



# Macroinvertebrate Community Indices for Auckland's Soft- bottomed Streams

December 2004    Technical Publication 303

Auckland Regional Council  
Technical Publication No. 303, 2004  
ISSN 1175-205X,    ISBN 1-877416-35-5  
[www.arc.govt.nz](http://www.arc.govt.nz)





CAWTHRON

Report No. 970

# Macroinvertebrate Community Indices for Auckland's Soft-bottomed Streams

Prepared for

Auckland Regional Council

John D. Stark<sup>1</sup>  
and John R. Maxted<sup>2</sup>

Cawthron Institute<sup>1</sup>  
98 Halifax Street East  
Private Bag 2  
NELSON  
NEW ZEALAND  
Phone: +64.3.548.2319  
Fax: +64.3.546.9464  
Email: [john.stark@cawthron.org.nz](mailto:john.stark@cawthron.org.nz)

Auckland Regional Council<sup>2</sup>  
21 Pitt Street  
Private Bag 92-012  
AUCKLAND  
NEW ZEALAND  
Phone: +64.9.366.2000  
Fax: +64.9.366.2155  
Email: [john.maxted@arc.govt.nz](mailto:john.maxted@arc.govt.nz)

*Information contained in this report may not be used without the prior consent of the client*

December 2004



## **PREFACE**

This document was a joint effort of the Auckland Regional Council (ARC) and the Cawthron Institute, including the approach, analyses, presentation of results, and staff time. Funding was provided by the ARC and the Cawthron Institute, with supplemental funding provided by the Ministry for the Environment. The Cawthron Institute worked under contract with the ARC, and contributed considerable staff time over and above the contract amount. The work utilized macroinvertebrate and physical habitat data collected as part of the ARC's Freshwater Ecology Programme, and water quality data collected as part of the ARC's Stream Water Quality Monitoring Programme. Catchment land use estimates were derived from the New Zealand Land Cover Database, version 2. New tolerance values for macroinvertebrate taxa followed a procedure developed by Dr. Bruce Chessman in Australia. Acknowledgments are included in Section 7.

### **Recommended Citation:**

Stark, J. D.; Maxted, J. R. 2004. Macroinvertebrate community indices for Auckland's soft-bottomed streams and applications to SOE reporting. Prepared for the Auckland Regional Council by the Cawthron Institute, Cawthron Report No. 970. 66 pages.



## EXECUTIVE SUMMARY

This report was commissioned by the Auckland Regional Council (ARC) with supplementary funding from the Ministry for the Environment. It describes the development of biotic indices for soft-bottomed (SB) streams in the Auckland region. The new indices are variants of New Zealand's MCI and SQMCI/QMCI. New taxon-specific tolerance scores were derived specifically for SB streams using an objective iterative rank correlation process developed by Dr Bruce Chessman in Australia. This process requires a macroinvertebrate dataset collected from sites covering a wide range of disturbance from reference conditions with native vegetation to severely degraded rural or urban catchments. Macroinvertebrate data collected under ARC's State of the Environment (SOE) Monitoring Programme intentionally covered the range of disturbance within a variety of land-uses (native forest, exotic forest, rural, urban), and was an ideal basis for deriving new index scores. The analysis used 179 samples from 41 SB sites collected between 2000 and 2004.

The terms MCI, SQMCI and QMCI are retained for the original hard-bottomed (HB) stream indices with the terms MCI-sb, SQMCI-sb and QMCI-sb coined for the new SB versions. In this report, we use the term QMCI-sb to refer to indices derived from relative abundance (RA) data using both coded abundance (SQMCI-sb) and full- or fixed-count (QMCI-sb) processing methods.

We assessed four aspects of the performance of the various indices. First, the MCI-sb and QMCI-sb provided a wider range of index values (from the highest quality to the most degraded sites) than the MCI and QMCI when applied to SB stream macroinvertebrate data. This improved the ability of the indices to discriminate between sites of slightly differing quality.

Second, the MCI-sb showed stronger significant rank and linear correlations than the MCI with many local- and catchment-scale environmental variables when applied to SB streams. Linear correlations were particularly strong with the percentage of developed land in the catchment ( $r^2 = 0.689$ , cf. 0.640 for the MCI), ARC's Habitat Quality Index ( $r^2 = 0.410$ , cf. 0.328 for the MCI) and ARC's Water Quality Index ( $r^2 = 0.542$ , cf. 0.472 for the MCI). In general, the MCI-sb performed better than the QMCI-sb, and we suspect this was because the QMCI-sb scores were more variable.

Third, using a subset of sites defined as "severely degraded" based upon non-biological criteria (including measures of habitat quality, and measures of water quality such as temperature and dissolved oxygen), we tested the ability of the various indices to assign severely degraded sites correctly to the bottom half of the score range. The MCI-sb correctly assigned 81% (MCI 25%), and the QMCI-sb assigned 99% (QMCI 13%) of the samples to the bottom half of the scoring range. The poor performance of the original indices indicated that they may underestimate environmental degradation in SB streams.

Fourth, using the SOE dataset we tested the sensitivity of the various indices to a wide range of disturbance across three major land use classes (forestry, rural, and urban). Both the MCI-sb and QMCI-sb provided greater separation between land use classes and levels of disturbance within each class, compared to the original HB versions.

We also developed quality thresholds (i.e., excellent, good, fair, poor) for interpretation of index values. An objective process based on the statistical distribution of index values at reference sites together with an estimate of the lowest practicable biotic index score

(obtained from a degraded urban stream concrete channel) defined the full scoring range. Intermediate classes were defined at uniform intervals within this range. Conveniently, the classification systems for the SB versions were the same as those published for the HB versions. Having the same interpretation for both SB and HB versions of the indices is a major advantage because it avoids confusion.

SOE data used to develop the new indices provided insights of interest to water managers, highlighting the value of linking research and monitoring objectives. Several examples are provided on the application of the new indices to the management of a variety of activities that affect SB streams in the Auckland region. To maintain (or restore) excellent biological communities, it is essential to have both high habitat quality and low catchment development. Some forestry and rural sites with excellent habitat quality had biological conditions that approached those in native bush catchments, although forestry sites had substantial habitat impacts due to sediment deposition. In fully urbanized catchments, it may not be possible to achieve conditions comparable to reference conditions unless poor water quality is improved, although substantial improvements in the biota can be achieved by improving habitat quality. The results support a range ARC's initiatives affecting streams including riparian zone management, land use planning, and control of pollution discharges.

#### Advice for Freshwater Ecologists

Well, there you have it. Combined with the national macroinvertebrate protocols (Stark et al., 2001), we now have a complete and standardised process for the assessment of SB and HB streams using macroinvertebrates - collecting, processing, QC, and now metrics and reporting for SB streams. We hope this assessment approach provides for accurate and consistent reporting of macroinvertebrate data in the Auckland region. The applicability of the MCI-sb (and variants) to SB streams in other regions of New Zealand will require evaluation. However, given that the MCI was developed using data from the Taranaki region and was found to perform well nationwide, it is our belief that the MCI-sb will also be applicable throughout New Zealand (although scores for taxa not encountered in the Auckland dataset will need to be derived). We plan to test the performance of the new indices using additional data collected from Auckland streams, and using existing data from SB streams elsewhere in New Zealand.

Proper interpretation of index scores requires understanding of the variability inherent in the data, particularly when single samples are collected. Single hand-net samples from SB streams provide robust estimates of the MCI-sb (two samples must differ by around 12 MCI-sb units for them to be considered significantly different - cf. about 11 for the MCI). The QMCI-sb performs relatively poorly in this respect. A detectable difference of nearly 1.4 (cf. 0.8 for the SQMCI), equivalent to nearly 28% of an average value (5.0) of this index could substantially affect the assignment of quality classes. We suspect that this reflects the greater between-replicate variability in community composition in SB stream samples. HB samples tend to be collected from relatively uniform stony riffles, whereas the protocol for SB streams involves sampling multiple habitats (submerged wood, bank margins and macrophytes) that vary in occurrence and proportion between sites and samples.

The aim of SOE monitoring is to obtain a "snapshot" of stream quality, but sampling normally is undertaken over a period of several weeks because it is not practical to get around all of the sites more quickly. Interpretation of the SQMCI and QMCI (and SB



variants) can be confounded by temporal changes in dominance (over the period of sample collection) which are unrelated to the pollution status of the site. We recommend caution when using quantitative or semi-quantitative biotic indices for SOE reporting.

We advocate the collection of relative abundance data for SOE monitoring because it is useful to know which taxa are dominant, and the additional information provides for a greater depth of ecological understanding. For reasons outlined above and for better precision in scoring and quality classification (than the quantitative or semi-quantitative versions), we suggest that the MCI (HB streams) and MCI-sb (SB streams) be used for SOE reporting.

In situations where resources (i.e., personnel, expertise, time or funding) are limited, considerable savings for SOE monitoring programmes could be made by processing SB stream samples to provide presence-absence data. SB samples are more costly to process than HB samples due to the amount of organic material collected. We still advocate noting which taxa are dominant (say the top three to five) because that is useful information for an ecologist. This approach would free-up resources that could be used for more intensive surveys aimed, for example, at understanding the impacts of particular activities or land-use practices on stream communities. Semi-quantitative and quantitative versions of the MCI are most suitable for synoptic surveys (when all samples are collected within a day or two) and upstream vs downstream compliance monitoring, where temporal changes in community composition are unlikely to confound interpretation. The SQMCI and QMCI are, of course, suitable for monitoring where temporal changes in community composition are of interest.

Although they were all developed for assessing organic enrichment, the MCI and SQMCI / QMCI (and their SB variants) really are different indices. High values of the MCI (i.e., > 120) indicate that the average tolerance score of the taxa present at the site is high (> 6) and the site would be classified as excellent. If that same site has a QMCI or SQMCI score that results in classification into a lower quality class (i.e., 5.0 – 6.0 = good), it means that the dominant taxon/taxa had a relatively lower score (< 6). The different quality class assignments are not really in conflict when one understands the basis of the indices. In this case, one could conclude that while conditions are not bad enough to exclude many clean-water taxa, there was sufficient enrichment to stimulate the proliferation of one (or more) taxa indicative of enriched conditions. In this respect the quantitative and semi-quantitative versions of the MCI may be more sensitive than the MCI, but this is also the reason why QMCI and SQMCI scores can also be more variable.

Finally, it must be appreciated that there is no single objective method for determining the ecological status of a site. The different quality class assignments that may arise from the use of the MCI-sb and the SQMC-sb/QMCI-sb should not be regarded as “a problem”. In our view, any of these indices, when used appropriately within a robust study design, with their high correlations with environmental variables, can provide a sound basis for water managers to make decisions aimed at maintaining or improving SB stream health in their regions.



## TABLE OF CONTENTS

<b>1.</b>	<b>INTRODUCTION .....</b>	<b>1</b>
<b>2.</b>	<b>DATA REQUIREMENTS .....</b>	<b>2</b>
2.1	Sampling sites, catchment data, and land-use classification .....	2
2.2	Macroinvertebrate Data.....	5
2.2.1	<i>Sample collection</i> .....	5
2.2.2	<i>Sample processing and data management</i> .....	5
2.3	Habitat Quality Index.....	6
2.4	Water Quality Indices .....	7
<b>3.</b>	<b>CALCULATION OF THE MCI, QMCI, AND SQMCI .....</b>	<b>7</b>
<b>4.</b>	<b>DEVELOPMENT OF BIOTIC INDICES FOR SB STREAMS .....</b>	<b>9</b>
4.1	Derivation of taxon scores (tolerance values) .....	9
4.2	Performance assessment of the new biotic indices for SB streams .....	16
4.2.1	<i>Performance Test 1: Expanded biotic index ranges</i> .....	17
4.2.2	<i>Performance Test 2: Relationships with environmental variables</i> .....	19
4.2.3	<i>Performance Test 3: Classification of “severely degraded” sites</i> .....	21
4.2.4	<i>Performance Test 4: Discrimination between land-use classes</i> .....	24
4.3	Determining quality thresholds .....	24
4.4	Precision of index estimates .....	25
<b>5.</b>	<b>APPLICATION OF THE MCI-sb AND QMCI-sb TO THE ASSESSMENT AND MANAGEMENT OF AUCKLAND STREAMS.....</b>	<b>27</b>
5.1	Predictive relationships with land-use.....	31
5.2	Predictive relationships with habitat quality (riparian management) .....	31
5.3	Predictive relationships with water quality .....	33
5.4	The role of the MCI-sb and QMCI-sb in managing SB streams .....	34
<b>6.</b>	<b>SUMMARY AND FUTURE DIRECTIONS .....</b>	<b>37</b>
<b>7.</b>	<b>ACKNOWLEDGMENTS .....</b>	<b>39</b>
<b>8.</b>	<b>LITERATURE CITED .....</b>	<b>40</b>

## LIST OF TABLES

<b>Table 1</b>	Convention used to combine LCDB2 data into four broad land-use categories for the catchments above each sampling site .....	4
<b>Table 2</b>	Criteria used to classify sites into seven land use classes using LCDB2 land-use data.....	4
<b>Table 3</b>	Taxon-specific tolerance scores for use with the MCI, SQMCI and QMCI for HB streams. ....	8
<b>Table 4</b>	Quality thresholds for interpretation of the MCI, SQMCI, & QMCI .....	9
<b>Table 5</b>	Derivation of new taxon scores for SB indices.....	12
<b>Table 6</b>	Differences in taxon scores between the original MCI, scores based on MCI and QMCI iterations and the resulting SB taxon scores.....	14
<b>Table 7</b>	Taxa showing greatest differences in tolerance scores between the original and SB versions of the MCI .....	15
<b>Table 8</b>	Comparison of selected descriptive statistics for the MCI, MCI-sb, QMCI and QMCI-sb.....	17
<b>Table 9</b>	Spearman rank correlation and Pearson linear correlation values between biological metrics and selected environmental variables.....	20
<b>Table 10</b>	Land use, physical habitat, and water quality criteria used to classify sites as “severely degraded” .....	22
<b>Table 11</b>	Ability of the original and proposed SB biotic indices to identify severely degraded (SD) sites correctly .....	23

