at places within the influence of land use activities.

4.4 Field Sampling

Field sampling methods followed the British Columbia Resource Inventory Standards Committee documents (*http://ilmbwww.gov.bc.ca/risc/pubs/aquatic/index.htm*) and the CABIN field and laboratory methods (Reynoldson et al 2003). Protocols and required measurements that were completed in the field were laid out on a data sheet (Appendix A) that was filled out at each site, thus standardizing the data collection process. Based on experience over the three years of sampling, a revised field sheet that is recommended for future use is listed in Appendix B.

Upon arriving at a potential site, a visual assessment was completed to determine its suitability for sampling. Of primary consideration was the requirement for adequate flow and the presence of riffle habitat. Ideal riffle habitats had fast-flowing water (0.2-0.8 m/s), 10 to 30 cm deep. Preference was given to wadeable sites; however, larger streams and rivers were sampled, with measurements and samples confined to the wadeable areas near the river margins. If continuous riffle habitat was not available, the stream was still assessed, with the benthic invertebrate sample collected from multiple riffles, moving in an upstream direction.

If a site was suitable for sampling based on the presence of riffle habitat and adequate flow, tasks were assigned and the crew members began sample collections and measurements. Water samples were collected before anyone entered the stream to avoid sample contamination caused by disturbance of the substrata. One crew member then collected the benthic invertebrate sample while another timed the sample collection. The bank full width was measured using survey tape for widths <3 m or a rangefinder if the stream width was >3 m. An assessment reach was defined as 6 times the bankfull width. This length was considered necessary to include enough habitat area to satisfy all measurements and observations noted on the field form. All crew members worked together to record notes and observations and complete the instream measurements. The average sampling time at each site was 1-2 hours. Including travel times between sites, an average of 3 sites were completed in an 8-10 hour field day by a single field crew, thus meeting scheduling requirements that were laid out at the start of the field season in each year (Section 4.3).

Methods used for each of the sample collections and measurements are listed in Table 2.

Table 2. List of habitat variables and method of measurement used in all years.

Variable	Method or Standard Used
Weather	
Current and recent weather	Documentation based on personal observation; choices include: storm, rain, showers, overcast, and sunny
General Site Information	
Location	A hand-held global positioning system was used to record latitude and longitude (in decimal degrees) and elevation. A site description was noted and site diagram completed.
Photographs	Taken looking upstream, downstream, across, and at the substrate (substrate photo included a 50cm quadrat for scale)
Water Quality	
Air Temperature	Recorded from thermometer or handheld meter
Water Temperature	Recorded from handheld meter (YSI model 63)
pH	Recorded from handheld meter (YSI model 63)
Specific Conductance	Recorded from handheld meter (YSI model 63)
Dissolved Oxygen	Recorded from handheld meter (Handy MK II)
Water Samples	Collected according to sampling procedures outlined in the B.C. Ambient Freshwater and Effluent Sampling Manual (RIC 1997). Sample bottles included: • 1 L for general ions (alkalinity, chloride, true colour, sulphate, total dissolved solids, total suspended solids and turbidity), • 250 mL for nutrients (ammonia, nitrite, nitrate, organic and total nitrogen; orthophosphorus, total phosphorus), • 250 mL for total organic carbon (H_2SO_4 preservative added in the field), and • 250 mL for total metals (collected in acid washed bottle, with HNO_3 preservative added in the field). All water samples were sent to Maxxam Analytics Inc. by overnight courier, to achieve the recommended 72hr holding time. Metals analysis was done using an ICPMS scan, and all parameters were analyzed according to methods in the B.C. Environmental Laboratory

Variable	Method or Standard Used
	Manual (Horvath 2005).
Benthic Invertebrates	
Sample	Each invertebrate sample was collected using a 400μm mesh kick-net, according to the timed procedure reported in the CABIN Invertebrate Biomonitoring Field and Laboratory Manual (Reynoldson et al 2003). The method was modified to include sampling from riffle sections only; other habitat types (pools, glides, etc.) were not sampled. The kick-net operator moved upstream in a zig-zag pattern, kicking substrate and collecting sample for 3 minutes. The 3 minute timer was stopped anytime the sampler did not have the net in the water (e.g. moving to an upstream riffle section or climbing over an instream log). After kick netting was completed, the sample was dispensed into a plastic bin, and large debris was cleaned and removed from the sample and excess water was strained off. The samples were placed in labelled 500 mL plastic jars. 10% buffered formalin was added to the jars to preserve the contents. The samples were sent to laboratories at the end of the sampling season for identification and enumeration of the invertebrates.
Vegetation Cover	
Percent Cover	An estimate of the % of the wetted surface area that is covered (within 1 m of the water surface) by each of the following: woody debris, boulders, undercut banks, deep pools, and overhanging vegetation. Modified from the Reconnaissance (1:20 000) Fish and Fish Habitat Inventory: Standards and Procedures - Version 2.0 (RIC 2001).
Macrophyte Coverage	An estimate of the % of the streambed that is covered by macrophytes (Reynoldson et al 2003).
Periphyton Coverage	An estimate of periphyton coverage in the running water, using a scale of 1 to 5 (explained in Appendix A).
Disturbance Indicators	
Misc. disturbance indicators	Documentation of the presence of scour, sediment wedges, extensive riffles and limited pools, unvegetated and mid-channel bars, multiple channels, eroding banks, isolated sidechannels, recently formed large woody debris (LWD) jams and LWD parallel to banks. Modified from RIC (2001).
Stream Channel Characteristics	
Gradient	Measured using a clinometer and reported in %
Habitat Units	An estimate of the % of the channel area within the reach that is occupied by pools, glides, riffles, and cascades.
Stream Widths	Three measurements of the wetted and bankfull width, taken from within the stream reach. Modified from RIC (2001). Values were later averaged to obtain wetted and bankfull width.
Stream Profile (for Discharge Measurement)	A discharge measurement from the kicked area, calculated from velocity and depth measurements made at 5-8 equidistant points across the stream, using a Swoffer 2001 current meter. Field

Variable	Method or Standard Used
	methods were based on procedures in the Manual of Standard Operating Procedures for Hydrometric Surveys in B.C. (RIC 1998), with velocity measurements taken at a depth of 0.6 x total depth. Discharge calculations followed RIC (1998).
Substrate	
Composition	A visual estimate of the % of the stream reach that is covered with various particle sizes (sand, gravel, pebble, cobble, boulder and bedrock), according to the Wentworth Scale. The estimate was made by each of three members of the field crew and average values were recorded.
Embeddedness	An estimate of how embedded the cobbles are in the surrounding fines (measured in the riffle habitat).
Pebble Count	A Wolman Pebble Count (Wolman 1954), where the intermediate diameter of 100 randomly-selected particles within the stream reach was measured using a ruler. From this data, SYSTAT 11 and MS Excel were used to calculate the median particle size (D50), geometric mean particle size (Dg) and the Fredle Index (FI).
Odours/Oils	Presence documented on field sheet (Appendix A)
Riparian Vegetation	
Vegetation Types	An estimate of the % of different vegetation types (bare soil, grass/herb, shrub, deciduous, coniferous) present at the site. Modified from RIC (2001). Included documentation of species present.
Structural Stage	An estimate of structural stage, according to RIC (2001).
Canopy Closure	An estimate of canopy closure, according to RIC (2001).
Mountain Pine Beetle Infestation	An estimate of the presence of pine trees and severity of the pine beetle infestation, at the site (riparian) and in the watershed.
Land Use	
Observations	Observations of land use, erosion, and NPS pollution at or near the site.
Rapid Bioassessment	
USEPA Rapid Bioassessment Protocol (RBP)	A scoring of 10 habitat parameters (epifaunal substrate/cover, embeddedness, velocity-depth combinations, sediment deposition, channel flow status, channel alteration, channel sinuosity, bank stability, bank vegetative protection and riparian vegetative zone width), using a scale of 1-20 according to the USEPA RBP Field Sheet (Barbour et. al. 1999). Also referred to as the Alaska Stream Condition Index (ASCI).

4.5 Laboratory Protocols for Benthic Invertebrate Counts

4.5.1 Sub-sampling

Raw invertebrate samples can contain up to several thousand animals and typically require subsampling procedures to facilitate enumeration. Vinson and Hawkins (1996) showed that taxa richness increases hyperbolically as a function of number of organisms identified in sub-samples (Figure 4). The relationship, known as a collector curve, can change according to method of sample sorting. One curve is based on single phase sorting wherein the entire sample is placed in a grid (e.g. Marchant Box, Marchant 1989) or other device and sub-samples are removed until some pre-defined number of individuals are removed (Rosenberg et al.1999). The second curve involves 2-phase sorting wherein large and possibly rare organisms are removed by picking from the entire sample or by passing the sample through a large mesh sieve (1 mm or larger) that retains large particulate matter. Common small animals adhered to that debris may either be picked out and returned to the sample for subsequent sub-sampling or they may be included in the large-rare group and that group can be called "macrobenthos". The collector curves produced from single phase sorting have a more gradual break in slope and tend to have lower richness for a given sample count than curves produced from 2-phase sorting. Because of this sorting effect, a recommendation from Vinson and Hawkins (1996) is that 2-phase sorting is preferable. It leads to data more closely representing whole sample richness than does single phase sorting. Nichols and Norris (2006) found that the method can introduce bias to rare taxa, which is not desirable in RCA methods, but as long as the same method is used for both reference site and test site samples, the method will not bias test results. Based on these findings the 2-phase sorting method was used in this study for both reference and test samples.

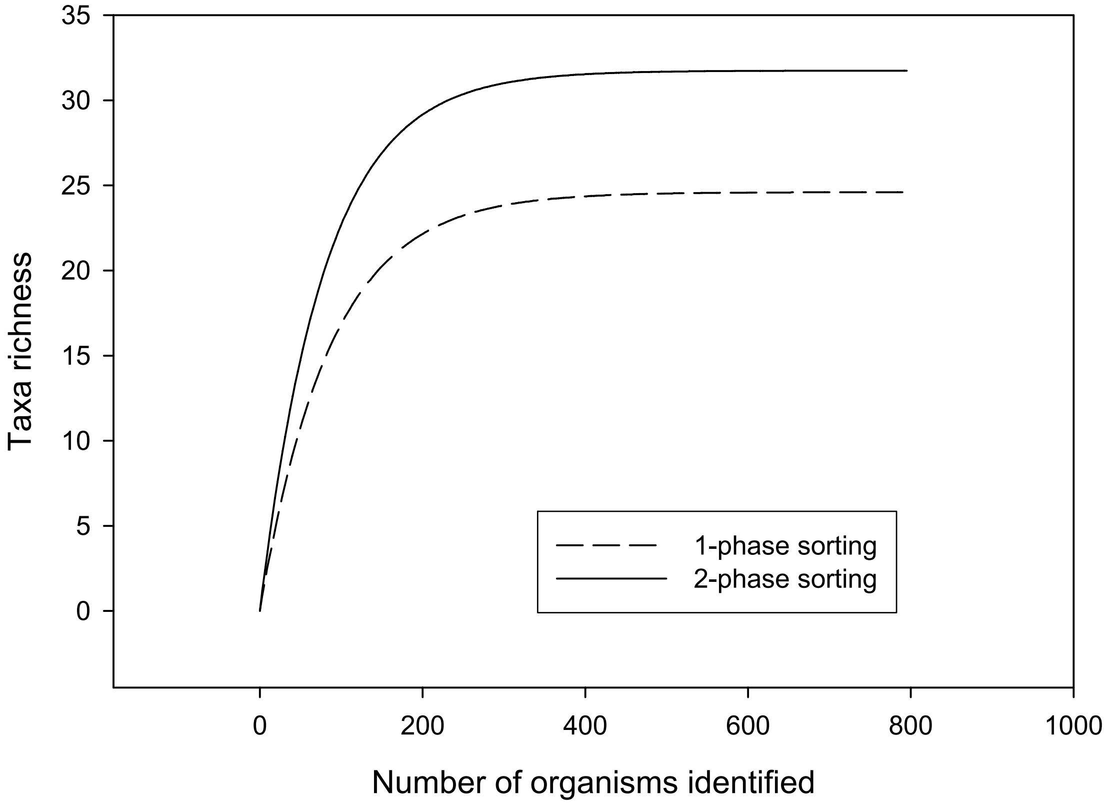

Figure 4. Collector curves showing the effect of 1–phase versus 2–phase sorting on the relationship between taxa richness and number of organisms identified. Data taken from Vinson and Hawkins (1996).

Vinson and Hawkins (1996) reported that richness from samples with counts >300 individuals will result in accurate statistical inferences. A count of 300 individuals will yield 88% of maximum attainable richness (Figure 4). In going from a count of 300 to a count of 600, the collector curves in Figure 4 show that the doubling of effort only results in a 6% increase in the estimate of richness for a sample. Somers et al. (1998) found that a count of 100 individuals was equally powerful to sample counts of 200 and 300 individuals for purposes of distinguishing littoral benthic communities using ANOVA on selected endpoints and a multivariate metric. Growns et al. (1997) found that the ability of multivariate procedures to detect differences in community structure between reference and polluted sites was optimal when using a 100 selective animal count. A selective count involves maximizing the number of taxa picked and limiting counts of individuals per taxa, particularly if they are dominant. Sovell and Vondracek (1999) found that richness increased with sample size (100 to 300 individuals) but several community metrics did not. Hence, assessment methods that rely on metrics may be able to get away with relatively small counts of 100 individuals and still be sensitive to detecting disturbed sites. For sample counts that are required for RCA, Reynoldson et al. (2001) referred to Rosenberg et al. (1999) who showed that a sub-sample count of 200 individuals was acceptable. To be conservative we selected a sub-sample count of 300 individuals. It provided data explaining close to 90% of community diversity with application of 2-phase sorting, based on the collector curves in Figure 4.

Test enumerations from 2004 resulted in some modification of the 300 sub-sample count procedure. A 2 mm size sieve was used to split the samples in the two-phase sorting. During processing of the test samples, <100 to approximately 1200 animals were retained on the sieve (macrobenthos), which was much more than anticipated. With the minimum sub-sample count of 300 individuals applied to the benthos passing the sieve (microbenthos), the resulting total count ranged from approximately 400 individuals to well in excess of 1000, in some cases reaching 1500 individuals. Since counts in excess of 300 were not expected to substantially improve richness (Figure 4), the procedure was revised. Large organic debris (e.g. leaves, twigs, etc.) that was retained on the sieve was picked clean and removed. Picked animals were returned to a “macrobenthos” tray. Remaining material on the sieve, including animals, were placed in the tray and partitioned into 2 or 4 parts. The smaller partition or no partitioning was selected when there were relatively few individuals present. Each of those parts was considered a macrobenthos sub-sample. One or more sub-samples were enumerated until a target of 200 animals was counted. If the target of 200 animals was reached before a sub-sample was completely sorted, that last subsample was sorted in its entirety. Total macrobenthos was the number of individuals, by taxa, enumerated in the sub-samples multiplied by the partitioned amount (e.g. a quarter fraction was multiplied by 4 to yield the total macrobenthos count). Sub-sampling of the microbenthos (animals passing the sieve) continued according to the method outlined above but the sub-sample count was reduced from 300 to 200 individuals. A maximum of 16 splits of the microbenthos was applied. With these changes, the actual count of all animals from an average sample dropped from up to 1500 animals using the original method (300 sub-sample count of microbenthos and total count of macrobenthos) to approximately 400 animals (approximately 200 individuals in each of the macrobenthos and microbenthos fractions). Time required to sort and enumerate an average sample using the revised procedure was 5 – 7 hours.

4.5.2 Sub-sampling devices

We addressed the question of what device is "best" for distributing a sample and for selecting sub-samples. The Marchant Box or Caton tray used with single phase sorting was used by Reynoldson et al. (2001). Other approaches involve size fractionation in sieves followed by volumetric sub-sampling using plankton splitters or other device that divides the sample into any number of units (several described by Glozier et al. (2002)). Another approach is selection of random sub-samples from a grid that is laid out in an open tray into which a sample is poured (Sovell and Vondracek

1999). Still another approach is counting of individuals that are found in sub-sample volumes that are sequentially poured from a sample until a given animal count is achieved (Vinson and Hawkins 1996, Somers et al. 1998). Plankton splitters were used in this project. We found that the Marchant Box has so many cells (100 in total) that time to "suck up" individuals out of many cells at low animal density can be a deterrent. The box also requires a shaking technique that may or may not achieve desired distribution of particles among cells. Thus, its effectiveness and the precision and accuracy of resulting data may vary between taxonomists. We found that use of a large plankton splitter avoided this problem, making it a more reliable device than the Marchant Box or Caton tray when several users were involved.

4.5.3 Quality assurance and quality control (QAQC)

A description of QAQC prepared by Glozier et al. (2002) was followed as a laboratory protocol. There were three components to QAQC; sorting efficiency, sub-sampling precision, and sub-sampling accuracy.

For sorting efficiency, the basic rule was that 10% of all samples from a given year of collections were resorted. A target for acceptable sorting was that >90% of the sample must be enumerated on the first sort. If this 90% efficiency was found, the animals that were found on the second sort were not included in the sample count. If the efficiency was <90% (e.g. >10% of the total count is found on a second sort), then all samples in the group of samples to which the test applied require resorting. Resorting was also required if an entire group of taxa (e.g. all ostracods) were missed in the first sort.

Precision is a measure of how close a count of animals in a particular sub-sample is to counts in the other sub-samples. For example if a count in one sub-sample is 289 and the count in a second sub-sample is 316, the precision is defined as 8.5% (289/316*100).

Accuracy is a measure of how close an enumeration is to the actual value of animal abundance. It is assumed that actual abundance in a given sample can be determined to within 10% based on the rule of acceptable sorting efficiency defined above. Determination of accuracy required enumeration of all sub-samples plus remaining sample. For example;

- A count in sub-sample A is 189 animals, representing 25% of the total sample volume for a total sample count of 756 animals;
- A count in sub-sample B is 216 animals, representing 25% of the total sample volume for a total sample count of 864 animals; and
- A count in the remaining material is 453 animals, for an actual total count of 858 animals in the selected sample.

Accuracy was then determined for each sub-sample. For example, accuracy for sub-sample A was -11.9% determined as (756-858)/858*100.

Based on the review by Glozier et al. (2002), acceptable accuracy was defined as ±20%. A measure of accuracy was made on 10% of all samples. (e.g. 10% of a complete set of samples collected in a given year) to coincide with the test of sorting efficiency. If counting error exceeded this 20% rule, the method was modified and tested until error was reduced to less than 20%. The methods described in section 4.5.1 and 4.5.2 achieved this level accuracy.

4.5.4 Summary of procedures

A summary of laboratory procedures used to enumerate the benthic invertebrates is as follows:

1. The sample was washed through a 2 mm mesh sieve to yield a macrobenthos fraction that was retained on the sieve and a microbenthos fraction that passed the sieve;
2. The microbenthos fraction was split into 4 to 16 parts using a large plankton splitter;
3. Sub-samples of microbenthos were enumerated until 200 animals were counted. If the target of 200 animals was reached part way through the sorting of a sub-sample, that sub-sample was sorted in its entirety;
4. If the estimated abundance of animals in the macrobenthos fraction was less than 200 animals, that fraction was enumerated in its entirety. If that fraction contained more than 200 animals, it was partitioned in a level tray into 4 parts. Animals were enumerated from successive sub-samples until 200 animals were counted. If the target of 200 animals was reached part way through the sorting of a sub-sample, that sub-sample was sorted in its entirety;
5. Sub-sample counts were rated by number of sub-samples to determine the total count of benthos. For example, if 1 of 8 microbenthos sub-samples was enumerated, the sub-sample count was multiplied by 8 to determine the count of microbenthos in the complete sample. The same approach applied to the macrobenthos. The sum of microbenthos and macrobenthos in the complete sample was the sample count;
6. QAQC procedures were applied to 10% of samples from a given year including tests of sorting efficiency and measurement of accuracy of sub-sampling method as described in Section 4.5.3.

4.6 Compilation of Biological Data

Family level enumerations of samples were compiled into spreadsheet files ready for modelling and analysis. Metadata accompanying each observation included local site name, site code, EMS number that was assigned for logging the data into the Provincial water quality database, date of collection and a unique sample identifier. This identifier was required as a single code for use in uniquely identifying a sample in statistical software. Although duplicate or triplicate samples were collected from a few reference sites over the three years, data from only one replicate sample from each site in any given year was compiled for model development. The others were considered test samples that were used in examining model accuracy and precision (Section 6.1). Each biological observation was matched to an accompanying and complete compilation of habitat data described in Section 4.7. Where this matching of data was incomplete, the observation was discarded from further consideration.

4.7 Compilation of Landscape and Site Specific Habitat Data

Many of the natural gradient variables and stressor gradient variables were measured in the field as described in Section 4.4 and compiled in a spreadsheet ready for the next stage of model development. Each observation (row of data) was matched to an accompanying and complete compilation of biological data described in Section 4.6. Where this matching of data was incomplete, the observation was discarded from further consideration.

Many other habitat variables describing landscape attributes at or upstream of a given sampling site were compiled using ArcGIS 9 geographic information systems (GIS) software developed by ESRI© (*http://esri.com/*) and added to the habitat file. All spatial datasets were accessed through the Province of B.C. spatial data directory known as the Land and Resource Data Warehouse (LRDW). The LRDW *'is the corporate repository for integrated land, resource and geographic data that supports a variety of business requirements for the natural resource sector, other government agencies, industry, and the public'(*www.lrdw.ca*).* The following groups of data were accessed:

1. Watersheds (basin and sub-basin areas)
    - Database name – WHSE_BASEMAPPING.WSA_WATERSHEDS
    - Scale – 1:50,000
    - Date of Creation – Unknown

The watershed atlas (WSA) provided digital representation of aquatic related features including boundaries for all third-order and greater watersheds. In some cases watersheds were less than third-order and were hand digitized 'on the fly' guided by contour line work. All watersheds were selected and/or created based on the sample site location acting as the mouth of the drainage path for the watershed. The watershed boundary extended to the height of land of the drainage network.

2. Water Related Features
   - Lakes database name – WHSE_BASEMAPPING.TRIM_EBM_WATERBODIES
   - Wetlands database name – WHSE_BASEMAPPING.TRIM_EBM_WETLANDS
   - Ice database name – WHSE_BASEMAPPING.TRIM_EBM_ICEMASSES
   - Scale of all databases – 1:20,000
   - Date of Creation of all databases – 2002-10-24

   Area based values were calculated based on quantity (number of features) and total area within the watershed. The watershed area occupied by these water features (lakes, wetlands, ice masses) was expressed as a percent of the total watershed area upstream of the sample site.

   - Rivers database name – WHSE_BASEMAPPING.TRIM_EBM_WATERCOURSES
   - Streams database name – WHSE_BASEMAPPING.TRIM_WATER_LINES
   - Scale of both databases – 1:20,000
   - Date of creation of both databases– 2002-10-24

   Watercourses (streams and rivers) were classed as polygons, and were calculated as an area. Streams were classed as definite (always flowing; fcode-GA24850000), or indefinite (sometimes not flowing; fcode-GA24850140). Definite streams (including rivers) and indefinite streams were expressed as drainage density within the watershed (stream length divided by watershed area). A stream length ratio was calculated as the length of definite streams divided by the total stream length in the watershed.

3. Geology
   - Database name – WHSE_MINERAL_TENURE.GEOL_BEDROCK_UNIT_POLY_SVW
   - Scale – 1:100,000
   - Date of Creation – 2005-01-01

   Geology of the watersheds was defined by area of geologic classes including sedimentary, intrusive, metamorphic, volcanic, and ultramafic rock. The area of each rock class was expressed as a percentage of the total watershed area. This information described the bedrock environment of the watershed. Information related to specific rock types and approximate ages were also available for the dataset.

4. Forest Harvesting
    - Database name – WHSE_FOREST_VEGETATION_RSLT_OPENING_POLY
    - Scale – 1:20,000
    - Date of Creation – 2003-11-27

This information described the actual area of forest harvesting (termed opening) conceptually viewed as an opening in the forest landscape canopy as a result of forest harvesting. This information was calculated as the number of openings, and total area. The number of openings was expressed as percent of the total watershed area. This information was based on forest licensee forest development plan (FDP) submissions which are uploaded to the results database and maintained by the Ministry of Forests and Range.

5. Environmental Monitoring Sites
    - Database name – WHSE_WASTE.BC_ENV_MONITOR_LOCNS
    - Scale – 1:50,000
    - Date of Creation – 2003-03-31

Environmental Monitoring Sites (EMS) are established and maintained by the Ministry of Environment. These data include permitted quantities of waste discharged from a particular operation. Actual operations were not evaluated due to lack of attribute information but included mining operations and most effluent discharges. This information was expressed as the number of permitted discharges in the watershed.

6. Land Use
    - Database name – WHSE_BASEMAPPING_BTM.PRESENT_LAND_USE_V1_SP
    - Scale - 1:50,000
    - Date of Creation – Unknown

Attributes that were used from this data included:

- Logging: This information represented forest harvesting operations not captured in openings database. This information was calculated as the quantity and total area and expressed as percent of total watershed area. Total forest harvesting with the watershed was estimated by summing openings (above) and logging.

- Alpine Area: This information described the area of the watershed mostly devoid of vegetation at higher elevations and expressed as the percent of the total watershed area.
- Avalanche Area: This information described the area of the watershed prone to avalanche activity and was expressed as the percent of total watershed area.
- Urban Area: This information described the area of the watershed classed as developed. It was expressed as percent of total watershed area.

7. Roads
    - Database name – WHSE_BASEMAPPING.TRIM_TRANSPORTATION_LINES
    - Scale – 1:20,000
    - Date of Creation – 2003-02-06

Road information was classed into three main categories including paved, gravel and block. Paved roads were associated with urban areas and highways. Gravel roads were associated with forest service roads (FSRs) and other non-paved active roads. Block roads were considered to be access roads for forest harvesting operations, but information was not available with respect to maintenance activity and deactivation status. Paved road fcodes included DA25050180, DA25050190, DA25100200, DA25100210. Gravel road fcodes included DA25000110, DA25000120. Block road fcodes included DA25150000, DA25150100, DD31700000, and DD09950000. Calculated values included total length of each road class within the watershed and it was expressed as road density (road length divided by watershed area) in the watershed.

A point file generated in ArcGIS to show intersecting streams (by class) and roads (by class) was used to locate stream crossings in a given watershed. While inaccuracies in the datasets may have resulted in an overestimation of stream crossings, this type of analysis was considered acceptable for present needs. Information related to stream crossing quality in terms of sedimentation and potential for runoff from the road surface was not included in this analysis. Data were expressed as total number of stream crossings in the watershed.

8. Forest Age
    - Database name – WHSE_FOREST_VEGETATION.VEG_COMP_POLY (joined to table WHSE_FOREST VEGETATION.VEG_COMP_LYR_R1_VW)
    - Scale – 1:20,000
    - Date of Creation – Unknown (based on annual updates)
    - Name – FC1 – aka Forest Cover, circa 2001

Forest age and class was derived from two datasets. The forest vegetation information is more commonly known as vegetation resource inventory (VRI). These datasets are replacing forest cover datasets which have been slowly phased out since 2001. Currently not all information from the forest cover datasets have been rolled over to the VRI datasets. For this project, some assumptions were made with respect to tree species and age class, based on the biogeoclimatic ecosystem classification (BEC) system spatial datasets. Only age classes greater than 6 (greater than 100 years old) were examined to define the percentage of undisturbed forested area within the watershed.

9. Parks and Protected Areas
    - Database name – WHSE_PARKS.PA_PROTECTED_AREA_POLY
    - Scale – 1:20,000
    - Date of Data – 2004-07-22

   This data provided watershed area set aside and free of mining exploration and resource extraction including parks and protected areas. The data were expressed as percent of total watershed area.

10. Range
    - Database name – WHSE_FOREST_TENURE.FTEN_RANGE_POLYGONS
    - Scale – 1:20,000
    - Date of Creation – 2003-11-27

    Range tenures are an important aspect of agriculture activity in a watershed. Unfortunately small scale operations are not captured effectively in this dataset, and include only forest tenures for range operations. This data was used to identify the maximum head of cattle that may be utilizing the range tenure area at any one time. These data provided a surrogate for impacts associated with agricultural waste and physical impacts of cattle crossings of streams within a given watershed. The data were expressed as percent of total watershed area.

11. Agricultural Land Reserve
    - Database name –
      WHSE_ADMIN_BOUNDARIES.ALC_AGRI_LAND_RESERVE_
      POLYGONS
    - Scale – 1:20,000
    - Date of data – 2006-04-30

    Agriculture land reserve information defined the proportion of the watershed used (or set aside) for agricultural purposes. This information did not identify extent of agricultural use nor type of agricultural activity. The data were expressed as percent of total watershed area.

12. Mining

- Database name – WHSE_MINERAL_TENURE.MINFIL_MINERAL_FILE
- Scale – 1:50,000
- Date of data – 2005-09-10

This dataset provided information related to mineral showings and prospects as follows:

- Showing: Mineral showings identified outcrops in the landscape, that have not been developed. This information provided an indication of the potential for new mining operations in the watershed.
- Prospect: Mineral prospects identified the number and area of prospective claims in the watershed, again providing an indication of potential new mining operations in the watershed.

These data were expressed as the number of showing/prospect points in the watershed.

