CAWTHRON

A User Guide for the Macroinvertebrate Community Index

Cawthron Report No. 1166

April 2007

A User Guide for the Macroinvertebrate Community Index

John D Stark [1]

John R Maxted [2]

Prepared for the

Ministry for the Environment

[1] Cawthron Institute
98 Halifax Street East, Private Bag 2,
Nelson, New Zealand
Ph. +64 3 548 2319
Fax. + 64 3 546 9464
www.cawthron.org.nz
john.stark@cawthron.org.nz

[2] South Florida Water Management District
PO Box 24680
West Palm Beach, Florida 33416-4680, USA
jmaxted@sfwmd.gov

Reviewed and approved for release by: _____

Rowan Strickland, Coastal & Freshwater Manager

Recommended citation:
Stark JD, Maxted JR 2007. A user guide for the Macroinvertebrate Community Index. Prepared for the Ministry for the Environment. Cawthron Report No.1166. 58 p.

CONTENTS

FIGURES

TABLES

APPENDICES

ACKNOWLEDGEMENTS

This report is based on many years' experience undertaking biomonitoring programmes, both in New Zealand and in the United States. We especially thank Auckland Regional Council for financial and in-kind support for the development of a soft-bottomed MCI for the Auckland region. We thank Mike Thompson from the Ministry for the Environment for organising the Ministry contract that has enabled this report to be prepared, and for his helpful critique on the draft report. The Ministry for the Environment, Chris Fowles (Taranaki Regional Council), Brett Stansfield (Hawke's Bay Regional Council), Trevor James (Tasman District Council), and Kevin Collier (Environment Waikato) provided helpful comments on a draft of this report. Nevertheless, the views expressed in this report are those of the authors and are not necessarily endorsed by the various reviewers.

PART ONE: BACKGROUND TO BIOTIC INDICES

1. INTRODUCTION

This report has been prepared under contract to the Ministry for the Environment. It aims to help users understand the new soft-bottom variant of the Macroinvertebrate Community Index (MCI), and how it and other members of the "family" of MCI biotic indices can best be used to measure the health of New Zealand streams. This is not a formal national guideline or protocol, nor is it a complete description of biotic indices or the design of monitoring programmes. The views expressed are based on the experience of the authors in developing and using these indices. We hope you will find our suggestions useful and that the guidance we provide will stimulate discussion, promote greater use of the macroinvertebrate community[1] in the assessment of hard-bottomed and soft-bottomed streams, and lead to a greater consistency of approach among users of the MCI.

2. WHAT ARE BIOTIC INDICES?

Traditionally, stream quality or "health" assessments were based on analysing water quality and focused on chemical data. The problem was that these measures reflect only the conditions at the moment the sample is taken, and only a defined set of parameters. In contrast, most macroinvertebrates (*e.g.* mayflies, caddisflies, true flies, snails) possess a life cycle of at least a year or more, do not move great distances, and are more or less confined to the area of stream being sampled. The macroinvertebrate community of a stream lives with the stresses and changes that occur in the aquatic environment, whatever their cause, including those that are due to human activities (such as nutrient enrichment from diffuse and point-source discharges) as well as natural events such as floods and droughts. They are ideal candidates for "biotic" (rather than chemical) measures of stream health.

Biological data can be complex and difficult to understand for laypeople, so various "biotic indices" have been developed to make them easier to understand. Biotic indices rely on the fact that biological communities are a product of their environment, in that different kinds of organisms have different habitat preferences and pollution tolerances. So when an organic effluent is discharged into a stream, intolerant organisms reduce in numbers or disappear, while those that can tolerate such stresses increase in number.

This principle is well illustrated in Figure 1, which shows three sites on the Waiongana River in Taranaki that were subjected to enrichment and pollution from both diffuse and point sources (*e.g.* dairy farmland, dairy factory, and a piggery) in 1981. The response of the stream

[1] Community: all of the different kinds of macroinvertebrates living in the same place.

macroinvertebrate community is shown visually by photographs of the macroinvertebrates collected in a hand-net sample from each site. The MCI quantifies the stream condition with a single number.