<b>Table 12</b>	Average standard deviation estimates derived from replicate data for four biotic indices.....	26
<b>Table 13</b>	Mean values for selected water quality variables at rural and urban SB streams in the Auckland region .....	33

## LIST OF FIGURES

<b>Figure 1</b>	Locations of SB sampling sites in the Auckland region .....	3
<b>Figure 2</b>	Boxplots of MCI, MCI-sb, QMCI, and QMCI-sb for SB stream samples in seven land-use classes in the Auckland region .....	18
<b>Figure 3</b>	Box plot depicting the approach to setting quality thresholds for interpretation of biotic index values .....	23
<b>Figure 4</b>	Box plot depicting the approach to setting quality thresholds for interpretation of MCI-sb values.....	25
<b>Figure 5</b>	MCI-sb and QMCI-sb versus % developed land in the catchment .....	28
<b>Figure 6</b>	MCI-sb and QMCI-sb versus ARC Habitat Quality Index .....	29
<b>Figure 7</b>	MCI-sb and QMCI-sb versus ARC Water Quality Index .....	30
<b>Figure 8</b>	Scatter plots of MCI-sb and QMCI-sb vs ARC Habitat Quality Index .....	32
<b>Figure 9</b>	Surface plot showing the relationships between MCI-sb, the percentage of developed land in the catchment, and the ARC Habitat Quality Index.....	36
<b>Figure 10</b>	Surface plot showing the relationships between QMCI-sb, the percentage of developed land in the catchment, and the ARC Habitat Quality Index. ....	36

## LIST OF APPENDICES

<b>Appendix 1</b>	SB stream sites in the Auckland region .....	43
<b>Appendix 2</b>	Catchment landuse data for 41 SB streams in the Auckland region .....	44
<b>Appendix 3</b>	Habitat quality data for 41 SB stream sites in the Auckland region .....	45
<b>Appendix 4</b>	Median water quality data for 14 SB streams in the Auckland region .....	46
<b>Appendix 5</b>	Macroinvertebrate taxon codes used in Appendix 6 .....	47
<b>Appendix 6</b>	Mean counts of 117 macroinvertebrate taxa at 41 SB stream sites in the Auckland region.....	48
<b>Appendix 7</b>	Macroinvertebrate metric values for 41 SB stream sites in the Auckland region.....	54
<b>Appendix 8</b>	Taxon-specific tolerance scores for MCI-based biotic indices in HB and SB streams.....	55
<b>Appendix 9</b>	Rank and linear correlations between MCI, QMCI, MCI-sb, and QMCI-sb and selected environmental variables .....	56

## 1. INTRODUCTION

The need to develop an assessment framework for collection and interpretation of macroinvertebrate data in soft-bottomed (SB) streams was first identified in 1999 (MfE 1999). The essential elements include standard procedures for sample collection and processing, data analysis tools (e.g., sensitive metrics), and quality thresholds (e.g. excellent, good, fair, poor) for interpreting and reporting results as part of the State of the Environment (SOE) monitoring. Since 1999, national protocols have been developed for sample collection, processing, and quality control in both hard-bottomed (HB) and SB streams in New Zealand (Stark et al., 2001). Further, the Macroinvertebrate Community Index (MCI) and the semi-quantitative MCI (SQMCI) were found to be sensitive to organic pollution in SB streams using a data from 10 sites in the Auckland region (Maxted et al., 2003). The need for further testing of metrics across greater variety of sites and the development of quality thresholds was identified. Robust, nationally applicable biotic indices for assessing the health of New Zealand's SB stream habitats are still required if the vision expressed by MfE is to be realized.

The Auckland Regional Council (ARC) contracted with Dr. John Stark to investigate the applicability of the MCI and SQMCI to Auckland SB streams using data from 179 samples collected using Protocol C2 (Stark et al. 2001) from 41 different sites in the Auckland region. This report documents the results of this investigation.

Preliminary results indicated that the ranges of MCI and SQMCI scores were in a narrower band between high and low disturbance in SB streams compared to HB streams, which severely constrained the ability to define ecological conditions accurately. This compression of the scores at the SB sites was due to the lower scores at reference sites and higher scores at severely disturbed sites compared with HB sites. Lower scores at reference sites were due to fewer sensitive EPT taxa in SB streams (Maxted et al., 2003). Higher scores at severely disturbed SB sites may be due to the added taxa found in the multiple habitats sampled using the SB protocol (i.e., submerged wood, streams banks and macrophytes), raising the lowest scores. Ideally, indices that provide lower scores for degraded SB sites, and higher scores for reference sites, are desirable because it increases the range of quality conditions that can be reported and improves the ability to discriminate between sites of similar quality.

Maxted et al. (2003) found that SB reference sites were not classified as "clean water" using the quality thresholds developed for HB streams (Stark 1998). They suggested that providing quality thresholds for interpreting the MCI and SQMCI for SB streams was likely to be more cost-effective than developing a different assessment approach (e.g., multivariate) or a different set of metrics. For example, whereas an MCI > 120 is indicative of reference condition in HB streams (Stark 1998), in SB streams MCI's > 100 seem to indicate reference conditions and few, if any sites had MCI values > 120 (Maxted et al., 2003). Clearly, new index scores or quality thresholds were needed for SB streams.

Our recent analyses of the Auckland data suggested that development of separate MCI, QMCI, and SQMCI indices for SB streams was feasible using an objective method for deriving taxon scores for biotic indices developed in Australia (Chessman and et al., 1997, Chessman 2003). This report documents how these new indices were developed and how they improve both the precision and accuracy of metric scores in defining ecological conditions and stressor relationships in Auckland SB streams.

Note that all references in this report to MCI, SQMCI, and QMCI refer to the original HB stream versions of these indices as described by Stark (1985, 1993, 1998). The versions developed for

SB streams are denoted MCI-sb, SQMCI-sb, and QMCI-sb. Since the SQMCI and QMCI have been shown to provide similar assessments (Stark 1998) any general discussion concerning the performance of the QMCI should apply equally to the SQMCI. We expect the SB versions to be no different in this respect.

## 2. DATA REQUIREMENTS

### 2.1 Sampling sites, catchment data, and land-use classification

Data collected as part of the ARC's Freshwater Ecology Programme were used to develop and test the new biotic indices. The 41 SB sites covered the full range of land uses and degrees of disturbance in the region to provide a comprehensive assessment for State of the Environment (SOE) reporting (Figure 1). This site distribution provided for a robust application of the Chessman process because the method requires a wide range in conditions in order to derive accurate taxon-specific tolerance scores.

Catchment boundaries for each site were delineated using elevation data (20m resolution) digitised from 1:50,000 scale topographic maps (NZ260 series) and ARCVIEW software on the ARC GIS system. Estimates of site elevations and distance from the sea were determined using the River Environment Classification digital database (NIWA 2004) on the ARC GIS system (see Appendix 1).

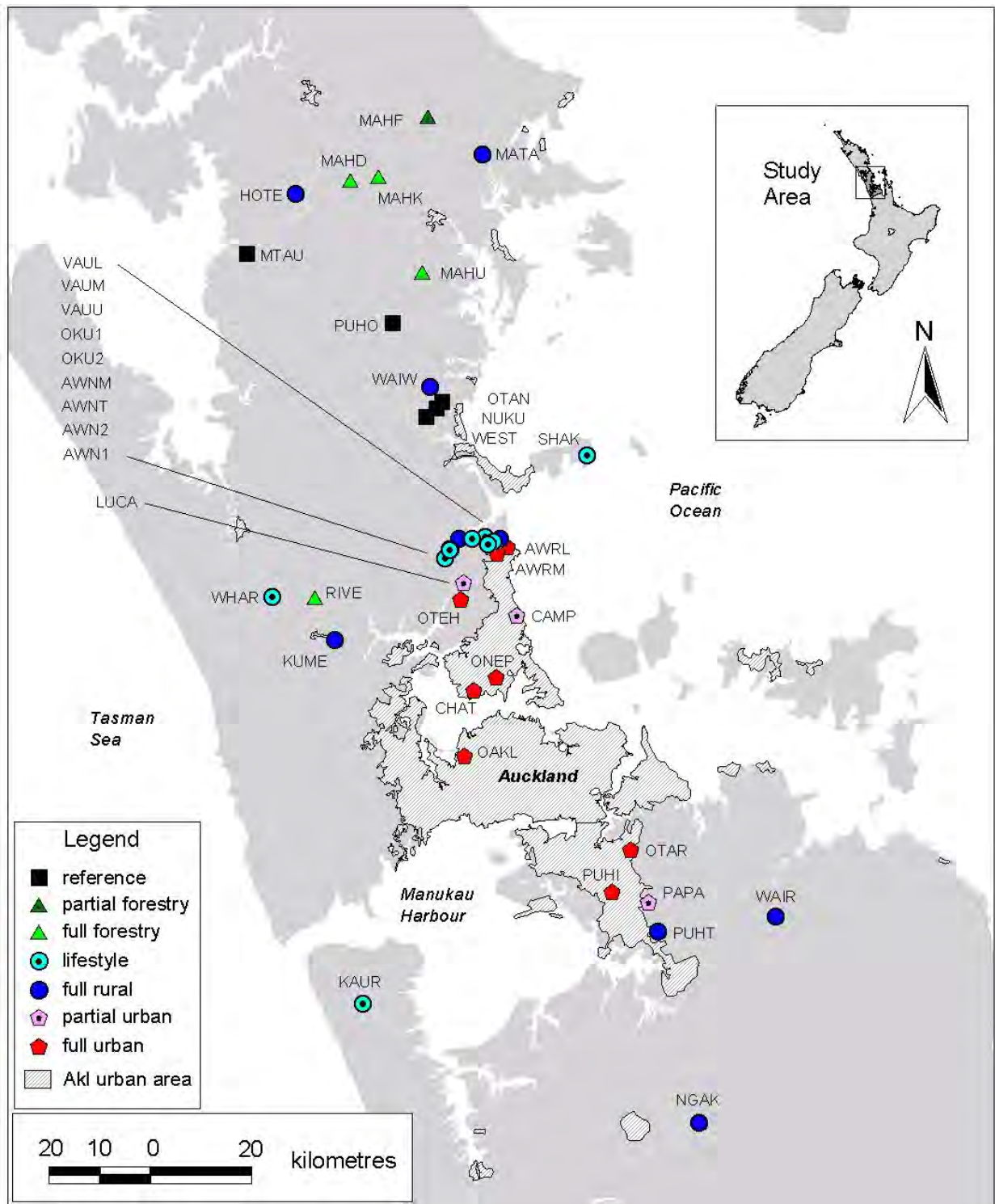
Land-use data were compiled for each catchment from the NZ Land Cover Data Base version 2 (LCDB2) (MfE 2004a). The data were derived from 2002 satellite imagery which was within the field sampling period 2000-2004. The 24 land-use categories found in the Auckland region (MfE 2004b) were combined into four broad land-use categories following the convention presented in Table 1. The area of open water (LCDB2 code 20) was deleted for each site. Minor land-use classes representing < 0.1% of the catchment area were placed in the category "OTHER". Land-use data for the catchments above each site are shown in Appendix 2.

Sites were classified into 4 broad land-use classes (LUC) for plotting and data interpretation using land-use and habitat quality criteria (Table 2). Sites classified REF (reference) had > 95% of the catchment in indigenous or regenerating native bush, < 1% forestry and urban land uses, < 5% rural, and habitat quality scores > 120 points (Table 2). Sites classified RUR (rural) had > 13% of the catchment in rural land use, < 30% forestry, and < 5% urban. Sites classified URB (urban) had > 7% urban land use and < 30% forestry. Sites classified FOR (forestry) had > 95% of the catchment in commercial pine trees (*Pinus radiata*), with one exception: the MAHF site had 30% of the catchment in native bush. Figure 1 shows the location of the 41 sites by LUC. Geographic co-ordinates are presented in Appendix 1.

The ARC SOE monitoring site network was designed to represent the full range of land use intensity, water quality, and habitat quality conditions. Sites were classified further into low and high levels of disturbance using quantitative and qualitative factors. This classification was used to assess the ability of the various biotic indices to separate sites by the level of disturbance within each LUC. Data analysis utilised box plots.

Low and high levels of disturbance in the RUR class were defined by the type of rural land use and the degree of riparian protection in the catchment. RUR1 sites were in a part of the region dominated by "lifestyle" land use (i.e., "partial rural" or "lifestyle rural"; PR) characterised by relatively low stock densities and relatively good riparian protection, including the fencing of stock out of waterways for much of the stream length. RUR1 sites also had HQI scores > 100 points, and generally were well shaded by native vegetation (Table 2). RUR2 sites were in

intensively-farmed catchments (i.e., “full rural”; FR), with HQI scores < 90 points indicating impacted riparian zones with limited shade and direct stock access to steams (Table 2).



