

A revised predictive model for bioassessment of streams in northwest
British Columbia using the reference condition approach:

Skeena model

August 3, 2011



A revised predictive model for bioassessment of streams in northwest
British Columbia using the reference condition approach:

Skeena model

Submitted to:

Environmental Protection Division
Skeena Region
Ministry of Environment

Prepared by:

Shauna A. Bennett, MSc. RPBio.
Bio Logic Consulting

August 3, 2011

Citation: Bennett, S.A. 2011. A revised predictive model for bioassessment of streams in northwest British Columbia using the reference condition approach: Skeena model. Report prepared by Bio Logic Consulting for Environmental Protection Division, Ministry of Environment, Skeena Region. 35p.

Cover image: Anweiler Creek. Photo by Jason Bennett.

Abstract

The reference condition approach (RCA) to stream bioassessment has been developed to overcome common impact assessment design difficulties (e.g. lack of suitable upstream reference site in an upstream-downstream approach) (Bailey *et al.* 1998, Bailey *et al.* 2004). RCA has been gaining momentum in Canada where tools are needed for screening and biomonitoring sites. Environment Canada has set up a publicly available database, Canadian Aquatic Biomonitoring Network (CABIN) for storing and sharing data, and running test site assessments.

Perrin *et al.* (2007) sampled 86 reference sites and 170 test sites in north-central and north western B.C. to build and compare invertebrate-based predictive bioassessment models based on abundance (Reynoldson *et al.* 1995, Reynoldson *et al.* 1997) and presence-absence (Parsons and Norris 1996) of taxa. In 2009, the original abundance-based predictive model was updated to include an additional 59 reference sites sampled in 2007 to 2009, to include a greater and more varied number of potential predictor variables in the environment data set, to allow a thorough quality assurance check of all the biological data in CABIN, and to re-run the GIS analysis with a quality assurance process in place for checking waypoint locations and catchment boundaries. Using the quality assured data set of 145 reference sites, the Skeena model was created. In the Skeena model, a combination of 6 predictor variables were best able to differentiate the four reference groups determined using biotic assemblage based clustering. The predictor variables were percent ice area, total precipitation in January, percent wetlands area, percent intrusive rock, total snow in January and percent lakes area. The probability of assigning a site to the proper group was 77%, measured by the overall jackknife classification success. Model accuracy was estimated by testing multiple samples collected from 7 reference sites, and in 71% of cases, the test samples were accurately assessed. Model precision was estimated by assessing the degree of mutual agreement between site assessments completed using independent observations (samples) collected from a single test site on the same day. Model precision was 75%.

The Skeena model has been uploaded to the CABIN website and is publicly available for use. The purpose of this document is to describe the revised predictive bioassessment model, called *Skeena model 2010* on CABIN, and to provide background information for distribution on the CABIN website to support the use of the revised model by consultants, government and industry biologists, foresters and resource managers that have some knowledge of CABIN and RCA.

Keywords: bioassessment, benthic invertebrates, reference condition approach, freshwater streams, CABIN, Skeena RCA

Table of Contents

<i>Abstract</i>	<i>iv</i>
<i>Table of Contents</i>	<i>v</i>
<i>Acknowledgements</i>	<i>vi</i>
1 Introduction	1
1.1 Purpose of this document	3
2 Methods	3
2.1 Study area and site descriptions	3
2.2 Field stream assessments	6
2.3 Watershed delineation and compilation of landcover data	7
2.4 Selection of reference sites	8
2.5 Model Development	8
3 Results and Discussion	11
3.1 Definition of reference groups	11
3.2 Predictor variables	15
3.3 Descriptive summary of the reference groups	16
3.4 Tests of Accuracy and Precision	19
3.5 Using RCA bioassessments for impact assessment	19
4 Important considerations	21
5 Recommendations	22
6 References	24
<i>Appendix A. Invertebrate Subsampling and Enumeration Procedures for Skeena (from Perrin et al. 2005).</i>	<i>26</i>
<i>Appendix B. Taxa list for Skeena model.</i>	<i>28</i>
<i>Appendix C. GIS variables, descriptions, source databases and source paths.</i>	<i>31</i>

Acknowledgements

Financial support for this project was provided by the B.C. Ministry of Environment. The work was an extension of a three year Forest Sciences Program (FSP) project originally completed with a large team of people both in the field, in various laboratories and in the office. These people were acknowledged in the original document created by Perrin *et al.* (2007), and in this report, I'd like mainly to acknowledge the authors that contributed to the original project including Chris Perrin, Simon Linke, A.J. Downie, Greg Tamblyn, Blair Ells, Ian Sharpe and Bob Bailey. This report is largely based on that original document (Perrin *et al.* 2007), and the Forest Investment Account (FIA) funded work that followed (Bennett 2008, Bennett 2009).

Greg Tamblyn, Impact Assessment Biologist for B.C. Environment was the contract manager for this project. Thank you to both Mr. Tamblyn and to Mr. A.J. Downie (Environmental Quality Section Head for B.C. Environment Skeena Region) for editing the draft document and providing input in developing a table of contents. Thank you to Leon Gaber, Provincial Bio-monitoring Science Specialist for B.C. Environment, and Stephanie Strachan, Freshwater Quality Monitoring Scientist for Environment Canada, for on-going scientific discussion regarding modeling and laboratory sorting inconsistencies between methods used in the Skeena and those recommended for CABIN.

Thank you to Simon Norris of Hillcrest Geographics, Victoria, B.C., for completing all GIS analyses and creating the related figures. A special thanks to Megan Hoole, GIS Analyst at B.C. Timber Sales, Jon Schulz, former Forest Planner at Coast Tsimshian Resources Limited Partnership, and Lance Loggin, Silviculture Forester at Westfraser Mills Ltd. for providing access to their spatial datasets for forest operations in Tree Farm Licence (TFL) areas.

1 Introduction

Environmental management in B.C. has largely adopted a ‘results based’ management philosophy. Consequently, to ensure the effectiveness of land use practices in protecting water quality and sustaining aquatic resources, new tools were needed to monitor and research the outcomes of management strategies as implemented. In 2004, the Forest Sciences Program (FSP) funded a 3 year program to develop an invertebrate-based stream bioassessment tool in northwest B.C. Benthic invertebrates are commonly the basis for stream bioassessment tools as they are ubiquitous, easy and inexpensive to sample, and generally include species sensitive to human disturbance (Davis 1994, Karr and Chu 1999).

A predictive model was built for north-central British Columbia in 2007 (Perrin *et al.* 2007) so that a study design known as the Reference Condition Approach (RCA) (Bailey *et al.* 2004, Sylvestre *et al.* 2005, Mazor *et al.* 2006) could be used for impact assessment. RCA has been developed to overcome common impact assessment design difficulties (e.g. lack of suitable upstream reference site in an upstream-downstream approach) (Bailey *et al.* 1998, Bailey *et al.* 2004). In the RCA, benthic invertebrate community data 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. The ability to compare a test site with an appropriate reference condition provides an effective environmental screening tool to monitor for impacts from land and water management activities. In British Columbia, the reference condition approach (RCA) to stream bioassessment is becoming more commonly employed (Reynoldson *et al.* 2001, Sylvestre *et al.* 2005, Mazor *et al.* 2006, Perrin *et al.* 2007, Perrin and Bennett 2010) and Environment Canada hosts a website with on-line training opportunities and an on-line database for sharing and storing data (Canadian Aquatic Biomonitoring Network or CABIN).

Beginning in 2004 in north-central and north western B.C. 256 stream sites were sampled. Sites that were not affected by anthropogenic disturbance were called **reference sites**. Sites affected by anthropogenic disturbance were called **test sites**. In total, 86 were classified as reference condition sites and were used to build a predictive bioassessment model called the *Skeena BEAST*¹ (Perrin *et al.* 2007). For model development, clustering and ordination techniques were used to assemble the reference sites

¹ The model was originally labeled using the ‘BEAST’ acronym (Reynoldson *et al.* 1995, Reynoldson *et al.* 1997). The updated version of the model has been renamed the *Skeena model 2010*.

into distinct groups based on the similarity of biological communities between samples from these sites. Habitat variables unaffected by anthropogenic disturbance (e.g. elevation) were compiled for each reference site. Variables that best discriminated between the reference groups were selected using discriminant function analysis (DFA) and termed *predictor variables*. For each test site, the predictor variable values were used to predict the reference group membership. To complete the RCA bioassessments, a test site invertebrate community was then compared to the range of communities within the predicted reference group. Full details of the original model building and validation processes are found in Perrin et al. (2007).

The Skeena model was developed for routine site quality testing in streams of northern British Columbia. The Skeena model can be used as a tool for assessment of streams for any number of land uses or industries including cumulative effects and point and non-point source. B.C. Environment has been working towards including RCA bioassessments as an important aspect of environmental effects monitoring programs for mining and other industries (Greg Tamblyn, pers. comm., June 28, 2011). The bioassessments can be used to track changes in site condition over time and where necessary, determine if improvements have been realized. Used in this way, it is an effective environmental screening tool that is scientifically defensible and can be used to quickly assess a site at relatively low cost. The tool promotes sustainable resource management by directly measuring the condition of the aquatic biota as an indicator of whether forest management practices, waste treatment systems, or mitigation associated with developments have adequately protected aquatic resources.

In 2007 and 2008, funding from the Forest Investment Account (FIA) was allocated to sample and test 59 sites in the Kalum and Skeena-Stikine Forest Districts using the Skeena model. During the test site assessment phase, a number of minor data inconsistencies and errors were discovered within the reference site dataset (both biological and habitat variables) that may have confounded the test site assessments (Bennett 2008). These included differences in taxonomic effort and data entry between reference site samples used to build the model. During the same time period, B.C. Environment partnered with a number of mining companies to increase the number of reference sites sampled in highly mineralized areas of the northwest. In 2009, the Skeena model was updated to include additional reference sites sampled in 2007 to 2009, to include a greater and more varied number of potential predictor variables in the environment data set, to allow a thorough quality assurance check of all the biological data in CABIN, and to re-run the GIS analysis with a quality assurance process in place for checking waypoint locations and catchment boundaries. Using the quality assured data set of 345 complete observations from 300 unique sites, the Skeena model was rebuilt.

1.1 Purpose of this document

The purpose of this document is to describe the revised predictive bioassessment model built for the north-central and north western B.C. in 2009 and outline the steps taken to build the model. The objective is to provide information to support use of the revised model by consultants, government and industry biologists, foresters and resource managers that have some knowledge of CABIN and RCA. For more detailed information on the RCA, please refer to the book published by Bailey *et al.* (2004) and the resource documents supplied on the CABIN website².