Upper reaches – intact riparian margin, good shade, very good water quality

High-density invertebrate community dominated by mayflies and caddisflies. MCI = 142.

5 km further downstream in dairy farmland and below a dairy factory discharge

Densities greatly reduced, few mayflies and caddisflies, chironomids dominant, with a few snails. MCI = 103.

Another 7 km further downstream below a piggery discharge, with thick green algal mats covering the river bed

Higher densities of chironomids and few other taxa. MCI = 51.

Figure 1. Three sites on the Waiongana River in Taranaki, sampled in October 1981, showing the macroinvertebrate community and MCI response to increasing enrichment

A single number that characterises the stream community is useful to the specialist biologist as well as to those non-specialists charged with managing stream health. Raw macroinvertebrate data are lists of scientific names and counts or relative abundances, which are meaningless to most people. A biotic index provides a single number that summarises this complexity (albeit with some loss of information), provides a measure of stream health, and can be related statistically to a wide range of physical, chemical, and biological measures. These relationships are fundamental to understanding how ecosystems work and respond to stressors. Although methods that use a number of variables provide one way to manage this complexity, a single index value has been shown to work well both in New Zealand and overseas. This simplicity has allowed the MCI to double as a tool for scientists to characterise complexity, and as a measure of stream health that is easily understood by non-scientists.

To give a formal definition, biotic indices are numerical expressions coded according to the presence of bioindicators differing in their sensitivity to environmental conditions (Graca & Coimbra 1998). They generally are specific to a type of pollution (usually organic enrichment). They involve assigning tolerance values to various types of organisms (or taxa), based on either generally accepted organism sensitivities to pollution and habitat disturbance (BMWP 1978), or on calculations based on the distribution of taxa at a range of stream sites, grouped (or ranked) according to the degree of human impact (Stark 1985; Chessman *et al.* 1997; Chessman 2003; Stark & Maxted 2004, 2007). Because biotic indices incorporate the pollution tolerances of indigenous taxa, they are regionally specific.

Most of New Zealand's freshwater macroinvertebrates are not found in other countries, so we cannot apply any of the biotic indices developed overseas in this country without first deriving tolerance values for local taxa. It is worth noting that tolerance values (Hilsenhoff 1977, 1987, 1988) have been variously referred to as "taxon scores" (Armitage *et al.* 1983; Stark 1985, 1993b, 1998), "quality values" (Chutter 1972), or "sensitivity grade numbers" (Chessman *et al.* 1997; Chessman 2003).

Biotic indices such as New Zealand's MCI (and its variants, see Stark 1985, 1993b, 1998; Stark & Maxted 2004, 2007) can be thought of as indicator species applied at the community level. An indicator species is one that is taken to be a measure of stream health. To a large extent, biotic indices were developed to overcome particular shortcomings of the indicator species approach. We know, for example, that good populations of the spiral-cased caddisfly *Helicopsyche* indicate that a stony stream is in excellent health, and that the mayfly *Zephlebia* is an indicator of a healthy soft-bottomed stream. However, there are healthy stony- and soft-bottomed streams that do not support populations of these taxa. Conversely, red bloodworm midge larvae (*Chironomus*) and tubificid oligochaetes are indicators of grossly enriched conditions, but they are not found in all highly polluted places and are also found occasionally (generally in low numbers) in high-quality environments. Another problem arises when the particular indicator organism is not found. This is not to say it was not present – just that it was not collected in samples – so in this case the indicator organism approach tells us nothing about stream condition.

Macroinvertebrates are found in almost all aquatic habitats, so by assessing entire communities, rather than one or two indicator species, whatever species are present can be used to convey information about the health of their habitats. There is no doubt that well-performing biotic indices could be produced based on a subset of the entire community, but the philosophy and value behind the Macroinvertebrate Community Index, as its name implies, is that the assessment is based on the entire macroinvertebrate community.