**Figure 1** Locations of SB sampling sites in the Auckland region.

**Table 1** Convention used to combine LCDB2 data into four broad land-use categories for the catchments above each sampling site (see Appendix 2 for site data): URB = urban, CROP = cropland, PASGRA = pasture & grassland, ETRSH = exotic trees and shrubs, NAT = indigenous and regenerating native forest, NEWFOR = newly harvested or planted pine, OLDFOR = mature pine, cedar, or gum, OTHER = minor categories.

le	Category (LCDB2)	Description	Grouping (LCDB2 codes)	Land-use category
	built-up area transport infrastructure	urban urban parkland, open space	<b>URB</b> (1+5)	<b>Urban</b>
	short rotation cropland vineyard orchard and other perennial crops	crops vineyard orchard	<b>CROP</b> (30+31+32)	<b>Rural</b>
	urban parkland/ open space high producing exotic grassland low producing grassland	grassy parks and ball fields pasture pasture	<b>PASGRA</b> (2+40+41)	<b>Rural</b>
	gorse and broom mixed exotic shrubland major shelterbelt deciduous hardwoods	retired pasture blackberry, brier linear exotic tree lines (>200m) willow and poplar	<b>ETRSH</b> (51+56+61+68)	<b>Forestry</b>
	Manuka and/or Kanuka broadleaved indigenous hardwoods indigenous forest	regenerating native bush native bush native bush	<b>NAT</b> (52+54+69)	<b>Native</b>
	afforestation (post LCDB 1) forest harvested pine forest - open canopy	recently planted pine recently harvested pine newly planted pine (5-15 yrs old)	<b>NEWFOR</b> (63+64+65)	<b>Forestry</b>
	pine forest - closed canopy other exotic forest	pre-harvest pine (> 20 yrs old) macrocarpa, gum	<b>OLDFOR</b> (66+67)	<b>Forestry</b>
	surface mine dump flaxland (wetland)	open excavations land fills wetlands	<b>OTHER</b> (3+4+47)	-

**Table 2** Criteria (**bold**) used to classify sites into seven land use classes using LCDB2 land-use data from Appendix 2: ranges of site values in *italics* and parentheses.

Land-use class	Class code	Sites N	Samples N	% catchment				ARC habitat score
				native	forestry	rural	urban	
reference	REF	5	30	> <b>95</b>	< <b>1</b>	< <b>5</b>	< <b>1</b>	> <b>120</b>
partial forestry	FOR1	1	4	(68)	<b>30</b>	< <b>5</b>	< <b>1</b>	(75)
full forestry	FOR2	4	14	(0-2)	> <b>95</b>	(0-2)	(0-2)	(54-90)
partial rural	RUR1	10	46	(14-86)	< <b>30</b>	(13-72)	< <b>5</b>	> <b>100</b>
full rural	RUR2	10	39	(1-55)	< <b>30</b>	> <b>35</b>	< <b>5</b>	< <b>90</b>
partial urban	URB1	3	9	(9-39)	< <b>30</b>	(33-58)	<b>10-25</b>	(93-113)
full urban	URB2	8	37	(2-30)	< <b>30</b>	(3-85)	> <b>40</b> *	(42-115)
Totals		41	179					

Low and high levels of disturbance in the URB class were defined by the percent of the catchment in urban land use. URB1 sites had 10-25% of the catchment in urban land use (i.e., “partial urban”; PU), whereas URB2 sites had > 40% urban land use (i.e., “full urban”; FU) (Table 2). Studies in the Auckland region have shown that adverse effects on the invertebrate



community due to urbanisation reach a maximum at approximately 25% impervious cover, which is equivalent to 40% of an area in urban land use (Allibone et al., 2001). The OTAR site with 7.6% urban land use was classified URB2 because the density and intensity of urban land use in the vicinity of the sampling site was extremely high (commercial and industrial) in this predominately rural catchment.

Low and high levels of disturbance in the FOR class were defined by how recently the site had been clear-cut. Sites with minimal disturbance had mature (> 30 years old) pine trees ready for harvest, and were defined as "mature forestry" (MF). Sites with maximum disturbance had been clear-cut within the last two years, and were defined as "immature forestry" (IF).

## **2.2 Macroinvertebrate Data**

### **2.2.1 Sample collection**

A total of 179 macroinvertebrate samples were collected from 41 SB sites during four consecutive summer seasons: years 2000 (18 samples; 18 Feb – 6 Apr), 2001 (28 samples; 5 – 21 Mar), 2002 (89 samples; 20 Feb – 5 Apr), 2003 (8 samples; 31 Mar – 9 Apr), and 2004 (36 samples; 15 Jan - 26 Mar). Triplicate samples were collected at 31 sites (76%) on 44 occasions representing 132 samples (74%): eight sites had replicate samples in more than one year. Collections at six sites were single samples repeated in subsequent years: four sites (MTAU, CAMP, HOTE, CHAT) had a single sample collected in a single year. The number of samples collected at each site appears in Appendix 7.

A 100 m reach of stream, that had relatively homogeneous habitat and no major tributaries, was selected at each site. Macroinvertebrate sample collection followed Protocol C2 of the national protocols (Stark et al. 2001). Samples were collected from submerged wood, bank margins, and macrophytes. These substrata have been found to be stable and the most likely to be colonised by pollution-sensitive taxa in SB streams in New Zealand (Stark et al. 2001). Sand, silt, mud, and coarse and fine detritus were avoided because they are difficult to sample without producing large volumes of detritus that make sample processing more costly and difficult to QC.

Submerged wood, bank margins, and macrophytes were sampled and combined in the 500 µm mesh sieve bucket for a total sample area of 3.0 m<sup>2</sup>. Pieces of submerged wood (30 to 300 mm diameter) were placed in or over the mouth of a sieve bucket. Water was poured over the substratum while moving a hand gently over the surface. Each 1 m length of submerged wood represented a sample area of 0.3 m<sup>2</sup>. Bank margins and macrophytes were sampled by aggressively jabbing the net into submerged structure over a distance of 1 m, followed by two or three cleaning sweeps to collect dislodged organisms. Each sweep represented a sample area of 0.3 m<sup>2</sup>. Substrata were selected in proportion to their abundance at the site, although preference was given to submerged wood due to its importance for epifaunal colonisation. Submerged wood and bank/macrophyte substrata were collected and processed separately in 2000 and 2001, and the data combined for a total surface area of 6.0 m<sup>2</sup>.

### **2.2.2 Sample processing and data management**

Samples collected in 2001, 2002, 2003, and 2004 were processed following Protocol P1 (coded abundance, CA) of the national protocols (Stark et al. 2001). Samples were separated into 2-4 size fractions using stacked 0.5-4.0 mm Endecott<sup>®</sup> sieves, and sorted in white-bottomed trays. One to five specimens of each taxon were removed, identified to the taxonomic level required for the Macroinvertebrate Community Index (MCI), Quantitative Macroinvertebrate Community Index (QMCI), and Semi-Quantitative Macroinvertebrate Community Index (SQMCI) (Stark 1998), and given one of the following abundance codes: rare (R) = 1 to 4 animals/sample;

common (C) = 5 to 19; abundant (A) = 20 to 99; very abundant (VA) = 100 to 499; and very, very abundant (VVA) = 500+ (Stark 1998).

Samples collected in 2000 were processed following Protocol P3 (full counts with subsampling option) of the national protocols (Stark et al. 2001). The entire contents of the samples were sorted, counted, and identified to the lowest practicable taxonomic level. This generally was genus or species for Ephemeroptera, Plecoptera, Trichoptera, Hemiptera, Megaloptera and Odonata, and a mixture of species, genus, and family for most other groups. The number of very abundant taxa (> 300 individuals) was estimated by random counting in 10% of the sample.

Prior to further analyses, the taxonomic data was reduced to the level required for the MCI, generally genus for most groups. The numerical full count data (2000) was converted to CA codes to provide a consistent format for the entire dataset. The full CA dataset was then converted to numerical values corresponding to the lower value for each CA category range (e.g., Abundant = 20). Data quality was checked and documented following the QC procedures contained in the national protocols (Stark et al., 2001).

### **2.3 Habitat Quality Index**

Habitat quality was quantified within a 100m reach at each sampling site using methods developed by the Auckland Regional Council (ARC). Methods developed in 2001 were used to score sites sampled in 2000, 2001, and 2002 (ARC 2001a). Twenty point scores were assigned to seven aquatic, stream bank, and riparian measures. Two measures (i.e., channel bank stability & riparian vegetation disturbance) were scored separately on the true left and true right sides of the channel. A habitat quality index (HQI) was calculated as the sum of all measures for a total possible score of 140 points. Low values of the HQI indicate poor condition and high values indicate excellent condition.

The riparian vegetative type (RVT) measure from the 2001 protocol was replaced with the channel shading (CS) measure in 2003. This change was implemented to eliminate the cultural bias associated with the RVT measure, and to focus on the ecologically relevant shading function that controls stream primary production. Habitat scores prior to 2003 were updated with the new CS variable using estimates of channel shading reported separately on field sheets.

#### Habitat Measures

1. aquatic habitat abundance (AHA) – % channel favourable for epifaunal colonization and fish cover: riffles, submerged wood, bank structure, snags, aquatic macrophytes.
2. aquatic habitat diversity (AHD) – variety of stable aquatic habitats (see AHA).
3. hydrologic heterogeneity (HH) - variety of hydrologic conditions: fast, slow, pool, riffle, run, chute.
4. channel alteration (CA) - degree of human-altered channel pattern (i.e., straightening, armouring) and profile (i.e., bed material, uniform depth).
5. bank stability (BS) - % channel showing evidence of erosion and bank failure (left and right banks scored separately).

6. channel shading (CS)<sup>1</sup> – percent of the channel shaded by canopy and overhanging vegetation of any kind.
7. riparian width (RW) – width of undisturbed riparian vegetation within 20 m of stream edge (left and right banks scored separately).

Habitat quality remained largely unchanged over the sampling period (2000-2004) at all sites, except MAHD which was clear-cut between 2002 (2 samples) and 2004 (1 sample). To reduce data variability between years, a single estimate of habitat quality was selected for each site. Updated (i.e., CS replacing RVT) 2002 habitat data was used for 28 sites (68%) because the 2002 invertebrate data represented 50% of all samples in the data record. Updated 2001 data was used for two sites. The remaining 11 sites sampled in 2003/2004 used the most recent habitat data. The final habitat quality scores (Appendix 3) provided consistent and representative data for the sampling period (2000-2004).

## 2.4 Water Quality Indices

Monthly water quality samples have been collected at 16 sites across the Auckland region since 1991, and the data summarised over the period 1991-2000 (ARC 2000). In this report, Wilcock and Stroud developed a water quality index (WQI-W&S) using median values for 7 parameters: black disc transparency, % DO saturation, faecal coliforms, ammoniacal-N, nitrate-N, total-P, and suspended solids. The median values were used to assign numerical rankings (1-16) to each site, and the rank values were summed for each site to derive index (WQI-W&S) scores.

We used a similar approach to develop a separate water quality index for SB streams (WQI-sb). The data record was updated to cover the period 1991-2003. There were 144 monthly values for most parameters. Three sites with HB substrata were eliminated from the dataset, and the WEST site (SB reference) sampled over 23 months between 2001 and 2003 was added to the dataset for a total of 14 SB sites. The black disk transparency parameter was deleted from the ranking due to data quality problems (i.e., many data values were censored - reported simply as > 1.0 m). Inclusion of turbidity instead of black disk produced a poorer performing WQI based upon regressions with biotic indices. Consequently, the WQI-sb was based on six parameters: % DO saturation, faecal coliforms, ammoniacal-N, nitrate-N, total-P, and suspended solids. Median values were calculated, and the sites ranked and summed following the procedure developed by Wilcock & Stroud (ARC 2000) (Appendix 4). The WQI-sb was calculated as 100 minus the rank sum so that increasing WQI-sb values indicated better water quality. The sensitivity of the biotic indices to water quality conditions was tested using the WQI-W&S and WQI-sb scores for each site.

## 3. CALCULATION OF THE MCI, QMCI, AND SQMCI

Macroinvertebrate Community Index (MCI) values were developed by Stark (1985, 1993, 1998) for assessing organic enrichment of HB streams based on sampling macroinvertebrates from riffle (preferably) or run habitats. The MCI relies on prior allocation of scores (between 1 and 10) to taxa (usually genera) of freshwater macroinvertebrates based upon their relationship to the degree of organic enrichment. Taxa that are characteristic of un-enriched conditions score more

---

<sup>1</sup> Prior to 2003 the sixth habitat measure was riparian vegetation type (RVT) - scored in descending order: native bush, native scrub, mixed native and exotic scrub, exotic trees (e.g., pine, willow, and poplar), planted garden, grassland, none; (left and right banks scored separately). This was replaced by channel shading (CS) from 2003 onwards.

highly than taxa that may be found predominantly in polluted conditions. These scores are given in Table 3. The MCI is calculated as follows:-

$$MCI = \frac{\sum_{i=1}^{i=S} a_i}{S} \times 20$$

where  $S$  = the total number of taxa in the sample, and  $a_i$  is the score for the  $i$ th taxon. The MCI ranges from 0 (when no taxa are present) to 200 (when all taxa score 10 points each) although MCI scores  $< 40$  or  $> 150$  are rare.

**Table 3** Taxon-specific tolerance scores for use with the MCI, SQMCI and QMCI for HB streams (Stark 1985, 1993, 1998; Winterbourn et. al. 2000).