2 Methods

2.1 Study area and site descriptions

Approximately 250 streams were sampled during 2004, 2005 and 2006 as part of the original B.C. Environment project to develop a multivariate predictive bioassessment tool (Perrin *et al.* 2007). Potential sites were selected using a mapping exercise combined with discussions with resource management professionals familiar with the area. Criteria for selection of sites included accessibility (most sites were accessed by vehicle with roughly 10% of sites reached by helicopter or boat) and level of disturbance. The goal was to capture a broad range of anthropogenic disturbance types and intensities, and an equal number of undisturbed sites (see Perrin *et al.* 2007 for details). Land use in the Skeena Region includes forestry, mining, recreation, agriculture, urbanization and other industrial developments. Insect infestations (e.g. mountain pine beetle) have caused widespread forest disturbance in some areas. An additional 50 sites were sampled in 2007, 2008 and 2009.

The project area was located in north-central and northwest British Columbia (project centroid - 127.51°W, 54.59°N) and covered a roughly triangular shaped area of approximately 187,000 km² (Figure 1). The 300 unique sites extended from roughly 175 km northeast of Prince George (-120.75°W) to the exposed fjords west of Kitimat and Kemano (near 130°W) and north west to approximately 90 km south southeast of Atlin (-133.27°W) on the Taku River. Latitude of the sites ranged from 53.01°N (remote drainages in the Kitlope Heritage Conservancy and Tweedsmuir Park in the south) to 58.82°N (near Atlin).

Sites were sampled from three terrestrial ecozones, including 173 sites in the Montane Cordillera ecozone, 117 sites in the Pacific Maritime ecozone and 10 sites in the Boreal Cordillera ecozone. Overall,

² <http://www.ec.gc.ca/rcba-cabin/default.asp?lang=En&n=72AD8D96-1>

the project area covered a broad range of landscapes, from the rugged coastal mountains to the interior rolling plateaus and plains of north-central British Columbia.

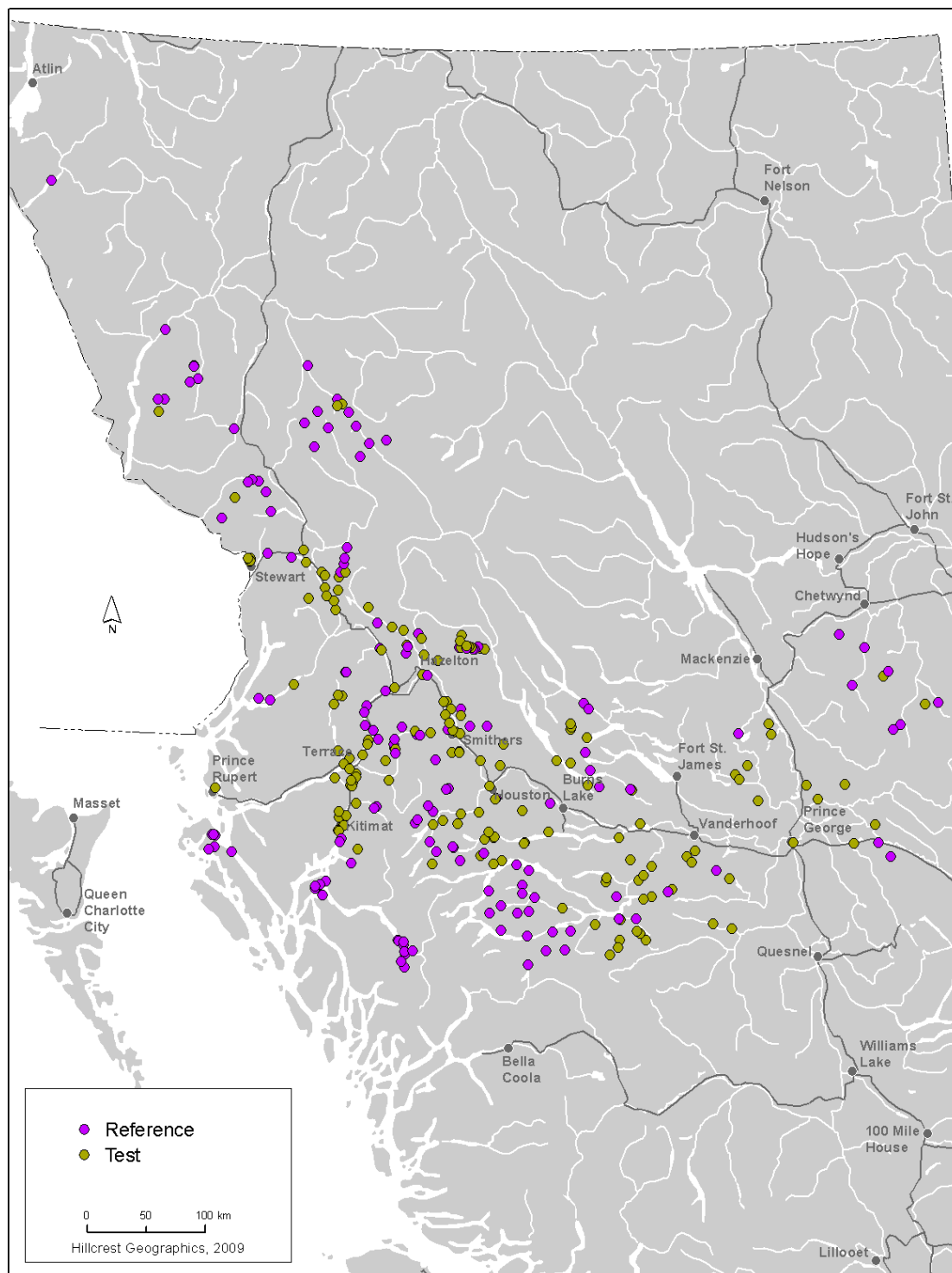


Figure 1 Stream bioassessment sites in northwest British Columbia.

2.2 Field stream assessments

All invertebrate samples were collected during the low flow, late summer period using a 400µm mesh kick-net, according to the timed procedure reported in Perrin *et al.* (2007) for the Skeena model and in Reynoldson *et al.* (2003) for CABIN. Field collection procedures followed those outlined in the *Canadian Aquatic Biomonitoring Network Field Manual* (Environment Canada 2010) and *The Canadian Aquatic Biomonitoring Network Field Manual* (Ministry of Environment 2009). **Laboratory sorting and subsampling procedures differed from the recommended CABIN protocols. Figure 2 (from Perrin *et al.* 2005) provides an overview of the subsampling and enumeration procedures used for Skeena model samples and the detailed steps have been provided in Appendix A. It is critical that this method is followed for any test samples that will be assessed using the Skeena model.**

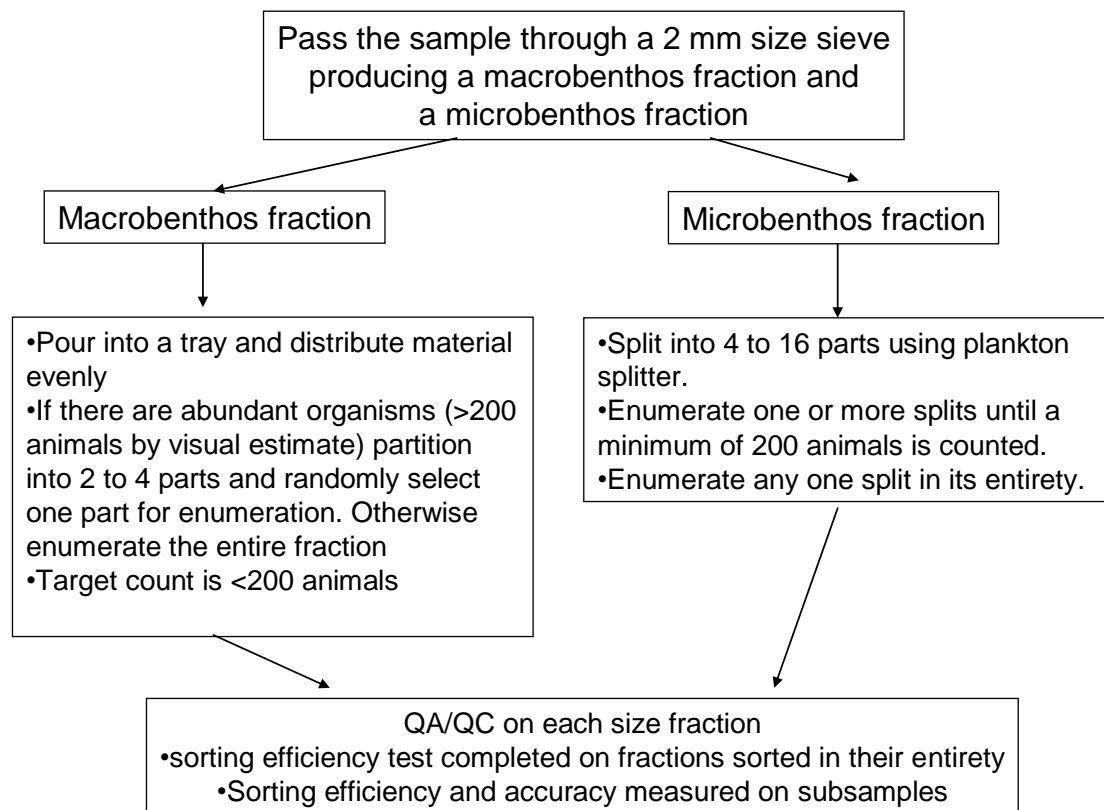


Figure 2. Flow chart of recommended procedures to enumerate kick net samples collected as part of the Skeena RCA analysis (from Perrin *et al.* 2005).

Benthic invertebrates in each sample were identified and entered to the CABIN database at the lowest possible taxonomic resolution. Data were exported from CABIN at the family level for model

building. In the Skeena Region samples, 99% of individuals in a given sample were typically identified to family level or lower. This is important since taxa identified at a higher resolution are not included in a family-level export, and therefore, the higher level taxa are not considered part of the community assemblage during the assessment. During interpretation of the test site assessments, it may be useful to consider what percentage of the overall test sample was identified to the family level or lower. While all individuals were entered to the CABIN database, not all taxa were used in the assessments. Taxa that were not included have been identified in Appendix B.

2.3 Watershed delineation and compilation of landcover data

All of the landcover data were compiled by Simon Norris of Hillcrest Geographics³. Once the waypoint for each site was confirmed, the watershed upstream of the sampling point was delineated. The watershed upstream of a sampling site was created from the Fresh Water Atlas (formerly Corporate Watershed Base) layer (first order watersheds) and the TRIM 25 metre Digital elevation model (25 m DEM), a digital model of the ground surface topography. It was a three step process that involved extracting all first order polygons upstream of the sample point, using the 25 m DEM to refine the boundary to terminate at the sample point, and then combining the first order watersheds into a single polygon (pers. comm., S. Norris, March 20, 2009) which was then given a unique identifier.