13. Extraction Sites

- Database name – WHSE_BASEMAPPING.TRIM_EXTRACTIONSITES
- Scale – 1:20,000
- Date of data – 2002-10-24

Extraction sites included the number of gravel and sand pits (fcode AG21550000), open pit mines (fcode AG17600000), tailings piles (fcode AG21275000), mines (fcode AG17750000) and tailings ponds within a watershed. These data provided an indication of the level of mining activity in the watershed. The data were expressed as percent of the total watershed area.

14. Elevation

- Database name – WHSE_BASEMAPPING.TRIM_CONTOUR_LINES
- Scale – 1:20,000
- Date of Creation – 2002-10-24

Elevation attributes were calculated from this data set. They included mean elevation, minimum elevation (elevation of the sampling site), maximum elevation, and relief. Relief was defined as maximum minus minimum elevation of the watershed.

15. Slope

- Database name\source –
  giswhse.env.gov.bc.ca\whse_np\corp\gdbc\cat\tdem_bc
- Scale – 25 meter resolution
- Date of Creation – Unknown

Two slope classes were identified as surrogates of the potential for surface movement, and stability. Slope class IV (50-70%, 27-35 degrees) were terrain polygons with moderate likelihood of landslide initiation events following timber harvesting and/or road building. Slope class V (greater than 70%, greater than 35 degrees) were terrain polygons with high likelihood of landslide initiation events following timber harvesting and/or road building. More information on these classes can be found at *http://www.for.gov.bc.ca/tasb/legsregs/fpc/fpcguide/iwap/iwap-toc.htm*. These data were derived from a digital elevation model (DEM) raster dataset. Spatial analysis tools in ArcMap were used to convert the raster image to slope polygons and they were re-classed to conform to Ministry slope classification standards. Slope classes IV and V (above) were expressed as polygon area within the total watershed area.

4.8 Assignment of Reference and Test Sites

4.8.1 Field component

During the initial site selection process (Section 4.1), reference sites were defined as having minimal disturbance at the local and watershed scale. The lack of development in potential reference condition watersheds was roughly confirmed using internet mapping tools including GoogleEarth, iMapBC (http://lrdw.bcgov/index.html), and FISHWizard (© 2005 Freshwater Fisheries Society of BC - *www.fishwizard.com*) maps which included orthophoto, Landsat and other land use related GIS layers.

Since minimally disturbed sites can be difficult to find, some were selected from within parks and protected areas, including Carp Lake, Entiako, Kitlope, Sugarbowl and Tweedsmuir provincial parks. However, sites from backcountry parks and protected areas that were accessed by helicopter and by boat were also required to increase the number of reference sites.

A rough list of reference sites was compiled prior to field sampling. They were selected using best professional judgement (BPJ), recommendations from other biologists, hydrologists, and foresters who were familiar with the area, and narrowed down using maps and FISHWIZARD satellite images. Further fine tuning of site selections was based on known access, river flows, and safety. The list was revised if new information from opportunistic site visits indicated that a site was wrongly assigned.

4.8.2 QAQC using stressor gradient analysis

Landscape scale descriptors of the stressor environment of each site (Table 3) were used in a Principal Component Analysis (PCA) that summarized the correlation of stressors among the sites sampled. This procedure resulted in the calculation of a Principal Component (PC) score along a stressor gradient for each site, where the degree to which each site is exposed to stressors like roads, forestry, and mining is integrated and quantified. The provisional, *a priori* classification of sites as either reference or test that was made using local knowledge was then cross-checked with the position of each site on the stressor gradient.

Any sites designated *a priori* as reference sites that had PC scores along the stressor gradient that exceeded those of test sites, or where some descriptors were missing, were flagged for further detailed review. Similarly, any sites classified *a priori* as test sites that had PC scores lower on the stressor gradient than confirmed reference sites were flagged for further review.

Table 3. List of variables used in the stressor gradient analysis for selection of reference and test sites.

Stressor Variable	**Description**
EMS_Q_Permit	Number of waste discharge permits (pulp mills, wastewater treatment plants, etc.)
Prox_Sampsite	Average distance of discharge permit site to sample site
Pct_Open	Percent of opening area in watershed
Pct_HL	Percent of watershed with historic logging
Grvl_RdDen	Gravel road density
Blk_RdDen	Block road density
Pav_RdDen	Paved road density
Grvl_StrmX	Number of gravel road stream crossings
Blk_StrmX	Number of block road stream crossings
Pav_StrmX	Number of paved road stream crossings
Pct_Agecl12	Percent of agecl12 area within the watershed
Pct_Park	Percent of park area in watersheds
Pct_Range	Percent of range area in watersheds
Max_Cows	Maximum number of livestock that may be using the range at any one time
Pct_Urban	Percent of urban area in the watersheds
Pct_ALR	Percent of ALR area in the watersheds

Stressor Variable	Description
Past_Prod	Mine sites, classed as Past Producer
Dev_Pros	Mine sites, classes as Developed Prospects
Showing	Mineral showings
Prospect	Mineral prospects
GrvlSnd_Pit	Number of gravel/sand pits
MOpen_Pit	Number of open pit mines
Tailing_Pile	Number of tailings piles
TrimExt_Mine	Number of extraction mines
Tailings_Ponds	Number of tailings ponds
Pct_Openbuf	Percent opening area in buffer area
Pct_slpcl4_Ar	Percent of slpcl4 area in buffer area
Pct_slpcl5_Ar	Percent of slpcl5 area in buffer area
LU_forest	0=absent, 1=present
LU_logging	0=absent, 1=present
LU_field/pasture	0=absent, 1=present
LU_mining	0=absent, 1=present
LU_ag	0=absent, 1=present
LU_indus	0=absent, 1=present
LU_resid	0=absent, 1=present
ASCI 1	Epifaunal Sub.
ASCI 2	Embeddedness
ASCI 3	Velocity-Depth
ASCI 4	Sediment Deposition
ASCI 5	Channel Flow Status
ASCI 6	Channel Alteration
ASCI 7	Channel Sinuosity
ASCI 8	Bank Stability (average)
ASCI 9	Bank Veg Protection (average)
ASCI 10	Riparian Vegetation (average)

Flagged sites were reviewed by people familiar with the sites and the associated landscape data. The site status was changed if review of all data suggested that the site had been assigned incorrectly to either reference or test status. The following criteria were used as a guide when considering changing the status of site from test to reference:

- absence of mining activity;
- % openings less than 5%, with % buffered openings less than 1%;
- total road density less than 0.25;
- "relatively" small number of stream crossings (esp. gravel crossings);
- Relatively low % range and or number of cattle within watershed; and
- absence of logging and other land uses observed in the field.

Using this iterative process of Quality Assurance, a defensible and rigorous set of reference sites was determined with the combination of local knowledge and the stressor gradient analysis.

4.9 Model Development and Site Testing

4.9.1 Overview

Two RCA models were built and compared. One called the Benthic Assessment of Sediment (BEAST; Reynoldson et al. 1995) followed the protocols used in development of the Fraser Basin (Rosenberg et al. 1999, Reynoldson et al. 2001) and Georgia Basin RCA models (Sylvestre et al. 2005), which followed from original application of RCA methods in the Great Lakes (Reynoldson et al. 1995). The principle of this model is that sites are classified into groups. If a test site does not show the typical biological assemblage of the group to which it belongs (based on environmental attributes), the site is assessed as stressed. The second model was called the Skeena River Assessment System (SkeenRIVAS), which followed the Australian (Parsons and Norris 1996) and UK (Wright 1995, Wright et al. 2000) RIVPACS protocols for model development and site testing. These models also classify sites into groups, but then predict the probability of occurrence of single taxa. If the number of those taxa matches the number of expected taxa, the site is deemed in reference condition. If not, it is labelled as stressed. While the mechanisms in both models are very similar, the actual test score is different. Although the BEAST approach has been commonly used in Canada, a comparison with SkeenRIVAS was considered important to determine if one model may be superior to the other with respect to precision and accuracy for routine site testing. This information was necessary for making recommendations on the selection of a model for future use.

Development of both the BEAST and SkeenRIVAS models and site testing followed six common steps:

1. Benthic invertebrates were identified and enumerated in samples collected from a large number of reference stream sites across the pre-defined project area (see Sections 4.1 through 4.6);
2. Variables describing natural habitat attributes (defined in Section 4.2) at the sites where the invertebrates were collected were measured or accessed from GIS data bases and the data were compiled in spreadsheets (see Section 4.7);
3. Groups of reference sites based on similarity of biological composition between samples from these sites were defined using clustering and ordination techniques;
4. Natural gradient habitat variables that best explained dissimilarities between the biological groups were selected using discriminant function analysis (DFA). Those selected variables were called predictor variables;

5. The predictor variables were used in another DFA to predict the probability that a test site belonged to each sample group.
6. The deviation between the composition of biota observed at a test site and the composition of biota expected at the test site if it was similar to that in a reference condition found in one (BEAST) or more (SkeenRIVAS) sample groups was measured to determine site status. The assignment of a test site to sample groups was based on probabilities of group membership defined in step 5.

4.9.2 BEAST

Family level invertebrate counts from reference site samples were compiled in PRIMER (Clarke and Gorley 2001, Clarke and Warwick 2001) and were fourth root transformed to down-weight the very abundant taxa and to allow the midrange and rarer taxa to exert some influence on the calculation of between – sample similarities. Similarities between every pair of samples were calculated using the Bray Curtis coefficient (Krebs 1999) to form a similarity matrix. A dendrogram was plotted using the group average linkage in the hierarchical, agglomerative clustering algorithm in PRIMER. The dendrogram was examined for obvious groupings of samples and a group label was assigned to each sample. The assignment of a sample to a group was assisted with interpretation of a non-metric multi-dimensional scaling analysis (MDS) that was run in PRIMER on the same similarity matrix that was used for the cluster analysis. MDS is a procedure for fitting a set of points in space such that the distances between points correspond as closely as possible to dissimilarities between objects. Output was displayed on two-dimensional or three-dimensional images called ordinations. An ordination had no scaling units and space between objects on the image provided a perspective of dissimilarities. A computation that accompanied each ordination was something called a “stress” value. Stress increased with reducing dimensionality of the ordination and it indicated if a 2-dimensional plot was a usable summary of the sample relationships. Where any two dimensional ordination had a stress value >0.2, interpretation of sample groups was done on the 3-dimensional ordinations. Any sample that was clearly separated from clusters of other samples on the cluster dendrogram and the ordinations was considered an outlier and it was removed from further model development.

Analysis of similarities (ANOSIM) was run in PRIMER to derive a statistic that indicated the degree of similarity of the benthic communities within and between the sample groups, excluding the outlier samples. The resulting *R* statistic was based on a non-parametric permutation procedure that was applied to the similarity matrix underlying the MDS. This procedure was a multivariate analogue of a standard one-way analysis of variance. The *R* statistic contrasted the observed differences in the composition of invertebrate families between sample groups with the composition of invertebrate families within sample groups using the equation:

$$R = \frac{\left(\overline{r_B} - \overline{r_W}\right)}{0.5M}$$

where $M = n(n-1)/2$ and n was the total number of samples, $\overline{r_W}$ was the average of all rank similarities among samples within groups, and $\overline{r_B}$ was the average of all rank similarities arising from all pairs of samples between groups. R can range from 0 in which there is no difference in community composition between groups (similarities between and within groups are approximately the same) to 1 in which all samples within groups are more similar to each other than they are to any samples from different groups. A test of significance defined by a probability value was applied to *R*. If all samples were randomly assigned to any group and *R* was re-calculated and this was done a very large number of times (default is 999 times), the *P* value was the probability of R being greater than the calculated *R* statistic.

R values for sample group pairs were then examined. For these comparisons, some rules were applied. For there to be clear differences, a group pair having $R \leq 0.2$ was considered to be almost identical, while a pair having $R > 0.2$ and $R \leq 0.4$ was considered weakly separable. A group pair having $R > 0.4$ and $R < 0.6$ was considered to indicate differences but also some overlap, and $R \geq 0.6$ was considered to indicate well separated sample groups. Any group pair showing an R value <0.6 was considered to have enough overlap to warrant merging of the sample groups having the overlap. This merging was not done, however, unless discriminant function analysis done in the next step of model development showed poor classification of samples to one or more sample groups.

Discriminant function analysis (DFA), run in Systat v11 (Systat 2004), was used to develop functions of habitat variables that best discriminated between the biological sample groups. The starting list of potential predictor variables was that shown in Table 1. The DFA was run using the habitat data only from reference sites, thus corresponding with the biological data. F-to-remove was set at >4 and tolerance (a measure of correlation between predictor variables) was set at >0.1 and no transformations were applied. Both forward and backwards stepping procedures were run, which provided a short list of predictor variables. A backwards stepping DFA was then run using that short list of variables. The model was accepted if it was significant ($P<0.05$) and all predictor variable tolerance values were in the range of 0.5 or greater. Tolerance less than this range was considered to indicate unacceptable correlation between any combination of predictor variables. Part of the output of a DFA in Systat is a classification test using the jackknife procedure in which an observation from a known group is cut out of the DFA and re-substituted back in to see how well the model is able to classify that site to a sample group. We accepted a model that achieved >60% correct classification to each sample group.

Poor classification to any one sample group using the jackknife procedure (e.g. <60% classification success) was justification to review the assignment of sample groups and possibly merge groups to improve the classification success. Where this outcome occurred, output from ANOSIM that tested the dissimilarity of biological communities between sample groups was reviewed. If ANOSIM showed $R<0.6$ for a given sample group pair and the jackknife misclassification success was poor (<60%) for any one of the same two groups, those groups were merged to form a single group. The DFA was then repeated using the same procedures outlined above, again checking for tolerance values. Where the tolerance value for one or more predictor variables was <0.5, the variables were reviewed for possible redundancy (and thus high correlation). Where possible high correlation was found between any two variables, the one having the lowest F-to-remove value was removed and the DFA was re-run. This iterative process of variable selection continued until all tolerance values were >0.5 and the jackknife reclassification success was >60%. DFA output provided a number of discriminant functions, also known as canonical variables, that were equal to one less than the number of sample groups. The function that individually explained most of the total dispersion was accepted as the final model.

It was recognized that any consolidation of sample groups, will invariably increase reclassification success because there are fewer chances of the model placing a sample in the wrong group. Conversely, by accepting a relatively large number of sample groups the model reclassification error can increase (Figure 5). One could argue that it is wrong to consolidate sample groups that appear in an initial dendrogram and ordination because it artificially increases reclassification success and it broadens the biological attributes of affected sample groups, potentially making the model less sensitive to detecting habitat disturbance. The view accepted for this project was that for the model to have acceptable error for management application (<60% reclassification was considered too low), the sample groups must have unique attributes that made them substantially dissimilar from each other.

Site testing proceeded by first running a DFA to assign a test site to a sample group based on attributes of the predictor variables. A complete estimation DFA was run in Systat v11 (Systat 2004) using data from the reference sites and all test sites. The variables selected for the DFA were those found to be the best discriminators of sample groups in the model development described above. Output provided probabilities of group membership for each test site. Each site was assigned to the sample group to which it had the highest probability of belonging.

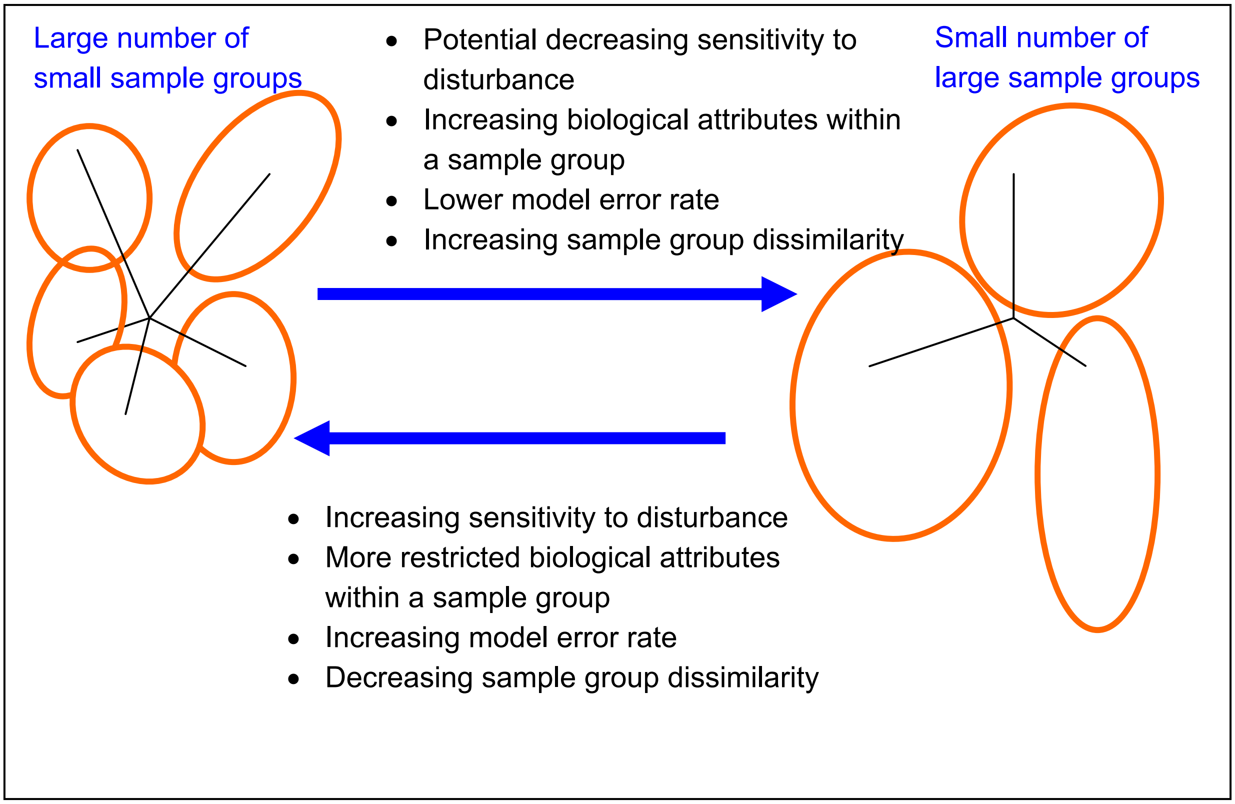

Figure 5. Schematic diagram of tradeoffs between modeling outcome and sensitivity to site disturbance to be considered in defining sample groups during development of a BEAST model.

An ordination of biological data from a given test sample and all reference site samples from a given sample group was run in PRIMER. Output included two dimensional and three dimensional ordination coordinates. If stress on the two dimensional ordination exceeded 0.2 (indicating that the two dimensional image was not a good representation of the sample similarities), only the three dimensional ordination coordinates were further examined. The reference sites and the single test site ordination coordinates were plotted along with 90%, 99%, and 99.9% probability ellipses in Systat (Figure 6). Where three dimensional ordination coordinates were used, three plots were produced (axis 1 versus axis 2, axis 1 versus axis 3, and axis 2 versus axis 3). In each of the plots, the coordinates for the test site were plotted to determine within which ellipse it lay. Stress was defined according to the following criteria:

a. Not stressed (test site laid inside the 90% ellipse in all three plots);

b. Slightly stressed (test site was situated between the 90% and 99% ellipses in at least one plot while it was within the 90% ellipse on the other plots);
c. Stressed (test site was situated between the 99% and 99.9% ellipse in at least one plot while it was inside of the 99% ellipse on the other plots);
d. Severely stressed (on at least one plot the test site was situated outside of the 99.9% ellipse); and
e. If a test site lay on top of a line delineating a probability ellipse, then that site was assigned a worst case rating (e.g. if the site laid squarely on the 90% ellipse, the site was considered slightly stressed).

Each test site was independently run through the ordination and plotting routine to assess site status.

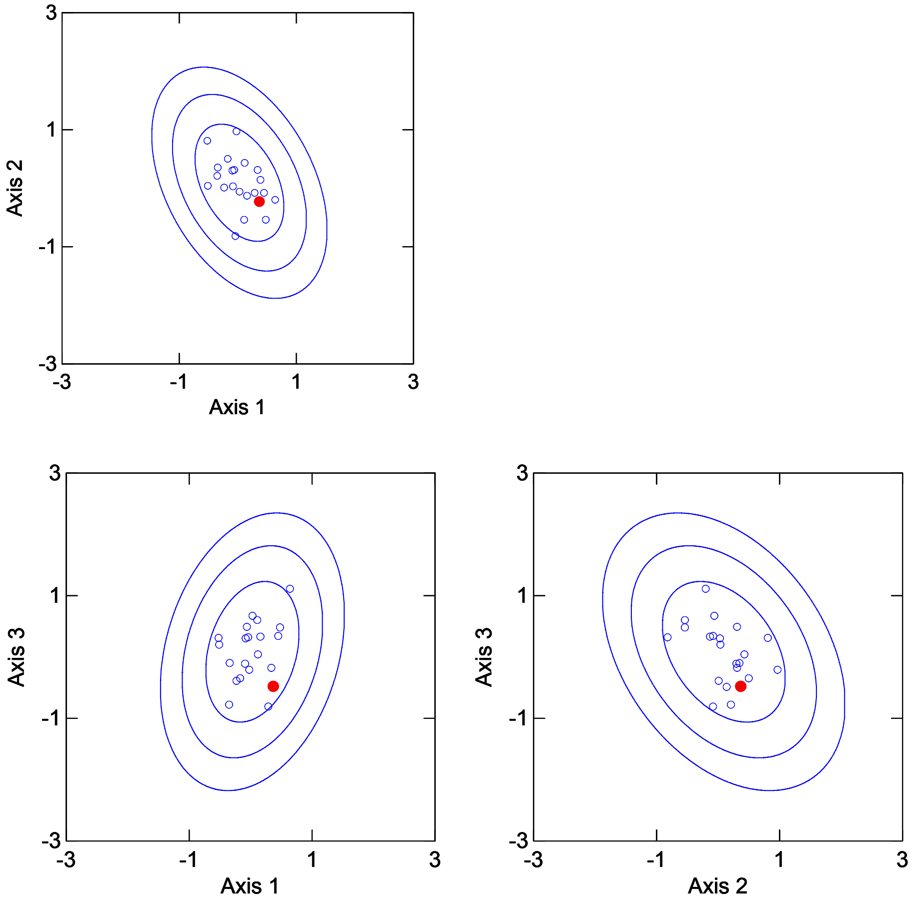

Figure 6. Ordination plots for a hypothetical test site shown in red and reference sites shown in blue belonging to a given sample group. The ellipses correspond to 90% (inner ellipse), 99% (middle ellipse), and 99.9% (outer ellipse) probabilities. The conclusion is that the test site is not stressed because it lay inside of the 90% ellipse in all plots.

4.9.3 SkeenRIVAS

Cluster analysis based on presence/absence of invertebrate taxa in reference site samples was conducted using the Bray-Curtis distance and the flexible beta clustering algorithm (beta=-0.1) in PCOrd (McCune & Mefford 1999).

A stepwise (forward/backward) discriminant function analysis (DFA) using the procedure called STEPDISC in SAS version 9 (SAS 2007) was used to select model parameters. A probability of 0.05 was used as the criterion for entry or removal of a variable at every step. This procedure identified which environmental variables best discriminated among the sample groups.

To predict the expected community from a certain combination of environmental variables at a test site, the discriminant functions were used to determine the standardized, multivariate distance of the site from the reference groups and predict the probability of membership of a test site to a sample Group. For example, in Figure 7, the test site is closest to Group 2 and second closest to Group 1. Hence, the probability of this test site belonging to Group 1 or 2 is higher than belonging to Groups 3 and 4. The weighted average probability of a taxon occurring at the test site was calculated by weighting the frequency of occurrence within groups by the probability of the site belonging to a group as shown in Table 4.

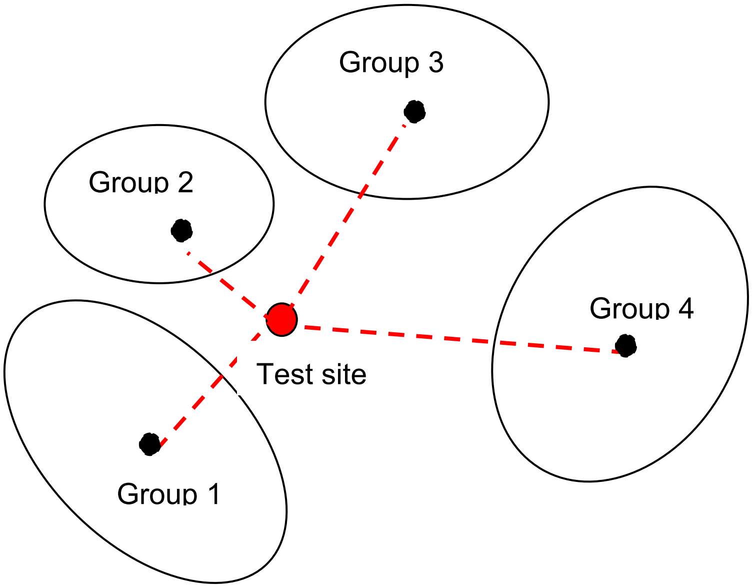

Figure 7. Graphical representation of the probability weighting of Group membership applied in SkeenRIVAS.

Table 4. Example calculation of the probability of a taxon occurring 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	0.3	60	18 (0.3 x 60)
2	0.4	80	32 (0.4 x 80)
3	0.1	60	6 (0.1 x 60)
4	0.1	10	1 (0.1 x 10)
			$\sum$ = Total Probability = 57%

Only taxa that had a 50% or greater probability of being present at a site were considered. The rationale here was to exclude taxa having a low chance of occurrence from the prediction, so that sampling variability would have a low impact on the sensitivity of the model. Enough taxa had to be included, however, to be able to measure a community's response to stress. A probability cutoff of 50% was considered adequate to achieve this sensitivity.

Site testing was done by comparing the observed number of taxa (O) at a site with the expected number (E) at that site. O was the number of taxa with more than a 50% chance of occurrence at a site, while E was the sum of the probabilities of those taxa predicted to occur at the test site. When all of the expected taxa occurred, the ratio of observed/expected (O/E) was close to one. In case of an unnatural change in the community, the number of observed taxa usually dropped and the O/E decreased. The acceptable range of O/E scores in SkeenRIVAS followed criteria reported by Simpson and Norris (2000), which was the range between the 10th and 90th percentiles of the reference sites. An O/E below the 10th percentile indicated an unnatural loss of taxa, and an O/E higher than the 90th percentile was judged to be richer than expected. The richer than expected sites were further reviewed for possible errors or site attributes that could explain relatively rich conditions.

Output was summarized using the same banding scheme that is used in the interpretation of AUSRIVAS output (Table 5, Figure 8). Band A was between the 10th and 90th percentiles. Band B started at the 10th percentile (typically about O/E=0.85) and had the same width as Band A. Band C had the same width again, whereas the width of Band D was determined as the difference between its start and an O/E of 0. Sites richer than reference (above the 90th percentile) were assigned to Band X.

Table 5. SkeenRIVAS banding scheme following that used in AUSRIVAS.

Band Label	Band name	Comments
X	Richer than reference	• More taxa found than expected • Potential biodiversity "hot-spot" • Possible mild nutrient enrichment
A	Reference	• Index value within the range of central 80% of reference sites
B	Stressed	• Fewer taxa than expected • Potential impact either on water quality or habitat quality or both, resulting in a loss of taxa
C, D	Severely stressed	• Many fewer taxa than expected • Loss of taxa due to substantial impacts on water and/or habitat quality • Severe impairment

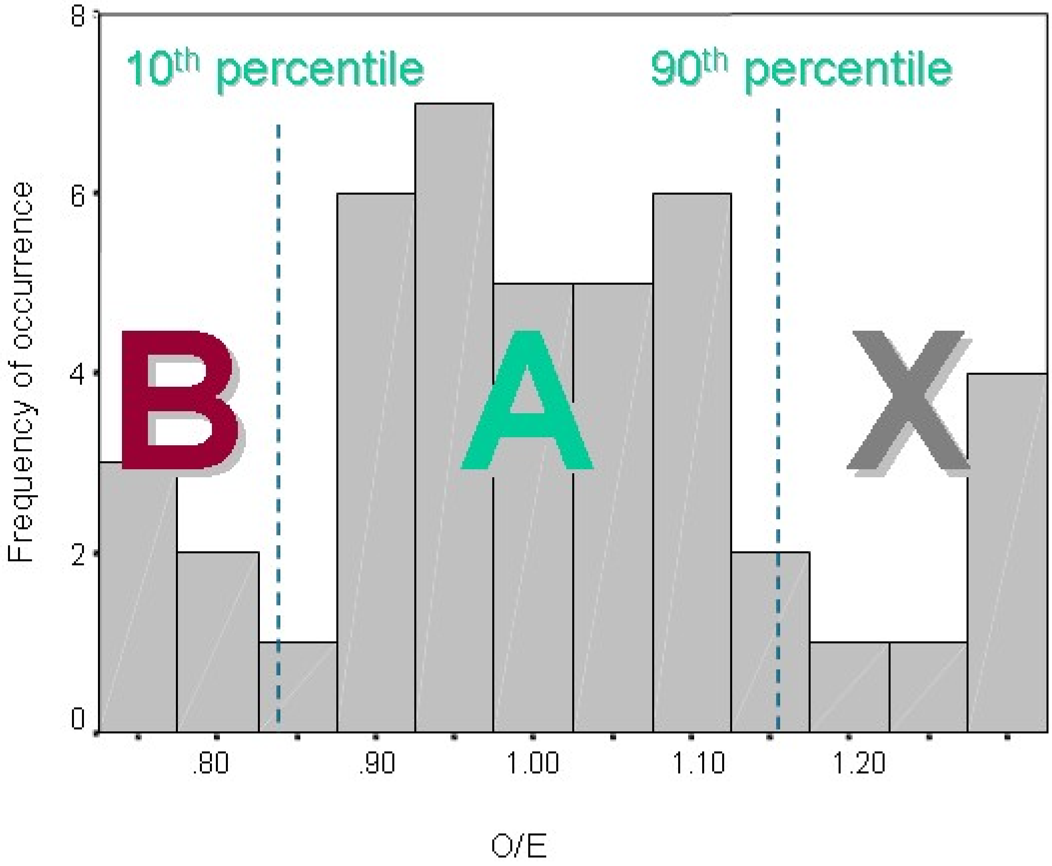

Figure 8. Illustration of the top Bands (A=reference condition, B= stressed, X= richer than reference) used in scoring a test site in SkeenRIVAS.