<b>INSECTA</b>		<b>Coleoptera</b>	<b>Diptera (cont.)</b>	<b>Trichoptera (cont.)</b>				
<b>Ephemeroptera</b>		<i>Antiporus</i>	5	<i>Nothodixa</i>	4	<i>Polyplectropus</i>	8	
	<i>Acanthophlebia</i>	7	<i>Berosus</i>	5	Orthoclaadiinae	2	<i>Psilochorema</i>	8
	<i>Ameletopsis</i>	10	<i>Copelatus</i>	5	<i>Parochlus</i>	8	<i>Pycnocentrella</i>	9
	<i>Arachnocolus</i>	8	Dytiscidae	5	<i>Paradixa</i>	4	<i>Pycnocentria</i>	7
	<i>Atalophlebioides</i>	9	Elmidae	6	<i>Paralimnophila</i>	6	<i>Pycnocentrodes</i>	5
	<i>Austroclima</i>	9	<i>Enochrus</i>	5	<i>Paucispinigera</i>	6	<i>Rakiura</i>	10
	<i>Austronella</i>	7	<i>Homeodytes</i>	5	Pelecoryhncidae	9	<i>Synchorema</i>	9
	<i>Coloburiscus</i>	9	Hydraenidae	8	<i>Peritheates</i>	7	<i>Tiphobiosis</i>	6
	<i>Deleatidium</i>	8	Hydrophilidae	5	Podonominae	8	<i>Triplectides</i>	5
	<i>Ichthybotus</i>	8	<i>Liodesus</i>	5	<i>Polypedilum</i>	3	<i>Triplectidina</i>	5
	<i>Isothraulius</i>	8	<i>Podaena</i>	8	Psychodidae	1	<i>Zelandoptila</i>	8
	<i>Mauiulus</i>	5	Ptilodactylidae	8	Sciomyzidae	3	<i>Zelollessica</i>	10
	<i>Neozephlebia</i>	7	<i>Rhantus</i>	5	Stratiomyidae	5	<b>Lepidoptera</b>	
	<i>Nesameletus</i>	9	Scirtidae	8	Syrphidae	1	<i>Hygraula</i>	4
	<i>Oniscigaster</i>	10	Staphylinidae	5	Tabanidae	3	<b>Collembola</b>	6
	<i>Rallidens</i>	9	<b>Neuroptera</b>		Tanypodinae	5	<b>ACARINA</b>	5
	<i>Siphlaenigma</i>	9	<b>Kempynus</b>	5	Tanytarsini	3	<b>ARACHNIDA</b>	
	<i>Tepakia</i>	8	<b>Diptera</b>		<i>Tanytarsus</i>	3	<i>Dolomedes</i>	5
	<i>Zephlebia</i>	7	Anthomyiidae	3	Thaumaleidae	9	<b>CRUSTACEA</b>	
	<b>Plecoptera</b>		<i>Aphrophila</i>	5	Tipulidae	5	Amphipoda	5
	<i>Acroperla</i>	5	<i>Austrosimulium</i>	3	<i>Zelandotipula</i>	6	Copepoda	5
	<i>Austroperla</i>	9	<i>Calopsectra</i>	4	<b>Trichoptera</b>		Cladocera	5
	<i>Cristaperla</i>	8	Ceratopogonidae	3	<i>Alloecentrella</i>	9	Isopoda	5
	<i>Halticoperla</i>	8	Chironomidae	2	<i>Aoteapsyche</i>	4	Ostracoda	3
	<i>Megaleptoperla</i>	9	<i>Chironomus</i>	1	<i>Beraeoptera</i>	8	<i>Paraleptamphopus</i>	5
	<i>Nesoperla</i>	5	<i>Corynoneura</i>	2	<i>Confluens</i>	5	<i>Paracalliope</i>	5
	<i>Spaniocerca</i>	8	<i>Cryptochironomus</i>	3	<i>Conuxia</i>	8	<i>Paranephrops</i>	5
	<i>Spaniocercoides</i>	8	<i>Culex</i>	3	<i>Costachorema</i>	7	<i>Paratya</i>	5
	<i>Stenoperla</i>	10	Culicidae	3	<i>Cryptobiosella</i>	9	Tanaidacea	4
	<i>Taraperla</i>	7	Diptera indet.	3	<i>Diplectrona</i>	9	<b>MOLLUSCA</b>	
	<i>Zelandobius</i>	5	Dixidae	4	<i>Ecnomina</i>	8	<i>Ferrissia</i>	3
	<i>Zelandoperla</i>	10	Dolichopodidae	3	<i>Edpercivalia</i>	9	<i>Glyptophysa = Physastra</i>	5
	<b>Megaloptera</b>		Empididae	3	Ecnominidae	8	<i>Gyraulus</i>	3
	<i>Archichauliodes</i>	7	Ephydriidae	4	<i>Helicopsyche</i>	10	<i>Hyridella</i>	3
	<b>Odonata</b>		Eriopterini	9	<i>Hudsonema</i>	6	<i>Latia</i>	3
	<i>Aeshna</i>	5	Harrisius	6	<i>Hydrobiosella</i>	9	<i>Lymnaea</i>	3
	<i>Antipodochlora</i>	6	Hexatomini	5	<i>Hydrobiosis</i>	5	<i>Melanopsis</i>	3
	<i>Austrolestes</i>	6	<i>Limnophora</i>	3	<i>Hydrochorema</i>	9	<i>Physa (= Physella)</i>	3
	<i>Hemicordulia</i>	5	<i>Limonia</i>	6	<i>Kokiria</i>	9	<i>Potamopyrgus</i>	4
	<i>Procordulia</i>	6	<i>Lobodiamesa</i>	5	<i>Neurochorema</i>	6	Sphaeriidae	3
	<i>Uropetala</i>	5	<i>Maoridiamesa</i>	3	<i>Oecetis</i>	6	<b>OLIGOCHAETA</b>	1
	<i>Xanthocnemis</i>	5	<i>Mischoderus</i>	4	Oeconesidae	9	<b>HIRUDINEA</b>	3
	<b>Hemiptera</b>		<i>Molophilus</i>	5	<i>Olinga</i>	9	<b>PLATYHELMINTHES</b>	3
	<i>Anisops</i>	5	Muscidae	3	<i>Orthopsyche</i>	9	<b>NEMATODA</b>	3
	<i>Diaprepocoris</i>	5	<i>Nannochorista</i>	7	<i>Oxyethira</i>	2	<b>NEMATOMORPHA</b>	3
	<i>Microvelia</i>	5	<i>Neocurupira</i>	7	<i>Paroxyethira</i>	2	<b>NEMERTEA</b>	3
	Saldidae	5	<i>Neolimnia</i>	3	<i>Philorheithrus</i>	8	<b>COELENTERATA</b>	
	<i>Sigara</i>	5	Neoscatella	7	<i>Plectrocnemia</i>	8	<i>Hydra</i>	3

The QMCI is calculated from count data as follows:-

$$QMCI = \sum_{i=1}^{i=S} \frac{(n_i \times a_i)}{N}$$

where  $S$  = the total number of taxa in the sample,  $n_i$  is the abundance for the  $i$ th scoring taxon,  $a_i$  is the score for the  $i$ th taxon (Table 3) and  $N$  is the total of the coded abundances for the entire sample.

The SQMCI is calculated in a similar manner to the QMCI except that coded abundances (assigned to the R, C, A, VA and VVA abundance classes) are substituted for actual counts: i.e.,

$$SQMCI = \sum_{i=1}^{i=S} \frac{(n_i \times a_i)}{N}$$

where  $S$  = the total number of taxa in the sample,  $n_i$  is the coded abundance for the  $i$ th scoring taxon (i.e. R=1, C=5, A=20, VA=100, VVA=500),  $a_i$  is the score for the  $i$ th taxon (Table 3) and  $N$  is the total of the coded abundances for the entire sample. The QMCI and SQMCI indices range from 0 to 10.

The interpretation of MCI, SQMCI, and QMCI values when applied to HB streams is given in Table 4. Four quality classes were provided by Stark (1998) corresponding to different levels of organic pollution (enrichment). In recognition of the fact that biotic indices, even those developed primarily to reflect organic enrichment, can respond to other forms of disturbance (e.g., sedimentation, toxic pollution), we now suggest that sites should be assigned to quality classes on an excellent – good – fair – poor scale (Table 4).

**Table 4** Quality thresholds for interpretation of the MCI, SQMCI, & QMCI.

Quality Class	Stark (1998) descriptions	MCI	SQMCI & QMCI
Excellent	Clean water	> 120	> 6.0
Good	Doubtful quality or possible mild pollution	100 - 120	5.0 - 6.0
Fair	Probable moderate pollution	80 - 100	4.0 – 5.0
Poor	Probable severe pollution	< 80	< 4.0

## 4. DEVELOPMENT OF BIOTIC INDICES FOR SB STREAMS

### 4.1 Derivation of taxon scores (tolerance values)

During the development of the MCI, Stark (1985) assigned taxon scores using a weighting procedure based upon the relative percentage occurrence of taxa at three site groups differing in enrichment status (i.e., clean and unenriched, slight to moderate pollution, moderate to gross pollution). Assignment of sites to the three site groups was fairly subjective – based upon knowledge of catchment land-use and the existence of diffuse and point-source discharges. Scores for less common taxa for which this procedure was unreliable (Stark 1985), or those added subsequently (Stark 1993, 1998) were assigned by professional judgment. Prior to 1985, taxon scores or tolerance values for most of biotic indices that had been developed overseas had been assigned by professional judgment.

More recently in Australia, however, Bruce Chessman devised a procedure whereby taxon scores could be derived objectively from a taxa by site data matrix provided the sites covered a wide (preferably a full) range of disturbance or stream health (Chessman et al. 1997). This procedure was improved subsequently by Chessman (2003).

The iterative process described by Chessman (2003) was used to derive taxon scores for 117 invertebrate taxa recorded in 179 samples collected in summer 2000 – 2004 from SB streams in the Auckland region. To avoid pseudo-replication, the dataset was reduced from a taxa-by-sample dataset to a taxa-by-site dataset by calculating mean values for each taxon from all samples collected at each site during the period 2000 - 2004. The resulting data file contained average counts of 117 taxa collected from 41 different SB stream sites (Appendices 5 & 6). All available SB data were used because we considered that the resulting scores would be more reliable than if a subset of the data were used, and to ensure that no taxa were omitted. (To derive scores for taxa omitted or not encountered in the SB data set, one must either repeat the iterative process using a data set containing the additional taxa, or assign the scores by professional judgement).

The level of taxonomic resolution (i.e., primarily generic) used for the existing MCI, SQMCI and QMCI was retained.

The iterative score derivation process proceeded as follows. MCI values were calculated for the SB stream dataset on an Excel spreadsheet. Spearman rank correlations ( $R_s$ ) were calculated between the MCI values and the abundances of all taxa across all 55 sites using STATISTICA 6.1. Since it is mathematically impossible for rare taxa with few occurrences to achieve large positive or negative correlations (Chessman et al. 1997), each  $R_s$  was expressed as a proportion of the maximum possible  $R_s$  for a taxon recorded from the same proportion of samples. The taxon with the highest positive adjusted  $R_s$  was assigned a score of 10, and that with the lowest negative adjusted  $R_s$  was assigned a score of 1. The remaining taxa were assigned integer scores between these extremes in proportion to their adjusted  $R_s$  values. The resulting taxon scores were pasted back into the Excel spreadsheet and a new set of MCI values were calculated. This procedure, which we termed the “binary process”, was repeated until the scores stabilized (i.e., there was no change in the scores for any of the taxa from one iteration to the next). This analysis was then repeated for the QMCI to derive another set of taxon scores. We refer to this procedure and the “quantitative process”.

At this point, we had three sets of taxon tolerance scores (i) the original scores (Stark 1985, 1993, 1998), (ii) scores derived from the iterations using the MCI, and (iii) scores derived from the iterations using the QMCI. Rank and linear correlations between environmental variables and MCI/QMCI values calculated using these three sets of scores were undertaken in order to determine whether new indices derived using MCI-based or QMCI-based iterations yielded the greater increase in  $R_s$  or  $r^2$ . The results were mixed, suggesting that some compromise between the taxon scores derived from the two iterative processes was warranted.

A variety of methods for determining scores for the MCI-sb and QMCI-sb based upon the scores derived by the binary and quantitative iterations were trialled. The approach that was adopted is described below. For each taxon, the scores derived by the MCI- and QMCI-based iterations were averaged. If the average was an integer, then that became the new taxon score. If the average was 1.5, 2.5, 3.5, ...8.5, or 9.5, then the score was rounded down or up to the integer nearest the original MCI score. For example, if the new scores from the binary and quantitative processes were 4 and 7, the mean value would be 5.5. If the original MCI score was 5, then the mean MCI-sb value of 5.5 would be rounded down to 5. *A priori* we decided to round taxon tolerance scores towards the original MCI scores because the MCI did perform acceptably in SB

streams and rounding away from the original score was more likely to result in poorer performing indices.

The apparent circularity of the iterative process developed by Chessman (2003) deserves some comment. The process requires that all sites (or samples) in the dataset are ordered initially from best to worst in terms of an environmental gradient. For most existing biotic indices (including the MCI) this normally has been an enrichment gradient. Ideally, this would be done independently of the biological data, but if there was an easy way of doing this perhaps there would be no need for biotic indices? Chessman (2003) used SIGNAL (the Australian MCI equivalent) to determine the initial site order noting that SIGNAL was well-proven as an indicator of stream health in Australia. Likewise, we have used the MCI (and QMCI) to determine the initial site order, because in New Zealand the MCI and QMCI have been shown to correlate well with indicators of enrichment (e.g., Quinn & Hickey 1990a), and to perform adequately in SB streams (Maxted et al. 2003).

Chessman's (2003) iterative procedure provided stable scores from the 4<sup>th</sup> to the 5<sup>th</sup> iteration for both the MCI- and QMCI-based processes. The resulting scores were averaged and integerised (if necessary) by rounding up or down towards the original MCI scores (Table 5). Calculation of SB versions of the MCI is exactly the same as that for the original MCI (and variants) except that the SB taxon tolerance scores are used (see Section 3.0, Table 5, Appendix 8).