3. NEW ZEALAND'S MCI-TYPE BIOTIC INDICES

3.1. Origin and development of the MCI

A preliminary version of the MCI (the IHQI, or Invertebrate Habitat Quality Index) was included in the Taranaki ringplain freshwater biological report (Taranaki Catchment Commission 1984), but it was the Water and Soil Miscellaneous Publication prepared under secondment in 1984 to the Water Quality Centre (Hamilton) (Stark 1985) that proposed New Zealand's Macroinvertebrate Community Index (MCI) and its quantitative variant (QMCI) for assessing organic enrichment in stony riffles[2]. The concept was derived from the United Kingdom's BMWP Score System (BMWP 1978), although genera are mainly used for scoring in New Zealand indices in contrast to families for the BMWP Score System. The MCI is analogous to the ASPT (Average Score Per Taxon) variant of the BMWP Score System (Armitage *et al.* 1983).

Subsequent research funded by the Public Good Science Fund through the Foundation for Research, Science and Technology (FRST) focused on characterising the performance and precision of the MCI and QMCI (Stark 1993b).

Stark (1993b) used macroinvertebrate data from both the North and South Islands to investigate the influences of sampling method, water depth, current velocity, and substratum on the MCI and QMCI. When calculated from macroinvertebrate samples collected by hand-net or Surber sampler from stony riffles, the MCI and QMCI are independent of depth, velocity, and substratum; a major advantage when assessing water pollution or enrichment. The statistical precision of MCI and QMCI values obtained in these ways was defined, along with two methods for detecting statistically significant differences between index values (Stark 1993b).

A more cost-effective variant of the QMCI called the Semi-Quantitative Macroinvertebrate Community Index, or SQMCI was developed in 1998 (Stark 1998). The SQMCI uses a five-point scale of coded abundances (*i.e.* Rare, Common, Abundant, Very Abundant, Very Very Abundant). This index produces values very similar to the QMCI, but at less than 40% of the cost, due to reduced numbers of replicate samples being required to achieve the desired

[2] Riffle: a shallow part of a stream or river with broken water flow.

precision, and savings in macroinvertebrate sample processing time. Stark (1998) also re-evaluated the statistical precision of the MCI and QMCI from hand-net and Surber samples, based on a larger sample database than was previously available. Similar information was provided for the SQMCI.

Recently, Stark & Maxted (2004,[3] 2007) developed new biotic indices for assessing the health of soft-bottomed streams. These indices are analogous to the MCI, SQMCI and QMCI, and are denoted by the addition of "-sb" to the respective index names (*i.e.* MCI-sb, SQMCI-sb and QMCI-sb). New Zealand appears to be the only country with qualitative, semi-quantitative and quantitative versions of the same biotic index, and different versions for hard- and soft-bottomed streams (Stark 1985, 1993b, 1998; Stark & Maxted 2004, 2007).

3.2. Assigning tolerance values to taxa

Most biotic indices require tolerance values to be assigned to macroinvertebrate taxa. These tolerance values are related in some way to stream condition or an environmental gradient; for example, from unmodified native forest (the reference condition), through to highly intensive urban or rural land use. Well-performing biotic indices have been developed using a variety of methods for deriving tolerance values, including:

- Professional judgement (Chutter 1972; Hilsenhoff 1977; Chessman 1995);
- Numerical proportioning applied to taxon occurrences and/or abundances along pollution gradients, or among site groups differing in their pollution status (Stark 1985; Chessman *et al.* 1997; Chessman 2003; Stark & Maxted 2004, 2007);
- Associating taxon occurrences or abundances with water quality data (Lawrence & Harris 1979);
- Canonical Correspondence Analysis (Suren *et al.* 1998; Davy-Bowker *et al.* 2005).