The watershed polygons were then used to run a GIS overlay. A complete list of the data layers compiled and variables calculated has been included in Appendix C. Variables were compiled using ArcGIS 9 geographic information systems (GIS) software developed by ESRI© (<http://esri.com/>). Spatial datasets were accessed through the Province of B.C. spatial data directories known as the Land and Resource Data Warehouse (LRDW) and the Integrated Land Management Bureau (ILMB) ARC warehouse, and the federal spatial database known as the Canadian Soil Information System (CANSIS). All GIS deliverables met GeoBC spatial data standards (ILMB 2008).

Spatial datasets for this project were selected based on what was available at the time of the model-build. As time goes on, more accurate spatial datasets may be available, but **it is critical to compile data for each watershed from the same sources and using the same methods that were used to compile data for the reference sites used to build the model.** Prediction of a test site to a reference group is based on the values for the predictor variables and data substitutions (e.g. using real precipitation data measured from within a watershed rather than 30 year climate normals from the CANSIS data layer) are not acceptable.

³ Simon Norris, Hillcrest Geographics, Victoria, B.C. phone 250.818.0114 or email snorris@hillcrestgeo.ca

2.4 Selection of reference sites

Rebuilding the model in 2009 created an opportunity to include some of the sites sampled in 2007 and 2008 as reference sites, thereby increasing the geographic coverage within the project area and including some reference sites from highly mineralized areas of northwestern B.C.

Selection of reference sites was done *a priori* to model building, so that selection was not biased by the biological assemblage at any given site. Sites were selected as reference condition for model-building if they met the following criteria (modified from Perrin *et al.* 2007):

1. Percent of watershed as urban land use must be less than 1 %.
2. Percent of watershed as agriculture land use (not including range lands) must be less than 10%.
3. Percentage of the watershed as forest harvest since 1980 must be less than 10%.
4. Road density within the watershed (km of roads per km² of watershed) must be less than 0.25.
5. Percentage of the watershed as forest burn since 1980 must be less than 20%.
6. Percentage of the watershed as mining land use must be zero.
7. Any watershed that had one or more “MINFILE” tags was individually checked by MOE personnel to exclude any with active mining claims or open adits.

From the 300 unique sites, 145 reference sites were selected for model-building using the selection criteria above.

2.5 Model Development

Only the basic steps in model development are outlined in this report as detailed methods are available in Perrin *et al.* (2007). Development of the RCA model and site testing were as follows (adapted from Perrin *et al.* 2007):

Family level invertebrate counts from reference site samples were exported from CABIN and imported into 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 routine in PRIMER called SIMPROF identified groups of samples that should be further divided into subclusters based on differences in community structure. Group assignments of individual samples were confirmed using a distance-based ordination (non-metric multi-dimensional scaling or NMDS) that was run in PRIMER from the same similarity matrix that was used for the cluster analysis. NMDS 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 but the relative distance between any pair of points reflects the degree of dissimilarity between the two community assemblages. 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.

An iterative, backwards-stepping discriminant function analysis (DFA) was run in Systat v11 (Systat 2004), to develop functions of habitat variables that best discriminated between the biological sample groups. Habitat variables were eliminated if the tolerance value (a measure of correlation between predictor variables) was less than 0.5 or the variable had low overall importance to the model (F value less than 4). 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. 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 removed from the DFA and re-substituted back in to see how well the model is able to classify that site to a sample group. A model that achieved $> 60\%$ correct classification to each sample group was considered acceptable. 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.

To complete the test site assessments for model evaluation, the community structure data for test sites were exported at the family level from CABIN and compiled into spreadsheet files containing all the reference site data for the four groups in the Skeena model. Invertebrate families present in the test sites but not the reference sites were added to the spreadsheets. Predictor variables for the test sites were compiled in an excel file containing all the reference sites and their *a priori* group assignment based on the Skeena model. A complete DFA was run in SYSTAT with the six Skeena RCA predictor variables and four *a priori* reference groups. Based on the predictor variable values, the model assigned each test site to a reference group. The reference group represents the natural range of expected invertebrate communities.

In PRIMER, an individual Bray-Curtis similarity matrix was created for raw, untransformed family-level invertebrate enumerations for each single test site and the samples for the reference sites in

its predicted group. An MDS 3-dimensional ordination plot was created and the ordination coordinates for the test site and its corresponding reference sites were saved and imported into SYSTAT. The reference sites and the single test site ordination coordinates were plotted along with 90%, 99% and 99.9% probability ellipses in SYSTAT (Figure 3). The confidence ellipses were drawn around the reference sites in SYSTAT to delineate four categories of stress:

1. Reference Condition (test site laid inside the 90% ellipse in all three plots)
2. Slightly stressed (test site was situated between the 90% and 99% ellipses in at least one plot but it was never outside or on the 99% ellipse in any plot)
3. Stressed (test site was situated between the 99% and 99.9% ellipse in at least one plot but it was never outside or on the 99.9% ellipse in any plot)
4. Severely stressed (on at least one plot the test site was situated outside or on the 99.9% ellipse)

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).

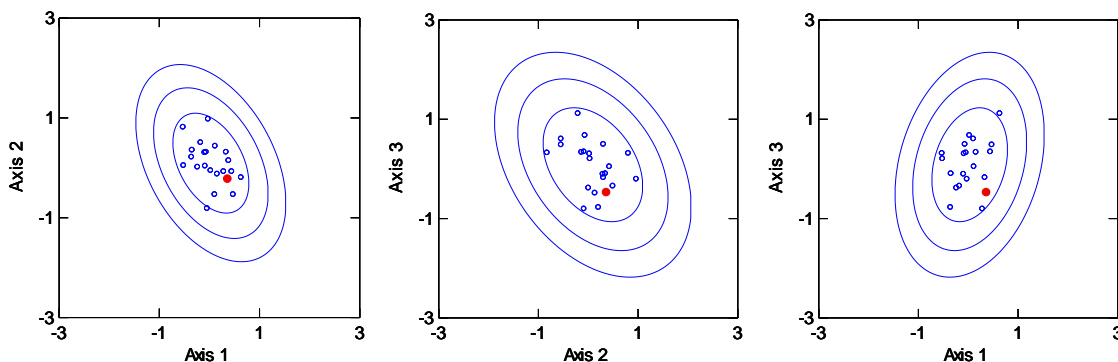


Figure 3. Ordination plots for a hypothetical test site shown as a solid red circle and reference sites shown as open blue circles belonging to a given sample group. The ellipses correspond to 90% (inner ellipse), 99% (middle ellipse), and 99.9% (outer ellipse) probabilities. The conclusion is that the test site is not stressed (Reference Condition) because it lay inside of the 90% ellipse in all plots.

As of June 2010, the Skeena model has been available on the CABIN website, and test site assessments can be carried out on-line.

3 Results and Discussion

3.1 Definition of reference groups

A total of 345 complete observations were compiled from all years of sampling (2004 to 2008), from 300 unique sites. Of those, 145 sites met the reference site selection criteria outlined in section 2.4 and were selected for clustering. A single observation from each of the 145 unique sites was used for model building.

Prior to clustering, the family level data were fourth root transformed. A cluster dendrogram (Figure 4) was created based on the Bray-Curtis similarities between each pair of reference sites. The dendrogram was sliced at 52% similarity to create 3 groups and 16 outliers. The large group was further split at 57% similarity into groups 2 and 4. A routine in PRIMER called SIMPROF (overlaid in Figure 4, SIMPROF tests for ‘no structure’ in subclusters) supported the differentiation of sites into the 4 groups and showed that there might be further structure present (*i.e.* subgroups) within groups 1, 2 and 3. A red line joining two or more sites or groups of sites indicates similarity between those sites, while a black line indicates some differences in community structure between those sites. For example, group 3 could be further split into two subgroups, but the number of samples in each subgroup (11 and 5) would be less than necessary to reliably define the biological reference condition for a group.

In Figure 5, the relative position of the sites based on the fourth-root transformed Bray Curtis similarity between each pair of sites is shown on a nonmetric multi-dimensional scaling (MDS) plot. Overall, clustering and ordination techniques produced four groups and 16 outliers from 145 reference condition sites in the Skeena and surrounding area that were geographically distributed as shown in Figure 6. The final number of sites included in each group is shown in Table 1.