The quality of the SkeenRIVAS model was determined by plotting the observed number of taxa against the expected number of taxa for all reference sites (Figure 9). An acceptable model had a regression line pass through or close to the origin, with slope close to 1 (Figure 9a). This outcome indicated that on average all reference sites had an O/E of 1 and there was no bias. That is, reference sites with few taxa were predicted to have few taxa and vice versa reference sites with many taxa were predicted to have many taxa. Figure 9b shows a biased model, where the intercept is negative and the slope is greater than one. Biased models cannot properly predict very high or very low values, leading to error-prone assessments of unusual assemblages. Linke et al. (2005) showed that biased models were those having a slope outside of the interval of 0.85 to 1.15, or an intercept outside the range of –1.5 to 1.5.

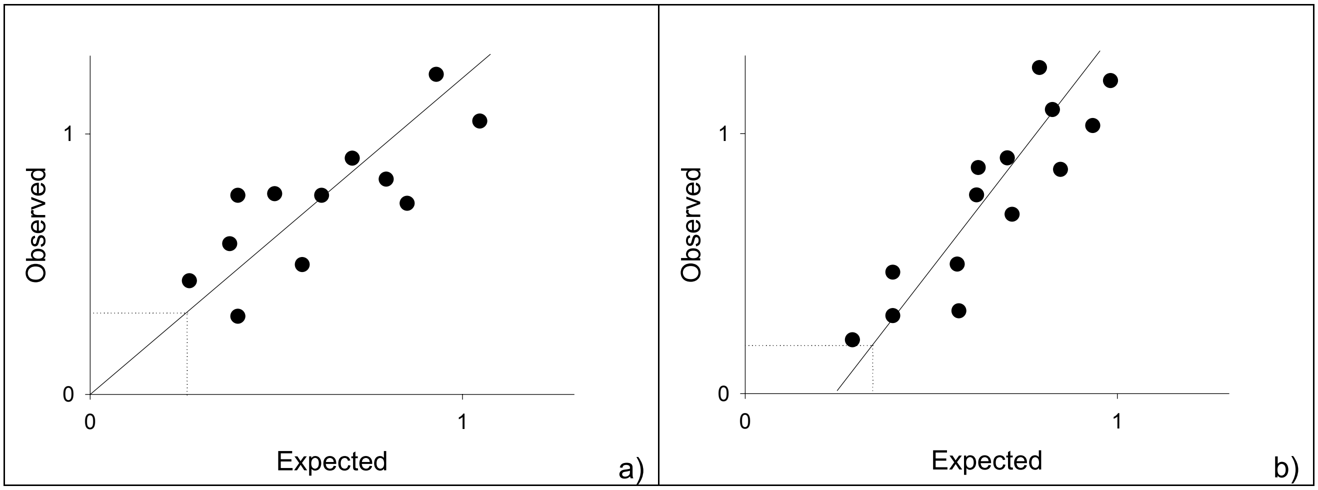

Figure 9. Example of regression lines used to assess the quality of the SkeenRIVAS model. The best model was one having the regression line pass through or close to origin as in plot "a". A poorer model having bias was one in which the regression line passed off of origin.

4.9.4 Differences between models

Differences between the modeling approaches are outlined in Table 6. SkeenRIVAS clustering to form sample groups was based on presence-absence data, which gave equal weight to all species present. BEAST clustering was based on abundance data, which gave more weight to rarer taxa. Initial groups identified using cluster analysis for BEAST were further evaluated using nonmetric multidimensional scaling (MDS) and a test of similarity between groups (ANOSIM in PRIMER), which was not done in SkeenRIVAS.

Using the discriminant functions, the distance between reference groups and each of the test sites was calculated. In BEAST, the test site was assigned to the closest reference group because that was the Group to which it had the highest probability of belonging. To assess the test site, an MDS was run to examine the similarity of the test site to reference sites in the selected group in ordination space. Confidence ellipses were drawn around the reference group (90%, 99% and 99.9%) to define thresholds for four categories of stress. A test site that fell within the 90% confidence ellipse was considered to be similar to reference condition while a test site that fell outside the 99.9% confidence ellipse was considered severely stressed. In SkeenRIVAS, the test site was compared to all the reference groups, although the distance from each reference group was used to weight the assessment. For each taxon, the observed to expected (O/E) ratio was calculated based on the distance from each reference group and the probability of that taxon occurring in that reference group (see Section 4.9.3). The final assessment of a test site depended on the O/E ratio for that site, compared to the bands that were created using the O/E ratios of the reference sites.

Table 6. Differences in RCA model development between BEAST and SkeeRivAs.

RCA MODEL DEVELOPMENT STEP/ COMPONENT	**BEAST**	**RIVAS**
Distance measure for quantifying biological similarity between pairs of samples	Bray-Curtis (relative abundance)	Bray-Curtis (presence-absence)
Demarcate reference sample groups using biological data	Agglomerative hierarchical clustering of abundance, in combination with nonmetric multidimensional scaling (MDS) and a test of similarity between sample groups (ANOSIM)	Agglomerative hierarchical (UPGMA) clustering
Find environmental features associated with each group	Both methods use discriminant function analysis with the *a priori* groups determined in above step (although groups may be different for each method because of differences in clustering and distance measures)	
Use discriminant functions to assign the probability of test sites belonging to each reference group	**Assumed discrete reference groups.** The Mahalanobis distance of a test site from each group centroid was calculated along with a probability of group membership. A test site was tested against the group to which is had the highest probability of membership.	**Assumed reference groups occur along a continuum.** The standardized distance of test site from each group centroid was calculated and used to create a weighted average probability of each taxon occurring at a test site. The final probability of a taxon occurring at a test site was a function of the distance from each reference group and the probability of the taxa occurring at sites in that reference group. .
Evaluation of test site	Ordinate test site against	Calculate a ratio of observed (O)

RCA MODEL DEVELOPMENT STEP/ COMPONENT	BEAST	RIVAS
	chosen reference group.	number of taxa to the expected (E) number of taxa based on all taxa that have a 50% probability of occurring.
Assessment Thresholds	Ellipses are drawn on plots to show the 90%, 99% and 99.9% confidence bands around the reference cloud.	The central 80% of reference O/E scores (mean ± 40%) are considered equivalent to reference. Bands are drawn in for sites below the 10th and above the 90th percentiles to depict assessment thresholds.
Final assessment	Depends on position of a test site relative to reference group confidence ellipses. The farther away a test site is from the reference cloud, the more it is considered to be 'stressed'.	Depends on the O/E score. Any site falling with an O/E in the central 80% about the mean reference O/E score is similar to a reference condition. The farther away a test site is from the central band of O/E scores, the more it is considered to be 'stressed'.

5 RESULTS

5.1 Assignment of Reference and Test Sites

A total of 256 complete observations were compiled from all years of sampling. Of this total, 86 observations (34% of the total) were from reference sites and 170 were from sites that were considered to have some degree of disturbance and were called "test" sites or were replicate reference sites that were used for tests of model accuracy and precision. The distribution of reference and test sites is shown in Figure 10.

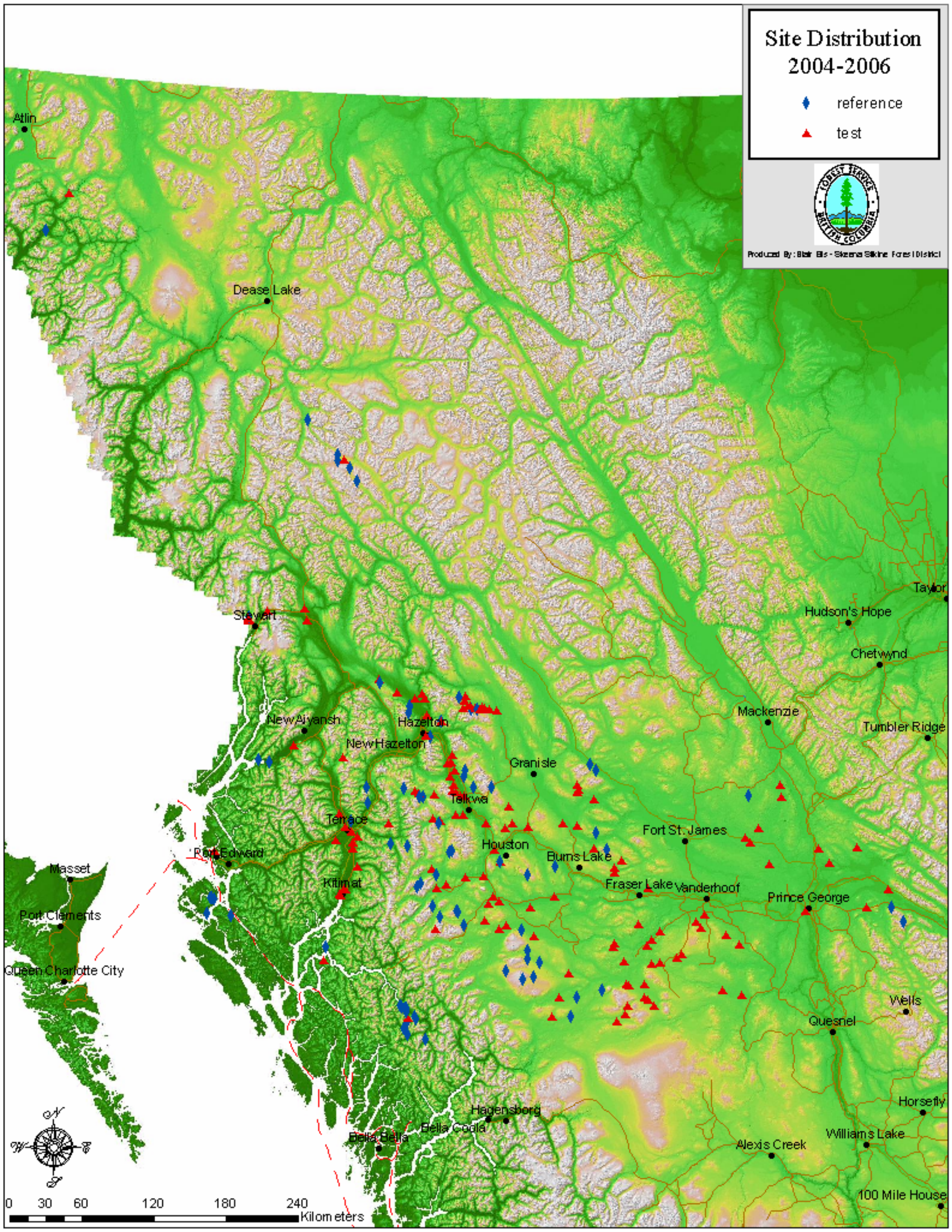

Figure 10. Distribution of reference and test sites that were sampled in 2004 – 2006.

5.2 Model Development and Site Testing

5.2.1 SkeenRIVAS

Eight sample groups were identified from the cluster dendrogram (Figure 11). Chaining was relatively high, but still below 10%. Apart from an outlier group (Group 8) and substantial overlap of Groups 2 and 3, separation in a multi-dimensional scaling plot was reasonable (Figure 12). Overlap of Groups 4 and 5 was also apparent in Figure 12, but these two groups did separate out when viewed in three dimensions.

The discriminant function analysis (DFA) resulted in a jackknife misclassification error rate of 39%. It is important to note that this misclassification error was dependent on the number of groups and is less robust than using a validation dataset to test the model.

The variables selected by the discriminant function analysis as being the best discriminators of sample groups and thus the best “environmental predictors” are listed in Table 7.

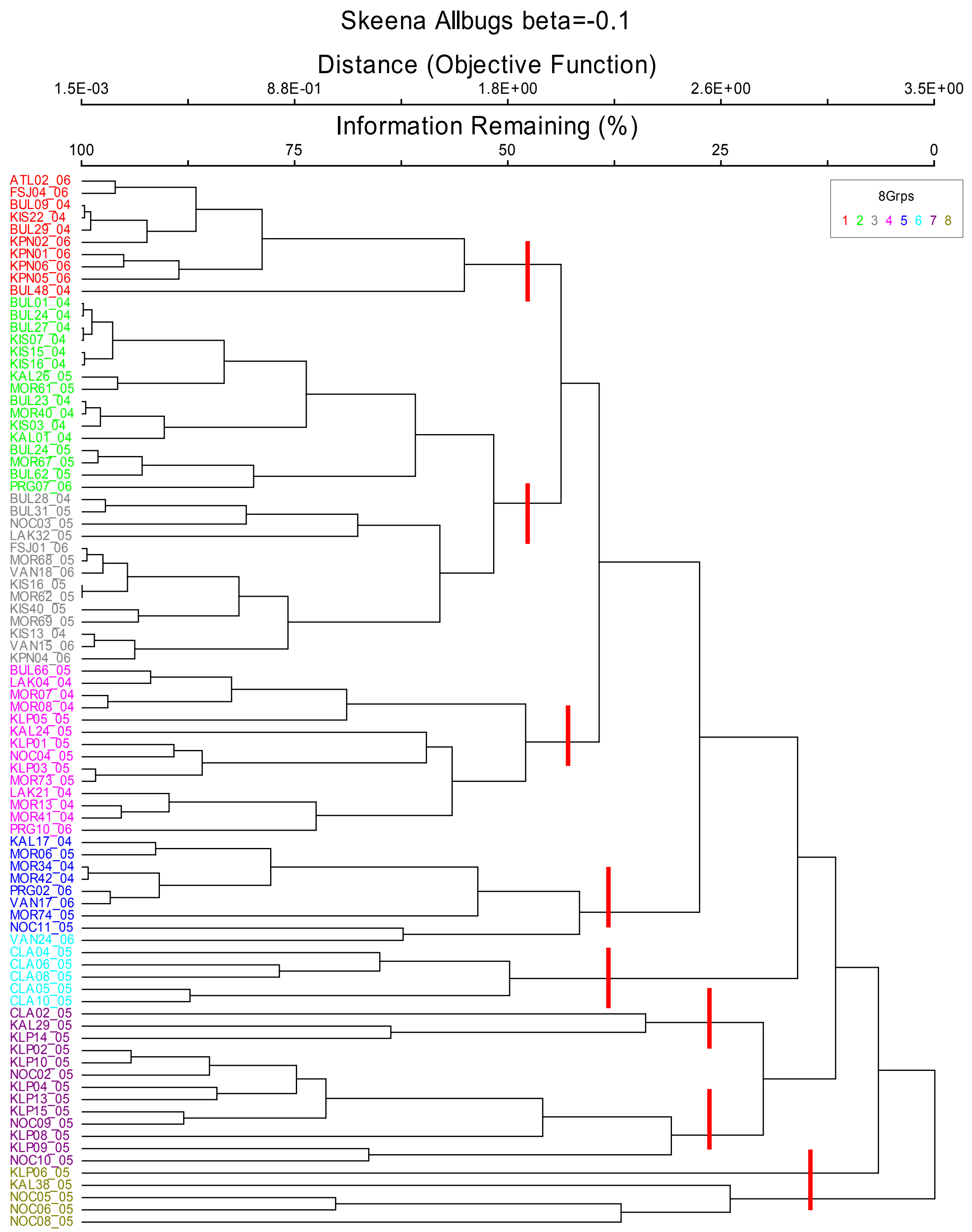

Figure 11. Dendrogram of reference sites using SkeenRIVAS methods.

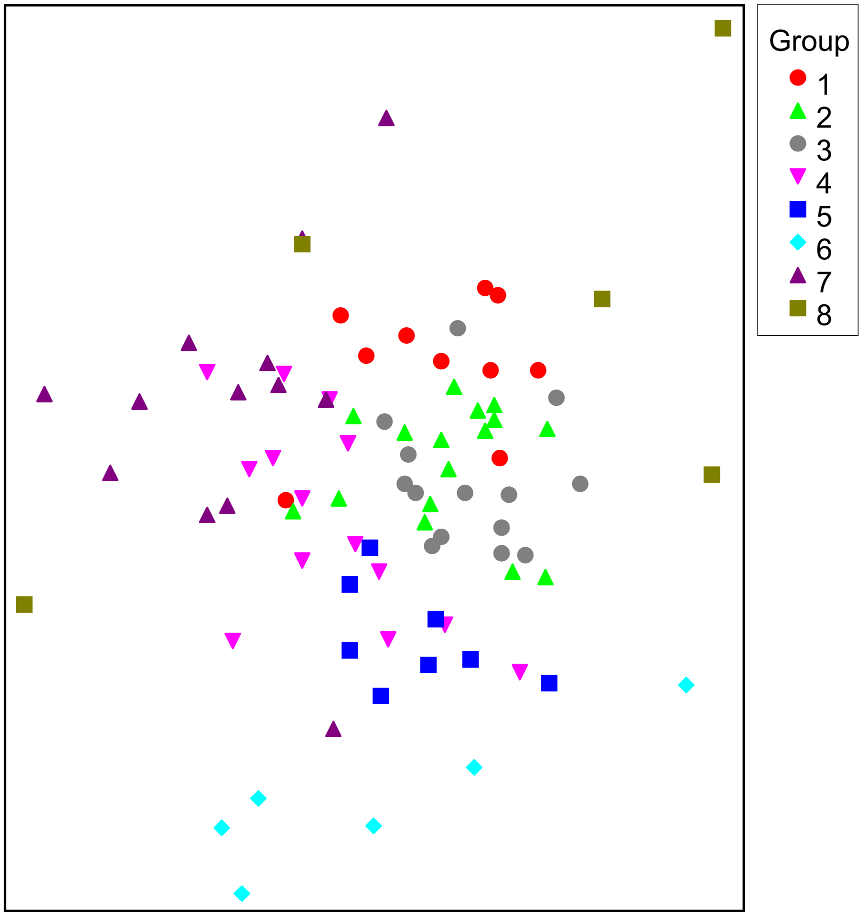

Figure 12. Non-metric multidimensional scaling plot to test for group separation. Stress was 0.17 indicating that the 2-D ordination was an acceptable representation of the multidimensional data.

Table 7. Predictor variables found by DFA as the best discriminators of the sample groups in SkeenRIVAS procedures.

X indicates selection of the variable in the final model	**Natural gradient variable**	**Description**
	Area (m2)	Area of watershed - square meters
	Pct_Wtlnd	Percent of wetland area in watershed
X	Pct_Lake	Percent of lake area in watershed
X	Pct_Rvr	Percent of river area in watershed
	Rvr_Lngth	Length of rivers (approximation)
X	Pct_Ice	Percent of ice in watershed
	Strdefin_Lgth	Definite stream length
X	SL_ratio	ratio of definite stream lgth to total stream length (does this include intermittent streams?)
	StrRiv_DrnDnsty	Drainage density of all rivers and streams
	Pct_tot_sedimentary	percent of watershed with sedimentary rock
X	Pct_tot_intrusive	percent of watershed with intrusive rock
	Pct_tot_volcanic	percent of watershed with volcanic rock
	Pct_other _rocks	percent of watershed with rock other than sedimentary, intrusive, volcanic
	Pct_Alpine	Percent of alpine area in watersheds
X	Pct_Ava	Percent of avalanche chute area in watersheds
X	Latitude	Latitude
	Longitude	Longitude
X	Min_Elev	minimum elevation of the watershed measured from the sampling site. This measure is the elevation of the sampling site
	Relief	Maximum minus minimum elevation
	Pct_SC4	Percent of watershed area with slope class 4 – define
	Pct_SC5	Percent of watershed area with slope class 5 - define
X	WT	water temperature at time of sampling
X	% gradient	gradient measured in percent
	% pools	percent pools
	% glides	percent glides
	% riffles	percent riffle
	% cascades	percent cascade
	BWAve	average bankfull width (from 3 field measurements)
	dom_substrate	dominant substrate particle size by percent composition
X	dominant Rveg	dominant riparian vegetation - define

Apart from a few outliers, prediction of O/E was very good. The regression r^2 was 0.54, which was considered high. In a review of AUSRIVAS models that have been in use in Australia for 10 years, only 4 out of 15 reviewed models have a higher r^2 (Linke et. al. 2005). Slope and intercept were within the boundaries set as optimal by Linke et al 2005, indicating no bias (Figure 13).

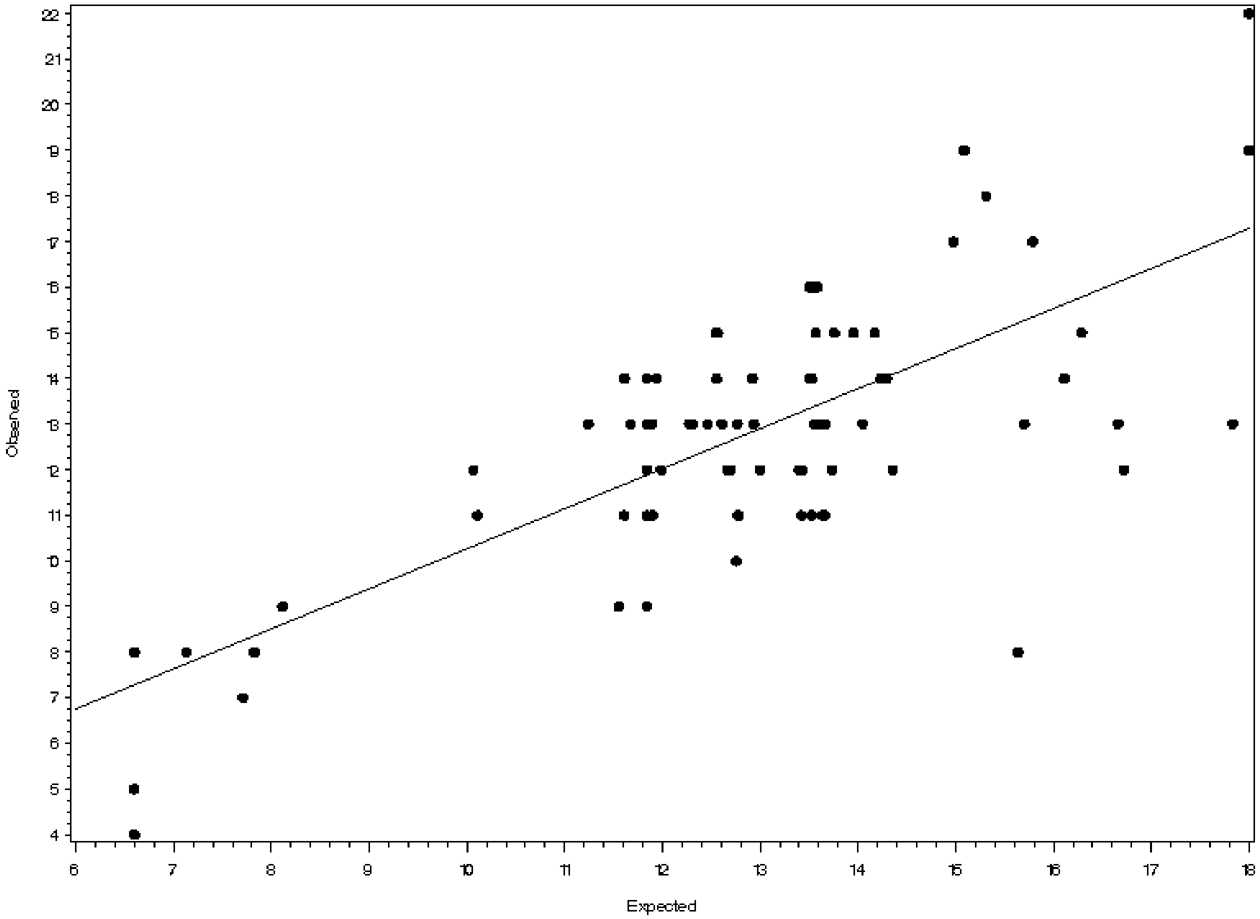

Figure 13. Regression analysis for observed vs. expected number of taxa in reference site samples. Intercept=1.48, slope=0.88, r^2=0.543.

A frequency distribution of O/E was centered around the mean (Figure 14), indicating a relatively even distribution of errors in the reference site assessment. However, there were two outliers having O/E scores $\leq$ 0.6. Once these were removed, the Band A cutoff was 0.82. As there were still a few suspicious sites and the bandwidth was wider than standard models (see Linke et al. 2005, van Sickle et al 2005), we introduced an extra category ‘Slightly impaired’ which was analogous to a BEAST category of the same name.

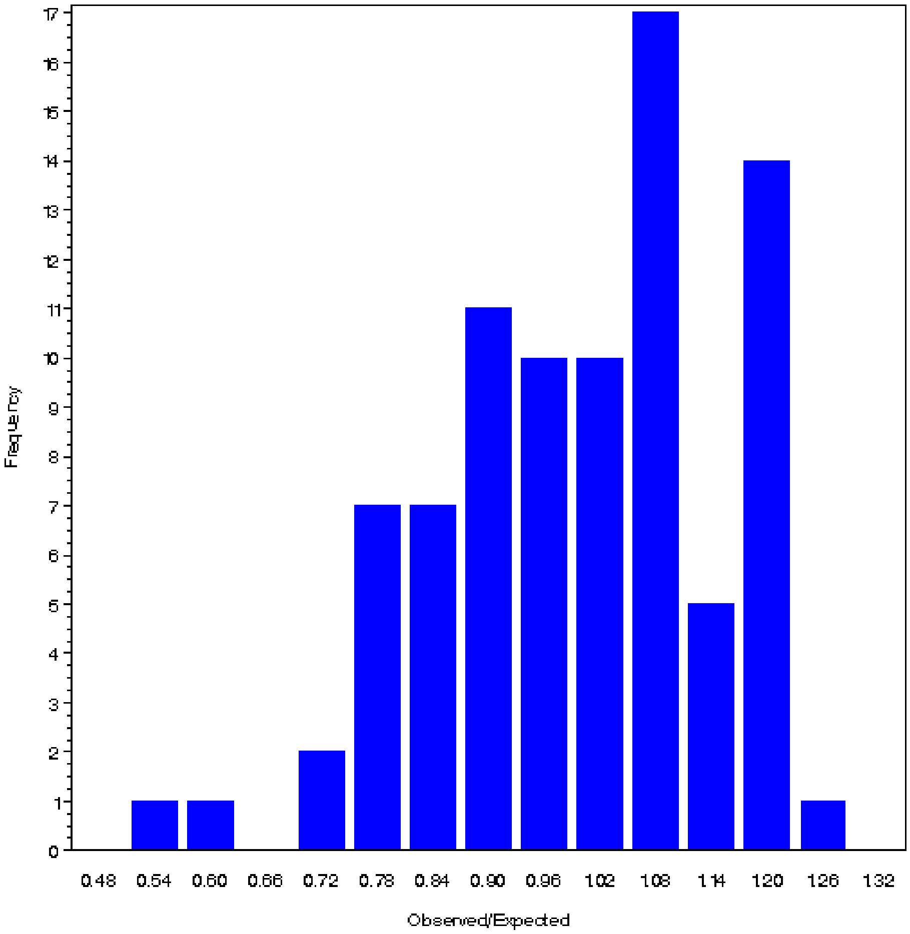

Figure 14. Frequency distribution of O/E on the reference sites. The 10th percentile after removal of the two outliers <0.66 was 0.82.

This change resulted in the following bands:

- Band A: 0.88 - 1.16 indicating the reference condition
- Band B: 0.82 - 0.88 indicating slightly stressed condition
- Band C: 0.6 - 0.82 indicating significantly stressed condition
- Band D: < 0.6 indicating severely stressed condition

5.2.2 BEAST

Clustering and the MDS revealed 6 sample groups separating at similarity levels of 42-60% and 5 outliers (Figures 15 and 16). Those outliers were KAL38 (E260426_05_1), KLP06 (E256735_05_1), NOC10 (E260480_05_1), VAN24 (E263889_06_1), and MOR74 (E260697_05_1). Group 1 samples were well separated from the others. All these samples were from Tweedsmuir Park. Groups 2 and 3 separated from each other and from the other groups approximately at the 50% level of similarity. Similarity between the other groups was slightly higher and appeared as a gradient on the MDS ordination with Group 6 being more similar to Groups 2 and 3 than it was to Group 5. Group 4 samples formed a bridge between Groups 6 and 5.

Analysis of similarities (ANOSIM) revealed a global *R* value of 0.77 and it was significant at <0.1%. Pairwise comparisons of the sample groups showed that Groups 6 and 2 had substantial overlap (R=0.5) but there was good dissimilarity among all other Group pairs (R=0.6 to 1).

Using the backwards and forwards stepping DFA procedures, a model was found having 5 predictor variables (pct_lake, pct_tot_volc rock, longitude, min_elevation, and relief) and high tolerance values (all >0.71). Reclassification success by the jackknife procedure was good for Groups 1 to 5 (60-80%) but it was poor for Group 6 (9%).