**Table 5** Derivation of new taxon scores for SB indices using binary (MCI) and quantitative (QMCI) iterations.

Taxon	Original Score	MCI iterations					QMCI iterations					Average	SB Score	
		1	2	3	4	5	1	2	3	4	5			
<i>Acanthophlebia</i>	5	9	9	9	9	9	9	9	9	9	9	9	9	9
<b>ACARINA</b>	5	6	6	5	5	5	5	5	5	5	5	5	5	5
<i>Acroperla</i>	5	6	5	5	5	5	5	5	5	5	5	5	5	5
<i>Ameletopsis</i>	10	10	10	10	10	10	10	10	10	10	10	10	10	10
Amphipoda	5	5	5	5	5	5	5	5	5	5	5	5	5	5
<i>Anisops</i>	5	3	3	3	3	3	3	3	3	3	3	3	3	3
Anthomyiidae	3	4	4	4	4	4	4	4	4	4	4	4	4	4
<i>Antipodochlora</i>	6	6	6	6	6	6	7	7	7	7	7	7	6.5	6
<i>Antiporus</i>	5	4	4	4	4	4	5	5	5	5	5	5	4.5	5
<i>Aoteapsyche</i>	4	6	6	6	6	6	6	6	6	6	6	6	6	6
<i>Aphrophilia</i>	5	6	6	6	6	6	6	6	6	6	6	6	6	6
<i>Arachnocolus</i>	8	8	8	8	8	8	7	7	7	7	7	7	7.5	8
<i>Archichauliodes</i>	7	7	7	7	7	7	7	7	7	7	7	7	7	7
<i>Austroclima</i>	9	6	6	6	6	6	6	5	5	5	5	5	5.5	6
<i>Austrolestes</i>	6	2	2	2	2	2	2	2	2	2	2	2	2	2
<i>Austronella</i>	7	5	4	4	4	4	5	5	5	5	5	5	4.5	5
<i>Austroperla</i>	9	8	8	8	8	8	8	8	8	9	9	9	8.5	9
<i>Austrosimulium</i>	3	4	4	4	4	4	4	4	4	4	4	4	4	4
Ceratopogonidae	3	6	6	6	6	6	7	8	8	8	8	8	7	7
Chironomidae	2	5	5	5	5	5	5	5	6	6	6	6	5.5	5
<i>Chironomus</i>	1	3	4	4	4	4	5	5	5	5	5	5	4.5	4
Collembola	6	5	5	5	5	5	6	6	6	6	6	6	5.5	6
<i>Coloburiscus</i>	9	8	8	8	8	8	7	7	7	7	7	7	7.5	8
<i>Copelatus</i>	5	4	4	4	4	4	4	3	3	3	3	3	3.5	4
Culicidae indet.	3	1	1	1	1	1	3	4	4	4	4	4	2.5	3
<i>Deleatidium</i>	8	6	6	6	6	6	5	5	5	5	5	5	5.5	6
Diptera	3	4	4	4	4	4	5	5	5	5	5	5	4.5	4
Dixidae	4	8	9	9	9	9	9	9	9	9	9	9	9	9
Dolichopodidae	3	8	8	8	8	8	8	8	8	8	8	8	8	8
<i>Dolomedes</i>	5	7	7	7	6	6	4	4	4	4	4	4	5	5
<i>Ecnomina</i>	8	9	9	9	9	9	8	8	8	8	8	8	8.5	8
Elmidae	6	7	7	7	7	7	7	7	7	7	7	7	7	7
Empididae	3	6	6	7	7	7	8	8	8	8	8	8	7.5	7
Enochrus	5	3	3	3	3	3	1	2	2	2	2	2	2.5	3
Eriopterini	9	8	8	8	8	8	8	8	8	8	8	8	8	8
<i>Ferrisia</i>	3	3	3	3	3	3	2	3	3	3	3	3	3	3
<i>Gyraulus</i>	3	2	2	2	2	2	3	3	3	3	3	3	2.5	3
<i>Harrisius</i>	6	5	5	5	5	5	4	4	4	4	4	4	4.5	5
<i>Helicopsyche</i>	10	8	8	8	8	8	8	8	8	8	8	8	8	8
<i>Hemicordulia</i>	5	1	1	1	1	1	2	2	2	2	2	2	1.5	2
Hexatomini	5	7	7	7	7	7	7	7	7	7	7	7	7	7
<b>HIRUDINEA</b>	3	2	2	2	2	2	2	2	2	2	2	2	2	2
<i>Hudsonema</i>	6	7	7	7	7	7	7	7	7	7	7	7	7	7
<i>Hydrobiosis</i>	5	6	6	6	6	6	6	6	6	6	6	6	6	6
Hydrophilidae	5	7	7	7	7	7	6	6	6	6	6	6	6.5	6
<i>Hygraula</i>	4	2	2	2	2	2	1	1	1	1	1	1	1.5	2
<i>Hyridella</i>	3	7	7	7	7	7	6	6	6	6	6	6	6.5	6
Isopoda	5	4	4	4	4	4	5	4	5	5	5	5	4.5	5
<i>Isothraulius</i>	8	7	7	7	7	7	7	7	8	8	8	8	7.5	8
<i>Latia</i>	3	6	6	6	6	6	7	7	6	6	6	6	6	6
<i>Limnophora</i>	3	5	5	5	5	5	6	6	6	6	6	6	5.5	5
<i>Limonia</i>	6	6	6	6	6	6	6	6	6	6	6	6	6	6
<i>Lobodimaesa</i>	5	6	7	7	7	7	9	9	9	9	9	9	8	8
<i>Lymnaea</i>	3	2	2	2	2	2	1	1	1	1	1	1	1.5	2
<i>Maoridiamesa</i>	3	5	5	5	5	5	5	5	5	5	5	5	5	5
<i>Mauiulus</i>	5	4	4	4	4	4	4	4	4	4	4	4	4	4



**Table 5** continued

Taxon	Original Score	MCI iterations					QMCI iterations					Average	SB Score	
		1	2	3	4	5	1	2	3	4	5			
<i>Megaleptoperla</i>	9	7	7	7	7	7	7	7	7	7	7	7	7	7
<i>Melanopsis</i>	3	3	3	3	3	3	3	3	3	3	3	3	3	3
<i>Microwelia</i>	5	5	5	5	5	5	4	4	4	4	4	4	4	4.5
<i>Mischoderus</i>	4	5	6	6	6	6	5	5	5	5	5	5	5	5.5
<i>Molophilus</i>	5	6	6	7	7	7	7	7	7	7	7	7	7	7
Muscidae	3	3	3	3	3	3	2	2	2	2	2	2	2	2.5
Nematoda	3	4	4	4	4	4	3	3	3	3	3	3	3	3.5
Nematomorpha	3	3	3	3	3	3	5	6	6	6	6	6	6	4.5
Nemertea	3	3	3	3	3	3	3	3	3	3	3	3	3	3
<i>Neolimnia</i>	3	5	5	5	5	5	4	4	4	4	4	4	4	4.5
<i>Neozephlebia</i>	7	7	7	7	7	7	7	7	7	7	7	7	7	7
<i>Nesameletus</i>	9	9	9	9	9	9	9	9	8	8	8	8	8	8.5
<i>Nesoperla</i>	5	7	7	6	6	6	3	2	2	2	2	2	2	4
<i>Neurochorema</i>	6	5	5	5	5	5	5	5	5	5	5	5	5	5
<i>Oecetis</i>	6	8	8	8	8	8	7	7	7	7	7	7	7	7.5
<i>Oeconesus</i>	9	6	6	6	6	6	6	6	6	6	6	6	6	6
Oligochaeta	1	4	4	4	4	4	4	4	4	4	4	4	4	4
<i>Olinga</i>	9	8	8	8	8	8	7	7	7	7	7	7	7	7.5
Orthoclaadiinae	2	3	3	3	3	3	4	4	4	4	4	4	4	3.5
<i>Orthopsyche</i>	9	7	7	7	7	7	6	6	6	6	6	6	6	6.5
Ostracoda	3	2	2	2	2	2	3	3	3	3	3	3	3	2.5
<i>Oxyethira</i>	2	2	2	2	2	2	2	2	2	2	2	2	2	2
<i>Paradixa</i>	4	8	8	8	8	8	8	8	8	8	8	8	8	8
<i>Paralimmophila</i>	6	7	7	7	7	7	7	7	7	7	7	7	7	7
<i>Paranephrops</i>	5	8	8	8	8	8	8	8	8	8	8	8	8	8
<i>Paratya</i>	5	4	4	4	4	4	3	3	3	3	3	3	3	3.5
<i>Paroxyethira</i>	2	4	4	4	4	4	4	4	4	4	4	4	4	4
<i>Paucispinigera</i>	6	7	7	7	7	7	8	8	8	9	9	9	9	8
<i>Physa</i>	3	1	1	1	1	1	1	1	1	1	1	1	1	1
Platyhelminthes	3	2	2	2	2	2	2	2	2	2	2	2	2	2
<i>Polypedilum</i>	3	7	8	8	8	8	7	7	7	7	7	7	7	7.5
<i>Polypsectropus</i>	8	8	8	8	8	8	8	8	8	8	8	8	8	8
<i>Potamopyrgus</i>	4	2	3	3	3	3	1	1	1	1	1	1	1	2
<i>Psilochorema</i>	8	8	8	8	7	7	7	7	7	7	7	7	7	7
Psychodidae	1	8	9	9	9	9	8	8	8	8	8	8	8	8.5
Ptilodactylidae	8	7	7	7	7	7	7	6	6	6	6	6	6	6.5
<i>Pycnocentria</i>	7	7	7	7	7	7	6	6	6	6	6	6	6	6.5
<i>Pycnocentroides</i>	5	4	4	4	4	4	5	5	5	5	5	5	5	4.5
<i>Rallidens</i>	9	4	4	4	4	4	5	5	5	5	5	5	5	4.5
<i>Rhantus</i>	5	1	1	1	1	1	2	3	3	3	3	3	3	2
Saldidae	5	5	5	4	4	4	3	3	3	3	3	3	3	3.5
Sciomyzidae	3	4	4	4	4	4	4	4	4	4	4	4	4	4
Scirtidae	8	7	7	7	6	6	6	6	6	6	6	6	6	6
<i>Sigara</i>	5	3	3	3	3	3	3	3	3	3	3	3	3	3
<i>Spaniocerca</i>	8	9	9	9	9	9	7	7	7	7	7	7	7	8
Sphaeriidae	3	4	4	4	4	4	2	2	2	2	2	2	2	3
Staphylinidae	5	6	6	6	6	6	7	6	6	7	7	7	7	6.5
Stratiomyidae	5	4	4	4	4	4	1	1	1	1	1	1	1	2.5
Tabanidae	3	7	7	7	7	7	7	7	7	7	7	7	7	7
Tanypodinae	5	7	7	7	7	7	5	5	5	5	5	5	5	6
<i>Tanytarsus</i>	3	4	4	4	4	4	5	5	5	5	5	5	5	4.5
<i>Tepakia</i>	8	8	8	8	8	8	7	7	7	7	7	7	7	7.5
Tipulidae	5	5	5	4	4	4	5	5	5	5	5	5	5	4.5
<i>Triplectides</i>	5	6	6	6	6	6	5	5	5	5	5	5	5	5.5
<i>Uropetala</i>	5	2	2	2	1	1	1	1	1	1	1	1	1	1
<i>Xanthocnemis</i>	5	2	2	2	2	2	1	1	1	1	1	1	1	1.5
<i>Zelandobius</i>	5	7	7	7	7	7	6	6	6	6	6	6	6	6.5
<i>Zelandoperla</i>	10	9	9	9	9	9	8	8	8	8	8	8	8	8.5
<i>Zelandoptila</i>	8	7	7	7	7	7	6	6	6	6	6	6	6	6.5
<i>Zelandotipula</i>	6	4	4	4	4	4	3	3	3	3	3	3	3	3.5
<i>Zephlebia</i>	7	8	8	8	8	8	9	8	8	8	8	8	8	8

**Table 6** Differences in taxon scores between the original MCI, scores based on MCI and QMCI iterations and the resulting SB taxon scores.

	MCI iterations		QMCI iterations		New SB scores		
	No. of taxa	%	No. of taxa	%	No. of taxa	%	
SB score higher than original by	8	1	0.9	-	-	-	-
	7	-	-	1	0.9	1	0.9
	6	-	-	-	-	-	-
	5	3	2.6	4	3.4	2	1.7
	4	5	4.3	7	6.0	6	5.1
	3	6	5.1	7	6.0	7	6.0
	2	13	11.1	9	7.7	7	6.0
	1	21	17.9	16	13.7	20	17.1
SB and original score equal	0	20	17.1	25	21.4	35	29.9
SB score lower than original by	1	27	23.1	15	12.8	17	14.5
	2	12	10.3	18	15.4	14	12.0
	3	4	3.4	9	7.7	5	4.3
	4	4	3.4	6	5.1	3	2.6
	5	1	0.9	-	-	-	-

The degree of agreement of the objectively derived taxon scores with the original MCI scores is presented in Table 6. The MCI-based iterations produced scores that were up to 8 units higher and 5 units lower than the original scores. The scores for 20 taxa (17.1%) were identical to the original scores, with the scores for 68 taxa (58.1%) within  $\pm 1$  units of the original scores. The QMCI-based iterations produced new scores that were up to 7 units higher and 4 units lower than the original scores, and although the scores for 25 taxa (21.4%) were the same, there were fewer taxa (56: 47.9%) within  $\pm 1$  of the original MCI scores.

Averaging the scores from the MCI and QMCI iterations (and rounding up or down towards the original MCI scores if required) produced the final scores for the SB versions of the MCI. This procedure produced scores that ranged from 7 units higher to 4 units lower than the original MCI scores, and better agreement overall with the original scores. The scores for 35 taxa (29.9%) were the same as the original scores, with the scores for 72 taxa (61.5%) within  $\pm 1$  units of the original scores.

The extent of agreement between taxon scores developed for SB streams and the original scores developed for HB streams is comforting for several reasons. Firstly, the original MCI, SQMCI, and QMCI do perform acceptably in SB streams (despite the fact that they were developed for HB streams), albeit with restricted ranges of index values and different thresholds for interpretation or classifying sites into quality classes. In developing SB versions of the MCI using a much more objective process for assigning scores than was used by Stark (1985, 1993, 1998), the similarity of many of the taxon scores suggests that the original MCI scores for many taxa were “not too far off the mark”.

The MCI, SQMCI, and QMCI have been shown to reflect water quality and stream health in many studies of HB streams (e.g. Quinn & Hickey 1990a) and some studies of SB streams (Maxted et al. 2003; Collier 2004). Although SB streams differ in many ways from HB streams, it would be surprising if indices developed for them did not show a high degree of similarity to those for HB streams, especially since many of the same macroinvertebrate taxa are present in both stream types.

**Table 7** Taxa showing greatest differences in tolerance scores ( $> \pm 2$ ) between the original and SB versions of the MCI. Shaded rows indicate taxa that were not included in the original Taranaki Ringplain dataset (Stark 1985).