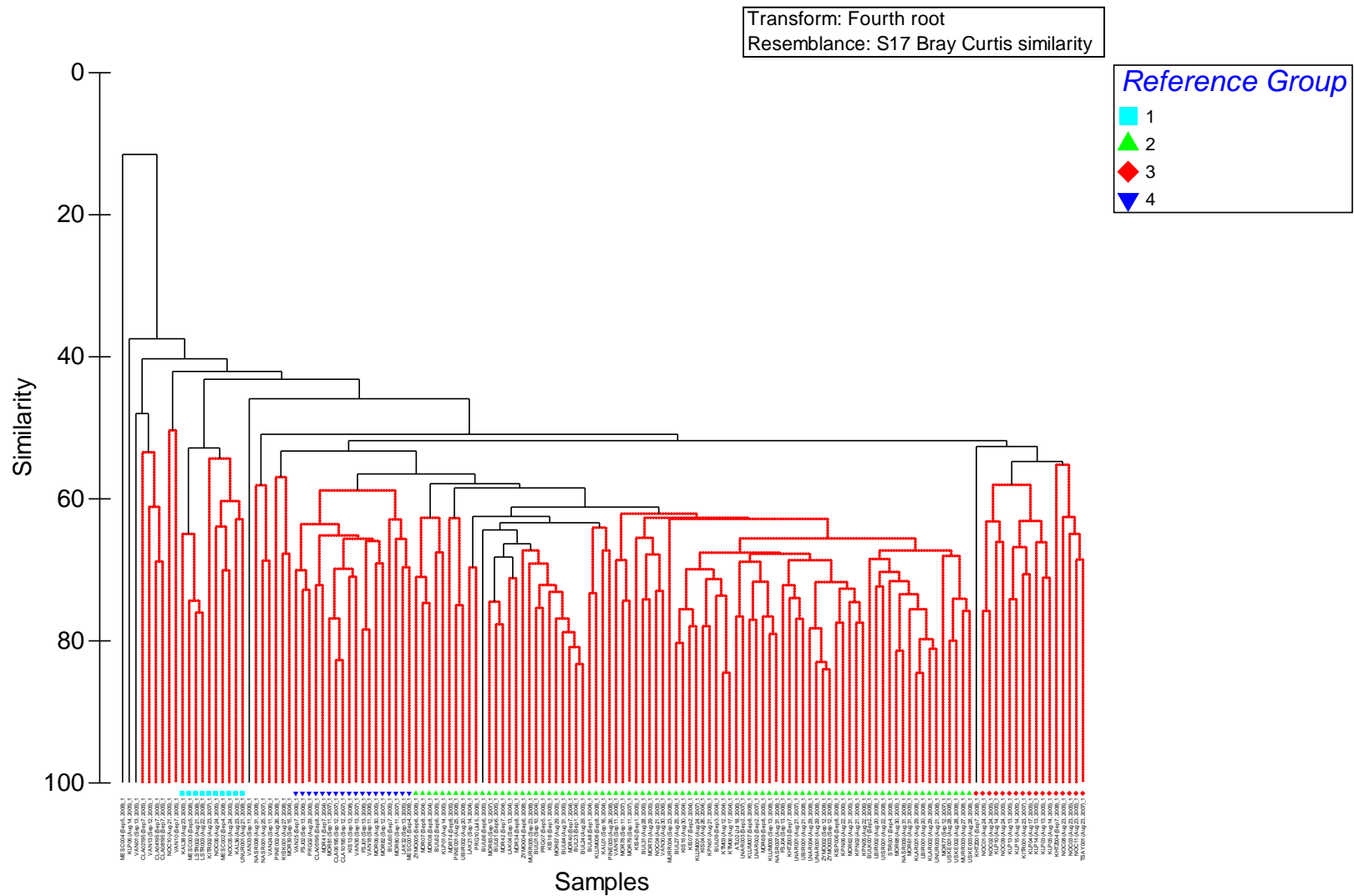


Figure 4 The cluster dendrogram based on fourth-root transformed family level community structure data for Skeena reference sites. SIMPROF results are overlaid where a red line joining two or more sites or groups of sites indicates similarity between those sites (no further groups), while a black line indicates some differences in community structure between those sites (more community structure based groups possible).

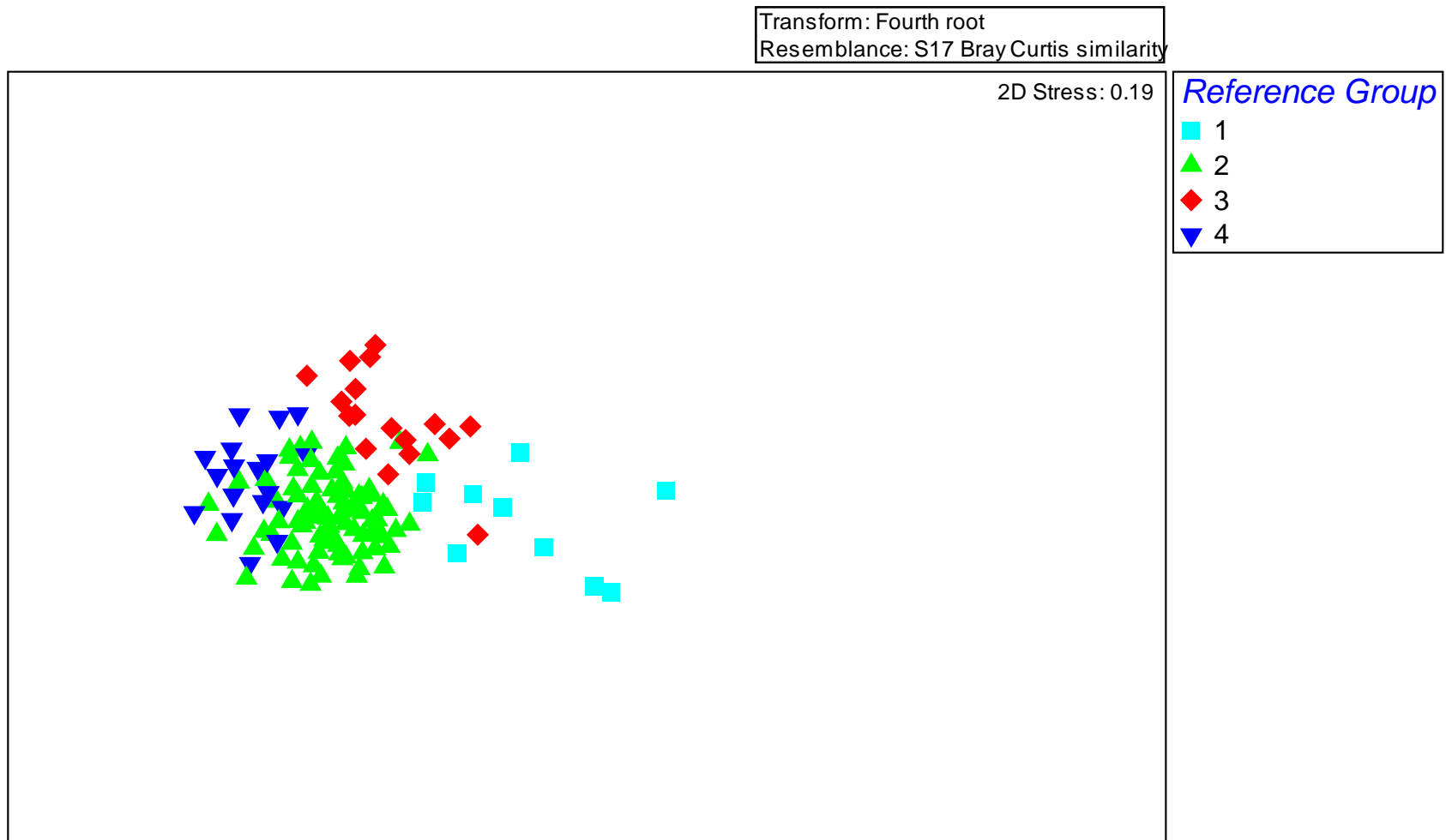


Figure 5. Skeena reference sites split into 4 groups and plotted with nonmetric MDS.

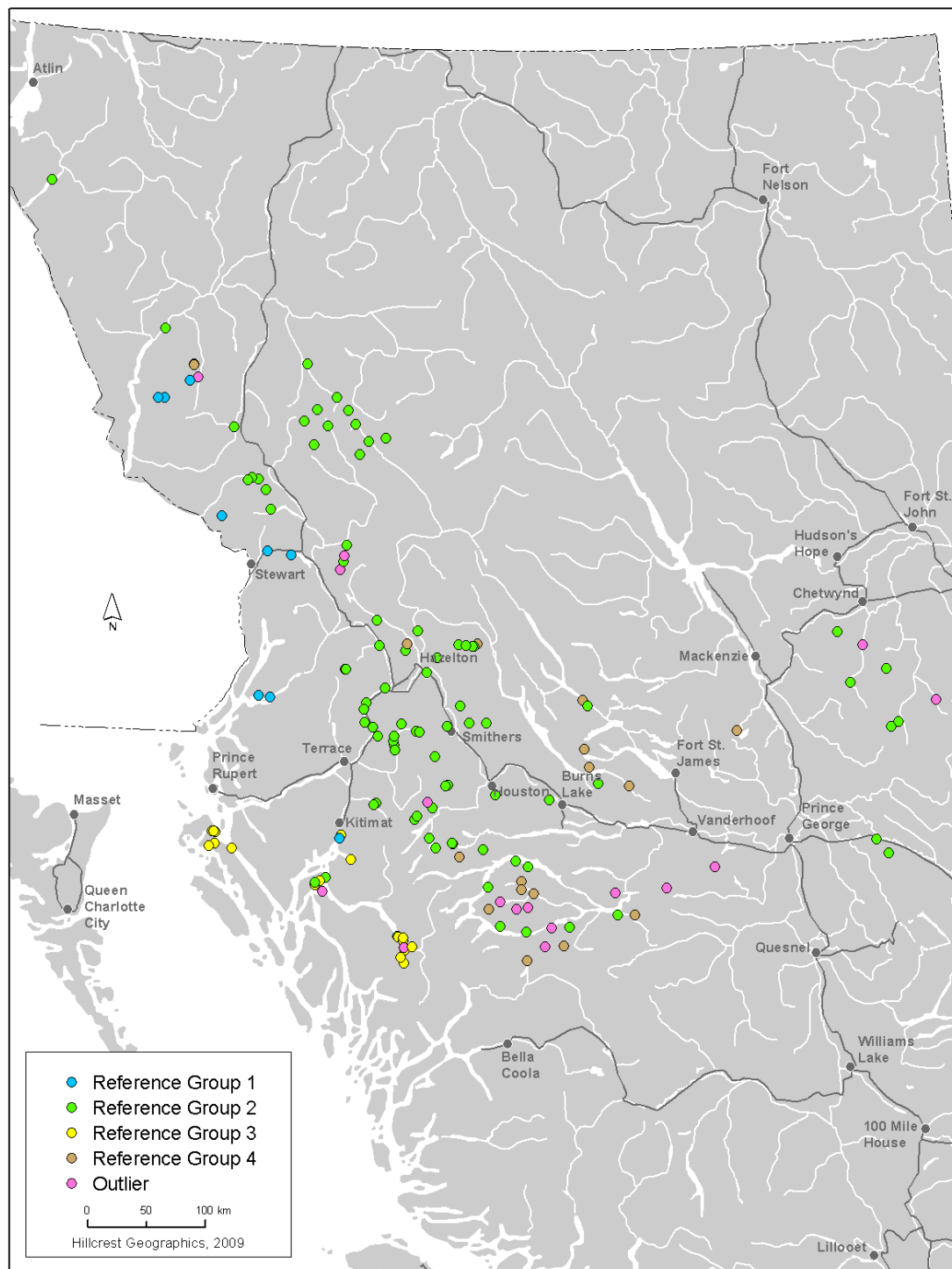


Figure 6 Distribution of reference sites in each of the four groups over the study area in north-western B.C.. Samples were collected over 5 years (2004 to 2008).

Table 1 Number of sites included in each group defined using cluster analysis and SIMPROF tests on fourth-root transformed community structure data for 145 sites.

Group	Group 1	Group 2	Group 3	Group 4	Outliers
Number of sites (n)	10	84	17	18	16

3.2 Predictor variables

A combination of six predictor variables (Table 2) was best able to differentiate the four *a priori* reference groups using discriminant function analysis (DFA). These were finalized as the predictor variables for the Skeena model and included percent ice area, total precipitation in January, percent wetlands area, percent intrusive rock, total snow in January and percent lakes area.