The combination of this poor reclassification success for Group 6 and the evidence from ANOSIM of relatively high similarity between Groups 2 and 6 was justification to merge Groups 6 and 2. Further iterations of the DFA and ANOSIM with samples reassigned to new groups resulted in four sample groups. ANOSIM indicated $R \geq 0.6$ among all final group pairs (Table 8). Group 1 was the same as before. Group 2 was the combination of the old Groups 2, 3, and 6. Group 3 was the old Group 4. Group 4 was the old Group 5. Iterations of the 4-Group DFA resulted in five predictor variables (pct_wtlnd, pct_lake, pct_tot_volc, longitude, and min_elev). The final model was highly significant ($P<0.001$) and the overall jackknife reclassification success was 0.78. All predictor variables had very high tolerance values (0.85 – 0.96), indicating virtually no correlation between predictor variables. The first canonical variable explained 75% of the total dispersion, which was very high. All F values ranged from 4 to 30. The order of predictor variables contributing most to least in discriminating the sample groups based on the F-to-remove values was Pct_lake > Min_elev > longitude > Pct_wtlnd > Pct_tot_volc (Table 9). This model met all of our test criteria and thus was accepted.

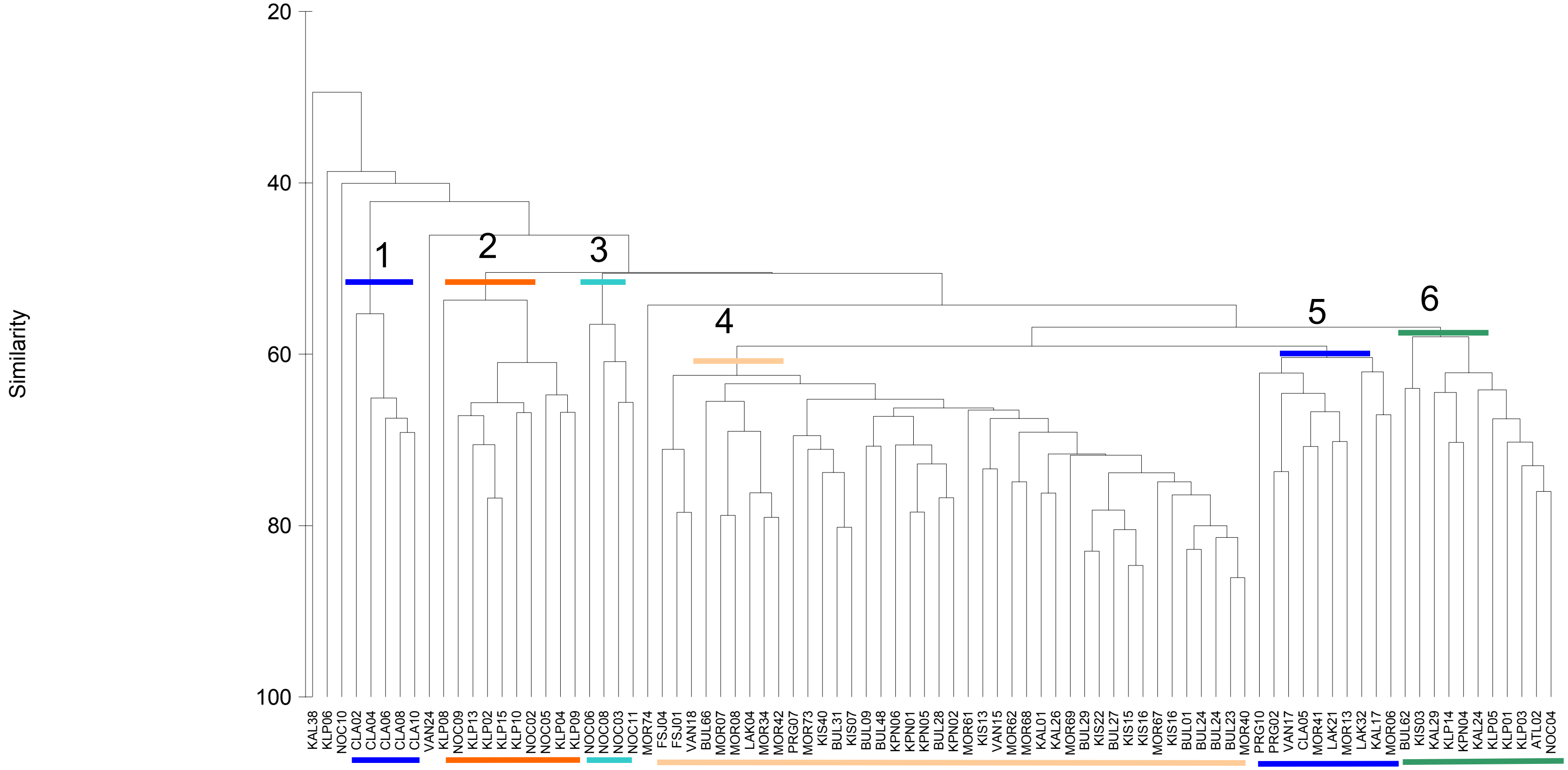

Figure 15. Cluster dendrogram of 4th root transformed abundance data from all reference sites. The image indicates 6 sample groups defined with colour coding and horizontal blue lines and 5 outliers (KAL38, KLP06, NOC10, VAN24, and MOR74).

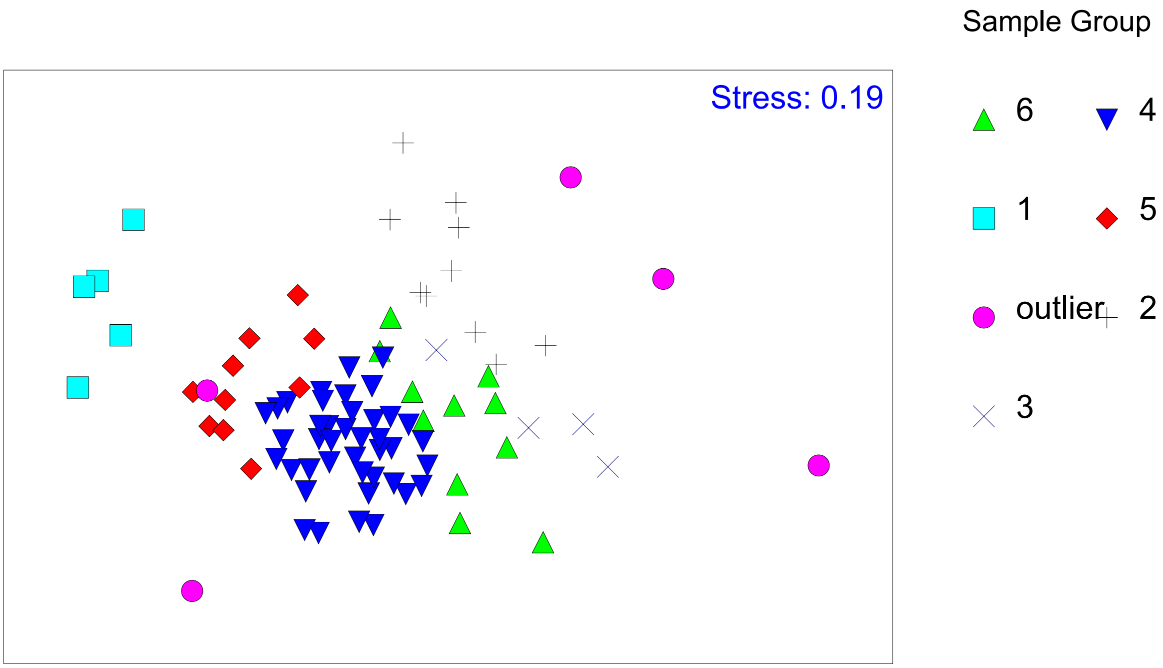

Figure 16. Ordination of samples by sample group defined in the cluster analysis in Figure 15.

Table 8. Results of pairwise tests of similarity among re-assigned sample groups using ANOSIM. Sample Group 1 was the original Group 1. Sample Group 2 was the combination of the original Groups 2, 3, and 6. Sample Group 3 was the original Group 4. Sample Group 4 was the original Group 5.

Sample group contrast	***R***	**Significance level (%)**
2 and 3	0.6	0.1
2 and 1	1	0.1
2 and 4	0.6	0.1
3 and 1	1	0.1
3 and 4	0.6	0.1
1 and 4	0.7	0.1

Table 9. Final model predictor variables with F-to-remove and tolerance values.

Predictor variable	F-to-remove value	Tolerance
Pct_wtlnd (Percent wetland by area in watershed)	4.2	0.85
Pct_lake (Percent lake by area in watershed)	30.37	0.96
Pct_tot_volc (Percent of watershed rock types by area comprised of volcanic rock)	4.04	0.94
Longitude	4.95	0.94
Min_elev (Sample site elevation)	16.71	0.86

There were clear differences in the composition and abundance among taxa between each of the sample groups (Figure 17). Group 1 samples (n=5) had a mean invertebrate abundance of 19,235 individuals/sample, which was several times that of any of the other groups. Invertebrates contributing most to the large numbers included the chironomids, Ephemeroptera (mayflies), Tricoptera (caddisflies), and Plecoptera (stoneflies). Unique to this group were large numbers of the Haplotaxida and Lumbriculida worms that comprised most of the “other” category in Group 1. Smaller abundances of the Arachnida (water mites), Coleoptera (beetles), and non-chironomid dipterans (true flies) comprised the remainder of this group. Group 2 (n=27) had the lowest mean invertebrate abundance among all groups of only 755 individuals/sample. Most common taxa in this group were the Ephemeroptera (mayflies), Plecoptera (stoneflies), and chironomids. Minor taxa were the Tricoptera (caddisflies), non-chironomid dipterans, water mites, and worms. Group 3 samples (n=41) had a mean animal abundance of 3,020 individuals/sample. Most common taxa in this group were again the Ephemeroptera (mayflies), Plecoptera (stoneflies), and chironomids as were found in Group 2 but they occurred in greater abundance. The Tricoptera (caddisflies) were more abundant than in Group 2. All other taxa were rare or absent. Group 4 (n=10) had moderate mean invertebrate abundance of 6,839 individuals/sample or approximately double the abundance found in Group 3 samples. Again, the Ephemeroptera (mayflies), Plecoptera (stoneflies), and chironomids were prominent. The Tricoptera (caddisflies) were less common but still occurred at an average of more than 6% of total abundance among all samples. The non-chironomid dipterans were relatively more common in Group 4 than in the other groups. Unlike Groups 2 and 3 but similar to Group 1, the Coleoptera (beetles) were present in Group 4. All other taxa were rare or absent in Group 4 samples.

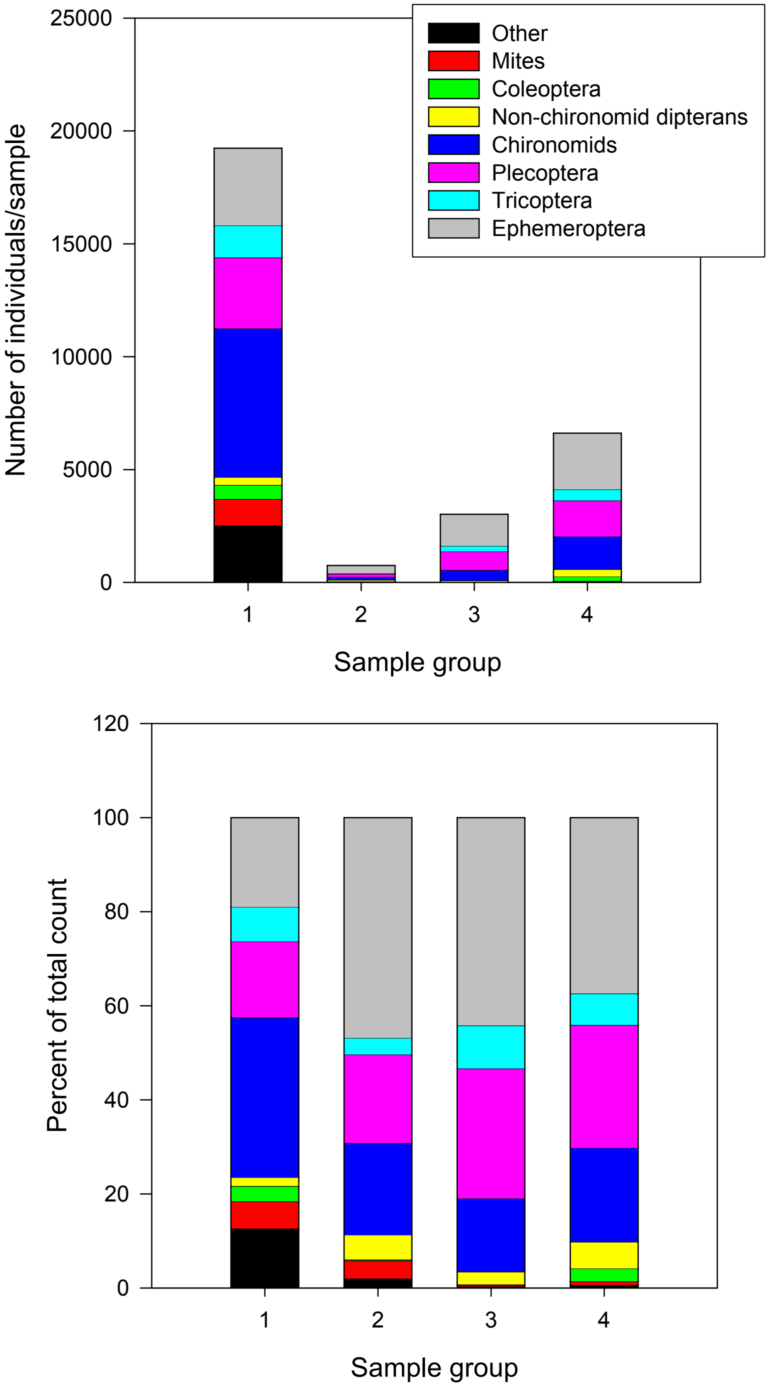

Figure 17. Taxonomic composition of BEAST sample groups from reference sites. Data are shown as mean number of individuals (top) and percentages (bottom) of the total count. The category called "other" included the Haplotaxida and Lumbriculida worms, Amphipoda (amphipods), Collembola (springtails), Hemiptera (true bugs), Megaloptera (alderflies), Odonata (damselflies), Hydrazoa (hydroids), Bivalvia (clams), and Gastropoda (gastropods).

The DFA indicated that five variables (percent wetland, percent lake, percent volcanic rock, longitude, and elevation of the sampling site) were the best discriminators of habitat between the sample groups. Table 10 shows the between-group differences of these attributes. Group 2 sites were mostly devoid of lakes and wetland higher in the watersheds and volcanic parent materials were largely absent. Most of these sites were close to sea level and were furthest west, located mainly along the coast. In contrast, Group 1 sites were at high elevations on the north-central plateau where volcanic parent materials were prominent, representing more than 70% of the watershed areas. Wetland and lakes represented a higher percentage of drainage areas than in any other sample group. Group 3 sites were mostly at moderate elevations (median of 880 m) covering a wide range of longitude mainly in the Central Interior and Sub-boreal Interior Ecoprovinces. Wetlands and lakes were present upstream of Group 3 sites but their areal extent was small, each occurring at a median of only 0.2% of drainage areas. Volcanic parent materials were common but patchy, as interpreted from the wide range of values of percent volcanic rock, occurring at a median of 31% of all rock types in the drainages. Group 4 sites occurred at moderate to high elevations again over the Interior ecoprovinces. The volcanic parent materials were patchy but common, occurring at a median of 35% of watershed area. The areal extent of wetlands in these drainages was relatively high (3.4% of watershed area) and greater than that of lakes (0.7% of watershed area), which was opposite of the differences between areas of wetlands and lakes found in the other groups.

Table 10. Median and range values for predictor variables among sample groups.

Predictor variable	**Natural gradient variable median and range**			
	Group 1	**Group 2**	**Group 3**	**Group 4**
Percent wetland	4.7 (0 to 5.3)	0 (0 to 16.1)	0.2 (0-5.2)	3.4 (0.3 to 9.2)
Percent lake	7.5 (0.8 to 7.6)	0.1 (0 to 4.2)	0.2 (0-3.9)	0.7 (0 to 3.6)
Percent volcanic rock	73.5 (60 to 100)	0 (0 to 48)	30.8 (0 to 100)	34.7 (0 to 100)
Longitude	-126.282 (-126.559 to -126.129)	-128.109 (-133.270 to -126.960)	-127.432 (-129.366 to -121.663)	-126.105 (-128.573 to -121.520)
Sample site elevation (m)	1360 (1020 to 1720)	20 (20 to 1320)	880 (200 to 1280)	990 (240 to 1280)

5.2.3 Results of site testing

Results of the site testing are shown in Table 11. The list includes results from each of the BEAST and SkeenRIVAS models.

Table 11. Results of site testing using the SkeenRIVAS and BEAST models. Site name was the common name for the sampling site. Site code was an identifying code. Ref-test indicates whether the site was from a reference (ref) or test (test) location or from a reference site used for model testing (r-test).

Site name	Site code	Ref-test	OE50	SkeenRIVAS assessment	BEAST group	BEAST assessment
Driftwood(dup)	BUL24	ref	0.94	not stressed	3	not stressed
Steep Canyon Ref. Rep 2	KIS15	ref	1.03	not stressed	3	not stressed
Steep Canyon Ref. Rep 3	KIS15	ref	1.18	richer than reference	3	not stressed
Compass Cr. #2	KIS22	ref	0.86	slightly stressed	3	not stressed
Compass Cr. #3	KIS22	ref	0.94	not stressed	3	not stressed
Keazoah	NOC06	ref	1.06	not stressed	2	not stressed
Capoose Creek	VAN17	ref	1.19	richer than reference	4	not stressed
Station Reference	KIS40	r-test	1.24	richer than reference	2	not stressed
Sibola 1.5 km	MOR06_04	r-test	0.86	slightly stressed	4	not stressed
Deep Cr D/S Bridge	BUL03	test	1.27	richer than reference	4	not stressed
Deep Cr D/S Bridge	BUL03_04	test	0.88	not stressed	4	not stressed
Sinclair Cr.	BUL07	test	0.98	not stressed	3	not stressed
Jonas Cr.	BUL10	test	0.81	stressed	3	not stressed
Howson Cr.	BUL11	test	0.82	slightly stressed	3	not stressed
Tenas Creek	BUL15	test	1.14	not stressed	3	slightly stressed
Goathorn above Tenas	BUL16	test	0.95	not stressed	3	stressed
Goathorn Cr.	BUL16_04	test	1.00	not stressed	3	not stressed
Goathorn above Tenas	BUL16_06	test	1.18	richer than reference	3	stressed
Chicken Cr	BUL19	test	0.91	not stressed	3	slightly stressed
Kathlyn Creek	BUL20	test	0.98	not stressed	3	slightly stressed
Unnamed @49 km	BUL31	test	0.92	not stressed	3	not stressed
Toboggan d/s	BUL33_04	test	1.00	not stressed	3	not stressed
Toboggan d/s	BUL33_05r1	test	0.99	not stressed	3	not stressed
Toboggan d/s	BUL33_05r2	test	0.95	not stressed	3	not stressed

Site name	Site code	Ref-test	OE50	SkeenRIVAS assessment	BEAST group	BEAST assessment
Gramophone Cr.	BUL37	test	1.04	not stressed	2	not stressed
Causqua Cr.	BUL40	test	0.59	severely stressed	2	slightly stressed
Corya Cr.	BUL41	test	0.84	slightly stressed	3	not stressed
Kwun	BUL42	test	0.99	not stressed	3	not stressed
Canyon Creek	BUL49	test	0.90	not stressed	2	not stressed
Sandstone	BUL50	test	0.82	stressed	1	slightly stressed
Sandstone	BUL50_04	test	1.06	not stressed	1	stressed
Coal Cr.	BUL51	test	1.07	not stressed	4	not stressed
Toboggan u/s	BUL52	test	0.97	not stressed	3	not stressed
Toboggan u/s	BUL52_r1	test	1.04	not stressed	3	not stressed
Toboggan u/s	BUL52_r2	test			3	slightly stressed
Toboggan u/s	BUL52_r3	test			3	stressed
Nichyeskwa 26k	BUL54	test	0.95	not stressed	3	not stressed
Glacier Gultch	BUL55	test	0.80	stressed	3	stressed
Willow	BUL56	test	0.97	not stressed	4	not stressed
Nichyeskwa	BUL57	test	1.09	not stressed	3	not stressed
Nichyeskwa 12k	BUL58	test	0.98	not stressed	3	slightly stressed
Nichyeskwa 8k	BUL59	test	1.18	richer than reference	3	not stressed
Nichyeskwa 14k	BUL60	test	0.93	not stressed	3	stressed
Nichyeskwa 18k	BUL61	test	1.07	not stressed	3	slightly stressed
Nichyeskwa 22k	BUL63	test	1.07	not stressed	3	not stressed
Nichyeskwa 34k	BUL64	test	1.14	not stressed	3	not stressed
Nichyeskwa Gate	BUL65	test	1.03	not stressed	3	not stressed
RC40 Suskwa FSR @ 1.5k	BUL68	test	1.13	not stressed	3	slightly stressed
Nichyeskwa 26.6k	BUL69	test	1.13	not stressed	3	not stressed
Pine Creek	BUL70	test	0.87	slightly stressed	3	not stressed
CLA12195	CLA12	test	1.00	not stressed	1	not stressed

Site name	Site code	Ref-test	OE50	SkeenRIVAS assessment	BEAST group	BEAST assessment
EUC07196	EUC07	test	0.81	stressed	4	stressed
EUC08196	EUC08	test	0.91	not stressed	4	stressed
RubyRock Creek	FSJ02	test	0.61	stressed	1	not stressed
Flemming Creek	FSJ03	test	0.75	stressed	4	slightly stressed
400 Rd @ 431	FSJ05	test	0.94	not stressed	4	not stressed
1600 Rd. Trib	FSJ06	test	0.88	slightly stressed	4	slightly stressed
Salmon River	FSJ07	test	0.91	not stressed	1	slightly stressed
Thornhill @ Skeena	KAL04	test	0.93	not stressed	2	slightly stressed
Anweiler @ bridge	KAL05	test	0.91	not stressed	2	not stressed
Luncheon d/s	KAL15	test	1.11	not stressed	2	not stressed
Spring	KAL20	test	1.03	not stressed	2	not stressed
North Kleanza	KAL23r1	test	1.02	not stressed	3	not stressed
North Kleanza (Dup)	KAL23r2	test	0.95	not stressed	3	not stressed
Ansedagon	KAL27	test	1.18	richer than reference	2	not stressed
Cascade cr. d/s Logan	KAL28	test	0.94	not stressed	2	not stressed
Cooper/Flecher	KAL30	test	1.11	not stressed	2	not stressed
Granite Creek d/s 1st Ave	KAL31	test	0.76	stressed	2	not stressed
Hannah Creek N	KAL32	test	1.17	richer than reference	2	slightly stressed
Hanna	KAL32_05	test	1.02	not stressed	2	not stressed
Hannah Creek N	KAL32_06	test	0.93	not stressed	2	slightly stressed
Salmon R. above Cascade	KAL33	test	0.61	stressed	2	stressed
Sockeye u/s	KAL34	test	1.17	richer than reference	2	not stressed
Surprise Cr.	KAL36	test	0.85	slightly stressed	2	not stressed
Surprise Cr. Dup.	KAL36	test			2	not stressed
Tintina S	KAL37_06	test	0.77	stressed	2	slightly stressed
Williams @ Hwy	KAL39	test	1.00	not stressed	2	not stressed

Site name	Site code	Ref-test	OE50	SkeenRIVAS assessment	BEAST group	BEAST assessment
Cataline Cr.	KIS02	test	0.87	slightly stressed	3	not stressed
McKuthcheon Cr	KIS04	test	1.07	not stressed	2	not stressed
Sterrit Cr.	KIS05	test	0.90	not stressed	2	not stressed
Pinenut Cr.	KIS06	test	0.87	slightly stressed	2	slightly stressed
Shegunia Trib 250m d/s	KIS09	test	0.99	not stressed	3	not stressed
Steep Canyon Cr.	KIS14	test	0.98	not stressed	2	not stressed
Helen @ 19 km	KIS18	test	0.89	not stressed	2	not stressed
Murder Cr	KIS21	test	1.05	not stressed	2	slightly stressed
Station d/s	KIS43	test	1.14	not stressed	2	not stressed
Allin d/s	KIS48	test	1.07	not stressed	4	not stressed
Didene-Fox Creek	KPN03	test	1.01	not stressed	3	not stressed
Kitimat River d/s Eurocan (750m d/s	KTM07	test	0.25	severely stressed	2	severely stressed
Kitimat River @ Rec. S.	KTM08	test	0.80	stressed	2	not stressed
Killutsal Creek	KTM09	test	1.27	richer than reference	2	not stressed
Kitimat R.	KTM10	test	1.03	not stressed	2	slightly stressed
Hirsh D/S	KTM11	test	0.76	stressed	2	not stressed
Hirsh U/S	KTM12	test	0.88	not stressed	2	not stressed
John Creek	LAK02	test	1.06	not stressed	4	not stressed
Coldwater Cr.	LAK03	test	0.92	not stressed	4	not stressed
Pinkut	LAK05	test	0.80	stressed	1	not stressed
Twain Cr	LAK11	test	0.95	not stressed	4	not stressed
Rat Cr.	LAK13	test	0.99	not stressed	3	not stressed
Roof Cr.	LAK14	test	1.09	not stressed	4	not stressed
Phantom Creek	LAK33	test	0.89	not stressed	3	slightly stressed
Tidsley Creek	LAK34	test	1.12	not stressed	3	not stressed
Wosket Creek (115)	LAK36	test	0.95	not stressed	3	not stressed
Nadina R	MOR12_04	test	1.09	not stressed	1	stressed

Site name	Site code	Ref-test	OE50	SkeenRIVAS assessment	BEAST group	BEAST assessment
Nadina	MOR12_05	test	1.14	not stressed	1	not stressed
Richfield Cr.	MOR20	test	0.93	not stressed	4	not stressed
Byman Cr.	MOR24	test	0.91	not stressed	4	not stressed
McQuarrie	MOR26	test	0.92	not stressed	1	not stressed
Buck 12km	MOR33	test	1.03	not stressed	4	not stressed
Bulkley @ Morice (upper)	MOR37	test	0.90	not stressed	4	slightly stressed
Shea Cr.U/S	MOR39	test	1.07	not stressed	4	not stressed
Lamprey Rec Site	MOR45r1	test	0.98	not stressed	4	stressed
Lamprey Rec Site	MOR45r2	test	0.88	slightly stressed	4	severely stressed
Lamprey Rec Site	MOR45r3	test	0.88	slightly stressed	4	stressed
Owen Cr Lower	MOR50	test	0.93	not stressed	4	slightly stressed
Guess Cr	MOR53	test	0.97	not stressed	4	not stressed
McBride Creek	MOR58	test	1.00	not stressed	1	slightly stressed
Isac 1	MOR59	test	1.02	not stressed	3	slightly stressed
Isac 2	MOR60	test	1.11	not stressed	4	not stressed
Nado Creek	MOR63	test	1.15	not stressed	3	not stressed
Peter Aleck Creek	MOR64	test	1.08	not stressed	4	not stressed
8km Andrew Bay	MOR65	test	0.89	not stressed	4	not stressed
Bergfar Field	MOR66	test	0.66	stressed	3	severely stressed
Pimpernell Ck	MOR70_05	test	1.19	richer than reference	4	not stressed
Shelford	MOR71	test	0.95	not stressed	4	not stressed
Shelford	MOR71r1	test	0.90	not stressed	4	not stressed
No Mans Creek	MOR72	test	1.20	richer than reference	3	not stressed
Haymeadow Creek	MOR75	test	1.15	not stressed	4	not stressed
Gate Creek	MOR76_05	test	1.11	not stressed	4	not stressed
Old Field Cr.	NOC07	test	0.81	stressed	2	slightly stressed

Site name	Site code	Ref-test	OE50	SkeenRIVAS assessment	BEAST group	BEAST assessment
Old Field Cr(dup)	NOC07	test				NOT DONE
West Creek D/S	PRG01	test	1.00	not stressed	4	slightly stressed
FREP 17	PRG03	test	0.88	not stressed	4	not stressed
Darby rd @ 6.5km	PRG04	test	0.74	stressed	4	slightly stressed
Seebach Creek @ 19k	PRG05	test	1.04	not stressed	4	not stressed
Prichard Creek	PRG06	test	0.84	slightly stressed	4	not stressed
Purden Outlet	PRG08	test	1.00	not stressed	1	not stressed
Frep 27	PRG09	test	0.81	stressed	4	not stressed
McMillan Creek	PRG11	test	0.49	severely stressed	4	severely stressed
Fraser Trib @ 60 Dup	PRG12	test	0.95	not stressed	4	not stressed
Fraser Trib @ 60	PRG12r1	test	0.97	not stressed	4	not stressed
Lucas Creek	VAN01	test	0.78	stressed	1	not stressed
Red 4000	VAN02	test	0.91	not stressed	4	slightly stressed
Marrilla @21 km	VAN03	test	0.95	not stressed	4	slightly stressed
Bird Lake Trib.	VAN04	test	1.00	not stressed	1	stressed
Bobtail Connector	VAN05	test	0.69	stressed	1	not stressed
Sinkut R.	VAN06	test	1.11	not stressed	4	not stressed
Corkscrew Creek	VAN07	test	0.90	not stressed	4	not stressed
Sinkut R. Trib	VAN08	test	0.84	slightly stressed	4	slightly stressed
Camp rd Trib	VAN09	test	0.97	not stressed	4	not stressed
Eco Reene-Meridian rd(Vanderhoof)	VAN10	test	0.54	severely stressed	4	stressed
Roy Creek	VAN13	test	0.66	stressed	4	slightly stressed
Big Bend Trib	VAN14	test	1.08	not stressed	4	not stressed
Cabin Creek	VAN16	test	0.97	not stressed	4	not stressed
Chedakuz Trib	VAN19	test	0.88	slightly stressed	4	not stressed

Site name	Site code	Ref-test	OE50	SkeenRIVAS assessment	BEAST group	BEAST assessment
Chedakuz Trib	VAN19r1	test	0.97	not stressed	4	not stressed
Chedakuz Upper Trib	VAN20	test	1.04	not stressed	4	not stressed
Cutoff Creek	VAN21	test	0.91	not stressed	4	not stressed
Davidson Creek	VAN22	test	1.02	not stressed	4	not stressed
Davidson Trib.	VAN23	test	0.94	not stressed	4	slightly stressed
Earhorn Creek	VAN25	test	1.09	not stressed	4	slightly stressed
Fawnie Trib Dup	VAN26	test	0.95	not stressed	4	not stressed
Fawnie Trib.	VAN26r1	test	0.74	stressed	4	not stressed
Fawnie Trib #2	VAN27	test	1.15	not stressed	4	not stressed
Greer u/s Bridge	VAN28	test	0.92	not stressed	4	slightly stressed
Matthews Creek	VAN29	test	0.97	not stressed	4	not stressed
Natalkuz Lk. Trib.	VAN30	test	1.09	not stressed	4	not stressed
Ormond Creek Below trib	VAN31	test	1.17	richer than reference	1	not stressed
Swanson u/s bridge	VAN32	test	1.31	richer than reference	4	not stressed
Tetachuk Trib	VAN33_05	test			4	stressed
Tetachuk Trib	VAN33_06	test	0.80	stressed	4	stressed
Alpha Cr.	NOC01	test	1.23	richer than reference	2	not stressed

6 DISCUSSION

6.1 Model Comparisons

Both BEAST and SkeenRIVAS were built for the purpose of objectively and accurately assessing the biological condition of test stream sites. Type I errors (incorrectly failing a test site), and Type II errors (incorrectly passing a test site) should be small and acceptable for management purposes (Bailey et al 2004). To ensure that the best model is uploaded to CABIN, the models were compared and contrasted to evaluate their suitability for routine use.