Taxon	Major group	Original MCI score	MCI-sb score	Difference
<i>Austrolestes</i>	Odonata	6	2	-4
<i>Rallidens</i>	Ephemeroptera	9	5	-4
<i>Uropetala</i>	Odonata	5	1	-4
<i>Austroclima</i>	Ephemeroptera	9	6	-3
<i>Hemicordulia</i>	Odonata	5	2	-3
<i>Oeconesus</i>	Trichoptera	9	6	-3
<i>Rhantus</i>	Coleoptera	5	2	-3
<i>Xanthocnemis</i>	Odonata	5	2	-3
Chironomidae	Diptera	2	5	3
<i>Chironomus</i>	Diptera	1	4	3
<i>Hyridella</i>	Mollusca	3	6	3
<i>Latia</i>	Mollusca	3	6	3
<i>Lobodimaesa</i>	Diptera	5	8	3
Oligochaeta	Oligochaeta	1	4	3
<i>Paranephrops</i>	Crustacea	5	8	3
<i>Acanthophlebia</i>	Ephemeroptera	5	9	4
Ceratopogonidae	Diptera	3	7	4
Empididae	Diptera	3	7	4
<i>Paradixa</i>	Diptera	4	8	4
<i>Polypedilum</i>	Diptera	3	7	4
Tabanidae	Diptera	3	7	4
Dixidae	Diptera	4	9	5
Dolichopodidae	Diptera	3	8	5
Psychodidae	Diptera	1	8	7

MCI-sb tolerance values for 24 taxa differed from the original MCI scores by 3 or more units (Table 7). Diptera (11 taxa) dominated the list of taxa whose tolerance scores changed most from the MCI to the SB version. Even more interesting is the fact that scores for dipteran taxa were higher for the SB index, suggesting that many of these Diptera may be associated more with finer sediments in HB streams rather than enrichment *per se*. Odonata (4 taxa) were next best represented and their scores were all lower for the SB versions of the MCI. The scores for three ephemeropterans, two molluscans, one coleopteran, one trichopteran, one crustacean, and one oligochaete also changed by three or more MCI units.

The greatest change was for Psychodidae (moth flies) where the score increased from 1 to 8. Stark (1985) considered that psychodids were characteristic of enriched habitats in HB streams. However, Winterbourn et al. (2000) noted that “little is known about their larvae or habitats” and that they “occur along stream margins in mud, and amongst decaying leaves.” Chris Fowles (Taranaki Regional Council, pers. comm.) has found them in leaf packs in pristine HB mountain streams on Mt Taranaki, so a higher score than 1 is almost certainly warranted.

Another taxon worthy of comment is the Oligochaeta, where the score increased from 1 to 4. There are many different segmented worms present in streams but it is primarily the Tubificidae that are indicative of grossly polluted (enriched) conditions. Other worms (e.g. the Lumbricidae or Lumbriculidae) can be found in pristine environments. There were few (if any) tubificids in the samples from Auckland SB streams, so a higher score for Oligochaeta was the result. The Oligochaeta is one group of invertebrates where scoring taxa at a finer taxonomic level could

result in better performing biotic indices. This would not be without some cost, however, as worms are often poorly preserved and fragmented in samples and can be difficult to identify.

A total of 12 taxa in Table 7 (50%) had their original MCI tolerance scores assigned by professional judgement because they were not present in the Taranaki Ringplain dataset that formed the basis for the original MCI (Stark 1985). This alone could be the reason scores derived from the SB dataset were different from those for the original MCI (let alone any differences that may exist in the habitat preferences of taxa between SB and HB streams).

With any numerical procedure for assigning scores to taxa – the simple method used by Stark (1985) or the more complex method used here – the scores that result reflect the pattern in the underlying dataset. If that pattern does not represent a valid environmental gradient (e.g., enrichment) then the resulting scores are unlikely to form a useful basis for a generally applicable biotic index.

The difficulty, of course, is determining what the true tolerance scores should be. Objective derivation of scores can produce scores that are an artefact of the dataset they are derived from if that dataset does not encompass a full range of conditions along a disturbance (enrichment) gradient. Scores allocated by professional judgement are not necessarily inaccurate but are highly depended upon one's knowledge of the habitat requirements or pollution tolerances of each taxon. For rarely encountered taxa, often there is not enough known to enable scores to be assigned without a high element of guesswork. Equally, if taxa are not well represented in datasets, then the objective score derivation procedure may not perform well either.

#### **4.2 Performance assessment of the new biotic indices for SB streams**

In the end, the performance of the resulting indices is the real test. This can be assessed simplistically, as Stark (1985) did, by determining whether or not assessments based upon biotic indices “make sense” or correspond with assessments made by experienced macroinvertebrate ecologists. More sophisticated tests of performance of biotic indices can be done by correlating index values with environmental data. Quinn & Hickey (1990a), for example, found “moderately strong” correlations between both the MCI and QMCI and indicators of enrichment when applied to run/riffle habitat in 88 New Zealand rivers, and concluded that these indices were more useful indicators of water quality than species diversity, species richness, and the EPT (i.e., the number of Ephemeroptera, Plecoptera and Trichoptera taxa) index.

Another way to test the performance of biotic indices is to apply them to an independent set of macroinvertebrate data collected from SB streams in the Auckland region. We could not do this because we decided to include all of the data from Auckland SB streams that was available to us in the iterative process for deriving taxon scores for the MCI-sb. We did this because we considered that taxon scores resulting from this process were likely to be more reliable if the dataset was large. Secondly, it was important to ensure that as many taxa as possible were included in the process otherwise their scores would not be derived objectively and would need to be assigned by “professional judgement”. The disadvantage of our approach is that we had no other macroinvertebrate data from Auckland's SB streams that we could use as an independent test of the performance of the new SB indices. This will be rectified when additional SB data are collected. These data will then provide an independent test of the ability of the MCI-sb and QMCI-sb to assign streams to appropriate quality classes.

Consequently, in this report we evaluated the performance of the MCI-sb and QMCI-sb by comparing them to the MCI and QMCI using four analyses. First, we tested the ability of the new SB indices to spread the scoring range, and provide greater separation between site scores.

Second, we examined linear and rank correlations between index values and a range of environmental variables at catchment (e.g., land-use, water quality) and local (e.g., habitat quality) scales. Third, we compared the ability of the HB and SB indices to assign “severely degraded” sites (based upon physical habitat and water quality) to the appropriate quality class. Finally, we examined the ability of the various biotic indices to discriminate between land use classes and degrees of disturbance.

#### 4.2.1 Performance Test 1: Expanded biotic index ranges

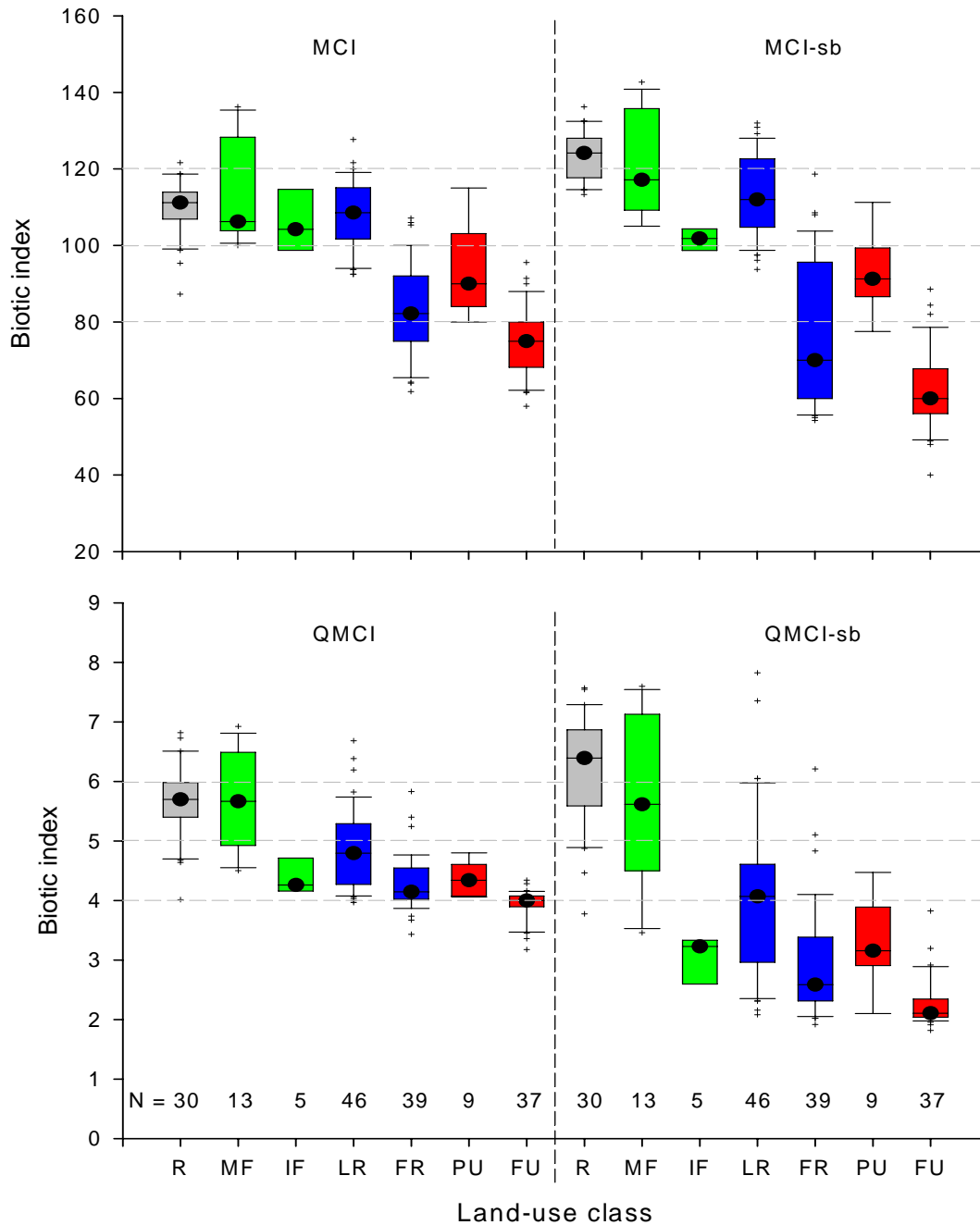
One undesirable feature of the MCI and QMCI when applied to SB streams was the reduced range of index values across the range of disturbance compared to the same indices applied to HB streams (Maxted et al., 2003). Box plots of the MCI, MCI-sb, QMCI, and QMCI-sb values arranged by land-use class illustrate the increased range of the SB versions compared to the original versions (Figure 2). It is also evident that the quality classes developed for HB streams (dashed lines; Stark 1998) may not be directly applicable to SB streams. For example, only one SB sample collected from reference sites achieved an “excellent” MCI score (> 120), while the majority of these samples achieved an “excellent” score using the MCI-sb. Quality classification is discussed further in Section 4.3.

Application of the MCI-sb to streams in Auckland expanded the range of index values by 38% compared with the MCI. The range increased by 68% for the quantitative index (QMCI-sb) (Table 8). There was little difference between mean and median MCI-sb and MCI values. In contrast, mean and median QMCI-sb values were 17% lower than QMCI values, suggesting that the QMCI when applied to SB streams scores a site somewhat better than it really is. The increased range in MCI-sb and QMCI-sb values between the best and worst SB stream sites is an advantage because it permits greater discrimination between sites.

**Table 8** Comparison of selected descriptive statistics for the MCI, MCI-sb, QMCI and QMCI-sb.

	N	Mean	Median	Minimum	Maximum	Range (max. – min.)
<b>MCI</b>	41	95.7	100.8	62	125	63
<b>MCI-sb</b>	41	96.8	101.1	47	134	87
<b>QMCI</b>	41	4.72	4.52	3.18	6.50	3.32
<b>QMCI-sb</b>	41	3.89	3.77	2.01	7.59	5.58

Box plots of MCI, MCI-sb, QMCI, and QMCI-sb values arranged by land-use illustrate the increased range of the MCI-sb and QMCI-sb compared with the original indices (Figure 3). It is evident that the quality classes for interpreting the MCI and QMCI in HB streams do not apply when these indices are applied to data from SB streams. Both the MCI and QMCI, for example, classify most SB reference sites as “good” rather than “excellent”. Quality classification is discussed further in Section 4.3.



**Figure 2** Boxplots of MCI, MCI-sb, QMCI, and QMCI-sb for SB stream samples in seven land-use classes in the Auckland region; box = 25<sup>th</sup> – 75<sup>th</sup> percentile, whiskers = 5<sup>th</sup> – 95<sup>th</sup>, + = outliers.  
 Key to land-uses: R = reference, MF = mature exotic forest, IF = immature exotic forest, LR = lifestyle (partial) rural, FR = full rural, PU = partial urban, FU = full urban. N = number of macroinvertebrate samples. Dashed horizontal lines indicate quality thresholds for interpretation of MCI and QMCI values (see Table 4). Key to land-use classes: ■ reference, ■ forestry, ■ rural, ■ urban.

#### 4.2.2 Performance Test 2: Relationships with environmental variables

The theory and practice of biological monitoring is dependent upon the fact that different macroinvertebrates vary in their pollution tolerances and habitat preferences. In other words, biological monitoring “works” if (and only if) macroinvertebrate communities are a product of their environment. If biotic indices reflect the composition of macroinvertebrate communities in some realistic manner, then it follows that they should be correlated with the underlying environmental variables that influence community composition. In reality it is not as simple as this because macroinvertebrate communities owe their character to a complex mix of environmental variables, which may interact synergistically or in a variety of different ways. The effects of environmental factors are unlikely to be linear – there may be threshold effects (i.e., no effect up to a point and a significant effect thereafter). There may also be important environmental factors at work that we do not appreciate and/or have no data for.

Despite these difficulties, we can obtain an indication of performance of the MCI-sb and QMCI-sb by undertaking rank and linear correlation analyses. For the new indices are to be an advance over the MCI and QMCI when applied to SB streams then we should expect some increase in the strength of their correlations with environmental variables.