Table 2. Predictor variables for the Skeena model.

Category / Variable	Description	Source Database ⁴
LAKES_AREA_PCT	Percent of watershed area as lakes	LRDW
WETLANDS_AREA_PCT	Percent of watershed area as wetlands	LRDW
ICE_AREA_PCT	Percent of watershed area as ice	LRDW
ROCK_INTRU_PCT	Percent of watershed area underlain by intrusive rock	LRDW
Monthly SNOW January	Total monthly snowfall (cm) 1961 - 1994	CANSIS
Monthly TOTAL P January	Total monthly precipitation (mm) 1961 - 1995	CANSIS

The probability of assigning a site to the proper group was 77%, measured by the overall jackknife classification success. As shown in Table 3, prediction success was very high for group 3 (94% correct) and lowest for group 1 (60% correct). The overall misclassification error rate for Skeena model was 23%. **To maximize success of predicting a test site to the most appropriate reference group, it is critical that the values for the predictor variables have been calculated for the test site according to the methods in section 2.3.**

⁴ Source path provided in Appendix C.

Table 3 Jackknifed classification matrix showing DFA model classification results by cross-validation. The number of reference sites in each group has been shown in brackets.

Reference Group	Group 1	Group 2	Group 3	Group 4	%correct
1 (n=10)	6	0	4	0	60
2 (n=84)	1	65	9	9	77
3 (n=17)	1	0	16	0	94
4 (n=18)	0	5	1	12	67
Total (n=129)	8	70	30	21	77

3.3 Descriptive summary of the reference groups

Group 1 samples (n=10) were characterized by very low invertebrate abundance (mean individuals per sample, 247 ± 194 SD) and richness (mean number of families, 11.4 ± 4.4 SD). The most abundant taxa in Group 1 were Chironomidae, the mayfly families Heptageniidae and Baetidae, and the stonefly family Chloroperlidae (Figure 7). On average, land cover in the 10 watersheds was nearly one third (28%) ice area, with almost no wetlands or lakes. The group 1 sites receive the second highest amount of precipitation in January (258 mm) but the least amount of snowfall in the same month (59 cm, Table 4). Parent material in the watershed included 55% intrusive rock, a slow weathering rock that likely does not introduce biologically significant amounts of nutrients to streams. Median catchment area was 177 km^2 for group 1 sites and ranged from 11 to 803 km^2 .

Table 4 Group mean values and standard deviation of the mean of the six Skeena predictor variables for each of the four reference groups.

Variable Category	Predictor Variable	Group 1	Group 2	Group 3	Group 4
		n=10	n=84	n=17	n=18
Land Cover	Ice Area %	29 ± 24	2 ± 5	3 ± 5	0.1 ± 0.3
Climate	Total Precipitation in January (mm)	258 ± 25	124 ± 74	331 ± 3	99 ± 52
Land Cover	Wetlands Area %	0.1 ± 0.2	0.9 ± 1.3	0.3 ± 0.7	3.5 ± 2.6
Bedrock Geology	Intrusive Rock %	55 ± 39	7 ± 16	37 ± 42	19 ± 22
Climate	Snow in January (cm)	59 ± 10	73 ± 18	71 ± 24	67 ± 15
Land Cover	Lakes Area %	0.1 ± 0.1	0.4 ± 1.0	0.6 ± 1.0	2.3 ± 2.9

Group 2 sites (n=84) were characterized by high invertebrate abundance (mean individuals per sample, 2843 ± 2556 SD) and moderately high taxa richness (mean number of families 17.6 ± 3.3 SD). Sample composition was dominated by mayflies (families Heptageniidae, Baetidae and Ephemerellidae), chironomids, stoneflies (families Nemouridae, Taeniopterygidae, and Chloroperlidae) and the caddisfly family Rhyacophilidae. On average, the area of land covered by ice in these watersheds was very low (2.1%), along with cover by wetlands (0.9%) and lakes (0.4%). Intrusive rock as a parent material was not predominant in these watersheds (7%). Total precipitation in January was relatively low compared with groups 1 and 3, while snow in January was highest in group 2 (73 cm). Median catchment area was 26 km^2 for group 2 sites, and ranged from 1 to 3792 km^2 .

Group 3 was made up of 17 reference sites that were located along the coastal areas (Figure 6). Invertebrates were lower in abundance (mean invertebrates per sample 1122 ± 883 SD) although mean family taxa richness was similar to group 2 (mean richness 16.6 ± 3.6 SD). Sample composition was dominated by mayflies where Baetidae, Heptageniidae and Ephemerellidae accounted for 56% of the total number of organisms in the sample. Stoneflies were also abundant making up 17% of the individuals (Nemouridae and Chloroperlidae), followed by Chironomids. There were very few caddisflies in these samples (2%) and water mites were relatively abundant (7%). These coastal reference sites received the most precipitation in January of all groups (331 mm) and the second highest amount of snow (71 cm). Median catchment area was 12 km^2 for group 3 sites, and ranged from 0.3 to 280 km^2 .

Invertebrates were the most abundant in group 4 samples (n=18, mean abundance 5619 ± 2955 SD) and taxa richness was the highest (mean richness 21.3 ± 3.9 SD). The composition of the samples was very similar to group 2 but with a lower relative abundance of caddisflies. Sample composition was dominated by mayflies (families Baetidae, Heptageniidae, Ephemerellidae and Leptophlebiidae), Chironomids, stoneflies (Nemouridae, Chloroperlidae, and Taeniopterygidae), the caddisfly family Rhyacophilidae, and the Dipteran family Simuliidae. Wetlands and lakes were common in the watersheds in group 4 (3.5% and 2.3% cover respectively, Table 1) while land covered by ice was not (0.1%). Both wetlands and lakes can serve to moderate flows and temperatures in-stream and can provide a steady source of nutrients and food to downstream stream networks. Intrusive rock was less common than for groups 1 and 3 (19% for group 4), and total precipitation in January was the lowest for these sites. Median catchment area was 47 km^2 for group 4 sites and ranged from 5 to 408 km^2 .

Overall, the EPT taxa (mayfly families Baetidae, Heptageniidae, Ephemerellidae and Leptophlebiidae; stonefly families Nemouridae, Taeniopterygidae, Chloroperlidae, Capniidae, Perlodidae, and Leuctridae; and caddisfly families Rhyacophilidae, Uenoidae, Hydropsychidae, Brachycentridae and Glossosomatidae) were very abundant in all groups ranging from 66% of individuals in group 1 samples to 77% of individuals in group 2 samples. However, caddisflies were less abundant in groups 3 and 1, along with a marked decrease in overall abundance of the macrobenthos. Generally, taxa richness and abundance increased along a gradient of increasing lakes and decreasing ice, precipitation and snowfall in January, and intrusive rock within a watershed.

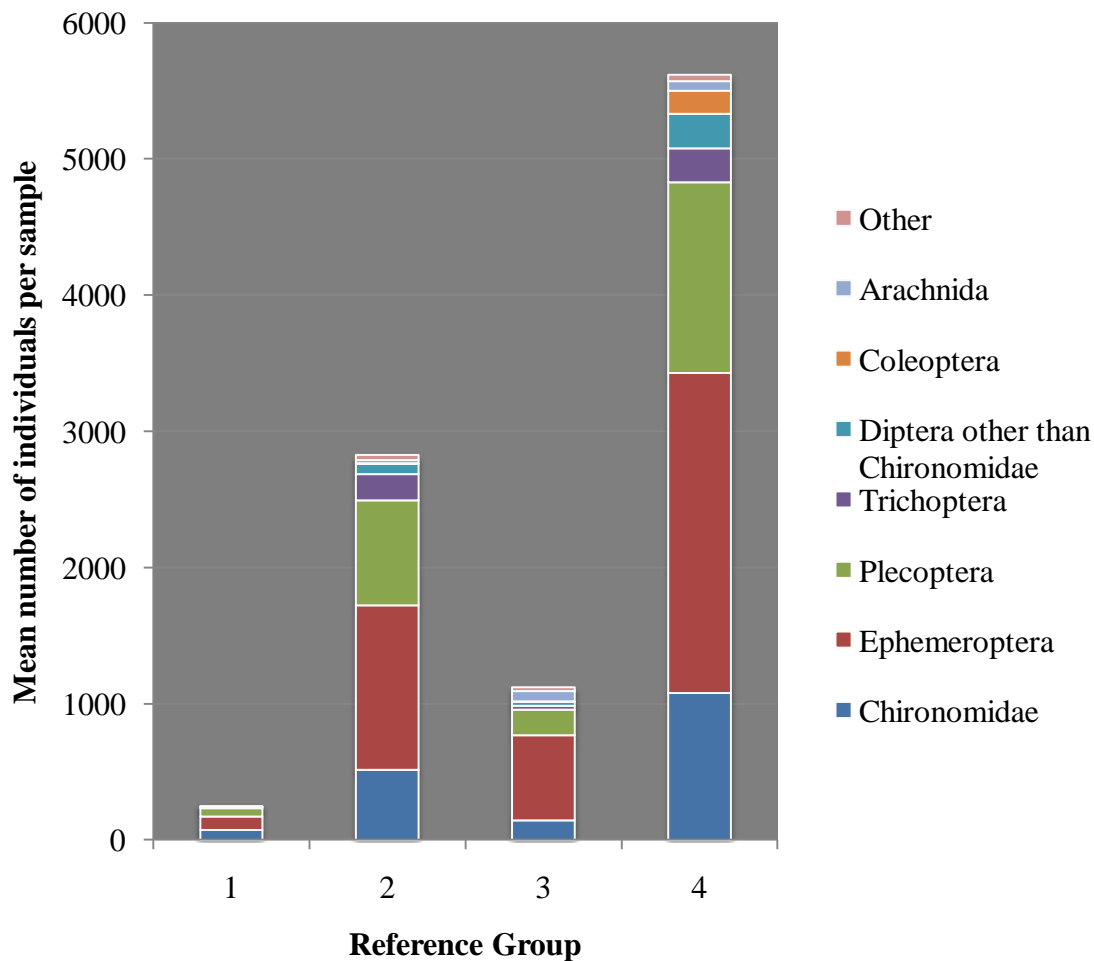


Figure 7 Taxonomic composition of reference sample groups. The mean number of individuals per sample for each reference group is shown for each taxonomic group. “Other” taxa include Oligochaetes (worms), Bivalves (clams), Platyhelminthes, Collembolla (springtails) and Crustacea.