The number of reference groups and the number of sites per reference group varied between models with 4 groups for BEAST and 8 groups in SkeenRIVAS as shown in Table 12. Reference site clustering in BEAST was based on abundance of invertebrates in contrast to SkeenRIVAS that was based on presence-absence of taxa. Additionally, different clustering techniques were used for each model. Both the number of reference groups and the number of sites per group can contribute to the accuracy of the model (Bowman and Somers 2005).

Table 12. Differences between the BEAST and SkeenRIVAS models including number of reference groups, number of predictor variables and misclassification error rates.

Model Parameter	BEAST	SkeenRIVAS
Number of Reference Groups	4	8
Number of Predictor Variables	5	8
Predictor Variables	• % wetland • % lake • % volcanic rock • longitude • elevation	• water temperature • drainage density (StrRiv_DrnDnsty) • Ratio of definite stream length to total stream length • % sedimentary rock • % lake • elevation • latitude • dominant riparian vegetation
Misclassification Error	22%	38%

Model performance can be partially examined using misclassification error rates (e.g. Sylvestre et al 2005), although this must be done with caution. Misclassification error occurs when the discriminant functions cannot assign a reference site to a proper group based on the selected predictor variables. Measurement of misclassification error is most useful when calculated using a set of validation data (independent data from reference sites that were not used to build the model) (Van Sickle et al 2006). For this project, all references site samples were needed to build the models so a validation dataset was not available. While model output usually includes two measures of misclassification error (resubstitution and cross-validation), cross-validation has been shown to be the more reliable method. While misclassification error can be used to compare models having the same number of sample groups, it is not useful for comparison of models having different numbers of groups as was the case in this study. Imagine re-allocating sites randomly into groups. A 2-group model, for example has a 'natural' random misclassification error of 50% (1:1). A 3-group model has a random error of 66% (2:1), while a 5-group model has an error of 80%. With fewer groups, the misclassification error invariably goes down and the probability of making a wrong choice declines. Table 12 shows that the misclassification error was 22% for BEAST and 38% for SkeenRIVAS. These error rates only confirm that the misclassification error was acceptable in both models. Other criteria were required for model comparisons.

We considered the number and types of predictor variables (PV) included in each model, as summarized in Table 13. An acceptable model should have PVs that are easy to measure, have low measurement error (e.g. sampling error), and do not require extrapolation of test data (Van Sickle et al 2006). Nearly all the PV's were GIS derived except for water temperature and dominant riparian vegetation that were used in SkeenRIVAS. All PV's were relatively easy to measure. Measurement of water temperature and dominant riparian vegetation would be expected to have low sampling error. Test sites had PV values outside the reference range for three of the eleven variables (percent lakes, water temperature and dominant riparian vegetation class) (Table 13). However, only a small number of test sites had values outside the reference range (n=3 for percent lakes, n=2 for water temperature, and n=1 for riparian vegetation) and the values were generally close to the maximum range for reference sites. Given the few number of test sites that would require extrapolation, this was not a concern for either of the models.

Van Sickle et al (2006) found that the optimal number of discriminant functions for a 5-group model was 7 (or 8 PVs), while 8 discriminant functions (or 9 PVs) were optimal for an 11-group model. Including more PVs led to model overfit and overly optimistic estimates of resubstitution error (Van Sickle et al 2006). Overfit models perform very well with model building data and have low resubstitution error rates, but perform poorly with new data due to the inclusion of spurious factors (Van Sickle et al 2006, Johnson and Omland 2004). Both BEAST and SkeenRIVAS had an acceptable number of predictor variables relative to the number of reference groups.

Table 13. Range of values within reference and test groups for predictor variables included in the BEAST and SkeenRIVAS models.

Model (B=BEAST, S=SkeenRIVAS)	**Predictor variable**	**Variable range**		**Test sites fall within reference range?**
		Reference sites (n=80)	**Test sites (n=150)**	
B	Percent wetland	0 to 16.1	0 to 9.2	yes
B	Percent volcanic rock	0 to 100	0 to 100	yes
B	Longitude	-121.52 to -133.27	-121.70 to -130.31	yes
B&S	Percent lake	0 to 7.6	0 to 14	no
B&S	Sample site elevation (m)	20 to 1720	20 to 1400	yes
S	Percent sedimentary rock	0 to 100	0 to 100	yes
S	Latitude	52.975 to 58.824	53.112 to 57.278	yes
S	Water temp (C.)	4.2 to 17.3	2.2 to 16.1	no
S	Drainage density	0.001 to 0.007	0.001 to 0.006	yes
S	Stream length ratio	0.35 to 1	0.4 to 1	yes
S	Dominant riparian vegetation class	2 to 5	1 to 5	no

Each model had 4 similar categories or bands of condition called reference or not stressed, slightly stressed, stressed, and severely stressed as summarized in Table 14. SkeenRIVAS included a fifth band for sites where the O/E ratio of observed to expected number of taxa at a site was greater than 1.16 (band X). Since higher than expected taxa richness could be natural (e.g. indicate a biodiversity hotspot) or anthropogenic (e.g. nutrient enrichment due to land use), sites scoring greater than 1.16 were not automatically failed but they were flagged for further review.

The distribution of test sites in each assessment category was similar for BEAST and SkeenRIVAS as shown in Figure 18. Sixty-nine percent of test sites were found to be similar to the reference condition with both models. Two percent of the sites were found to be severely stressed sites using both models. The remaining 29% of test sites were distributed among the enriched (X), slightly stressed, and stressed categories using SkeenRIVAS and between the slightly stressed and stressed categories using BEAST. Site testing using SkeenRIVAS resulted in a larger number of stressed sites than was found by BEAST, while BEAST found a larger number of sites categorized as slightly stressed than did SkeenRIVAS.

Table 14. Test site classifications for BEAST and corresponding SkeenRIVAS classification bands.

Ordination of test site relative to BEAST probability ellipses	**BEAST Assessment**	**SkeenRIVAS classification bands (based on O/E scores of the central 80th percentile of sites)**	**SkeenRIVAS Assessment**
		X > 1.16	Possibly enriched (or of high biodiversity)
Inside 90%	Not Stressed	A: 0.88 to 1.16	Reference
Between 90% and 99%	Possibly Stressed	B: 0.82 to 0.88	Slightly Stressed
Between 99% and 99.9%	Stressed	C: 0.6 to 0.82	Significantly Stressed
Outside 99.9%	Severely Stressed	D: < 0.6	Severely Stressed

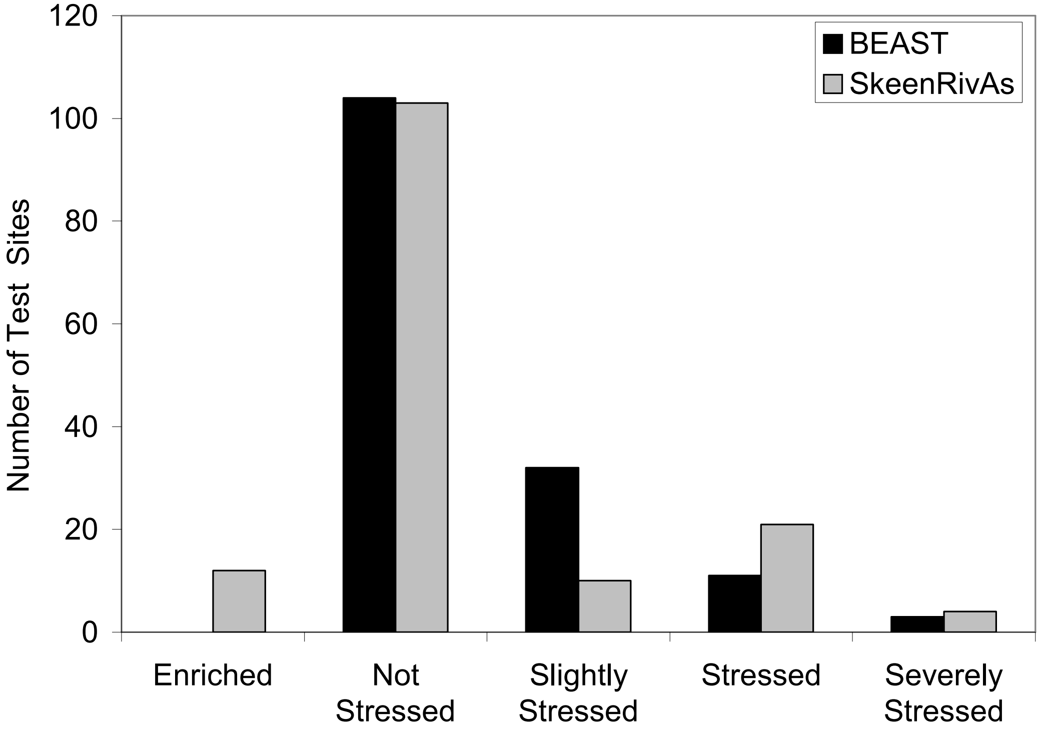

Figure 18. Distribution of 150 test site assessments using BEAST and SkeenRIVAS models.

Since the bands of condition were unequal and calculated using different reference site descriptive statistics (probability ellipses for BEAST, percentile values for SkeenRIVAS), the comparison of site assessments was simplified to an evaluation of agreement of passes and fails between methods. Any site that was enriched, not stressed or slightly stressed was passed. Any site that was stressed or severely stressed was failed. Using this pass or fail system, 90% of test sites passed using BEAST and 83% of test sites passed using SkeenRIVAS as shown in Table 15.

Table 15. Summary of test site passes and failures using BEAST and SkeenRIVAS. Pass was defined as 'not stressed', 'slightly stressed' or 'enriched'. Fail was defined as 'stressed' or 'severely stressed'.

Model	**Number of Test Sites**		**Percent of Test Sites**	
	Pass	**Fail**	**Pass**	**Fail**
BEAST	135	15	90%	10%
SkeenRIVAS	125	25	83%	17%

The models were in agreement for 126 of 150 sites (84%). For 24 sites, there was disagreement as shown in Table 16. However, test site assessments for eight of the sites were found to be close to a classification band boundary for either BEAST (BUL60, VAN04 and VAN26) or SkeenRIVAS (BUL10, BUL40, BUL50, PRG09 and NOC07), suggesting that they might not be true mismatched assessments but an artifact of the categorical stress assignments. This finding highlights the need for data review before accepting test site results. Another site that was likely not a true mismatch was BUL16. The invertebrate community at BUL16 had increased taxa richness and a greater number of individuals when compared with samples taken from the same site in 2004. The enriched community caused the site to fail in BEAST. The same site was assessed as enriched with SkeenRIVAS, which proved to be a similar assessment, rather than a model disagreement. Removing these 9 questionable sites decreased the percent disagreement from 16% to 10%.

In Australia, a number of mismatched assessments discovered during a comparison of two models lead to identification of a large scale pattern that suggested bias in one model for a particular geographic area (Norris et al 2002). There were no obvious patterns suggesting model bias in BEAST or SkeenRIVAS based on the set of mismatched sites from the Skeena data set. However, when the mismatched assessments were summarized by BEAST reference group as shown in Table 17, 16% of the sites were mismatched overall, but 53% of the assessments for test sites in group 1 disagreed between BEAST and SkeenRIVAS. It is possible that the small number of reference sites in Group 1 (n=5) resulted in less accurate site assessments for test sites predicted to that group. The test site has more influence on the ordination if there are fewer number of reference site than if there are many reference sites. Bowman and Somers (2005) recommend a minimum of 20 sites per reference group to accurately describe the biological community. Since the true condition of the test sites was unknown, we cannot make any conclusions about which model was more accurate for sites where there were disagreements.

Table 16. Test site assessment category mismatches of two categories between BEAST and SkeenRIVAS. Sites assessments in red text were close to the threshold between two assessment classes.

Site Code	Site Name	Year	BEAST		SkeenRIVAS	
			Group	Assessment	OE50	Assessment
BUL16	Goathorn above Tenas	2006	3	FAIL (stressed)	1.18	PASS (enriched)
MOR12	Nadina	2004	1	FAIL (stressed)	1.09	PASS (not stressed)
BUL50	Sandstone	2004	1	FAIL (stressed)	1.06	PASS (not stressed)
VAN04	Bird Lake Tributary	2005	1	FAIL (stressed)	1.00	PASS (not stressed)
MOR45	Lamprey @ Rec Site	2004	4	FAIL (stressed)	0.98	PASS (not stressed)
BUL60	Nichyeskwa 14k	2005	3	FAIL (stressed)	0.93	PASS (not stressed)
EUC08	Euchiniko River	2005	4	FAIL (stressed)	0.91	PASS (not stressed)
BUL50	Sandstone	2005	1	PASS (slightly stressed)	0.82	FAIL (stressed)
BUL10	Jonas	2004	3	PASS (not stressed)	0.81	FAIL (stressed)
PRG09	FREP Site 27	2006	4	PASS (not stressed)	0.81	FAIL (stressed)
NOC07	Old Field	2005	2	PASS (slightly stressed)	0.81	FAIL (stressed)
LAK05	Pinkut	2004	1	PASS (not stressed)	0.80	FAIL (stressed)
KTM08	Kitimat River @ Rec. Site	2005	2	PASS (not stressed)	0.80	FAIL (stressed)
VAN01	Lucas	2005	1	PASS (not stressed)	0.78	FAIL (stressed)
KAL37	Tintina	2006	2	PASS (slightly stressed)	0.77	FAIL (stressed)
KAL31	Granite Creek d/s 1st Ave	2005	2	PASS (not stressed)	0.76	FAIL (stressed)
KTM11	Hirsh d/s Landfill	2006	2	PASS (not stressed)	0.76	FAIL (stressed)
FSJ03	Flemming	2006	4	PASS (slightly stressed)	0.75	FAIL (stressed)
VAN26	Fawnie Tributary	2006	4	PASS (not stressed)	0.74	FAIL (stressed)
PRG04	Darby Rd at 6.5km	2006	4	PASS (slightly stressed)	0.74	FAIL (stressed)
VAN05	Bobtail Connector	2005	1	PASS (not stressed)	0.69	FAIL (stressed)
VAN13	Roy	2005	4	PASS (slightly stressed)	0.66	FAIL (stressed)
FSJ02	RubyRock	2006	1	PASS (not stressed)	0.61	FAIL (stressed)
BUL40	Causqua	2004	2	PASS (slightly stressed)	0.59	FAIL (severely stressed)

Table 17. Summary of two category assessment mismatches between BEAST and SkeenRIVAS, organized by BEAST reference group.

BEAST group	Number of reference Sites	Total number of test sites predicted to Group	Disagreement expected (based on 16% overall disagreement)	Disagreement observed	Percent mismatches within group
1	5	15	2.4	8	53%
2	27	34	5.4	6	18%
3	41	40	6.4	3	8%
4	12	61	9.8	7	11%
ALL	85	150	24	24	16%

Overall, the agreement between the two models was good. Site assessment precision and accuracy would likely improve if new reference sites similar to those that clustered into Group 1 were included to expand sample size in that Group.

In lieu of an evaluation of model accuracy with a known stressor gradient, we allowed agreement of test results between the models and known condition of selected sites to indicate accuracy. Conditions at sites where this agreement was found was reviewed to determine if results matched the expected site conditions based on the presence of a known stressor. Replicate test sites below a known stressor (e.g. below a mine site) was one class of sites where this judgement was used. We found 8 stressed and severely stressed sites for which there was model agreement (Table 18). For five of the eight sites (EUC07, KTM07, PRG11, MOR66 and BUL55) the results matched *a priori* expectations. The Euchiniko river (EUC07) was sampled at a cattle crossing and land use within the watershed included logging, range use and agriculture. McMillan Creek (PRG11) is an urban influenced creek that runs through the City of Prince George, while the Kitimat River site (KTM07) has both industrial and urban influences. Berg (MOR66) and Glacier Gulch (BUL55) Creeks are located downstream of old adits that discharge metals-laden mine water. The stressed condition of the remaining three sites did not meet *a priori* expectations. One site, Tetachuk tributary (VAN33) was downstream of a park area with no obvious source of anthropogenic stress. The second site, Eco Reene-Meridian Rd was a small stream with very low flow (0.01 m^3/s discharge) adjacent to logging and downstream of an ecological reserve. At the third site (KAL33), sampling error was suspected. Only 29 individuals were collected in the sample taken from the Salmon River site (KAL33). All three sites would need to be re-sampled to verify the findings. Overall, five out of eight stressed site assessments were reasonable, two were unexpected and one appeared to be due to sampling error.

Table 18. Eight most impacted streams for which test results by BEAST and SkeenRIVAS were in agreement.

Site Name	Site code	Year	BEAST		SkeenRIVAS	
			Group	Assessment	$O:E_{50}$	Assessment
Kitimat River 750m d/s Pulp Mill Outfall	KTM07	2005	2	Severely Stressed	0.3	Severely Stressed
McMillan Creek	PRG11	2006	4	Severely Stressed	0.5	Severely Stressed
Eco Reene-Meridian Rd	VAN10	2005	4	Stressed	0.5	Severely Stressed
Salmon River above Cascade Creek	KAL33	2005	2	Stressed	0.6	Stressed
Berg Creek Far Field	MOR66	2005	3	Severely Stressed	0.7	Stressed
Glacier Gulch	BUL55	2005	3	Stressed	0.8	Stressed
Tetachuk Tributary	VAN33	2006	4	Stressed	0.8	Stressed
Euchiniko River EUC07	EUC07	2005	4	Stressed	0.8	Stressed

Three known reference sites that were left out of the model building dataset (KIS40, MOR06 and NOC01) were used to test the models against known reference conditions. As shown in Table 19, BEAST correctly found that all of the sites were not stressed. SkeenRIVAS found two sites to be enriched (not stressed) and one site, MOR06, was 'slightly stressed'. However, the O:E_{50} value for MOR06 was very close to the threshold between 'not stressed' and 'slightly stressed' condition bands. Therefore, we accept that both models correctly assessed the 3 known reference sites.

Table 19. Summary of assessment results for known reference sites using BEAST and SkeenRIVAS.

Site code	Year	BEAST Group	BEAST assessment	SkeenRIVAS assessment
KIS40	2004	2	not stressed	Enriched
MOR06	2004	4	not stressed	Slightly Stressed
NOC01	2005	2	not stressed	Enriched

At 16 sites, multiple samples were collected during each site visit as shown in Table 20. Habitat and GIS variables were collected only once at each site, eliminating any chance of GIS or sampling error leading to the wrong reference group prediction within BEAST (i.e. that one of the multiple samples would be predicted to a different BEAST group than the other ones). Variation of assessment results among the multiple samples collected at a single site was low. For known reference sites, the first sample was used for model building and the additional samples were run through the model as test sites. In all cases, both BEAST and SkeenRIVAS accurately assessed the additional samples as 'not stressed' as expected. At 10 of 11 test sites, all samples tested the same using BEAST. At one site, BUL52, two samples were considered similar to the reference condition and one sample was not. One person collected all three samples and there was nothing in the field notes to indicate that the third sample was different in any way from the first two in terms of sample collection. Investigation of the raw data found that the third sample had fewer individuals and not as many unique taxa. It is possible that the difference reflected a natural habitat gradient at the site captured by the sampling, or sampling error. Overall, there was very good agreement between tests on replicate samples using BEAST.

Similarly, there was very good agreement between tests on replicates using SkeenRIVAS. There was one out of nine test sites, VAN26, where two separate samples resulted in different assessments. SkeenRIVAS results for extra samples collected at BUL52 and KAL36 were not available. It is likely that the results for BUL52 would have been similar to the BEAST results, since taxa richness was lower in the third sample collected at BUL52. These results show that both models provided consistent results across multiple samples collected at a given site.

Temporal variation was examined to evaluate the accuracy of SkeenRIVAS. Although there were a number of sites sampled in more than one year, most of them were test sites and any change in community structure over time was potentially confounded by a change in anthropogenic activities. Two reference sites were sampled in more than one year. Since O:E_{50} values were calculated for all reference and test sites, we can compare the SkeenRIVAS scores over time at the two reference sites as shown in Table 21. Although the sample size was very small (n=2), the O:E50 score was very stable at both reference sites, varying at BUL24 by 0.02 and at KIS16 by 0.05.

Table 20. Summary of results for multiple samples collected at a single site in a given year. (MB denotes that the sample was used for model building, P= pass, F = fail). Disagreements between multiple samples at a single site for a given model are shaded in yellow with a box.

Site code	**Site assignment (ref or test)**	**Year**	**BEAST assessments**				**SkeenRIVAS assessments**		
			BEAST Group	**Sample 1**	**Sample 2**	**Sample 3**	**Sample 1**	**Sample 2**	**Sample 3**
BUL24	ref	2005	3	MB	P		MB	P	
KIS15	ref	2004	3	MB	P	P	MB	P	P
KIS22	ref	2004	3	MB	P	P	MB	P	P
NOC06	ref	2005	2	MB	P		MB	P	
VAN17	ref	2006	4	MB	P		MB	P	
BUL16	test	2006	3	F	F		P	P	
BUL33	test	2005	3	P	P		P	P	
BUL52	test	2004	3	**P**	**P**	**F**	P	Not available	
KAL23	test	2005	3	P	P		P	P	
KAL32	test	2006	2	P	P		P	P	
KAL36	test	2005	2	P	P		P	Not available	
MOR45	test	2004	4	F	F	F	P	P	P
MOR71	test	2005	4	P	P		P	P	
PRG12	test	2006	4	P	P		P	P	
VAN19	test	2006	4	P	P		P	P	
VAN26	test	2006	4	P	P		**F**	**P**	

Table 21. Temporal variation at two reference sites using SkeenRIVAS.

Site Code	**SkeenRIVAS OE$_{50}$ Value**	
	2004	**2005**
BUL24	0.85	0.87
KIS16	0.98	1.03

There are some important differences in model structures to consider when using either of the two models. If a test site is predicted to the correct BEAST reference group using DFA (low misclassification error), the reference group fauna is well characterized (# of sites per group > 20) (Bowman and Somers 2005), and the faunal residual variation around the group centroid is relatively small (e.g. high faunal similarity) (Bailey et al 2004) then the classification of the test site will be more accurate than with SkeenRIVAS. This outcome is likely because the cutoff is sharper in BEAST for defining 'stressed' versus 'not stressed' conditions. If the test site classification to a BEAST reference group is incorrect, the test site assessment is likely to be more accurate with SkeenRIVAS than with BEAST. A test site classified to the wrong BEAST reference group is likely to fail unless by chance there is substantial overlap between the correct reference group and the misclassified reference group, which is something we tried to avoid in building the BEAST model. To explain it another way, in BEAST, the misclassification error was 22%. This means that on average one fifth of test sites could be predicted to the wrong group. In these cases, the site assessment will be in error, and if the reference groups are well separated and the test site is similar to reference condition, it will likely be failed. However, the 78% of the time when the correct group is predicted, the model is very accurate with test site assessment. The cut-offs are very clear for determining a change from reference condition. Since we don't know if a test site is predicted to the wrong group, it is useful to have another method to corroborate the results.

There are some minor concerns about the accuracy of assessments for test sites assigned to Group 1 and possibly Group 4 in BEAST. The number of reference sites in each of these groups (n=5 in group 1, n=12 in group 4), was less than the minimum 20 sites recommended by Bowman and Somers (2005) for precise characterization of the biological community representing the group. If the reference community is not precisely defined, inaccurate test site assessments may occur. Too small of a reference group can affect the model precision in two ways. First the misclassification error is likely to be higher, and second the variability within each faunal group will not be well explained (Bailey et al 2004).

SkeenRIVAS essentially has a built in 'safety net' that reduces its' susceptibility to misclassification error. The safety net is the weighted probability approach that considers the distance of a test site from each of the reference groups and does not rely on strict classification to a single group. However, the downside of the weighted probability approach to reference group prediction is that SkeenRIVAS is not as sensitive for detecting slight stress compared to BEAST, as was found in this study (Figure 18).

6.2 Model Selection for Upload to CABIN

All evaluations and comparisons between the two models suggest that either model would be a good choice for routine use. Since the CABIN website that will host the selected model is not set up to accept SkeenRIVAS, BEAST is recommended for immediate use in the Skeena Region. Since most of the money spent in an assessment program goes towards collection and processing of samples and data, there is good value in using more than one assessment method, particularly since they use the same data. All results shown here highlight the advantages of using both models. We therefore recommend that CABIN be modified to include SkeenRIVAS as well as BEAST for future use.

After more test and reference site data has been collected, the modelling should be reworked keeping some reference sites aside as a validation data set.

6.3 Attributes of Sample Groups in the Selected Model, BEAST

A comparison of the biological and habitat attributes of the four reference groups in BEAST can provide insight into similarities and differences in the ecological functioning among sites between the reference groups.

The relatively abundant and diverse invertebrates in Group 1 coincided with a very high proportion of volcanic parent materials in the watersheds. Phosphorus is abundant in volcanic rock. It weathers relatively rapidly, thereby contributing to phosphorus enrichment of surface streams (Murphy et al. 1983, Peterson and Grimm 1992, Perrin 1998). Phosphorus is important because it can limit biological productivity (Wetzel 2001) in coastal and Interior streams of British Columbia (Bothwell 1989, Perrin et al. 1987, Johnston et al. 1990) and greatly modify benthic invertebrate abundance and composition (Perrin and Richardson 1997, Mundie et al. 1991, Deegan et al. 1997). We hypothesize that phosphorus loading that occurs naturally from the weathering of volcanic rock, can contribute to a greater abundance of invertebrates than may be found in the relative absence of volcanic parent materials, largely due to the uptake of phosphorus for biological production. Sample Group 1 that had the highest invertebrate density was from sites having a very high percentage of volcanic rock in the watershed. The smallest average number of invertebrates per sample were in Group 2 samples that all came from watersheds having no volcanic parent materials. Higher numbers were found in Groups 3 and 4 where the relative extent of volcanic rock, by area, was between the extremes of Groups 1 and 2.

While total phosphorus concentration measured in water samples that were collected at the time of the biological sample might be expected to follow the gradient of volcanic rock, this relationship was not found. Mean total phosphorus concentrations in each of the four sample groups, in order, were 4.4 μg/L, 11 μg/L, 8 μg/L , and 17 μg/L respectively (Appendix C). All concentrations were low and many measurements were near the detectable limit of analytical laboratories, which can introduce substantial error to reported concentrations. The higher TP concentration at Group 4 sites did not necessarily mean that those sites were relatively rich in TP. Bothwell (1989) and Perrin and Richardson (1997) showed that known additions of inorganic phosphorus and/or nitrogen failed to show up in wet chemical analyses of water samples in mesocosm scale experiments. It was caused by the biological demand for the nutrients that resulted in periphyton sequestering the available nutrient pool even under enriched conditions, thus shifting the nutrient load from a simple ionic state to complexes in organic matter. Grazing by invertebrates can further advance the complexation of nutrients, while fish add another layer (Deegan et al. 1997). This process of element uptake and transformation is one of the main reasons why biomonitoring can often reveal more about habitat quality than can chemical measurements of water alone, particularly where there is a narrow range of concentrations of chemical analytes that are near detection limits of laboratories.