Statistical analyses using linear regression or linear correlation require data to be normally distributed if the results are to be valid. Normality can be assessed visually by inspection of histograms – normally distributed data conform to a bell-shaped distribution. Water quality and hydrologic data often are not normally distributed, but are described as skewed. The skewness indicates the departure from the normal curve. Data are positively skewed if the curve has a longer “right tail” and negatively skewed if the “left tail” is longer. Most event-related water quality data are positively skewed reflecting many relatively low values and a few very high values (Grabow et al. 1998). Faecal coliforms and nutrient concentrations (e.g. nitrate) in streams are classic examples. Skewness is calculated easily in a spreadsheet or as an option under descriptive statistics in a statistics package. Data that are normally distributed have skewness close to zero. As a rule of thumb, Grabow et al. (1998) suggested that an absolute value of skewness  $> 1.0$  indicates a degree of skewness (i.e., non-normality) in the data that should be addressed by applying a transformation. Grabow et al. (1998) indicated that the most commonly used transformation for water quality and hydrologic data is the logarithmic transformation. If data contain zeros, then a  $\log_{10}(x+1)$  transformation must be applied.

Metric performance was assessed by comparing Spearman rank correlations ( $R_s$ ) and Pearson linear correlations ( $r^2$ ) between environmental variables and the MCI and QMCI with similar correlations involving the MCI-sb and QMCI-sb (Appendix 9). All correlations were run using STATISTICA 6.1 with statistical significance at the  $p < 0.05$  level. All Pearson linear correlations were run first on untransformed data and then on  $\log_{10}(x+1)$  transformed data. To be considered a valid significant linear correlation we used the rule of thumb that the absolute value of skewness should be  $< 1.0$  (approximately).

Most reliance was placed on the rank correlation analyses because biotic indices are most often used to determine whether the “ecological health” of one site is better or worse than that of another, or to assign sites to quality classes. These are effectively ranking exercises. Secondly, rank correlation analysis is not affected by skewed or non-normal data, whereas linear regressions require non-normal data to be transformed prior to analysis if robust results are to be obtained. Many water quality variables (such as nutrient concentrations or faecal coliforms) have highly skewed log-normal distributions.

Linear regressions or correlations would be required in order to derive predictive linear relationships such as the response of MCI-sb as the level of catchment development increases. In such cases, data transformation would be required if data were non-normally distributed.

Appendix 9 presents the detailed results of rank and linear correlation analyses between the MCI, QMCI, MCI-sb, and QMCI-sb and selected catchment, land-use, habitat quality and water quality variables. Table 9 summarises these correlations for key environmental variables (including those variables that comprise the HQI and WQI-sb).

**Table 9** Spearman rank correlation (Rs) and Pearson linear correlation ( $r^2$ ) values between biological metrics and selected environmental variables. All correlations are statistically significant ( $P < 0.05$ ). NS = non-significant ( $P > 0.05$ ). Shaded cells denote the index with the highest correlation with each environmental variable. \* denotes  $\log_{10}(x+1)$  transformed variable used for linear correlations. Bold values are higher within metric pairs (e.g., MCI and MCI-sb) for each variable.

	N	Spearman Rank Correlation (Rs)				Pearson Linear Correlation ( $r^2$ )			
		MCI	MCI-sb	QMCI	QMCI-sb	MCI	MCI-sb	QMCI	QMCI-sb
<b>Land-use</b>									
NAT	41	0.476	<b>0.586</b>	0.578	<b>0.584</b>	0.289	<b>0.420</b>	0.460	<b>0.487</b>
RUR	41	-0.471	<b>-0.475</b>	<b>-0.474</b>	-0.453	0.169	<b>0.183</b>	0.227	<b>0.257</b>
URB*	41	<b>-0.627</b>	-0.625	<b>-0.616</b>	-0.541	0.409	<b>0.450</b>	<b>0.363</b>	0.278
DEVPER	41	-0.818	<b>-0.834</b>	<b>-0.844</b>	-0.808	0.640	<b>0.689</b>	<b>0.675</b>	0.632
<b>Habitat quality</b>									
AHA	41	NS	<b>0.327</b>	NS	NS	NS	NS	NS	NS
AHD	41	<b>0.548</b>	<b>0.548</b>	<b>0.522</b>	0.451	0.225	<b>0.231</b>	NS	NS
HH	41	0.559	<b>0.615</b>	0.497	<b>0.527</b>	0.292	<b>0.347</b>	0.238	<b>0.263</b>
CA	41	0.721	<b>0.737</b>	0.620	<b>0.637</b>	0.577	<b>0.636</b>	0.298	<b>0.330</b>
BS	41	NS	NS	NS	NS	NS	NS	NS	NS
CS	41	0.486	<b>0.568</b>	0.436	<b>0.472</b>	0.289	<b>0.354</b>	0.205	<b>0.240</b>
RW	41	0.489	<b>0.554</b>	0.462	<b>0.496</b>	0.304	<b>0.363</b>	0.259	<b>0.302</b>
HQI	41	0.557	<b>0.632</b>	0.508	<b>0.525</b>	0.328	<b>0.410</b>	0.230	<b>0.268</b>
<b>Water quality</b>									
DOSAT	14	<b>0.546</b>	NS	<b>0.570</b>	NS	NS	NS	NS	NS
FC*	14	-0.688	<b>-0.750</b>	<b>-0.659</b>	-0.589	0.476	<b>0.582</b>	0.309	<b>0.314</b>
NH4	14	NS	NS	<b>-0.578</b>	NS	NS	NS	NS	NS
NO3*	14	-0.684	<b>-0.754</b>	-0.666	<b>-0.793</b>	NS	NS	NS	NS
TP*	14	-0.575	<b>-0.588</b>	<b>-0.603</b>	NS	<b>0.343</b>	0.342	<b>0.292</b>	NS
SS	14	NS	NS	NS	NS	NS	NS	NS	NS
WQI-sb	14	0.735	<b>0.741</b>	<b>0.785</b>	0.583	0.472	<b>0.542</b>	<b>0.409</b>	0.345

The correlations in Table 9 enable the performance of the various biotic indices to be compared in several ways. Of paramount importance is whether or not the MCI-sb and QMCI-sb perform better when applied to SB streams than the original MCI and QMCI. If this was not the case, then the newly developed indices would be of highly dubious value.

The HQI and WQI-sb were developed independently from the biotic indices and included seven and six components, respectively, that we expected to show some relationships to biological communities. However, several of these components showed little or no correlation with the various biotic indices (Table 9). BS (Bank Stability) – a component of the HQI - and SS (Suspended Solids) – a component of the WQI-sb - were not correlated significantly with any of the biotic indices. AHA (Aquatic Habitat Abundance) was rank-correlated only with MCI-sb,



and NH<sub>4</sub> (ammoniacal nitrogen) was rank-correlated significantly only with the QMCI. Better performing habitat quality and water quality indices may be possible by revising the component variables that comprise them (i.e., by omitting those variables that did not have significant correlations with biotic indices). Generally, the HQI and WQI variables correlated better with biotic indices than the component measures that made up the HQI and WQI.

The MCI-sb had stronger rank correlations than the MCI for 14 of 16 environmental variables and was a very close second for the land-use variable URB\*. The only environmental variable where the MCI had a superior rank-correlation to the MCI-sb was DOSAT. The MCI-sb had stronger linear correlations than the MCI for 12 of 13 environmental variables in Table 9. The exception was TP\*, where the difference in  $r^2$  values was insignificant (0.001). These results suggest that the MCI-sb does perform better than the MCI when applied to macroinvertebrate data from SB streams.

The situation with the QMCI and QMCI-sb was less clear-cut. The QMCI-sb had stronger linear correlations with eight of 12 environmental variables, and stronger rank correlations with only seven of 16 variables (Table 9). These results suggest that, on balance, the QMCI-sb performed slightly better than the QMCI when applied to macroinvertebrate data from SB streams, but that the QMCI was still a well-performed index.

The second issue of importance is whether the MCI-sb performs better than the QMCI-sb. Given the costs of quantitative (or coded-abundance) sampling and sample processing compared with collection of presence-absence data required for the MCI/MCI-sb, the answer to this question could have considerable implications for the costs of ongoing monitoring.

In terms of linear correlations with environmental variables the MCI-sb was better than the QMCI-sb for 10 of 12 comparisons. The two cases where the QMCI-sb performed better were the correlations with the percentage of native forest (NAT) and rural (RUR) land-use in the catchment. In terms of rank correlations, the MCI-sb performed better than the QMCI-sb for 14 of 15 environmental variables. However, it should be noted that the QMCI had the highest rank correlation for five variables (DEVPER, DOSAT, NH<sub>4</sub>, TP\*, & WQI-sb), and the MCI had the highest rank correlations for two variables (URB\* & AHD). The highest rank correlation of all was between the QMCI-sb and the median nitrate concentration (Table 9).

Using data from 88 streams from around New Zealand Quinn & Hickey (1990a) found that the MCI had high rank correlations ( $R_s$  -0.56 - -0.61) with indicators of enrichment (total Kjeldahl nitrogen, A270F, chlorophyll *a*, periphyton ash-free dry weight). The QMCI also was rank correlated significantly with these same environmental variables ( $R_s$  -0.39 - -0.49), although the correlations were not quite as strong. Although we have no data for the indicators of enrichment considered by Quinn & Hickey (1990a), the rank correlations we found with median nitrate concentrations (MCI -0.684, MCI-sb -0.754, QMCI -0.666, QMCI-sb -0.793) are somewhat higher than those determined by Quinn & Hickey (1990a) for the MCI and QMCI. Furthermore, the performance of the MCI-sb and QMCI-sb was even better (Table 9).

### **4.2.3 Performance Test 3: Classification of “severely degraded” sites**

The performance of the various biotic indices was tested using a subset of sites defined as “severely degraded” (SD) using physical and water quality criteria. Non-biological measures such as water quality and physical habitat provide an independent basis to test the ability of the various biotic indices to assigning low scores to these SD sites. Our hypothesis was that the expanded score range provided by the SB indices would improve the ability to discriminate

between quality classes. An accurate index should score these SD sites in the bottom half of the score range.

A total of 15 sites (68 samples) were selected as SD sites (Table 10). As described below, the four criteria used were habitat quality (11 sites), datasonde water quality (4 sites), monthly grab sample water quality (8 sites), and % urban land use (7 sites). Thirteen sites met at least two of these criteria. One site (PUHI) met all criteria.

**Habitat Quality** – Sites with HQI scores < 75 were defined as SD. These sites lacked shade and would likely have wide diel swings in DO, temperature and pH, and exceedences of effect thresholds. These streams often have direct cattle access, show evidence of channelisation, and generally have a limited variety of beneficial aquatic habitats (e.g., pools, submerged wood).

**Datasonde WQ** – Continuous recording datasonde data for temperature and DO were available. A site was defined as SD if the DO < 2.0 mg/l or the temperature > 24 °C for more than 5 consecutive days. Adverse effects on a wide variety of aquatic fauna are likely if these thresholds are exceeded (Richardson et al. 1994; Quinn et al. 1994, Dean & Richardson 1999).

**Table 10** Land use, physical habitat, and water quality criteria used to classify sites as “severely degraded”, including sites with fully urbanised catchments (> 45%URB), ARC habitat quality index (HQI) scores < 75, exceedences of 2.0 mg/l dissolved oxygen (DO) and 24°C temperature from datasonde deployments, and exceedences of 4.0 mg/l DO, 24°C temp (T), and 0.031 mg/l total zinc (Zn) from monthly grab samples over a 5 year period (1998-2003).

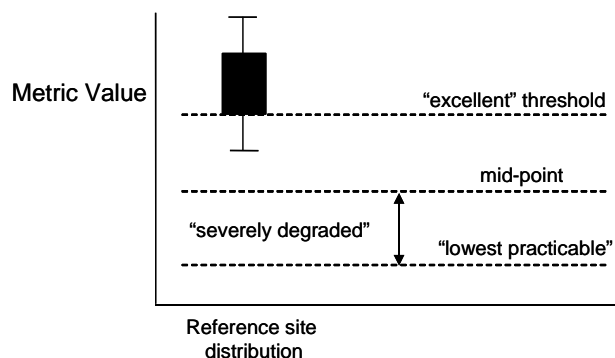
Site code	class	(samples)	(> 45%)	(< 75)	Sonde DO+T	<u>Monthly grab (1998-2003)</u>		
		N	%URB	HQI		DO	Temp	Total Zn
AWRL	URB	8	78	60	yes <sup>1</sup>			
AWRM	URB	9	81		yes <sup>1</sup>			
CHAT	URB	1	73					
OAKL	URB	4	80	71				yes <sup>2</sup>
ONEP	URB	3	64					
OTEH	URB	4	55			yes <sup>2</sup>		yes <sup>3</sup>
PUHI	URB	4	45	46		yes <sup>2</sup>	yes <sup>2</sup>	yes <sup>3</sup>
OTAR	URB	4		57		yes <sup>3</sup>		
NGAK	RUR	4		30	yes <sup>1</sup>			
VAUL	RUR	10		42	yes <sup>1</sup>	yes <sup>2</sup>		
PAPA	RUR	4		44			yes <sup>2</sup>	
HOTE	RUR	1		52			yes <sup>2</sup>	
WAIW	RUR	4		57			yes <sup>2</sup>	
KUME	RUR	4		58				
MAHU	FOR	4		56				
<b>Total</b>	<b>15 sites</b>	<b>68 samples</b>						

<sup>1</sup> = numerous exceedences on multiple days. <sup>2</sup> = low frequency (1-10 exceedences, 1998-2003). <sup>3</sup> = high frequency (10-25 exceedences, 1998-2003)

**Monthly WQ** – Monthly grab samples have been collected since 1991 at 13 SB sites and since 2002 from one SB site (WEST). The number of measurements that exceeded effect criteria for DO (4.0 mg/l), temperature (22 °C), and total Zn (0.031 mg/l) were compiled for each site. Data were compiled for the period 1998-2003 to match the conditions antecedent to the biological sampling period. The DO and temperature effects criteria were less stringent compared to the datasonde data analysis because these single samples likely missed the most extreme time of day. The effect level for total Zn was the 80% protection level published for Australia and New

Zealand (ANZECC 2000). The number of measurements exceeding each criterion were placed into low (1-10) and high (10-25) frequency classes.