3.4 Tests of Accuracy and Precision

Site assessment accuracy was examined by assessing seven sites with a known stream condition (not stressed) (Table 5, Bennett 2008). One observation from each reference site was used for model-building (MB) while any additional samples collected at these sites were run as test sites. In 5 of 7 cases, the test samples were correctly assessed as *not stressed* and in 2 cases, the test samples were incorrectly assessed as *slightly stressed*. This resulted in a site assessment accuracy of 71%.

Site assessment precision was examined by collecting more than one sample at a single test site on the same date (Table 5, Bennett 2008). Samples were collected by two different kick-net operators for the purpose of assessing the precision, or degree of mutual agreement, between site assessments using the two separate observations. There was agreement in all 7 cases when the assessment was reduced to a simple pass (P) or fail (F). In two cases site assessments using the two separate observations differed by more than one band (*e.g.* slightly stressed (SL) versus stressed (S)).

Table 5. List of sites and test site assessments for determining accuracy and precision.

PRIMER CODE	Site assignment (ref or test)	Year	BEAST Assessments			
			BEAST GROUP	Sample 1	Sample 2	Sample 3
BUL24	ref	2005	2 P		MB	
KIS15	ref	2004	2 MB			P
NOC06	ref	2005	1 MB		P	
KAL36	ref	2005	3 MB		SL	
TSAY001	ref	2007	3 MB		P	
UNUR003	ref	2008	2 P		MB	
USKE003	ref	2008	2 SL		MB	
BUL33	test	2005	2 P		P	
BUL52	test	2004	2 F (SL)		F (S)	F (SS)
KAL23	test	2005	2 P		P	
KAL32	test	2006	4 P		P	
MOR45	test	2004	4 F (S)		F (S)	F (SL)
MOR71	test	2005	4 P		P	
NOC07	test	2005	3 F (SL)		F (SL)	
KHTZ002	test	2008	3 P		P	

3.5 Using RCA bioassessments for impact assessment

Reference Condition Approach bioassessments using the Skeena model have wide application. The Skeena Region of the Ministry of Environment has used RCA bioassessments as an impact assessment tool to evaluate aquatic ecosystem integrity in the vicinity of a range of anthropogenic activities including solid waste management facilities (landfills), agriculture and

mining, and to monitor for cumulative effects. Several mining companies within the Skeena Region have collected baseline data for environmental assessment applications using CABIN protocols with the goal of testing the viability of the Reference Condition Approach as an assessment tool in their situations (e.g. highly mineralized areas often above treeline). First Nations in the region have also collected baseline data related to stream diversions for independent power production. Furthermore, Ministry of Environment has proposed the use of CABIN protocols for monitoring the effects of pulp mills, fertilizer application, urbanization / stormwater, sewage discharge, cumulative effects resulting from the construction of the Northern Transmission Line, and State of the Environment Reporting (Greg Tamblyn, pers. comm., June 28, 2011). In the Kalum Forest District, the stream bioassessment tool is being considered as a tool for monitoring water protection objectives set out in land use plans under the Forest and Range Practices Act (FRPA) (Ian Smith, pers. comm., January 13, 2011).

4 Important considerations

There are several very important things to keep in mind when conducting bioassessments and interpreting results:

- RCA bioassessment should be used as a screening tool and in a weight-of-evidence approach to impact assessment. It is important to consider other lines of evidence (e.g. invertebrate metrics, water chemistry, sediment chemistry) before drawing final conclusions and taking management actions.
- The probability of making a type I error (finding a test site to be impaired when it is actually in reference condition) for Skeena RCA bioassessments is 10% because using a 90% probability ellipse to define the reference condition automatically leaves 10% of the reference sites outside the ellipse. Reducing the type I error rate by using a different ellipse results in a trade-off with type II error rates, and may not be desirable from a management perspective. For more information on errors, please see Downie 2011.
- Spatial datasets for this project were selected based on what was available at the time of the model-build. As time goes on, more accurate spatial datasets may be available, but for site-testing purposes, it is critical to compile data for each test watershed from the same sources and using the same methods that were used to compile data for the reference sites used to build the model. Prediction of a test site to a reference group is based on the values for the predictor variables and data substitutions (e.g. using real precipitation data measured from within a watershed rather than 30 year climate normals from the CANSIS data layer) are not acceptable.
- **Sorting and subsampling procedures are different from standard CABIN protocols and are provided in Appendix A (from Perrin *et al.* 2005). Figure 2 (from Perrin *et al.* 2005) provides an overview of the subsampling and enumeration procedures used for Skeena reference site samples and the detailed steps have been provided in Appendix A. It is critical that this method is followed for any test samples that will be assessed using the Skeena reference sites.** Because of this, if a test site falls in a geographical area covered by more than one RCA model (e.g. some parts of the Nechako Plateau fall within the geographic range of both the Fraser model and Skeena model), the decision to use the Skeena model **MUST** be made prior to the sorting and enumeration of invertebrates samples so procedures specific to this model can be used.

- Test site analyses in CABIN are run on untransformed biological data, not fourth-root transformed data as were used for model-building. Comparison of assessment results for 94 sites found that results matched in 68 cases (72% agreement) regardless of transformation (S. Bennett, unreported data).

5 Recommendations

- Regional calibration of bioassessment tools is critical to ensure that land use effects are not masked by naturally occurring regional gradients (e.g. climate, geology, topography) (Bennett 2010). As the number of sampled reference sites within the area is expanded in coming years, the model should be rebuilt. The frequency of model rebuilds should depend on the number of new reference sites available to add. It is difficult to recommend certain types of streams to target for reference sampling since the group definitions would change with another model rebuild. As it stands now, groups 1, 3 and 4 each consist of fewer than 18 reference sites, and would likely benefit from additional reference sites. One approach might be to choose additional reference sites with predictor variable properties similar to those for groups 1, 3 and 4. Another approach for choosing additional reference sites might be to choose those reference sites with predictor variable properties similar to anticipated test sites. For example, if a number of test sites at a proposed mine are identified, it might be worthwhile to also sample some reference sites with similar values for % ice area, total precipitation in January, % wetlands area, % intrusive rock, snow in January and % lakes area.
- Resampling reference sites over time allows us to monitor temporal change in the environment, and keep the reference site database current. Ideally, some reference sites would be resampled every year, with a number of permanent sites sampled at a regular interval (1-3 years), in order to monitor the temporal signal (e.g. changes at reference sites over time). The frequency of reference site resampling in other areas of B.C. is varied. In the Fraser Basin, 10% of reference sites are sampled annually, including four permanent sites which are sampled every year (Stephanie Strachan, pers. comm., March 30, 2011). In the Skagit Basin, annual resampling of 20% of reference sites (10 sites out of 49 used to build the model) was recommended (Perrin and Bennett 2010). Permanent annual sites should ideally be in parks or other areas protected from land development; and nearby climate and hydrological monitoring (e.g. Water Survey of Canada hydrological station) would be a great benefit. With the large number of reference sites included in the Skeena

model, resampling 10% each year may be unrealistic due to budget constraints. In this case, it may be worthwhile to develop a rotating sampling list based on time since last sampling effort, with a target of resampling 10% of reference sites every 2 - 3 years, to keep the reference database current.

- For each test site that fails an assessment, the next step would be to identify stressor variables (e.g. nutrient concentrations, road density, etc.) that are correlated with changes in the macrobenthos between the test site and the reference sites. It's possible that some test sites that fail an assessment are simply outliers due to some unknown or unmeasured factor or combination of factors, or they may have failed due to a Type I error. Although there are some methods available to do this (e.g. LINKTREE analysis in PRIMER), it has not been tested thoroughly for the Skeena dataset. Further work will benefit from researching and testing these methods with the Skeena dataset.

6 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.
- Bennett, S.A. 2008. Bioassessment of streams in Northwest BC using the Skeena RCA. Report prepared by Bio Logic Consulting for West Fraser Mills Ltd. 52p.
- Bennett, S.A. 2009. Bioassessment of streams in Northwest BC using the Skeena RCA. Report prepared by Bio Logic Consulting for Gitxsan Forest Enterprises and Westfraser Mills Ltd. 66 p.
- Bennett, S.A. 2010. Stream invertebrates in northwest B.C.: an assessment of the relative importance of forest harvesting and environment factors at local and landscape scales. MSc thesis. University of British Columbia, Vancouver, B.C.
- Clarke, K.R. and R.N. Gorley. 2001. PRIMER v6: 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.
- Davis, W.S. 1994. Biological Assessment and Criteria: Building on the Past. In *Biological Assessment and Criteria: Tools for Water Resource Planning and Decision Making*. W.S. Davis and T.P. Simon (editors). CRC Press, Inc., Boca Raton, Florida, USA.
- Downie, A.J. 2011. Using impact simulations to evaluate the power of Skeena reference condition approach stream bioassessments. MSc thesis. Royal Roads University, Victoria, B.C.
- Environment Canada. 2010. Canadian Aquatic Biomonitoring Network Field Manual Wadeable Streams, March 2010. 52p.
- Integrated Land Management Bureau (ILMB) 2008. GeoBC File Geodatabase Standards for Data Creation, Publication and Distribution. Updated 22.90.2008. Created by the BC Integrated Land Management Bureau with input from GeoBC, ILMB LIM, MOFR Geomatics, BC Ministry of Agriculture and Lands.
- Karr, J.R., and E.W. Chu. 1999. *Restoring Life in Running Waters: Better Biological Monitoring*. Island Press, Covelo, California. 206 pp.
- Krebs. C.J. 1999. *Ecological Methodology*. Addison Wesley Longman
- Mazor, R.D., T.B. Reynoldson, D.M. Rosenberg and V.H. Resh. 2006. Effects of biotic assemblage, classification, and assessment method on bioassessment performance. *Can. J. Fish. Aquat. Sci.* **63**: 394-411.

- Ministry of Environment. 2009. The Canadian Aquatic Biomonitoring Network Field Manual. Resources Information Standards Committee, Victoria, B.C.
- 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..D. Dolecki and S. Salter. 2005. A recommended approach for sorting and sub-sampling benthic invertebrate samples for use in the Skeena RCA analysis. Report prepared by Limnotek Research and Development Inc. for B.C. Ministry of Water Land and Air Protection. Smithers, B.C. 9pp.
- Perrin, C.J., 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.
- Perrin, C.J. and S. Bennett. 2010. Quality of streams in the upper Skagit River watershed using the reference condition approach. Report prepared by Limnotek Research and Development Inc. for Skagit Environmental Endowment Commission. 73p.
- 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., D.M. Rosenberg, and V.H. Resh. 2001. Comparison of models predicting invertebrate assemblages for bio-monitoring in the Fraser River catchment, British Columbia. *Can. J. Fish. Aquat. Sci.* 58:1395-1409.
- Reynoldson, T.B., R.H. Norris, V.H. Resh, K.E. Day, and D.M. Rosenberg. 1997. The reference condition: a comparison of multimetric and multivariate approaches to assess water-quality impairment using benthic macroinvertebrates. *J.N. Am. Benthol. Soc.* 16(4): 833-852.
- Reynoldson, T.B., R.C. Bailey, K.E. Day, and R.H. Norris. Biological guidelines for freshwater sediment based on Benthic Assessment of SedimentT (the BEAST) using a multivariate approach for predicting biological state. *Australian Journal of Ecology* 20: 198-219.
- 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 Software, Inc. 2004. Systat for Windows, version 11. SYSTAT Software, Inc., Chicago, IL, U.S.A.