Group 2 sample sites were unique in having little or no wetland, lakes and volcanic parent materials, and they occurred mostly at low elevations close to the Pacific Coast. These attributes were associated with low numbers in a community of mayflies, stoneflies and chironomids in the kick net samples. Coastal streams can be extremely nutrient deficient (Stockner and Shortreed 1978, Perrin et al. 1987). Flows can be flashy, responding mainly to Pacific storm events that maintain relatively high and variable flows in winter and relatively low flows in summer. In the relative absence of wetlands and lakes upstream of Group 2 sample sites, hydrologic buffering was limited to water transport mechanisms in forest soils that can lead to very low flows in late summer. The presence of lakes and wetlands as was found in Groups 1 and 4 and to a smaller extent in Group 3 can moderate flashiness in stream flows and provide a reservoir to maintain a baseflow in the late summer period. The absence of this potential moderating effect of lakes and wetlands and the likelihood of extreme nutrient deficiency were natural characteristics that potentially contributed to the low abundance of invertebrates among all of the Group 2 samples.

In contrast, Groups 3 and 4 samples were from sites headed by lakes and wetlands and were characterized by greater numbers of invertebrates and more diverse communities than were found in Group 2 samples. The Groups 3 and 4 sites were situated more in the drier Interior ecoprovinces where snowmelt hydrology, modified by water yield from lakes and wetlands, can supply water to maintain base flows during extended dry periods in late summer. Group 4 sites in particular had a greater proportion of lakes compared to the Group 2 sites. Given that percent lake, by area of watershed, was the strongest discriminator of sample groups (Table 9), it is apparent that some attribute of lake outflow was important in defining the natural stream habitats across north-central British Columbia.

Lakes and wetlands can be important in several ways. They provide a reservoir of water and thus moderate flow extremes and most importantly maintain downstream flows at times of extended dry periods in late summer (when sampling for this project was done), thus maintaining the integrity of stream benthic communities. Lakes and wetlands can also be sources of organic matter and food that occurs as outwash in the form of seston and plankton (Richardson and Mackay 1991). In summertime, streams draining lakes are relatively warm and they cool over the downstream gradient in forested landscapes due to groundwater inflows (Mellina et al. 2002). In contrast, streams without headwater lakes and wetlands warm over the downstream gradient (Mellina et al. 2002). These variations in elevational temperatures, flows, and food can alter detritus processing rates (Buzby and Perry 2000) and alter abundance and composition of invertebrate communities according to taxa-specific temperature optima (Vannote and Sweeney 1980).

The structure of BEAST sample groups had similarities to the biogeoclimatic ecosystem classification (BEC) system in British Columbia (Demarchi 1996, http://www.for.gov.bc.ca/hre/becweb/index.html). A principle of BEC is to use vegetation, soils, and topography to infer the regional climate and to identify geographic areas that have relatively uniform climate. This information is used for mapping, to organize ecological information, to decide on what tree species to plant following harvest or other deforestation, understanding landscape vegetation associations, planning land use, setting biodiversity objectives and policy, wildlife management, wildfire management, etc. (Mah et al. 1996). In BEC, vegetation, like invertebrates in BEAST, is considered the best integrator of the combined influence of environmental factors affecting a site. Vegetation units are determined by grouping biological data and the areas delineated by these groups are called biogeoclimatic units. BEAST does the same thing. It is a result of classifying groups of biological samples within which combinations of taxa have more similarities to samples within a group than to samples from other groups.

The five stream habitat variables that best discriminated the sample groups in BEAST were closely alligned with gradients of climate and landforms that are the basis of BEC. They were strong predictors, having almost no correlation with each other and providing a model that explained 75% of the total dispersion among sample groups. For a large regional scale model, these are excellent statistics. The importance of elevation and longitude suggests that invertebrate communities in north-central British Columbia adjust to the gradient from a low elevation coastal climate to a higher elevation drier climate of the eastern slope of the coast range to the more continental climate of the Central Interior plateau. Elevation, in particular, was the second strongest discriminator of the sample groups, making it very important in the discriminating ability of the BEAST model. High abundance and diversity of invertebrates (Group 1) appears to be associated to localized and discrete landforms hosting volcanic rock that can introduce growth-limiting nutrients via weathering. The climatic and landform gradient from the coast to the Interior coincided with a shift from low overall numbers of animals in samples from communities dominated by mayflies, stoneflies and chironomids (Group 2) to higher overall numbers from communities having the same orders of invertebrates and increased prevalence of caddisflies, true flies, and beetles (Groups 3 and 4). In this regard the BEAST model is a line of evidence that biological attributes in aquatic ecosystems of British Columbia may be determined by climate and landforms that determine vegetation associations on the landscape.

7 EXTENSION PLAN

Extension activities from this project involved hosting of workshops and conferences, preparation of progress reports, training sessions, and presentations to industry and interest groups. An extension plan was prepared in the first project year and delivered as follows.

Annual international bioassessment workshops were held at the University of British Columbia. One of these was incorporated into the annual North American Benthological Society conference held at UBC in May, 2004. The bioindicator portion of this conference was attended by several hundred biologists from around the world, and served as an idea generator to fine tune project extension efforts. The next 2 annual bioassessment workshops were focussed on western North American applications, and were attended by more than 50 practitioners in each year. These venues were invaluable in promoting communications among bioindicator system developers, researchers and prospective users. In all instances, potential users of the assessment information participated, and gained experience, which led to commitments for use of the RCA. The proceedings from these workshops were posted on the UBC Riparian Laboratory web site for a period of 12 months.

Reports included 2 annual progress reports and the present report. They were submitted to Forest Sciences Program (FSP), in fulfillment of funding requirements. Progress reports to the B.C. Ministry of Environment Impact Assessment Biologists have also been delivered in the form of presentations at the last three annual meetings. Further publications using data and findings from this project by authors of this report are expected in 2007 and 2008. These publications will serve to distribute project information to other researchers, and inform natural resource managers about the progress in bioassessment system development in British Columbia.

Training sessions included two Canadian Benthic Inventory Network (CABIN) database management workshops (Environment Canada) attended by a total of more than 50 biologists in Vancouver, and a field work training program attended by six biologists who were associated with the field component of year 3 of the project.

Extension forums have utilized teleconference / PowerPoint presentation and face to face workshop formats to present project results in case study format, and discuss impact assessment methods built into the indicator system. In 2005, a Forest and Range Extension Partnership (FORREX) monitoring conference held in Victoria was used to inform forest practitioners about the project, and its interim outputs. A contact list for email information dissemination was established using the attendee list from this conference, and has since been expanded. In 2006, over 30 forest management practitioners participated in teleconference based focus group sessions aimed at fine tuning how indicator reporting is done to meet their needs. This was followed by 2 face-to-face workshops. The first was internal to the B.C. Ministry of Environment (MOE), and was aimed at gaining executive support for bioassessment funding in the Province. This session was attended by MOE managers and Section Heads from throughout the Province and it resulted in an increase in the proportion of environmental monitoring and reporting dollars available for this work on a Province wide basis. The second was part of a watershed sensitivity classification session in Smithers, attended by forest management practitioners and researchers from all over the Province. The emphasis was on Mountain pine beetle infestation affected forests, and accelerated harvesting rates in this large area. This workshop was also used as a focus group session to fine tune outputs from the project, and resulted in changes to the geographic information system (GIS) analysis approach used in the project. In 2007, a forest management related monitoring and assessment workshop held in March was used as the final focus group session in which the case study approach to indicator results dissemination was used to get feedback for the final project report.

Extension notes will be published in 2007 issues of FORREX's Streamline periodical describing the intended uses for the new indicator system.

Methods on how to use the RCA as an indicator system are part of the present report (Section 8 and 9). They have been submitted to the Resource Inventory Science Committee (RISC) for review and approval as provincial standards for use by others. The guidelines, once approved will be available on the RISC website.

The Forest Practices Board has received a case study of forest harvesting effects on streams in the Mountain pine beetle affected area of the province. This will be published as part of the 2007 State of the Forest Report, and used to showcase the system and its applicability in forest management.

A proposal for another 3 year indicator development project has been submitted for FSP funding, to commence in 2007. This proposal is to capitalize on the first 3 year project, and expand the system for province wide implementation. The extension plan developed in year 1 of the first project will be expanded for use in the second project.

8 CASE EXAMPLES OF APPLICATION OF RCA IN THE SKEENA REGION

8.1 Introduction to Case Examples

Resource managers, land use planners and regulators in British Columbia are seeking cost-effective and meaningful indicators of stream health to support decisions aimed at protecting aquatic resources. As the Province moves toward results-based management, new tools are required to assess the effectiveness of current practices in protecting and restoring water quality and aquatic ecosystem integrity. To date, no such effective and practical tool exists that is integrated into operational or strategic level adaptive management and resource decision making systems in British Columbia.

Through the following case examples, we demonstrate how the Reference Condition Approach (RCA) can fill an existing void in aquatic ecosystem assessment by providing natural resource practitioners with a convenient, affordable and flexible assessment and monitoring tool. Once the reference site database and models are in place, new users can simply “piggy-back” on the existing infrastructure.

The RCA is a rapid bioassessment tool suitable for a wide range of scales and assessment purposes. It provides an approach for overcoming traditional design problems such as finding suitable upstream controls in a control-impact design. A thorough description of how RCA compares to traditional impact assessment designs can be found in Section 2 of this report. When used as a landscape level screening tool for non-point source pollution or cumulative effects, it can reveal the biological condition of a large number of potentially widespread test sites. Continued use of the RCA over time allows trend monitoring and the ability to manage adaptively. To determine the cause of the stress or trend, a closer look at the data (chemical, physical and biological) and an analysis of the types and extent of land use in the watershed may be required. Focussed assessment may also be necessary. The RCA is also effective at assessing the effects of point source pollution[1] and operational-level resource management practices. As potential or actual causes of impairment are identified, the range of remedial possibilities is considered on a sliding scale of consequence. For example, if impairment is minor but persistent, then improvements in implementing best management practices may suffice. If impairment is significant and values are high, then causal activities may need to be substantially altered, mitigated, or abated through an adaptive management framework.

We have chosen four case examples to show the applicability of the RCA approach to assessing site quality at different scales and for different resource development activities:

1. LRMP – strategic or “state of the resource” level monitoring of cumulative effects from multiple land uses;

---

[1] Point sources include industrial effluent discharges from mines, pulp-mills, smelters, individual roads, and municipal discharges from storm water or sewage outfalls

2. Toboggan Creek – watershed specific assessment of cumulative effects;
3. Forest Stewardship Plan and Certification – landscape level assessment of forest management activities; and
4. Equity mine – assessment of point source discharge of mine effluent.

8.2 Land and Resource Management Plans

A land and resource management plan (LRMP) is a sub-regional strategic land use plan. LRMPs are created through multi-stakeholder, consensus-based planning processes that establish broad direction for land and resource management for a given area. LRMPs set goals, objectives, strategies, and in newer LRMPs, resource management direction and targets for Crown land in British Columbia.

LRMPs now cover the majority of British Columbia. One of the most recently completed plans covers the Morice Timber Supply Area (TSA), located in northwestern B.C. (Figure 19). The Morice LRMP area lies along the western edge of the Interior Plateau, in a transition zone between the Interior and the coast. The northern and eastern regions of the area are characterized by rolling topography, while glacier studded mountains dominate the southwest (Horn and Tamblyn 2000). The small communities of Houston, Granisle and Topley are found within the plan boundaries. First Nation traditional territories that overlap the Morice LRMP area include those of the Lake Babine Nation, the Wet'suwet'en Nation, the Yekooche First Nation, the Cheslatta-Carrier Nation, the Carrier-Sekani Tribal Council, the Skin Tyee Band and the Nee-Tahi-Buhn Band.

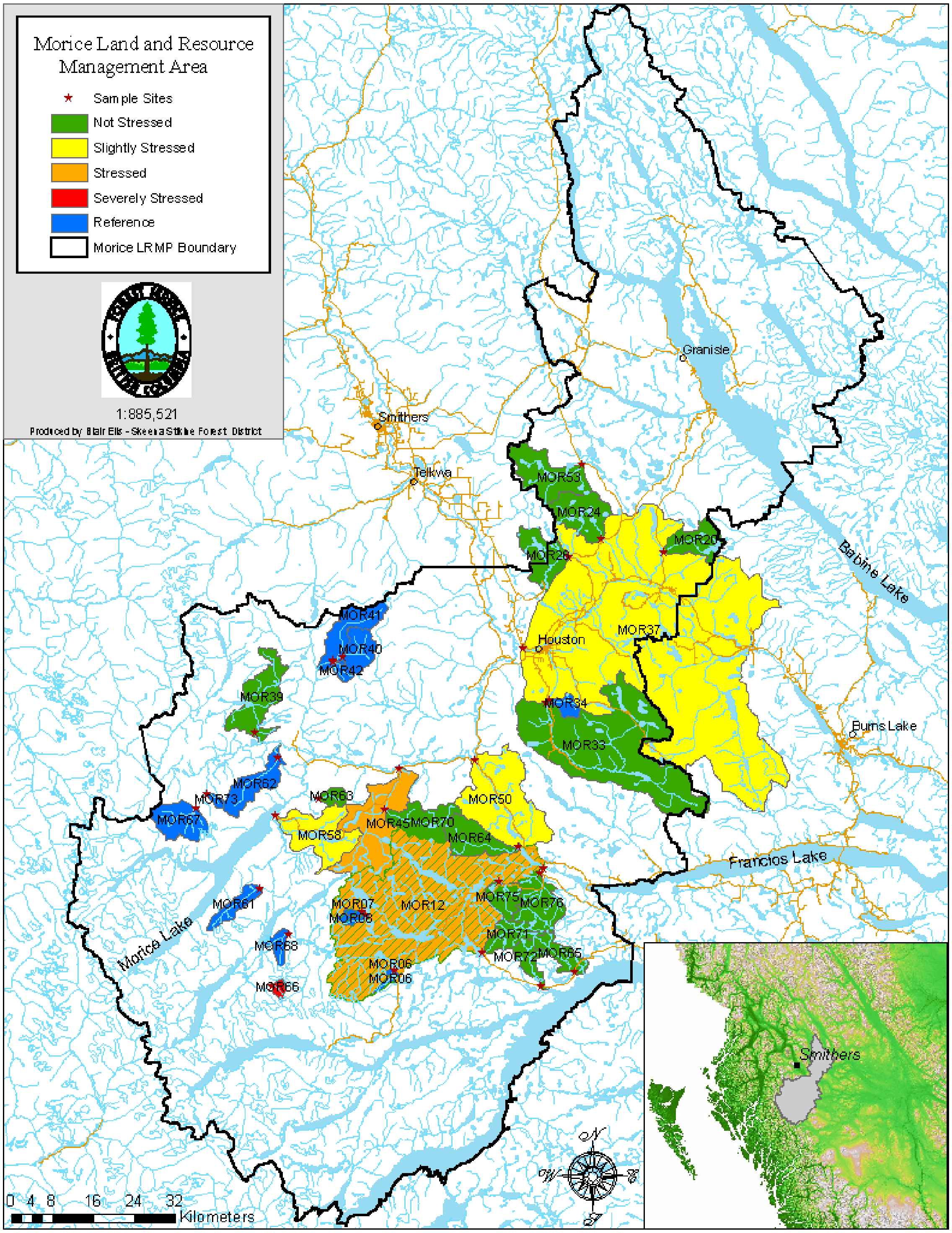

Figure 19. Location and biological condition (BEAST results) of sites sampled within the Morice LRMP area. One watershed shown with orange and green stripes indicates it was stressed in one year and not stressed in another year.

Land uses and natural resource values within the Morice LRMP area are diverse and rich. Forestry is the most widespread resource development activity. Valley bottoms, particularly along the Bulkley River are primarily private land; cattle ranching and hay production are concentrated in this area. In summers, cattle range extensively through crown forest land. Virtually the entire LRMP area contains high or very high metallic mineralization and both historic and current mining activities are evident. Fisheries resources in the major rivers including the Morice, Bulkley and Babine are highly valued by recreational anglers, First Nations and commercial fishers. The area is world-renowned for its quality steelhead and Chinook salmon angling experiences. Guide-outfitting, trapping, and hunting are among a wide range recreational and tourism activities that occur within the plan area.

The Morice LRMP was crafted to help ensure the sustainability of the resources within its geographic boundaries. To verify that the plan is achieving this goal, the Integrated Land Management Bureau is currently involved in a process to determine the status of key public resources within the plan area. The government also anticipates developing a program to monitor the effectiveness of the plan in protecting the area's water quality and aquatic ecosystems. The Reference Condition Approach offers an opportunity to assess both current conditions and trends related to cumulative effects of resource development activity on aquatic ecosystems.

During the 3 years of this study, 20 test sites were sampled within the Morice LRMP area (Figure 19). One test site (Nadina River - MOR12) was sampled in both 2004 and 2005. The majority of streams sampled in the Morice LRMP area showed little or no biological impairment (Figure 19). According to the BEAST model, 14 (70%) sites were not stressed, 3 (15%) were slightly stressed, 1 (5%) was stressed, and 1 (5%) was severely stressed. The remaining site, the Nadina River, appeared to be stressed in 2004, but not stressed in 2005. SkeenRIVAS indicated that the Nadina River was not stressed in either year. Given the weight-of-evidence provided by the two models, the Nadina River appeared to be within its range of natural biological condition. However, as this river is vulnerable to reaching relatively high temperatures[2] (MSRM 2004), future assessments should include monitoring stream temperature in conjunction with benthic invertebrates.

The slightly stressed sites included the upper Bulkley River (Bulkley at Morice - MOR37), McBride Creek (MOR58) and Owen Creek (MOR50). The upper Bulkley River watershed has a high degree of agriculture, particularly beef cattle production, in addition to logging, mining, and both a highway and railway running along the valley bottom. The upper Bulkley has been identified for area-specific management by the Morice LRMP, in part due to extensive loss of native riparian vegetation, past damage to the integrity of aquatic habitat and ecosystems, and concerns about water flows and water temperature (MSRM 2004). The McBride and Owen creek watersheds have had

---

[2] The temperature was 15.0°C when sampled on August 30, 2005.

significant forest harvesting in the past and according to MSRM (2004), temperature concerns exist in these watersheds.

The RCA provided evidence that the Lamprey (Lamprey Rec. Site - MOR45) watershed had one of the most stressed aquatic ecosystems within the plan area. The watershed has experienced a significant amount of historic logging and is identified as a potentially temperature sensitive system (MSRM 2004).

The test sites results within the Morice River watershed (including McBride, Owen, Lamprey, Pimpernel and Shea) are highly correlated with the level of historical land use. The watersheds with the highest amounts of land use (road densities, stream crossings and forest harvesting) - Lamprey and McBride, followed by Owen (Tamblyn 2005) - exhibited benthic invertebrate communities the furthest from reference condition. The Pimpernel Creek watershed (MOR70), a tributary to Lamprey had significantly less land development activity than Lamprey Creek and was found to be unstressed. Likewise, Shea Creek, a tributary to Gosnell Creek, a relatively undeveloped watershed, was in reference condition.

The most highly stressed site in the LRMP area is “Bergfar” (MOR66), a tributary to Bergeland Creek. This severely stressed site is approximately 700m downstream of a discharge from a mine adit in the Sibola Mountains. The substrate at the site was coated with a white precipitate. Water samples taken at the time of benthic invertebrate sampling indicated poor water quality at the site. The pH was low (5.36) and several chemical analytes (dissolved sulphate, and dissolved and total aluminum, cadmium and copper) significantly exceeded B.C. or federal water quality guidelines.

The RCA provides a cost-effective means to assess and report the status of aquatic ecosystems at a landscape or watershed scale, which is fundamental to monitoring the effectiveness of LRMPs[3]. Sites can be reassessed over time to determine trends and the overall effectiveness of the strategies or management direction in LRMPs.

8.3 Toboggan Creek Watershed

The Toboggan Creek watershed, located approximately 10 kilometres north of Smithers, B.C. (Figure 20), is a rural watershed sourced in the glaciers of the Hudson Bay Mountain Range. The 112 km^2 watershed is home to forestry, mineral exploration, rural housing development, a major highway corridor, a rail line, recreation, a fish hatchery, and beef and dairy cattle production. This valuable salmon stream provides water for drinking, residential use, irrigation and livestock watering. Because of the high valued water resources and the widespread and varied land use, the Ministry of

---

[3] Should RCA be adopted for LRMP effectiveness monitoring, the objectives of the monitoring program would need to be clearly stated and an appropriate monitoring program designed to meet the objectives.

Environment is developing site specific Water Quality Objectives to protect the water quality and aquatic ecosystems of Toboggan Creek.

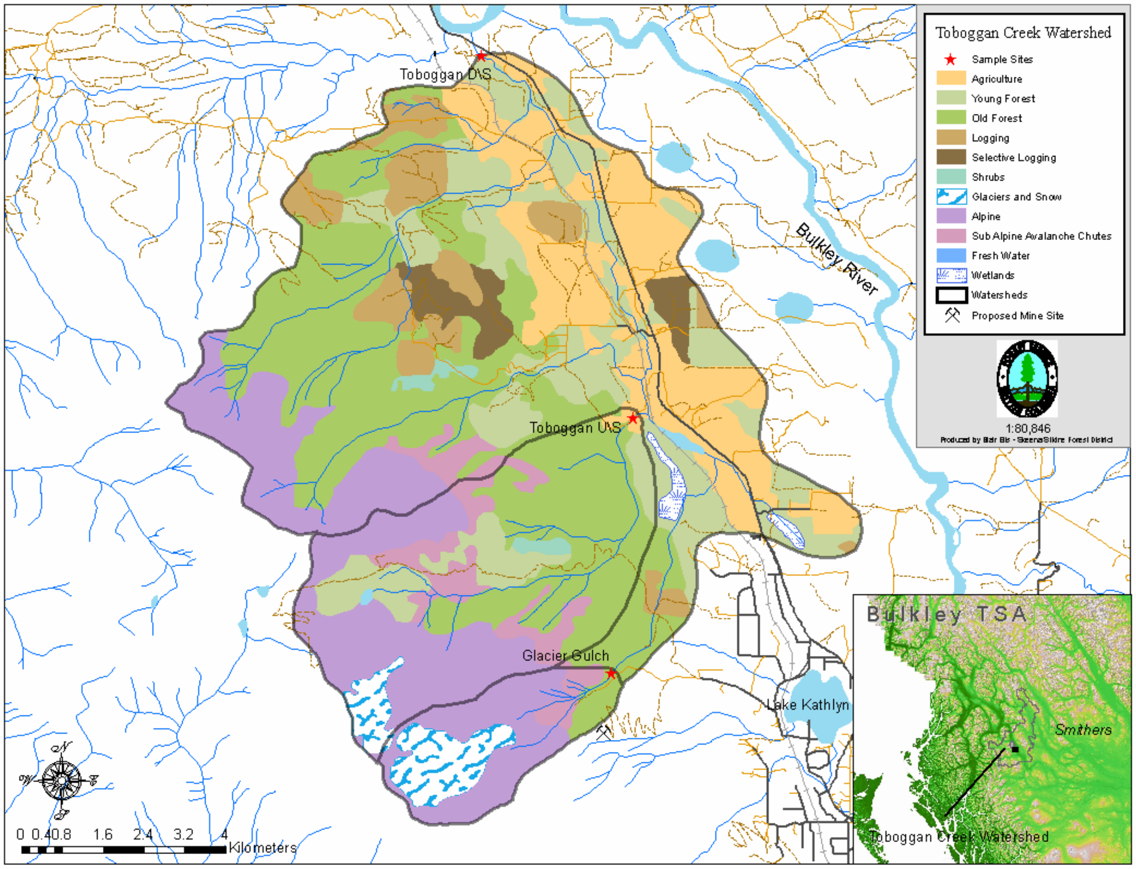

Figure 20. Dominant land uses and location of sampling sites within the Toboggan Creek watershed.

Over the 3 years of the RCA project, 3 sites were sampled within the Toboggan Creek watershed (Table 11). The site furthest downstream ("Toboggan d/s" – BUL33) is located near the creek mouth, several hundred metres upstream of Highway 16. The second site, "Toboggan u/s," is situated approximately half way up the watershed above the confluence with Glacier Gulch. This site is located in the upper part of a field used to grow hay without any fertilizers or pesticides. Although not strictly a reference site in the context of this study, little development activity is found upstream of "Toboggan u/s." The third site, the glacial headed "Glacier Gulch", is situated directly downstream from a discharge of adit water from an active mineral exploration site.

Results from the RCA assessments using both the BEAST and SkeenRIVAS models were consistent for all sites (Table 22). Both the upstream and downstream sites on Toboggan Creek were not stressed according to the RCA techniques. However, the Glacier Gulch site was stressed.

Table 22. BEAST, SkeenRIVAS and Index of Biological Integrity results for test sites within the Toboggan Creek watershed.

Site name	Site code	Year	BEAST	SkeenRIVAS	IBI*
Toboggan d/s	BUL 33	2004	Not stressed	0.99 – not stressed	Good
		2005	Not stressed	0.99 – not stressed	Fair
Toboggan u/s	BUL 52	2004	Not stressed	1.04 – not stressed	Good
		2005	Not stressed	0.97 – not stressed	Good
Glacier Gulch	BUL55	2005	Stressed	0.80 - stressed	Poor

*These IBI scores were preliminary at the time of writing. The results have not yet been calibrated to the sampling methodology used in the RCA.

One might expect to find the Toboggan downstream (d/s) site more stressed than the upstream (u/s) site due to the level of human activity, particularly agriculture and rural development, in the lower half of the watershed. According to Remington and Lough (2005), who conducted a full year of monthly water quality sampling at 8 sites in the watershed, some aspects of water quality worsened along a downstream gradient of the Toboggan Creek mainstem. They concluded that upper Toboggan Creek was meso-eutrophic, while the lower site was slightly eutrophic. Looking in more detail at the water quality data collected by Remington and Lough (2005), algal biomass and the concentrations of ammonia, phosphorus, orthophosphate, and coliforms were higher at Toboggan d/s than Toboggan u/s. Total nitrogen and nitrate+nitrite concentrations were higher at the upstream site. B.C. water quality guidelines were rarely exceeded at either site. Although the water chemistry and periphyton data indicate possible degradation at the downstream site, the benthic invertebrate communities, according to both models utilized in our study, suggest that the water quality is not causing significant effects on the benthic invertebrate assemblages in the stream.

In the past, a multimetric approach to bioassessment has been developed and used in the Skeena-Stikine Forest District (Bennett 2004). The results of the Benthic Index of Biological Integrity (IBI) generally agreed with those of the RCA (Table 22) for the two Toboggan Creek sites in 2004 and 2005[4]. In both years, the IBI scores (calculated using methods outlined in Bennett 2004) indicated that the upstream site was in good condition (on a continuum from poor to excellent). However, the results for the downstream site differed between the years with a slight decrease in IBI from good in 2004 to fair in 2004. This categorical result is somewhat misleading as the actual score

---

[4] The interim results of the Benthic Index of Biological Integrity were included in this example for comparison purposes to demonstrate how RCA and BIBI can be used together in a weight of evidence assessment approach. However, the BIBI results are tentative given the need to calibrate the results to the kicknet field method used for RCA sampling.

in 2005 was at the upper end of the fair category. In general though, the RCA and IBI can be complementary tools in areas where IBI has been calibrated for a given watershed or geographic area. In trend monitoring, the individual metrics used to calculate the IBI score may also be compared over years to assist with assessment of a watershed.

BEAST, SkeenRVIAS and IBI all indicated that Glacier Gulch (BUL55) was stressed, despite the fact that it was in the upper headwaters of the Toboggan Creek watershed. This site was directly downstream of an effluent discharge from a mineral exploration site. Fortunately, for the sake of interpreting results, water quality samples were collected by the mining company at BUL55 and at a site upstream (control) for several months before we collected benthic invertebrates. These results indicated that although some metal concentrations were relatively high at BUL55, they were below B.C. Water Quality Guidelines for aquatic life. The exception was total aluminum, which was unusually high at BUL55 (>5000 ug/L; BC guideline = 100 ug/L) in the two months before we sampled benthic invertebrates. However, the aluminum concentrations at the control site were even higher, indicating the high aluminum levels were natural. Whether this natural flush of alumimium affected the benthic invertebrate community at BUL55 is unknown. There is, however, no strong evidence that the mine discharge is responsible for the stressed biological condition of this site. We recommend that BUL55 be sampled again and that another RCA site be established upstream of the discharge to confirm whether or not the stream is a naturally stressed site due to high concentrations of metals in runoff from undeveloped ore deposits upstream of the mine adit.

In conclusion, although some signs of water quality degradation exist in the Toboggan Creek mainstem due to the cumulative effects of human development, the aquatic ecosystem as represented by the benthic invertebrate community is in reference condition. Current management practices for land use in the watershed appear to be adequate; nonetheless, residents and industry should be encouraged to adopt best management practices to help ensure that further development in the watershed minimizes stress on the aquatic ecosystem. Ongoing monitoring is important to ensure Toboggan Creek remains healthy. The Ministry of Environment has taken an important step toward ensuring monitoring by proposing Water Quality Objectives for the watershed. We recommend that the Ministry adopt the RCA to supplement the proposed physical, chemical and periphyton parameters included in their Water Quality Objectives for the Toboggan Creek watershed. The marginal additional cost will provide significant information to assess the biological condition of the watershed.