Urban land use - Sites with > 45% of the catchment in urban land use were defined as severely degraded. There is extensive literature on the poor water quality conditions in fully urbanised catchments, and while no attempt has been made here to summarise this information, data from the ARC water quality network confirms this conclusion. Fully urban sites have the poorest water quality conditions in the region, and often exceed effects levels for DO, temperature, and total Zn, have poor water clarity, and high faecal coliforms (ARC 2000). Two sites (ONEP, CHAT) were included in the SD group based upon this criterion alone.



**Figure 3** Box plot depicting the approach to setting quality thresholds for interpretation of biotic index values.

The performance of the various biotic indices was assessed by determining the proportion of samples taken from SD sites that were assigned low biotic scores. Their classification as “severely degraded” using physical variables would then be confirmed using the biotic index. “Low scores” were defined as < 50% of the expected range between “excellent” and the “lowest practicable” score for each index. The upper end of the expected range (i.e., “excellent”) was defined as the 25<sup>th</sup> percentile of the reference site distribution (Figure 3). Sites scoring above this value would be defined as “excellent”. The lower end of the expected range was defined as the “lowest practicable” score observed at the most degraded site. We used data from an urban concrete channel (not part of the present study) to define this lower range [unnamed tributary of Wairau Creek (off Goldfield Road, North Shore City)]. The SD class thresholds appear in Table 11.

**Table 11** Ability of the original and proposed SB (**bold**) biotic indices to identify severely degraded (SD) sites correctly. SD defined as < 50% of the range between “excellent” (E) and the “lowest practicable” (LP) scores; E defined as the 25<sup>th</sup> percentile of the samples from SB reference sites. See Table 10 for site names and non-biological criteria used to select SD sites.

	MCI	MCI-sb	QMCI	QMCI-sb
25 <sup>th</sup> percentile of reference samples (n = 32)	105.5	<b>118.5</b>	5.4	<b>6.1</b>
Excellent class threshold (E)	100	<b>120</b>	5.5	<b>6.0</b>
Severely degraded class threshold (50% range; “E” to “LP”)	70	<b>80</b>	3.8	<b>4.0</b>
Lowest practicable score (worst site) (LP)	40	<b>40</b>	2.0	<b>2.0</b>
Number (Percentage) of samples correctly assigned as “SD”	17(25)	<b>55(81)</b>	9(13)	<b>67(99)</b>

A substantially higher proportion of samples from severely degraded sites had SB index scores in the lower half of the expected range compared to the original indices, indicating an improvement in assessment reporting (Table 11). For the presence absence metrics, the MCI-sb correctly assigned 81% of the SD samples compared to 25% for the MCI. There was a wider difference for the quantitative versions: 99% for the QMCI-sb compared to 13% for the QMCI.

#### **4.2.4 Performance Test 4: Discrimination between land-use classes**

The distribution of individual sample data in Figure 2 illustrates the improved separation between land use classes and degrees of disturbance within each class for the MCI-sb and QMCI-sb (cf. MCI & QMCI). In general, MCI-sb and QMCI-sb scores for the reference sites (REF) were higher, and scores for the land use classes were lower. The variability between samples within a class and subclass (i.e., size of the box) was generally higher for the MCI-sb and QMCI-sb indices (cf. MCI & QMCI).

The MCI-sb and QMCI-sb indices showed a greater sensitivity to disturbance compared to the MCI and QMCI (Figure 2). The primary improvement was in the high disturbance classes where the SB versions had substantially lower distributions for immature forestry (IF), full rural (FR), and full urban (FU) subclasses. QMCI-sb index values were also substantially lower for the low disturbance subclasses (lifestyle rural, LR; and partial urban, PU). There was little difference between the MCI and MCI-sb. Maxted (2003) also found that relative abundance data were necessary to detect subtle effects (LR subclass) in rural areas. The lower scores using the MCI-sb caused little change in the quality classifications, but the QMCI-sb changed classifications to “poor”, as compared to “fair” using the QMCI.

All four indices are consistent with Quinn et al.’s (2004) findings that late-rotation forestry (MF sites, Figure 2) support similar stream invertebrate communities to native forest (REF) sites. They also found that logging disturbance can degrade biodiversity values unless buffers of undisturbed riparian forest are retained. We also found that Auckland streams with recent (within 1-year) logging disturbance (IF sites, Figure 2) had decreased biotic index values (cf. MF & IF, Figure 2). However, the MCI-sb and QMCI-sb values suggest that the severity of logging disturbance is greater than the MCI and QMCI values suggest. In this case, we believe that the MCI-sb and QMCI-sb provide the more realistic assessment of ecological condition.

The original indices appear to have underestimated the severity of adverse effects compared to the SB versions in two general areas. First, the QMCI-sb scores were substantially lower (cf. QMCI) for all land use classes and degrees of disturbance except for the mature exotic forest (MF) class. Second, the MCI-sb scores for the high disturbance subclasses were substantially lower (cf. MCI) in rural (FR) and urban (FU) areas.

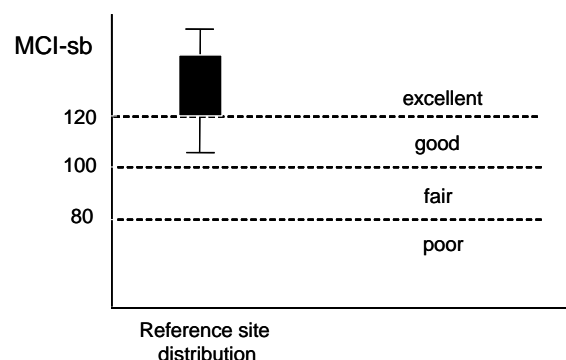
### **4.3 Determining quality thresholds**

Interpretation of biotic index values ideally requires quality thresholds to be developed so that sites can be assigned, based upon index values, to one of several quality classes that convey information about the pollution status or “health” of the stream to water managers and the general public. The interpretation of MCI, SQMCI and QMCI values provided by Stark (1998) was developed from an earlier “subjective assessment” based on knowledge of water quality, catchment land-use and the existence of diffuse or point-source discharges provided by Stark (1985). In this report, we have used an objective procedure based upon the statistical distribution of biotic index values at reference sites to define four quality classes (Figure 3). This procedure requires good characterisation of the reference sites and an estimation of the lowest practicable biotic index value (which can be estimated or taken from the most degraded site in the region).

The criteria used to define SD sites (Section 4.2.3) were used to define quality classes for the MCI-sb and QMCI-sb. The “excellent” quality class was set at the 25<sup>th</sup> percentile of the reference site biotic index distribution. Thus, 75% of all reference samples have higher index values than this threshold and are assigned to the “excellent” quality class. The SD class threshold (midpoint between “excellent” and “lowest practicable”) was set as the threshold between “fair” and “poor”. The range between the “excellent” and SD thresholds was bisected to set the threshold between the “good” and “fair” classes.

The classification results for the MCI and QMCI (and the SB versions) are shown in Table 11. The resulting thresholds for the MCI-sb are illustrated in Figure 4. The 25<sup>th</sup> percentile of the reference site distribution for the MCI-sb was 118.5, and was rounded up to 120. Sites scoring > 120 were assigned to the “excellent” class. The midpoint between this value and the “lowest practicable” score (40) was 80, and defined the “poor” threshold. Sites scoring < 80 were assigned to the “poor” class. Bisecting the range between these two thresholds defined the threshold between “good” and “fair”. Sites scoring between 100 and 120 were defined as “good”, while scores between 80 and 100 were assigned to the “fair” class (Figure 4). Application of a similar procedure to the QMCI-sb provided the following quality classes: excellent > 6.0, good 5.0-6.0, fair 4.0-5.0, and poor < 4.0

Quality classes derived objectively for the MCI-sb and QMCI-sb were identical to those proposed by Stark (1998) for the MCI and QMCI (see Table 4). This is another advantage of the SB versions of the MCI and QMCI. The MCI-sb and QMCI-sb have the same thresholds for data interpretation (when applied to SB streams) as the MCI and QMCI (when applied to HB streams), which avoids confusion.



**Figure 4** Box plot depicting the approach to setting quality thresholds for interpretation of MCI-sb values.

#### 4.4 Precision of index estimates

The precision of index values must be known to permit detection of significant differences if single samples are collected per site or occasion. Detection of statistically significant differences and assignment of sites to quality classes are affected by sample variance. For example, scores just above a threshold may not be significantly different from scores just below the threshold. Replicate samples (n = 3) collected using Protocol C2 (Stark et al. 2001) from SB streams on 35 occasions provided the basis for analysis of variance. We prepared two analyses of variance using the standard deviation and a calculation of “detectable differences”.

Mean and standard deviation values were calculated for each sampling event, and summarized for each of the 4 quality class score ranges (Table 12). The variance was similar for the MCI (+/- 6.34) and MCI-sb (+/- 5.47), representing 5-7% of the index scores. The variance was two times higher for the QMCI-sb (+/- 0.65) compared to the QMCI (+/- 0.33), representing 7-18% of the index scores for the QMCI-sb, and 4-10% for the QMCI. The variances were highest for the middle score classes for all metrics except the MCI which had variances proportional to the score class.

**Table 12** Average standard deviation estimates derived from replicate data (n = 3) for four biotic indices, for the four quality class score ranges. SB versions in bold.

RANGE	N	MCI	N	MCI-SB	RANGE	N	QMCI	N	QMCI-SB
> 120	1	7.03	<b>10</b>	<b>5.63</b>	> 6	2	0.23	<b>5</b>	<b>0.44</b>
100 -120	19	7.02	<b>10</b>	<b>6.13</b>	5 - 6.0	9	0.43	<b>5</b>	<b>0.98</b>
80 -100	7	6.23	<b>5</b>	<b>6.31</b>	4 - 5.0	18	0.46	<b>4</b>	<b>0.83</b>
< 80	8	5.09	<b>10</b>	<b>3.80</b>	< 4	6	0.19	<b>21</b>	<b>0.35</b>
<b>Overall</b>	35	6.34	<b>35</b>	<b>5.47</b>	all	35	0.33	<b>35</b>	<b>0.65</b>

Stark (1998) adapted a procedure involving ANOVA developed by Narf et al. (1984) (for Hilsenhoff's Biotic Index) to calculate "detectable differences" for the MCI, SQMCI, and QMCI based upon replicate samples from HB streams. Narf et al. (1984) derived the following equation:

$$DD = \frac{2.124S}{\sqrt{N}}$$

where  $N$  is the sample size (i.e., the number of replicates: 1 for a single sample), and  $S$  is the best estimate of biotic index standard deviation, and  $DD$  is the "detectable difference". For single hand-net samples ( $N = 1$ ),  $DD$  is the difference in biotic index values required for them to be statistically significant.

Using the above equation Stark (1998) determined that MCI values calculated from single hand-net samples needed to differ by 10.83 MCI units if the difference was likely to be statistically significant. The equivalent value for the SQMCI was 0.83. These detectable differences represent approximately 11% and 17% of average values of the MCI (100) and SQMCI (5.0) respectively.

The best overall estimates of biotic index standard deviation that we have from SB streams are 5.47 for the MCI-sb and 0.65 for the QMCI-sb (Table 12). Substituting these estimates into the equation above suggests that MCI-sb estimates from single samples need to differ by 11.6 units and QMCI-sb by 1.38 units for the differences to be statistically significant. These detectable differences represent approximately 12% and 28% of average values of the MCI-sb (100) and QMCI-sb (5.0) respectively.

These analyses indicate that single samples provide estimates of the MCI-sb with similar precision to those of the MCI. However, the high between-replicate variability of the QMCI-sb raises considerable concern about the ability of this index to provide accurate or consistent assessments of stream health. For example, a QMCI-sb value of 5.5 would assign a site to the "good" quality class, but the high  $DD$  of 1.38 means that once cannot be certain that the site should not be in the "excellent" ( $5.5 + 1.38 = 6.88$ ), or "fair" ( $5.5 - 1.38 = 4.12$ ) classes (Table 4). If the QMCI-sb is to provide reliable assessments, then we believe that replicate sampling

may be required and/or particular care taken to ensure that all samples are as similar in composition as possible. This might require some departure from Protocol C2, by, for example, sampling only one habitat (such as woody debris) if it is in common to all sampling sites. This is likely to be difficult given the nature of SB stream habitats.

Given the nature of SB streams and the multi-habitat sampling protocol, greater between-replicate variability in samples from SB streams and less precise biotic index estimates is not altogether surprising. On the other hand, these estimates for SB streams were based on only 35 samples (from sites where replicate sampling had been undertaken) compared with 827 samples from HB streams (Stark 1998). Additional replicate sample data from SB streams may enable more accurate estimates of precision to be obtained.

## **5. APPLICATION OF THE MCI-sb AND QMCI-sb TO THE ASSESSMENT AND MANAGEMENT OF AUCKLAND STREAMS**

Relationships between biotic indices can be useful in the water management industry and can help address the implications of, for example, changes in land-use on stream communities. The primary objective of indices such as the MCI-sb and QMCI-sb is to summarise complex data on macroinvertebrate communities down to single numbers that convey information about the “health” or pollution status of a waterway. By plotting index values against information on land-use, habitat quality, and water quality we can gain further insight into the factors that may influence stream health. Once causes are identified, then water managers can determine what may be required to implement stream restoration.

Figures 5 – 7 are scatterplots of the MCI-sb and the QMCI-sb against the percentage of developed land in the catchment, ARC’s Habitat Quality Index and a Water Quality Index. Graphs like these, when compared and contrasted, convey a great deal of information to water managers. It is beyond the scope of this report to discuss these graphs in detail, but we consider that it is useful to highlight some important points. If nothing else, the fact that useful insights may be gained from graphs such as these indicates that the indices we have developed are likely to be useful to water managers.