Appendix A. Invertebrate Subsampling and Enumeration Procedures for Skeena (from Perrin *et al.* 2005).

Laboratory sorting and subsampling procedures differed from the recommended CABIN protocols. The following excerpt from Perrin *et al.* (2005) provides the details of the subsampling and enumeration procedures used for Skeena samples. It is critical that this method is followed for any test samples that will be assessed using the Skeena model.

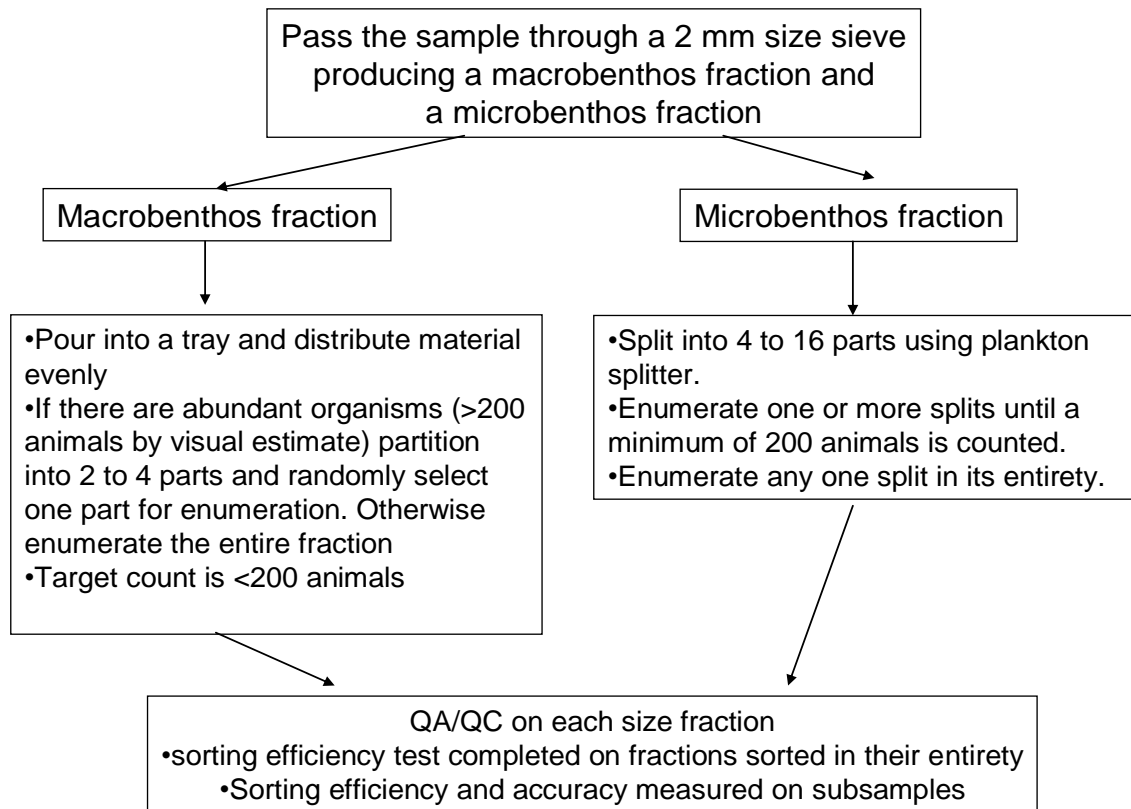


Figure 8. Flow chart of recommended procedures to enumerate kick net samples collected as part of the Skeena RCA analysis (from Perrin *et al.* 2005).

The recommended steps are as follows:

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

Appendix B. Taxa list for Skeena model.

Taxa shaded grey were not included in the assessment.

Phylum	Class	Order	Family
Annelida	Clitellata	Haplotaxida	Enchytraeidae
Annelida	Clitellata	Haplotaxida	Lumbricidae
Annelida	Clitellata	Haplotaxida	Naididae
Annelida	Clitellata	Haplotaxida	Tubificidae
Annelida	Clitellata	Haplotaxida	(blank)
Annelida	Clitellata	Lumbriculida	Lumbriculidae
Annelida	Clitellata	Lumbriculida	(blank)
Annelida	Clitellata	Rhynchobdellida	Glossiphoniidae
Annelida	Clitellata	(blank)	(blank)
Annelida	Oligochaeta	(blank)	(blank)
Annelida	(blank)	(blank)	(blank)
Arthropoda	Branchiopoda	Cladocera	Bosminidae
Arthropoda	Branchiopoda	Cladocera	Chydoridae
Arthropoda	Branchiopoda	Cladocera	Daphniidae
Arthropoda	Branchiopoda	Cladocera	(blank)
Arthropoda	Entognatha	Collembola	Isotomidae
Arthropoda	Entognatha	Collembola	Poduridae
Arthropoda	Entognatha	Collembola	Sminthuridae
Arthropoda	Entognatha	Collembola	(blank)
Arthropoda	Insecta	Coleoptera	Amphizoidae
Arthropoda	Insecta	Coleoptera	Carabidae
Arthropoda	Insecta	Coleoptera	Curculionidae
Arthropoda	Insecta	Coleoptera	Dryopidae
Arthropoda	Insecta	Coleoptera	Dytiscidae
Arthropoda	Insecta	Coleoptera	Elmidae
Arthropoda	Insecta	Coleoptera	Halipidae
Arthropoda	Insecta	Coleoptera	Hydrophilidae
Arthropoda	Insecta	Coleoptera	Staphylinidae
Arthropoda	Insecta	Coleoptera	(blank)
Arthropoda	Insecta	Diptera	Athericidae
Arthropoda	Insecta	Diptera	Blephariceridae
Arthropoda	Insecta	Diptera	Ceratopogonidae
Arthropoda	Insecta	Diptera	Chironomidae
Arthropoda	Insecta	Diptera	Deuterophlebiidae
Arthropoda	Insecta	Diptera	Dixidae
Arthropoda	Insecta	Diptera	Dolichopodidae
Arthropoda	Insecta	Diptera	Empididae
Arthropoda	Insecta	Diptera	Ephydriidae
Arthropoda	Insecta	Diptera	Oreoleptidae
Arthropoda	Insecta	Diptera	Pelecorhynchidae
Arthropoda	Insecta	Diptera	Phoridae
Arthropoda	Insecta	Diptera	Psychodidae

Phylum	Class	Order	Family
Arthropoda	Insecta	Diptera	Sarcophagidae
Arthropoda	Insecta	Diptera	Sciomyzidae
Arthropoda	Insecta	Diptera	Simuliidae
Arthropoda	Insecta	Diptera	Stratiomyidae
Arthropoda	Insecta	Diptera	Tabanidae
Arthropoda	Insecta	Diptera	Tanyderidae
Arthropoda	Insecta	Diptera	Tipulidae
Arthropoda	Insecta	Diptera	(blank)
Arthropoda	Insecta	Ephemeroptera	Ameletidae
Arthropoda	Insecta	Ephemeroptera	Baetidae
Arthropoda	Insecta	Ephemeroptera	Ephemerellidae
Arthropoda	Insecta	Ephemeroptera	Heptageniidae
Arthropoda	Insecta	Ephemeroptera	Leptohyphidae
Arthropoda	Insecta	Ephemeroptera	Leptophlebiidae
Arthropoda	Insecta	Ephemeroptera	(blank)
Arthropoda	Insecta	Homoptera	Aphididae
Arthropoda	Insecta	Homoptera	Cicadellidae
Arthropoda	Insecta	Hymenoptera	(blank)
Arthropoda	Insecta	Lepidoptera	(blank)
Arthropoda	Insecta	Megaloptera	Sialidae
Arthropoda	Insecta	Neuroptera	(blank)
Arthropoda	Insecta	Odonata	Gomphidae
Arthropoda	Insecta	Plecoptera	Capniidae
Arthropoda	Insecta	Plecoptera	Chloroperlidae
Arthropoda	Insecta	Plecoptera	Leuctridae
Arthropoda	Insecta	Plecoptera	Nemouridae
Arthropoda	Insecta	Plecoptera	Peltoperlidae
Arthropoda	Insecta	Plecoptera	Perlidae
Arthropoda	Insecta	Plecoptera	Perlodidae
Arthropoda	Insecta	Plecoptera	Pteronarcyidae
Arthropoda	Insecta	Plecoptera	Taeniopterygidae
Arthropoda	Insecta	Plecoptera	(blank)
Arthropoda	Insecta	Trichoptera	Apataniidae
Arthropoda	Insecta	Trichoptera	Brachycentridae
Arthropoda	Insecta	Trichoptera	Glossosomatidae
Arthropoda	Insecta	Trichoptera	Hydropsychidae
Arthropoda	Insecta	Trichoptera	Hydroptilidae
Arthropoda	Insecta	Trichoptera	Lepidostomatidae
Arthropoda	Insecta	Trichoptera	Leptoceridae
Arthropoda	Insecta	Trichoptera	Limnephilidae
Arthropoda	Insecta	Trichoptera	Philopotamidae
Arthropoda	Insecta	Trichoptera	Polycentropodidae
Arthropoda	Insecta	Trichoptera	Rhyacophilidae
Arthropoda	Insecta	Trichoptera	Uenoidae
Arthropoda	Insecta	Trichoptera	(blank)
Arthropoda	Malacostraca	Amphipoda	Crangonyctidae
Arthropoda	Malacostraca	Amphipoda	Gammaridae