8.4 Forest Stewardship Plans / Certification

Forest management in British Columbia, under the *Forest and Range Practices Act (FRPA),* has shifted from a prescriptive regulatory environment to “results” based management over the past several years. Forest managers require appropriate monitoring tools to measure the effectiveness of forest practices in meeting a range of goals, including protecting water quality. Presently, aquatic ecosystems are commonly managed using "defaults" established in FRPA that act as proxies for maintaining water quality. Monitoring aquatic ecosystems is typically conducted using GIS-based surrogate variables (e.g. riparian management area statistics, road crossings, road densities, etc.). Geomorphically-based watershed assessments may be conducted once GIS indicators reach upper target limits. Bioassessments have generally not been used due to uncertainties around the ability to detect impairment and difficulties finding paired watersheds or suitable upstream control sites for more standard sampling designs.

The Reference Condition Approach (RCA) has widespread application for both government and forest licencees in monitoring and assessing aquatic ecosystems. It can be used for strategic-level monitoring in State of the Forest reporting or sustainable forest management plans, as well as for operational-level assessments under forest stewardship plans or the Forest and Range Evaluation Program (FREP). It can also meet the needs of assessment for forest licencee certification. Assessing the response of ecological communities to forest development provides the meaningful feedback required to evaluate whether current practices are sufficiently protecting our streams and rivers.

Forest stewardship plans (FSP) are legal plans that describe specific results that will be achieved relative to objectives set by Government. Certification is a methodical tool that can be used by companies to support FSP strategies, to illustrate compliance with strategies and results identified in FSP, and to demonstrate to product buyers and the public that their products are the result of sustainable practices. Certification can be a rigorous process, consisting of a system of principles, objectives, performance measures, and targets that must be met. Several types of certification exist including International Standards Organization (ISO), Canadian Standards Association (CSA), Forest Stewardship Council (FSC), and Sustainable Forest Initiative (SFI).

A basic principle of certification is the requirement to evaluate the effectiveness of forestry activities in meeting environmental objectives and to progressively reduce environmental impacts. This case example illustrates how aquatic ecosystem monitoring using RCA could be used by a forest company to fulfill some of its requirements in its forest stewardship plan, and for obtaining and maintaining certification.

Pacific Inland Resources (PIR), a division of West Fraser Timber Company Ltd., is located in Smithers, B.C. in the heart of our study area. PIR operates in a number of chart areas, the bulk of which are located in the Bulkley Timber Supply Area (Figure 21). PIR has ISO 14001 certification - achieved through creating and following environmental management systems which outline how the company manages its processes and activities. Its Forest Stewardship Plan details the company's environmental objectives and the monitoring program to determine if the objectives are being met. The plan relies on GIS-based indicators (e.g. equivalent clearcut area, road densities etc.) of development within watersheds with objectives under FRPA, including Fisheries

Sensitive Watersheds, Community Watersheds and watersheds with specific requirements contained in landscape unit plan (Baxter, Pers. Comm.). Based on best available science, maximum targets, which should not be exceeded, are established within the FSP for each indicator. Once a target is met or surpassed, PIR conducts a geomorphic-based watershed assessment (i.e. an Interior Watershed Assessment Procedure - IWAP) and may use locally developed tools such as the Stream Crossing Quality Index (Beaudry 2006) to determine if damage to aquatic resources may be occurring. PIR is interested in adopting a bioassessment technique, allowing it to meet several objectives including testing its FSP and IWAP targets to confirm they are appropriate to maintain the integrity of aquatic ecosystems.

Sixteen test sites within PIR's Bulkley TSA chart areas were sampled in this study (Figure 21). Three of those including Goathorn Creek (BUL16), Toboggan Creek u/s (BUL33), and Toboggan Creek d/s (BUL52) were each sampled in two different years. All sampled streams were tributaries to the Telkwa, Bulkley and Nichyeskwa rivers.

Assessment results using the BEAST model are shown in Figure 21. Table 23 compares the results for the BEAST and SkeenRIVAS models. The results generally agreed in 15 of 19 assessments. To assist in drawing conclusions about the biological condition for some of the sites and because IBI has been used by PIR in some its watersheds in the past, preliminary IBI results were considered (Table 23) (using methods described in Bennett 2004). The "combined assessment" column is a conclusion based on the weight-of-evidence provided by the three methods.

Based on the "combined assessment", 16 of 19 sites (84%) were found to be similar to reference condition (not stressed or slightly stressed). The three remaining sites Jonas Creek (BUL10), Causqua Creek (BUL40) and Nichyeskwa 14K (BUL60) were stressed. Any stressed sites should be re-sampled to confirm the results before remedial action is undertaken. If necessary, site specific experiments may be required to unequivocally define the cause of the impairment. Additional geoscience-based assessments might be considered to aid in deciding how to improve the condition of these streams once cause of the impairment is determined. A new phase in the international development of RCA includes advances in stressor gradient analysis that will further aid in the assignment of cause of site stress. These advances are expected to lower costs in activities beyond the initial RCA screening to determine cause of site impairment.

To date, we have used the RCA for watershed-level screening where the condition of a stream is unknown prior to resource development. PIR is also interested finding a biomonitoring tool that can be used, under certain circumstances, to assess the effects of localized harvesting[5]. RCA could be used in this type of assessment possibly

---

[5] Such assessments would be selective and would not be conducted everywhere PIR has objectives for watershed management or everywhere harvesting is proposed.

by defining impairment thresholds (currently 90%, 99% and 99.9% probability ellipses). Forest managers may then select target thresholds of tolerance (e.g. any one of the probability bands in BEAST or a band of O:E scores in SkeenRIVAS) that should be achieved using stream protection measures during harvesting activities and should be achieved in protecting receiving waters over the longer term following harvesting.

A significant advantage of using the RCA is that it provides more information per dollar than other assessment techniques. For example, once the RCA reference model exists, sampling can focus on test sites of interest. An astute manager, however, may also want to sample a few control areas as well as test sites as a means of testing model performance in partnership with regulatory agencies. As a rough guide, lab fees for the habitat measurements and enumeration of the benthic invertebrates at a test site may be $625 plus the cost of crews to complete the field sampling, consultation on the layout of sampling, GIS analysis, interpretation, and reporting. Conventional monitoring using a BACI design (Section 2) would require the collection of many more samples at replicated control and treatment sites extended over months or years at a cost of more than $4000 per site plus all of the same added costs.

Although RCA is new to many forest managers, benthic invertebrates have been used successfully for monitoring streams draining forestry operations in the past. The Benthic Index of Biological Integrity (IBI) (Bennett 2004) was used by staff at the Kispiox Forest District prior to its amalgamation into the Skeena-Stikine Forest District. Norm Bilodeau, Forest Officer with the former Kispiox Small Business Forest Enterprise Program, found IBI to be adaptable and very cost effective (Bilodeau, pers. comm.). The entire Kispiox operating area was assessed for approximately $15,000 per year (which included 3 invertebrate samples collected at each site compared to a single sample at each RCA site). Bilodeau also used IBI in experimentation to determine the effects of various silvicultural treatments on the health of aquatic ecosystems. He felt that IBI was tailor-made for forestry applications. The RCA would likely be similar, although it has the additional advantage of not requiring an index calibrated to a geographic area.

In conclusion, the RCA is a scalable, innovative and cost-effective method to assess the health of streams in watersheds with forestry activities. It can be used as a screening tool at a strategic or landscape level to help determine the effectiveness of forest stewardship plans. At an operational or stand level, it can help monitor the success of individual practices in protecting aquatic ecosystems. Furthermore, the RCA can identify where focussed assessment is required to determine causes of environmental stress and can be adapted for stand-level trials when a company is interested in testing or developing best management practices. Overall, the RCA is a valuable monitoring and assessment tool that will demonstrate environmentally responsible practices to both government regulators and certification bodies.

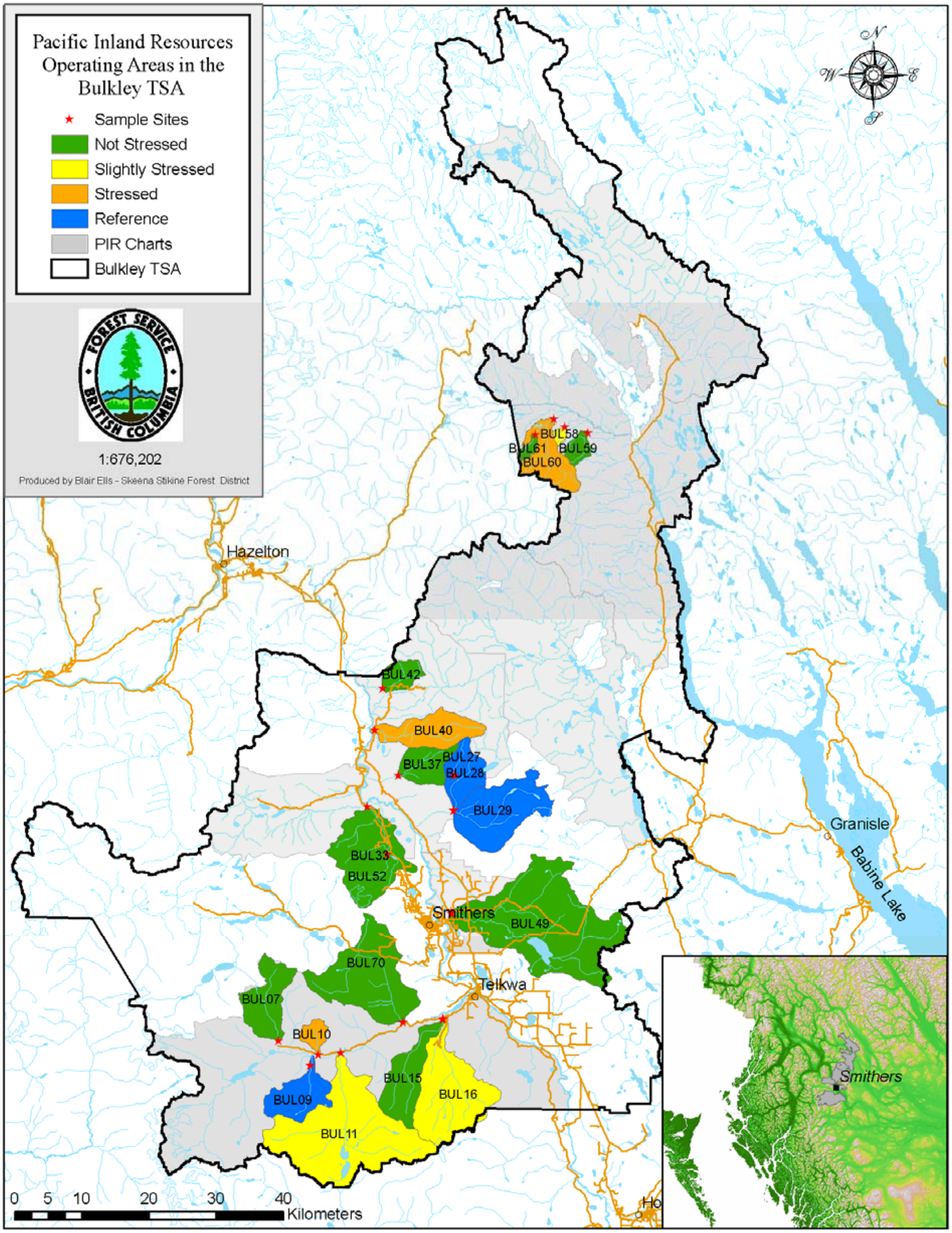

Figure 21. Location and biological condition of sites sampled in the operating areas of Pacific Inland Resources (Bulkley Timber Supply Area only).

Table 23. Conclusions from RCA models and a Benthic Index of Biological Integrity (IBI) for test sites within PIR chart areas.

Site Name	Site Code	Year	Reference Condition Approach		IBI Assessment*	Combined Assessment
			BEAST	SkeenRIVAS		
Canyon Creek	BUL49	2006	not stressed	not stressed	Good	not stressed
Causqua Cr.	BUL40	2004	slightly stressed	severely stressed	Poor	stressed
Goathorn above Tenas	BUL16	2004	not stressed	not stressed	Poor	not stressed
Goathorn above Tenas	BUL16	2006	stressed	richer than reference	Good	slightly stressed
Gramophone Cr.	BUL37	2004	not stressed	not stressed	Fair	not stressed
Howson Cr.	BUL11	2004	not stressed	slightly stressed	Poor	slightly stressed
Jonas Cr.	BUL10	2004	not stressed	stressed	Poor	stressed
Kwun	BUL42	2004	not stressed	not stressed	Good	not stressed
Nichyeskwa 12k	BUL58	2005	slightly stressed	not stressed	Poor	slightly stressed
Nichyeskwa 14k	BUL60	2005	stressed	not stressed	Very Poor	stressed
Nichyeskwa 18k	BUL61	2005	slightly stressed	not stressed	Excellent	not stressed
Nichyeskwa 8k	BUL59	2005	not stressed	richer than reference	Excellent	not stressed
Pine Creek	BUL70	2006	not stressed	slightly stressed	Good	not stressed
Sinclair Cr.	BUL07	2004	not stressed	not stressed	Good	not stressed
Tenas Creek	BUL15	2006	slightly stressed	not stressed	Good	not stressed
Toboggan d/s	BUL33	2004	not stressed	not stressed	Good	not stressed
Toboggan d/s	BUL33	2005	not stressed	not stressed	Fair	not stressed
Toboggan u/s	BUL52	2004	not stressed	not stressed	Good	not stressed
Toboggan u/s	BUL52	2005	not stressed	not stressed	Good	not stressed

* These IBI scores were preliminary at the time of writing. The results have not yet been calibrated to the sampling methodology used in the RCA.

8.5 Point Source Discharges

While the RCA has broad application in monitoring non-point source disturbance over the landscape, it has equal application to monitoring point source discharges. Point sources include industrial and municipal discharges including tailings ponds, sewage treatment plants, storm water sewers, pulp mills and smelters to name a few.

When assessing the effects of point source discharges, sampling design will vary somewhat from that used for larger scale landscape monitoring. Reference sites should be sampled as general practice as part of activities to test model performance. In addition to sampling obvious sites that are close to a point source discharge, a gradient design, where test sites are placed at increasing intervals away from a discharge point, may be appropriate.

Sampling should be conducted in late summer or early fall to be consistent with the sampling periods used to build the existing RCA models. Should sampling be required at other times of the year, reference sites should be sampled and run through the RCA model to determine if it can accurately predict stream condition at those different times. If this testing shows that the existing model is performing well (e.g. it shows the reference sites are mostly in an unstressed condition), the present model can be used for site testing. If the model fails to predict a reference condition among several control site samples, it should not be used for site testing. In this case, the benthic invertebrate data can still be used but it must be analysed using an alternative design.

Mine sites are good examples where RCA can be applied for routine monitoring of point source discharges. In British Columbia, those discharges are permitted under the *Environmental Management Act.* Each permit specifies chemical and physical discharge limits and a monitoring program is required to assess the effects of the discharge on the aquatic environment. Permits generally require toxicity tests and regular water sampling. However, intermittent water quality grab samples may miss infrequent or pulsed and potentially toxic discharges that have deleterious effects on aquatic ecosystems. Because benthic invertebrate communities assimilate the cumulative chemical and physical disturbances in a watershed, they are often considered an essential component of an environmental effects monitoring program.

The following case example illustrates the potential use of the Reference Condition Approach at the Equity Mine, a closed silver mine in north-central British Columbia having a history of acid rock drainage. In 2002 and again in 2006 as part of the environmental effects monitoring program at the Equity Mine, benthic invertebrate abundance and community composition was used with several other chemical measurements to assess condition of drainage streams (Perrin 2007). Because the present sampling techniques were not used before mine development (thus eliminating the possible use of a BACI design) multiple lines of evidence in upstream to downstream comparisons have been used to define stream condition. This design is weak but it has been strengthened with experimental evidence from testing of treated mine drainage in an on-site mesocosm facility (Perrin et al. 1992). The approach of linking findings from well controlled experiments to multiple lines of evidence from on-site monitoring has resulted in powerful analyses to assess the effectiveness of treatment used at the mine to neutralize the acid drainage and control downstream transport of metals (cadmium, copper and zinc) and sulphate. In the monitoring component, one control site and two treatment sites on each of 2 main drainage streams (Foxy Creek and Buck Creek) have been sampled in a cycle once every 4 years (Figure 22). Results including temporal and spatial comparisons have shown that the mine is doing a good job in protecting downstream condition by collecting acid drainage and treating it with lime before discharge (Perrin 2007).

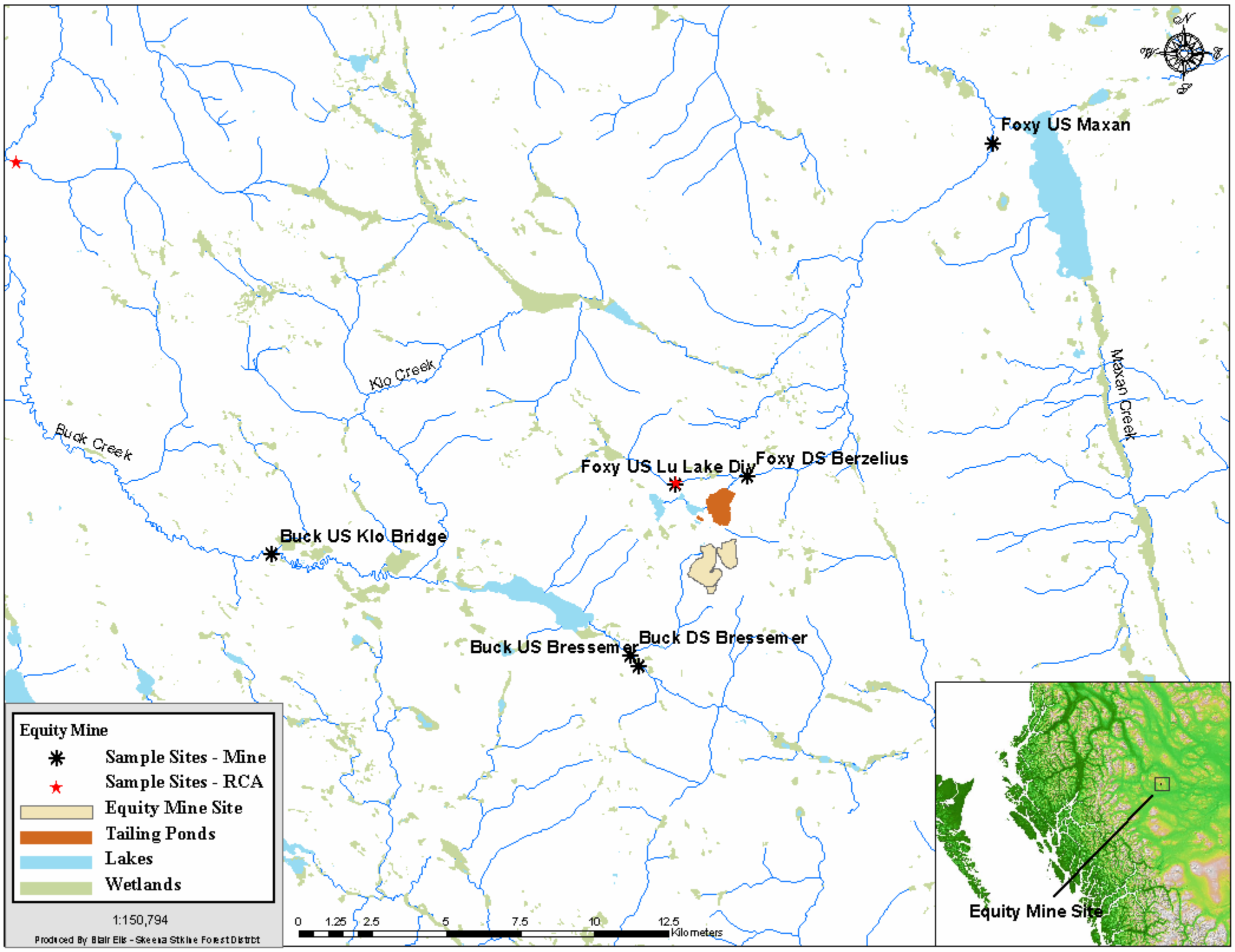

Figure 22. Locations of 2002 and 2006 sampling sites at the Equity Mine (Perrin 2007).

The Equity mine could replace the existing monitoring design with the RCA and achieve objectives of environmental effects monitoring. A difficulty with the present sampling layout at Equity is the absence of multiple control sites that would improve statistical rigor. This challenge is presently overcome with reference to site-specific experiments that assist with the interpretation of monitoring data and the use of multiple lines of evidence to support conclusions. The RCA has built-in control data in the model, thus reducing the concern about the lack of reference information in routine monitoring. Furthermore, the RCA is known to be good at assessing biological condition of streams along a gradient of trace metal concentrations. Sloane and Norris (2003) showed strong relationships between RCA results (AusRIVAS) and concentrations of metals in water and sediment.

Another advantage of the RCA at Equity would be a reduction in the ongoing monitoring cost. The cost of the benthic invertebrate monitoring, analysis and reporting that was completed in 2006 was approximately $50,000. The senior author of the present report estimates that the cost of a single round of sampling and interpretation would drop to $30,000 if RCA was applied. The saving would mainly be due to the difference in effort between the present use of basket samplers and RCA kick sampling methods. It would also be due to less effort required for data analysis and reporting using the RCA methods.

One disadvantage of switching to the RCA would be the loss of quantitative comparisons of new data to historical data. This loss would occur because of the change from the present use of basket samplers (Perrin 2007) to sample collection using a kick net. Data from the two methods are not comparable. Analysis of site condition over time has been an important part of stream assessments at Equity. Hence, any decision to change to the RCA would mean that monitoring would shift to pass and fail assessments with only qualitative reference to past results from what are presently quite detailed analyses of temporal and spatial variation using multivariate and univariate statistics.

Bioassessment is currently used across British Columbia as part of environmental effects monitoring programs at mines. These programs generally rely on the comparison of several biological endpoints (e.g. Bray Curtis distance, Simpson's diversity index, taxa richness) in a variety of control-impact study designs. The RCA could be used in place of this approach and is in fact one of the options outlined in the *Metal Mining EEM Guidance Document* (Environment Canada 2002) for monitoring at mines in British Columbia.

9 RECOMMENDATIONS AND USE OF THE RCA MODEL

9.1 Use of the model on the CABIN website

The BEAST model will be uploaded to the CABIN website in the near future and thereby made available for routine site testing. Instructions on using on-line tools and interpretation of the output is presently provided in a short course that is offered by Environment Canada at various locations in British Columbia on a frequency determined by demand. People completing the course are provided with a username and password to gain access to a secure area of the CABIN website that is used for site testing. In 2007, an on-line course is expected to be made available by Environment Canada to improve training logistics and access. Prospective users should visit the CABIN website at *http://cabin.cciw.ca* for updated information on the status and availability of this course. The B.C. Ministry of Environment is also planning for a centralised GIS service for all users from which GIS data will be accessible to users who provide information on test site locations. Again users can check for updates on availability of this service by visiting the CABIN website.

All predictor variables of the BEAST model that will be uploaded to CABIN are derived from GIS databases. That means users will be required to gain access to those databases to compile data for the predictor variables when running a site test. At present, there is no automated process on the CABIN website to gain access to spatial data in remote databases. Users must first derive values for each of the predictor variables and enter them when prompted on the CABIN website during the site testing process. Prompts to allow this manual data entry are not yet written into CABIN but discussion with representatives of Environment Canada suggest this facility may be available on CABIN in the near future. We recommend that GIS staff of the B.C. Ministry of Environment to work with GIS staff of Environment Canada or designated working group to develop software that will facilitate access to those databases and support user defined calculations of spatial attributes for use in running the Skeena BEAST model. This process would allow people who are not familiar with GIS techniques to quickly and easily run site tests on the CABIN website.

Although site testing using BEAST on the CABIN website requires only 5 GIS-derived predictor variables, it is important that the full suite of measurements that are listed on the field forms, including analyses of chemical analytes in water samples, be completed at each site being tested. Most of these additional variables will be “stressor variables” (defined in Section 4.2) that can be used after the test on CABIN to provide preliminary insight into possible cause of site disturbance. Definitive cause of stress, complete with tests of hypotheses, will most likely require more detailed monitoring or experimentation, either on-site or in a lab, if it is not immediately apparent. If that stress is unacceptable, either from a regulatory point of view or based on reference to environmental guidelines (e.g. CCME 2006) these additional data can help in defining the extent of remediation or compensation that may be required (e.g. defining the magnitude of decline in a nutrient discharge that is required to reduce eutrophication in a river or defining the extent of change in suspended solids concentration that is required from work to stabilize roads). Hence at no time should users delete measurements at test sites, even if they are not going to be used directly during on-line site testing using CABIN.

9.2 Revised field sheets

Three years of work in northwestern B.C. watersheds has resulted in a revised field sheet that future users can follow. That field sheet is included in Appendix B.

9.3 Further model development

The models that were developed in this project should be frequently updated to ensure they remain valid and to maintain their accuracy and precision. As new reference site data are acquired either by re-sampling existing sites or sampling new sites, the statistics can be easily re-run to compile new models that remain as up to date as the data will allow. If predictor variables change substantially over time, it will be useful to examine those shifts, particularly with respect to possible influence of climate change and change in progression of the Mountain Pine Beetle infestation. If this respect, RCA modeling should be considered an ongoing process.

There presently is weak coverage of reference sites over the range between Burns Lake and Prince George (Figure 10). Part of the reason for this gap was the difficulty in finding sites in that area that had not been disturbed in some way. This area is important with respect to range and forest management, particularly with regards to tracking the Mountain Pine Beetle infestation. It is recommended that additional effort be placed on adding reference sites in that area, since the RCA may be a valuable tool for examining ecological changes with progression of the infestation and land use activity. Sites already established as part of the Fraser Basin model (Sylvestre et al. 2005) in that area may be possible candidates either with respect to adding existing data or to use those sites for future sampling. Given that the beetle infestation is well established in that area *http://www.for.gov.bc.ca/hfp/mountain_pine_beetle/MPB_Magnitude_Maps_2001to2006.pdf*), a reference condition may be regarded as an infected area that does not have other stresses. Test site sampling may be used to follow time course response of surface waters to rapidly changing ecological dynamics over the landscape.

It is recommended that new reference sites be added to provide a validation data set for testing model accuracy and precision. This testing was completed on a limited scale in this project but it should be expanded as models are updated in future years.

We recommend that new reference sites be sampled that correspond to the sample groups that presently have a small sample size (e.g. Group 1 and possibly Group 5). Small sample size may be a factor in contributing to misclassification error in the models. Adding observations to the small sample groups may reduce that error.

Finally, we recommend that researchers continue to be engaged in future development of the RCA in British Columbia. This involvement will ensure that the models remain “cutting edge” and reflect the current state of knowledge internationally. This approach will also place British Columbia at the forefront in the rapidly changing science of bioassessment and its application to sustainable land management.

10 LIST OF REFERENCES

Bailey, R.C., M. G. Kennedy, M.Z. Dervish and R.M. Taylor. 1998. Biological assessment of freshwater ecosystems using a reference condition approach: comparing predicted and actual benthic invertebrate communities in Yukon streams. *Freshwater Biology* 39: 765-774.

Bailey, R.C., R.H. Norris, and T.B. Reynoldson. 2004. *Bioassessment of Freshwater Ecosystems: Using the Reference Condition Approach*. Kluwer Academic Publishers, Massachusetts, USA.

Barbour, M.T., J. Gerritsen, B.D. Snyder, and J.B. Stribling. 1999. Rapid Bioassessment Protocols for Use in Streams and Wadeable Rivers: Periphyton, Benthic Macroinvertebrates and Fish, Second Edition. EPA 841-B-99-002. U.S. Environmental Protection Agency; Office of Water; Washington, D.C. *http://www.epa.gov/owow/monitoring/rbp/download.html*

Beaudry, P. and Associates Ltd.. 2006. Results of the Stream Crossing Quality Index (SCQI) Survey for the Nichyeskwa Watershed, Skeena Stikine Forest District. Prepared for the Trustees of the Babine Watershed Monitoring Trust. Prince George, B.C.

Bennett, S. 2004. Benthic Invertebrate Index of Biological Integrity: 2003 Field Season Results for Streams in the Skeena Stikine Forest District. Prepared for Community Futures Development Corporation of Nadina, Smithers, B.C.

Bennett, S. and K. Rysavy. 2003. Guidelines for calibrating a benthic invertebrate multimetric index of biological integrity (B-IBI) for streams in British Columbia. Report prepared by Bio Logic Consulting Ltd. for B.C. Ministry of Sustainable Resource Management. Victoria, B.C. 43p.

Bothwell, M.L. 1989. Phosphorus-limited growth dynamics of lotic periphyton diatom communities: areal biomass and cellular growth rate responses. Can. J. Fish. Aquat. Sci. 46:1293-1301.

Bowman, M.F. and K.M. Somers. 2005. Considerations when using the reference condition approach for bioassessment of freshwater ecosystems. *Water Qual. Res. J. Canada* 40(3): 347-360.

Buzby, K.M. and S.A. Perry. 2000. Modeling the potential effects of climate change on leaf pack processing in central Appalachian streams. Can. J. Fish. Aquat. Sci. 57: 1773-1783.

Campbell, P.G.C. and P.M. Stokes. 1985. Acidification and toxicity of metals to aquatic biota. Can.J. Fish. Aquat. Sci. 42: 2034-2049.

CCME. 2006. Canadian Environmental Quality Guidelines. Canadian Council of Ministers of the Environment. Winnipeg. Manitoba.

Clarke, K.R. and R.N. Gorley. 2001. PRIMER v5: User Manual/Tutorial. PRIMER-E: Plymouth. UK.

Clarke, K.R. and R.M. Warwick. 2001. Change in marine communities: an approach to statistical analysis and interpretation, 2nd edition. PRIMER-E: Plymouth UK.

Croft, C. 2004. Development and calibration of the Morice and Lakes multimetric benthic index of biological integrity - Year 2: A biomonitoring study for assessing stream condition in the Morice and Lakes TSA's. Report prepared by Gartner Lee Ltd.

Deegan, L.A. and B.J. Peterson. 1992. Whole-river fertilization stimulates fish production in an arctic tundra river. Can. J. Fish. Aquat. Sci. 49:1890-1901.

Deegan, L.A., B.J. Peterson, H. Golden, C.C. McIvor, and M.C. Miller. 1997. Effects of fish density and river fertilization on algal standing stocks, invertebrate communities, and fish production in arctic tundra river. Can. J. Fish. Aquat. Sci. 54: 269-283.

Demarchi, D.A. 1996. An introduction to the ecoregions of British Columbia. Wildlife Branch. B.C. Ministry of Environment. Victoria, B.C.

Demarchi, D.A., R.D. Marsh, A.P. Harcombe, and E.C. Lea. 1990. The Environment. In. Campbell, R.W., N.K. Dawe, I. McTaggart-Cowan, J.M. Cooper, G. W. Kaiser, and M.C.E. McNall. 1990. The Birds of British Columbia. Volume 1. UBC Press. Vancouver, B.C.

Environment Canada. 2002. Metal mining guidance document for aquatic environmental effects monitoring. *http://www.ec.gc.ca.eem*

Glozier, N.E., J.M. Culp, and D. Halliwell. 2002. Revised guidance for sample sorting and sub-sampling protocols for EEM benthic invertebrate community surveys. National Water Research Institute. Environment Canada. Saskatoon Sask.

Green, R. H. 1999. Review of Potentially Applicable Approaches to Benthic Invertebrate Data Analysis and Interpretation. Discussion Paper prepared for Aquatic Effects Technology Evaluation (AETE) Program, Ottawa, Ontario. 38p.

Growns, J.E., B.C. Chessman, J.E. Jackson, and D.G. Ross. 1997. Rapid assessment of Australian rivers using macroinvertebrates: cost and efficiency of 6 months of sample processing. J. N. Am. Benthol. Soc. 16: 682-693.

Hickey, C.W. and W.H. Clements. 1998. Effects of heavy metals on benthic macroinvertebrate communities in New Zealand streams. Environ. Toxicology and Chemistry. 17: 2338-2346.

Horn, H. and G.C. Tamblyn. 2000. Morice Planning Area Background Report: An Overview of Natural, Cultural, and Socio-Economic Features, Land Uses and Resources Management. Prepared for the Prince Rupert Interagency Management Committee, Smithers, B.C.

Horvath, S (editor). 2005. British Columbia Environmental Laboratory Manual — For the Analysis of Water, Wastewater, Sediment, Biological Materials and Discrete Ambient Air Samples. 2005. Water and Air Monitoring and Reporting Section, Ministry of Environment.

Johnson, J.B. and K.S. Omland. 2004. Model selection in ecology and evolution. *TRENDS in Ecology and Evolution* 19(2): 101-108.

Johnston, N.T., C.J. Perrin, P.A. Slaney, and B.R. Ward. 1990. Increased juvenile salmonid growth by whole-river fertilization. Can. J. Fish. Aquat. Sci. 47(5): 862-872.

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.

Krebs. C.J. 1999. Ecological Methodology. Addison Wesley Longman.

Linke, S., R. Norris, D. P. Faith, and D. Stockwell. 2005. ANNA: A new prediction method for bioassessment programs. Freshwater Biology 50:147-158.

Mah, S., S. Thomson, and D. Demarchi. 1996. An ecological framework for resource management in British Columbia. Environmental Monitoring and Assessment 39: 119-125.

Marchant, R. 1989. A sub-sampler for samples of benthic invertebrates. Bull. Aust. Soc. Limnol. 12: 49-52.

McCune, B., and M. J. Mefford. 1999. Multivariate Analysis of Ecological Data. MjM Software, Gleneden Beach, Oregon, U.S.A.

Mellina, E., R.D. Moore, S. G. Hinch, J.S. Macdonald, and G. Pearson. 2002. Stream temperature responses to clearcut logging in British Columbia: the moderating influences of groundwater and headwater lakes. Can. J. Fish. Aquat. Sci. 59: 1886-1900.

Ministry of Sustainable Resource Management. 2004. Morice Land and Resource Management Plan: Final Land Use Recommendation. Smithers, B.C.

Mundie, J.H., K. Simpson, and C.J. Perrin. 1991. Responses of stream periphyton and benthic insects to increases in dissolved inorganic phosphorus in a mesocosm. Can. J. Fish. Aquat. Sci. 48(11): 2061-2072.

Murphy, T.P., K.J. Hall, and I. Yesaki. 1983. Coprecipitation of phosphate with calcite in a naturally eutrophic lake. Limnol. Oceanogr. 28: 58-69.

Nichols, S.J. and R.H. Norris. 2006. River condition assessment may depend on the sub-sampling method: field live-sort versus laboratory sub-sampling of invertebrates for bioassessment. *Hydrobiologia* 572: 195-213.

Norris, R., P. Liston and S. Linke. 2002. River condition: managing different assessments. Presented at North American Benthological Society 50th Annual Meeting. Pittsburgh, PA.

Norris, R.H. and C.P. Hawkins. 2000. Monitoring river health. *Hydrobiologia.* 435: 5-17.

Parsons, M. and R.H. Norris. 1996. The effect of habitat-specific sampling on biological assessment of water quality using a predictive model. Freshwater Biology. 36: 419-434.

Perrin, C.J. 2007. Benthic invertebrate and periphyton monitoring in drainages near the Equity Mine, 2006. Report prepared by Limnotek Research and Development Inc. for Goldcorp Inc. Equity Mine. 50p plus appendices.

Perrin, C.J. 2006. Periphyton and benthic invertebrate monitoring for water use planning in the Coquitlam River, 2006. Report prepared by Limnotek Research and Development Inc. for BC Hydro. 51p.

Perrin, C.J. 2001. Trophic structure and function in the Cheakamus River for water use planning. Report prepared by Limnotek Research and Development Inc. for BC Hydro and Resort Municipality of Whistler. 66p.

Perrin, C.J. 1998. Phosphorus transport and periphyton accrual in the Cheakamus River. Report prepared by Limnotek Research and Development Inc. for Resort Municipality of Whistler and BC Hydro. 126p.

Perrin, C.J. and C.A. Blyth. 1998. An ecozone classification for lakes and streams of British Columbia; Version 1.0. Report prepared by Limnotek Research and Development Inc. and AXYS Environmental Consulting Ltd. for Ministry of Environment, Lands and Parks. Water Quality Branch. Victoria, B.C. 95p plus map.

Perrin, C.J., M.L. Bothwell and P.A. Slaney. 1987. Experimental enrichment of a coastal stream in British Columbia: effects of organic and inorganic additions on autotrophic periphyton production. Can. J. Fish. Aquat. Sci. 44(6): 1247-1256.

Perrin, C.J. and J.S. Richardson. 1997. N and P limitation of benthos abundance in the Nechako River, British Columbia. Can. J. Fish. Aquat. Sci. 54: 2574-2583.

Perrin, C.J., B. Wilkes, and J.S. Richardson. 1992. Stream periphyton and benthic insect responses to additions of treated acid mine drainage in a continuous-flow on-site mesocosm. Environmental Toxicology and Chemistry. 11: 1513-1525.

Peterson, C.G. and N.B. Grimm. 1992. Temporal variation in enrichment effects during periphyton succession in a nitrogen-limited desert stream ecosystem. J. N. Am. Benthol. Soc. 11: 20-36.

Reice, S.R. and M. Wohlenberg. 1993. Monitoring freshwater benthic macroinvertebrates and benthic processes: measures for assessment of ecosystem health. In: Rosenberg, D.M. and V.H. Resh. (Ed). 1993. Freshwater biomonitoring and benthic macroinvertebrates. Chapman and Hall. New York.

Remington, D. and J. Lough. 2005. Water Quality Assessment and Objectives for Toboggan Creek and Tributaries: Technical Report. Prepared for the Ministry of Water, Land and Air Protection, Smithers, B.C.

Resources Inventory Committee (RIC) 1997. B.C. Ambient Freshwater and Effluent Sampling Manual. Prepared by B.C. Environment.

Resources Inventory Committee (RIC) 1998. Manual of Standard Operating Procedures for Hydrometric Surveys in B.C. Prepared by Ministry of Environment, Lands and Parks, Resources Inventory Branch for the Aquatic Inventory Task Force Resources Inventory Committee. November 2, 1998. Version 1.1

Resources Inventory Committee (RIC) 2001. Reconnaissance (1:20 000) Fish and Fish Habitat Inventory Standards and Procedures. Prepared by B.C. Fisheries Information Services Branch for the Resources Inventory Committee April, 2001. Version 2.0.

Reynoldson, T.B., R.C. Bailey, K.E. Day, and R.H. Norris.1995. Biological guidelines for freshwater sediment based on BEnthic Assessment of SedimenT (the BEAST) using a multivariate approach for predicting biological state. Australian Journal of Ecology 20: 198-219.

Reynoldson, T.B., C. Logan, T. Pascoe and S.P. Thompson. 2003. CABIN (Canadian Aquatic Biomonitoring Network) Invertebrate Biomonitoring Field and Laboratory Manual. National Water Research Institute, Environment Canada.

Reynoldson, T.B., R.H. Norris, V.H. Resh, K.E. Day, and D.M. Rosenberg. 1997. The reference condition approach: a xcomparison 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.

Richardson, J.S. and C.J. Perrin. 1994. Effects of the bacterial insecticide, *Bacillus thuringiensis* var. *kurstaki* (Btk) on a stream benthic community. Can. J. Fish. Aquat. Sci. 51:1037-1045.

Richardson, J.S. and R.J. Mackay. 1991. Lake outlets and the distribution of filter feeders: an assessment of hypotheses. Oikos 62: 370-380.

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. FRAP Rep. No. DOE-FRAP 1998-32. Fraser River Action Plan, Environment Canada, Vancouver, B.C.

Rosenberg, D.M. and V.H. Resh. (Ed). 1993. Freshwater biomonitoring and benthic macroinvertebrates. Chapman and Hall. New York.

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.

SAS. 2007. SAS version 9. SAS Institute Inc. NC.

Simpson, J. C., and R. H. Norris. 2000. Biological assessment of river quality: development of AUSRIVAS models and outputs. Pages 125-142 in J. F. Wright, D. W. Sutcliffe, and M. T. Furse, editors. Assessing the biological quality of fresh waters; RIVPACS and other techniques. Freshwater Biological Association, Ambleside, Cumbria, U.K.

Sloane, P.I.W. and R.H. Norris. 2003. Relationship of AUSRIVAS-based Macroinvertebrate Predictive Model Outputs to a Metal Pollution Gradient. J. N. Am. Benthol. Soc. 22(3):457-471.

Somers, K.M., R.A. Reid, and S.M. David. 1998. Rapid biological assessments: how many animals are enough? J.N. Amer. Benthol. Soc. 17: 348-358.

Sovell, L.A. and B. Vondracek. 1999. Evaluation of the fixed count method for rapid bioassessment protocol III with benthic macroinvertebrate metrics. J. N. Amer. Benthol. Soc. 18: 420-426.

Stewart-Oaten, A. W.M. Murdoch, and K.R. Parker. 1986. Environmental impact assessment: “Pseudoreplication” in time? Ecology. 67: 929-940.

Stockner, J. and K.R.S. Shortreed. 1978. Enhancement of autotrophic production by nutrient addition in a coastal rainforest stream on Vancouver Island. J. Fish. Res. Board. Can. 35: 28-34.

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. Environmental Conservation Branch, Environment Canada, Vancouver B.C. EC/GB/04/81. 54p.

SYSTAT 11. 2004. SYSTAT Software Inc. Richmond, CA.

Tamblyn, G.C. 2005. Morice Watershed-based Fish Sustainability Plan: A Plan to Conserve and Protect Morice Watershed Fish Populations. Prepared for Fisheries and Oceans Canada and Ministry of Water, Land and Air Protection. Smithers, B.C.

Underwood, A.J. 1994. On beyond BACI: sampling designs that might reliably detect environmental disturbances. Ecological Applications 4: 3-15.

Van Sickle, J., C. P. Hawkins, D. P. Larsen, and A. T. Herlihy. 2005. A null model for the expected macroinvertebrate assemblage in streams. Journal of the North American Benthological Society 24:178–191.

Van Sickle, J., D.D. Huff, and C.P. Hawkins. 2006. Selecting discriminant function models for predicting the expected richness of aquatic macroinvertebrates. *Freshwater Biology* 51: 359-372.

Vannote, R.L. and B.W. Sweeney. 1980. Geographic analysis of thermal equilibria: a conceptual model for evaluating the effect of natural and modified thermal regimes on aquatic insect communities. Am. Nat. 115: 667-695.

Vinson, M.R. and C.P. Hawkins. 1996. Effects of sampling area and sub-sampling procedure on comparisons of taxa richness among streams. J. N. Am. Benthol. Soc. 15(3): 392-399.

Wetzel, R.G. 2001. Limnology: Lake and river ecosystems. Academic Press. San Diego. CA. 1006p.

Wolman. M. 1954. A method of sampling coarse bed material. American Geophysical Union, Transactions, 35:951-956.

Wright, J. F. 1995. Development and use of a system for predicting macroinvertebrates in flowing waters. Australian Journal of Ecology 20: 181-197.

Wright, J.F., D.W. Sutcliffe and M.T. Furse. 2000. Assessing the Biological Quality of Freshwaters. RIVPACS and Other Techniques. Freshwater Biological Association, Ambleside, UK. 373pp.

Personal Communication:

Baxter, Alan. April 2007. Planning Forester. Pacific Inland Resources, West Fraser Timber Company Ltd., Smithers, B.C.

Bilodeau, Norm. April 2007. Regional Manager, Northern Region, Water Stewardship Division, Ministry of Environment. Prince George, B.C.

11 APPENDIX A: FIELD DATA FORM USED IN ALL YEARS

Stream Name		EMS #
Date	Time	Field Crew

A. Weather Conditions

Now: ☐ storm (heavy rain) **Past 24 hours:** ☐ storm (heavy rain)
☐ rain (steady rain) ☐ rain (steady rain)
☐ showers (intermittent) ☐ showers (intermittent)
☐ overcast ☐ overcast
☐ clear/ sunny ☐ clear/ sunny

Has there been a heavy rain in the past 7 days? ☐ Y ☐ N

B. General Site Information

GPS Unit #	GPS Datum **NAD83**	Elevation (m)
Latitude (decimal degrees)	Longitude (decimal degrees)	Waypoint Name
Site Description		

Sample Site Diagram (draw a diagram of the site and indicate areas sampled; include a scale)

Map Scale: 0 ← ---→ ____m

Photos:

☐ Field Sheet ☐ Upstream ☐ Downstream ☐ Across ☐ Substrate

C. Water Quality

Field Measurements:

Air Temp (ºC)	Water Temp (ºC)	pH	Spec. Conductance (µS)	D.O. (mg/L)

Water Samples:

☐ General Ions (1L) ☐ Nutrients (250mL) ☐ TOC (250mL) ☐ Metals (Do not rinse)

☐ *Preserved (H_2SO_4)* ☐ *Preserved (HNO_3)*

D. Benthic Invertebrates

Sample riffle zones using 400µm Kicknet for 3 minutes

Sample Number	1	2 (QA/QC)
Operator Name		
Typical Depth (cm)		
Number of sample jars filled		

Notes:

E. Coverage

Overall Percent Cover (estimate % of the wetted surface area that is covered by the following, within 1m of the water surface)

Woody debris	
Boulders	
Undercut banks	
Deep pools	
Overhanging Vegetation	

Macrophyte Coverage (circle the one which describes the amount of stream bed covered by macrophytes)

0% 1-25% 26-50% 51-75% 76-100%

(Note: Include moss, but add comments)

Periphyton Coverage (circle the one which best describes the presence of periphyton ***in running water***)

1. Rocks not slippery at all, no colour
2. Rocks slightly slippery, light yellow-brown in colour
3. Rocks have noticeable slippery feel, slippery to walk on, may be some patches of green/brown algae
4. Rocks are very slippery, can rub algae off with finger, and may be numerous large clumps of algae, dark brown colour
5. Rocks mostly obscured by algal mat, extensive algal mass (may be in long strands) brown or black colour

F. Disturbance Indicators

Indicate the presence of the following disturbance indicators at the site:

Bed Characteristics

☐ Extensive areas of scour
☐ Large extensive sediment wedges
☐ Extensive riffle zones
☐ Extensive areas of (unvegetated) bar
☐ Elevated mid-channel bars
☐ Limited pool frequency and extent

Channel Pattern

☐ Multiple channels (braiding)

Banks

☐ Eroding banks
☐ Isolated sidechannels or backchannels

Large Woody Debris

☐ Most LWD parallel to banks ☐ Recently formed LWD jams

Stream Name	Date

G. Stream Characteristics

☐ Glacial ☐ Clear ☐ Stained ☐ Other___________________________

Gradient: _______________ (report % using clinometer)

Habitat Units Present (estimate % of visible channel area occupied by each; consider 6x bankfull width)

_____ Pools _____ Glides _____ Riffles _____ Cascades _____ Other

Stream Widths (measure wetted width and bankfull width at 3 different locations within the 6x bankfull; Upstream end measurement must be from within kicked area)

Stream Location	Wetted Width (m)	Bankfull Width (m)
Downstream		
Middle		
Upstream		

Stream Profile at Upstream End (from somewhere in the kicked area, measure depth and velocity at 5-8 equidistant points across the stream)

Current meter used for velocity measurements: ________________________

Wetted Width (m)									
Tape reading (m)									
Depth (cm)	0								0
Velocity (m/s)	0								0

H. Substrate Characteristics

% Composition (estimate the relative % composition of each substrate type within the reach)

____sand	____gravel	____pebble	____cobble	____boulder	_____bedrock
(< 2mm)	(2-4mm)	(4mm – 6cm)	(6cm – 26cm)	(>26cm)	

Embeddedness (circle the one which describes how embedded the cobbles are in the riffle zone)

Not Embedded ¼ ½ ¾ Completely Embedded

Odours and Oils (indicate the presence of the following in the substrate)

Odours: ☐ None ☐ Sewage ☐ Petroleum ☐ Anaerobic (H_2S) ☐ Chemical ☐ Other____

Oils: ☐ Absent ☐ Slight ☐ Moderate ☐ Profuse

I. Riparian Vegetation

Vegetation Types (estimate the % of each vegetation type present at the site (total 100%))

Unvegetated (bare soil visible)		Deciduous Forest (trees >5m tall)	
Grass/Herb		Coniferous Forest	
Shrub (may include			

grasses/herbs growing beneath shrubs)		

Dominant Species Present

Structural Stage (indicate the structural stage of the dominant vegetation)

☐ Non-vegetated or initial stage following disturbance, with less than 5% cover
☐ shrub / herb stage, less than 10% tree cover
☐ pole-sapling stage, with trees overtopping the shrub layer, usually less than 15-20 years old
☐ young forest (30- 80 years) - forest canopy is differentiating into distinct layers
☐ mature forest - well developed understory

Canopy Closure (circle the proportion of the surface area of the stream covered by the projecting riparian canopy; hint - stand in the middle of the stream and look up!)

0% 1-20% 21-40% 41-70% 71-90% >90% covered

Mountain Pine Beetle Infestation - Forest Health (estimate riparian by considering 30m from stream and estimate watershed based on observations en route to site)

	RIPARIAN	WATERSHED
% Trees that are Pine		
% Pine Trees that are “Red”		
% Pine Trees that are “Grey”		

J. Land Use

Predominant Surrounding Land Use

☐ Forest
☐ Logging
☐ Field / Pasture
☐ Mining
☐ Agricultural
☐ Commercial / Industrial
☐ Residential
☐ Other

Local Watershed Erosion

☐ Heavy
☐ Moderate
☐ Light

Local Watershed NPS Pollution

☐ Obvious sources
☐ Some potential sources
☐ No evidence

☐ None comments: ____

Stream Name	Date

Pebble Count (zig-zag through the benthic sampling area, stopping every 2 steps to select and measure pebble diameter – record the intermediate diameter to the nearest 0.1cm)

Pebble #	Diameter (cm)	Pebble #	Diameter (cm)	Pebble #	Diameter (cm)	Pebble #	Diameter (cm)
1		26		51		76	
2		27		52		77	
3		28		53		78	
4		29		54		79	
5		30		55		80	
6		31		56		81	
7		32		57		82	
8		33		58		83	
9		34		59		84	
10		35		60		85	
11		36		61		86	
12		37		62		87	
13		38		63		88	
14		39		64		89	
15		40		65		90	
16		41		66		91	
17		42		67		92	
18		43		68		93	
19		44		69		94	
20		45		70		95	
21		46		71		96	
22		47		72		97	
23		48		73		98	
24		49		74		99	
25		50		75		100	

12 APPENDIX B: RECOMMENDED FIELD DATA FORM FOR FUTURE USE

B.C. Ministry of Environment

Benthic Macroinvertebrate Stream Assessments for Application of the Reference Condition Approach (RCA)

Field Sheet for CABIN Assessments

Stream Name		EMS #
Date	Time	Field Crew

A. Weather Conditions

Now:	**Past 24 hours:**
☐ storm (heavy rain)	☐ storm (heavy rain)
☐ rain (steady rain)	☐ rain (steady rain)
☐ showers (intermittent)	☐ showers (intermittent)
☐ overcast	☐ overcast
☐ clear/ sunny	☐ clear/ sunny

Has there been a heavy rain in the past 7 days? ☐ Y ☐ N

B. General Site Information

Potential Reference Site? ☐ Y ☐ N

GPS Unit #	GPS Datum	Elevation (m)
Latitude (decimal degrees)	Longitude (decimal degrees)	Waypoint Name
Site Description		

Sample Site Diagram (draw a diagram of the site and indicate areas sampled; include a scale)

Map Scale: 0 ← --→ ____m

Photos:

☐ Field Sheet ☐ Upstream ☐ Downstream ☐ Across ☐ Substrate (use grid)

C. Water Quality

Field Measurements:

Air Temp (ºC)	Water Temp (ºC)	pH	Spec. Conductance (µS)	D.O. (mg/L)

Water Samples:

☐ General Ions (1L) ☐ Nutrients (250mL) ☐ TOC (250mL) ☐ Metals (Do not rinse)

☐ *Preserved* (H_2SO_4) ☐ *Preserved* (HNO_3)

D. Benthic Invertebrates

Sample riffle zones using 400µm Kicknet for 3 minutes

Sample Number	1	2 (QA/QC)
Operator Name		
Typical Sampling Depth (cm)		

Number of sample jars filled		

Notes:

E. Stream Channel Characteristics

Gradient: ______________ (report % using clinometer)

Habitat Units Present (estimate % of visible channel area occupied by each; consider 6x bankfull width)

_____ Pools _____ Glides/Runs _____ Riffles _____ Cascades/Rapids _____ Other

Stream Widths (measure wetted width and bankfull width at 3 different locations within the 6x bankfull; Middle measurement should be from within kicked area)

Stream Location	Wetted Width (m)	Bankfull Width (m)
Downstream		
Middle		
Upstream		

Stream Profile at Upstream End (from somewhere in the kicked area, measure depth and velocity at 5-8 equidistant points across the stream)

Current meter used for velocity measurements: _______________________

Wetted Width (m)										
Tape reading (m)										
Depth (cm)	0									0
Velocity (m/s)	0									0

F. Macrophytes and Periphyton

Macrophyte Coverage (circle the one which describes the amount of stream bed covered by macrophytes; include moss, but add comments)

0%   1-25%   26-50%   51-75%   76-100%

Periphyton Coverage (circle the one which best describes the presence of periphyton ***in running water***)

1 Rocks not slippery at all, no colour

2 Rocks slightly slippery, light yellow-brown in colour

3 Rocks have noticeable slippery feel, slippery to walk on, may be some patches of green/brown algae

4 Rocks are very slippery, can rub algae off with finger, and may be numerous large clumps of algae,
dark brown colour

5 Rocks mostly obscured by algal mat, extensive algal mass (may be in long strands) brown or black
colour

G. Cover

Overall Percent Cover (observe the entire stream reach and visually estimate % coverage of the wetted surface area, within 1 m of the water surface; consider the cover types listed below) ________%

Cover Types and Amount (indicate abundance of each cover type by ticking the appropriate category)

Cover Type	None	Trace	Moderate	Abundant
Woody debris				
Boulders				
Undercut banks				
Deep pools				
Overhanging Vegetation				

H. Riparian Vegetation

Dominant Riparian Class (indicate which class describes the dominant vegetative cover in the riparian)

☐ Unvegetated (bare soil present)
☐ Grass/Herb
☐ Shrub (may include grasses and herbs beneath)
☐ Deciduous Forest
☐ Coniferous Forest
☐ Mixed Forest

Structural Stage (indicate the structural stage of the dominant vegetation)

☐ Non-vegetated or initial stage following disturbance, with less than 5% cover

☐ Shrub / herb stage, less than 10% tree cover
☐ Pole-sapling stage, with trees overtopping the shrub layer, usually less than 15-20 years old
☐ Young forest (30- 80 years) - forest canopy is differentiating into distinct layers
☐ Mature forest - well developed understory

Canopy Closure (circle the proportion of the surface area of the stream covered by the projecting riparian canopy; hint - stand in the middle of the stream and look up!)

0% 1-25% 26-50% 51-75% 76-100%

I. Substrate Characteristics

% Composition (estimate % composition of each substrate type within the reach – Wentworth Scale)

____sand	____gravel	___pebble	____cobble	____boulder	_____bedrock
(< 2mm)	(2-4mm)	(4mm – 6cm)	(6cm – 26cm)	(>26cm)	

Embeddedness (circle the one which describes how embedded the cobbles are in the riffle zone)

Not Embedded ¼ ½ ¾ Completely Embedded

Wolman Pebble Count (zig-zag along the entire stream reach, stopping every 2 steps to select and measure pebble diameter – record the intermediate diameter to the nearest 0.1cm)

Pebble #	Diameter (cm)	Pebble #	Diameter (cm)	Pebble #	Diameter (cm)	Pebble #	Diameter (cm)
1		26		51		76	
2		27		52		77	
3		28		53		78	
4		29		54		79	
5		30		55		80	
6		31		56		81	
7		32		57		82	
8		33		58		83	
9		34		59		84	

Pebble #	**Diameter (cm)**	**Pebble #**	**Diameter (cm)**	**Pebble #**	**Diameter (cm)**	**Pebble #**	**Diameter (cm)**
10		35		60		85	
11		36		61		86	
12		37		62		87	
13		38		63		88	
14		39		64		89	
15		40		65		90	
16		41		66		91	
17		42		67		92	
18		43		68		93	
19		44		69		94	
20		45		70		95	
21		46		71		96	
22		47		72		97	
23		48		73		98	
24		49		74		99	
25		50		75		100	

B.C. Ministry of Environment

Benthic Macroinvertebrate Stream Assements for Application of the Reference Condition Approach (RCA)

Recommended Additional Observations

Stream Name	Date

Site Information:

Directions for accessing the site

Stream Type:

☐ Glacial ☐ Clear ☐ Stained ☐ Other________________________

Disturbance Indicators:

Indicate the presence of the following disturbance indicators at the site:

Bed Characteristics

☐ Extensive areas of scour
☐ Large extensive sediment wedges
☐ Extensive riffle zones
☐ Extensive areas of (unvegetated) bar
☐ Elevated mid-channel bars
☐ Limited pool frequency and extent

Channel Pattern

☐ Multiple channels (braiding)

Banks

☐ Eroding banks
☐ Isolated sidechannels or backchannels

Large Woody Debris

☐ Most LWD parallel to banks
☐ Recently formed LWD jams

Substrate Characteristics:

Odours and Oils (indicate the presence of the following in the substrate)

Odours: ☐ None ☐ Sewage ☐ Petroleum ☐ Anaerobic (H_2S) ☐ Chemical ☐ Other____

Oils: ☐ Absent ☐ Slight ☐ Moderate ☐ Profuse

Land Use

Surrounding Land Use (consider what is visible from the sample site, or known/suspected to occurring upstream; tick all that apply)

☐ None (Forest) ☐ Range ☐ Agriculture (incl. fields) ☐ Residential
☐ None (Park) ☐ Mining ☐ Commercial / Industrial ☐ Recreation

Notes:

Local Watershed Erosion (visible at sample site)

☐ Heavy
☐ Moderate Comments:
☐ Light
☐ None

Local Watershed NPS Pollution

☐ Obvious sources
☐ Some potential sources Comments:
☐ No evidence

Mountain Pine Beetle:

Mountain Pine Beetle Infestation - Forest Health (estimate riparian by considering 30m from stream and estimate watershed based on observations en route to site)

	RIPARIAN	WATERSHED
% Trees that are Pine		
% Pine Trees that are "Red"		
% Pine Trees that are "Grey"		

13 RAW DATA APPENDICES

Raw data appendices are available on CD or via file transfer from the Ministry of Environment in Smithers, British Columbia.