Phylum	Class	Order	Family
Arthropoda	Malacostraca	Ostracoda	(blank)
Arthropoda	Maxillipoda	Calanoida	(blank)
Arthropoda	Maxillipoda	Copepoda	(blank)
Chelicerata	Arachnida	Araneae	(blank)
Chelicerata	Arachnida	Oribatei	Halacaridae
Chelicerata	Arachnida	Oribatei	Hydrozetidae
Chelicerata	Arachnida	Oribatei	Oribatidae
Chelicerata	Arachnida	Oribatei	Trhypochthoniidae
Chelicerata	Arachnida	Oribatei	(blank)
Chelicerata	Arachnida	Prostigmata	Aturidae
Chelicerata	Arachnida	Prostigmata	Feltriidae
Chelicerata	Arachnida	Prostigmata	Hydryphantidae
Chelicerata	Arachnida	Prostigmata	Hygrobatidae
Chelicerata	Arachnida	Prostigmata	Lebertiidae
Chelicerata	Arachnida	Prostigmata	Limnocharidae
Chelicerata	Arachnida	Prostigmata	Mideopsidae
Chelicerata	Arachnida	Prostigmata	Oxidae
Chelicerata	Arachnida	Prostigmata	Pionidae
Chelicerata	Arachnida	Prostigmata	Sperchonidae
Chelicerata	Arachnida	Prostigmata	Stygothrombidiidae
Chelicerata	Arachnida	Prostigmata	Torrenticolidae
Chelicerata	Arachnida	Prostigmata	(blank)
Chelicerata	Arachnida	(blank)	(blank)
Cnidaria	Hydrozoa	Hydroida	Hydridae
Mollusca	Bivalvia	Veneroida	Sphaeriidae
Mollusca	Bivalvia	Veneroida	Unionidae
Mollusca	Bivalvia	(blank)	(blank)
Mollusca	Gastropoda	Basommatophora	Ancylidae
Mollusca	Gastropoda	Basommatophora	Lymnaeidae
Mollusca	Gastropoda	Basommatophora	Physidae
Mollusca	Gastropoda	Basommatophora	Planorbidae
Mollusca	Gastropoda	Heterostropha	Valvatidae
Mollusca	Gastropoda	(blank)	(blank)
Nemata	(blank)	(blank)	(blank)
Platyhelminthes	Turbellaria	Tricladida	Planariidae
Platyhelminthes	Turbellaria	Tricladida	(blank)
Platyhelminthes	Turbellaria	(blank)	(blank)
Tardigrada	(blank)	(blank)	(blank)

Appendix C. GIS variables, descriptions, source databases and source paths.

Category / Variable	Description	Source Database	Source Path
Ecological Classification			
ECOREG	Ecoregion	CANSIS	http://sis.agr.gc.ca/cansis/nsdb/ecostrat/gis_data.html
ECOZONE	Ecozone	CANSIS	http://sis.agr.gc.ca/cansis/nsdb/ecostrat/gis_data.html
Hydrology			
WATERSHED_AREA_HA	Area of sample site watershed	n/a	Skeena Bioassessment Watersheds, derived from CWB_WATERSHED_POLY and TRIM DEM
STREAMS_LENGTH_KM	Length of streams within watershed	LRDW	WHSE_BASEMAPPING.CWB_STREAM_NETWORKS
STREAMS_ORDER_MAX	Stream order at sampling site	LRDW	WHSE_BASEMAPPING.CWB_STREAM_NETWORKS
LAKES_AREA_HA	Area of lakes	LRDW	WHSE_BASEMAPPING.CWB_LAKES_POLY
RIVERS_AREA_HA	Area of river polygons	LRDW	WHSE_BASEMAPPING.CWB_RIVERS_POLY
WETLANDS_AREA_HA	Area of wetlands	LRDW	WHSE_BASEMAPPING.CWB_WETLAND_POLY
ICE_AREA_HA	Area of ice	LRDW	WHSE_BASEMAPPING.CWB_GLACIERS_POLY
LAKES_MIN_DIST_METRES	Distance from sample site to upstream lake outlet (where applicable)	LRDW	WHSE_BASEMAPPING.CWB_LAKES_POLY and WHSE_BASEMAPPING.CWB_STREAM_NETWORKS
WETLANDS_MIN_DIST_METRES	Distance from sample site to upstream wetlands outlet (where applicable)	LRDW	WHSE_BASEMAPPING.CWB_WETLAND_POLY and WHSE_BASEMAPPING.CWB_STREAM_NETWORKS

Category / Variable	Description	Source Database	Source Path
Forest Cover			
FOR_YOUNG_HA	Area of forest, less than 140 years old, greater than 6 metres in height and NOT classified as Recently Logged or Selectively Logged	LRDW	WHSE_FOREST_VEGETATION.VEG_COMP_LYR_R1_POLY, TFL data from licencees, WHSE_BASEMAPPING.BTM_PRESENT_LAND_USE_V1_SVW
FOR_AC_GT6_HA	Area of forest, age class 6, 7, 8, & 9 (100 + years old)	LRDW	WHSE_FOREST_VEGETATION.VEG_COMP_LYR_R1_POLY, TFL data from licencees
FOR_OLD_HA	Area of forest, old growth (>140 yrs and greater than 6 metres in height)	LRDW	WHSE_FOREST_VEGETATION.VEG_COMP_LYR_R1_POLY, TFL data from licencees
FOR_BURNPAST20YRS_HA	Area of forest burned post 1988	LRDW	WHSE_FOREST_VEGETATION.VEG_COMP_LYR_R1_POLY and WHSE_BASEMAPPING.BTM_PRESENT_LAND_USE_V1_SVW
FOR_BURNPRE2000_HA	Area of forest, burned before 2000	LRDW	WHSE_FOREST_VEGETATION.VEG_COMP_LYR_R1_POLY and WHSE_BASEMAPPING.BTM_PRESENT_LAND_USE_V1_SVW
FOR_HARVPOST1980_HA	Area of forest, harvested since 1980	LRDW	WHSE_FOREST_VEGETATION.VEG_COMP_LYR_R1_POLY; WHSE_FOREST_TENURE.FTEN_CUT_BLOCK_POLY_SVW; WHSE_FOREST_VEGETATION.RSLT_OPENING_SVW
Land Cover			
AGRICUL_HA	Area classified as agricultural	LRDW	WHSE_BASEMAPPING.BTM_PRESENT_LAND_USE_V1_SVW
ALPINE_HA	Area classified as alpine	LRDW	WHSE_BASEMAPPING.BTM_PRESENT_LAND_USE_V1_SVW
AVALAN_HA	Area classified as avalanche chute - below the tree line, devoid of forest growth due primarily to snow avalanches. Usually herb or shrub	LRDW	WHSE_BASEMAPPING.BTM_PRESENT_LAND_USE_V1_SVW

Category / Variable	Description	Source Database	Source Path
	covered.		
BARREN_HA	Area classified as rock barrens, badlands, sand and gravel flats, dunes and beaches where unvegetated surfaces predominate, except where these conditions occur in ALPINE areas.	LRDW	WHSE_BASEMAPPING.BTM_PRESENT_LAND_USE_V1_SVW
MINING_HA	Area classified as mine/extraction site	LRDW	WHSE_BASEMAPPING.BTM_PRESENT_LAND_USE_V1_SVW
URBAN_HA	Area classified as Urban	LRDW	WHSE_BASEMAPPING.BTM_PRESENT_LAND_USE_V1_SVW
ROADS_LENGTH_KM	Length of roads	LRDW	WHSE_BASEMAPPING.DRA_DIGITAL_ROAD_ATLAS_LINE_SP
MINFILE_NUM	Number of Minfile mines	LRDW	WHSE_MINERAL_TENURE.MINFIL_MINERAL_FILE
Topography			
ELEV_MAX	Elevation (metres) of sampling site	ILMB arcwhse	P:\corp\arcwhse\gdbc\tdem_<250kMAP_TILE>
ELEV_MIN	Maximum elevation of watershed	ILMB arcwhse	P:\corp\arcwhse\gdbc\tdem_<250kMAP_TILE>
SLOPE_LT30_HA	Area of hillslope gradient of 0 to 30%	ILMB arcwhse	P:\corp\arcwhse\gdbc\tdem_<250kMAP_TILE>
SLOPE_30_50_HA	Area of hillslope gradient of 30 to 50%	ILMB arcwhse	P:\corp\arcwhse\gdbc\tdem_<250kMAP_TILE>
SLOPE_50_60_HA	Area of hillslope gradient of 50 to 60%	ILMB arcwhse	P:\corp\arcwhse\gdbc\tdem_<250kMAP_TILE>

Category / Variable	Description	Source Database	Source Path
SLOPE_GT60_HA	Area of hillslope gradient of 60% and greater	ILMB arcwhse	P:\corp\arcwhse\gdbc\tdem_<250kMAP_TILE>
Administration			
PARKS_NUM	Number of provincial parks	LRDW	WHSE_PARKS.PA_PROTECTED_AREA_POLY
PARKS_HA	Area of provincial parks	LRDW	WHSE_PARKS.PA_PROTECTED_AREA_POLY
Bedrock Geology			
ROCK_SEDIM_HA	Area underlain by sedimentary rock	LRDW	WHSE_MINERAL_TENURE.GEOL_BEDROCK_UNIT_POLY_SVW
ROCK_INTRU_HA	Area underlain by intrusive rock	LRDW	WHSE_MINERAL_TENURE.GEOL_BEDROCK_UNIT_POLY_SVW
ROCK_VOLCA_HA	Area underlain by volcanic rock	LRDW	WHSE_MINERAL_TENURE.GEOL_BEDROCK_UNIT_POLY_SVW
ROCK_METAM_HA	Area underlain by metamorphic rock	LRDW	WHSE_MINERAL_TENURE.GEOL_BEDROCK_UNIT_POLY_SVW
ROCK_ULTRA_HA	Area underlain by ultramafic rock	LRDW	WHSE_MINERAL_TENURE.GEOL_BEDROCK_UNIT_POLY_SVW
Climate Normals			
Monthly TMIN	Average daily minimum air temperature (deg C) 1961 – 1990	CANSIS	http://sis.agr.gc.ca/cansis/nsdb/ecostrat/district/climate.html
Monthly MAX	Average daily maximum air temperature (deg C) 1961 - 1991	CANSIS	http://sis.agr.gc.ca/cansis/nsdb/ecostrat/district/climate.html
Monthly TMEAN	Average daily mean air temperature (deg C) 1961 - 1992	CANSIS	http://sis.agr.gc.ca/cansis/nsdb/ecostrat/district/climate.html
Monthly RAIN	Total monthly rainfall (mm) 1961 - 1993	CANSIS	http://sis.agr.gc.ca/cansis/nsdb/ecostrat/district/climate.html

Category / Variable	Description	Source Database	Source Path
Monthly SNOW	Total monthly snowfall (cm) 1961 - 1994	CANSIS	http://sis.agr.gc.ca/cansis/nsdb/ecostrat/district/climate.html
Monthly TOTAL P	Total monthly precipitation (mm) 1961 - 1995	CANSIS	http://sis.agr.gc.ca/cansis/nsdb/ecostrat/district/climate.html