3.2.1. Deriving tolerance values for new biotic indices

For the MCI, tolerance values were determined initially by a weighting procedure based on the relative percentage occurrence of taxa at three site groups differing in their enrichment status (*i.e.* clean and un-enriched, slight to moderate pollution, moderate to gross pollution) (Stark 1985). Tolerance values for less common taxa, for which this procedure was unreliable (Stark 1985), or those added subsequently (Stark 1993b, 1998) have been assigned by professional judgement.

Stark & Maxted (2004, 2007) used an iterative rank correlation procedure developed by Chessman (2003) (hereafter referred to as the "Chessman process") to derive tolerance values

[3] Note that the MCI-sb described by Stark & Maxted (2004) is a preliminary version that is **not** the same as the final version (Stark & Maxted 2007). We have simplified the tolerance value derivation process and have derived tolerance values to the nearest 0.1 (rather than integers) to improve the performance of the MCI-sb and to reduce the possibility of confusion between the HB (integer) and SB (nearest 0.1) tolerance values.

for the MCI-sb using data primarily from the Auckland region, with data from other regions (Northland, Waikato, Bay of Plenty, Hawke's Bay, Taranaki, and Otago) to provide tolerance values for taxa that were not recorded in the Auckland data set. It is worthwhile pausing here to explain this useful procedure for objectively deriving tolerance values.

A prerequisite for using the Chessman process is a macroinvertebrate data set that covers the full range of disturbance, from the best to the worst sites in the region. The resulting tolerance values will be derived in response to the dominant gradient within the data set. Often, in natural systems the gradient is confounded by a variety of variables. In other words, it is due to a complex of interacting environmental factors, which may include enrichment, sedimentation effects (bed sediments tend to become less coarse progressively downstream), altitude, water temperature and other water quality variables, changes to riparian vegetation and condition, and the effects of stream order (velocity, depth). Most biotic indices developed from real-world data sets respond to this complex of factors.

Our implementation of the Chessman process proceeded as follows. First, the sites or samples need to be ordered from best to worst in terms of the environmental gradient of interest. We used MCI values calculated by a user-defined function (or a macro) on an Excel spreadsheet. Spearman rank correlations (r_s) were calculated between the MCI values and the abundances of all taxa across all samples using STATISTICA 7.1.[4] Because it is mathematically impossible for rare taxa to achieve large positive or negative correlations (Chessman 2003), 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 adjusted positive r_s was assigned a tolerance value of 10, and the taxon with the lowest adjusted negative r_s was assigned a tolerance value of 0.1. The remaining taxa were assigned tolerance values (rounded to the nearest 0.1) between these extremes in proportion to their adjusted r_s values. The resulting tolerance values were pasted back into the Excel spreadsheet and new biotic index values were calculated for each sample. This procedure was repeated until the tolerance values stabilised (*i.e.* no tolerance values changed from one iteration to the next), and these became the tolerance values that were adopted for the new biotic index.

The Chessman process entails an apparent circularity, since all samples in the data set are ordered from best to worst using the MCI. Ideally, this would be done independently of the biological data, but if there were an easy way of doing this there would be no need for biotic indices. Chessman (2003) used SIGNAL to determine the initial site order, noting that SIGNAL was a proven indicator of stream health. We used the MCI, because in New Zealand the MCI has shown high correlations with indicators of organic enrichment (*e.g.* Quinn & Hickey 1990), and it performs adequately in soft-bottomed streams (Maxted *et al.* 2003). The final set of tolerance values derived by this process is not overly sensitive to the starting condition if there is a strong environmental gradient in the data set. If there is more than one

[4] Note that Excel cannot normally be used to calculate Spearman rank correlations, not only because it does not have a function to do so, but also because the work-around (involving linear correlation of ranks) using Excel's RANK function provides incorrect results because it does not handle tied ranks correctly. A solution to this problem is presented here: http://udel.edu/~mcdonald/statspearman.html. However, given the number of rank correlations required when using the Chessman process, a spreadsheet-based approach would be tedious in the extreme.

strong gradient in the data, say enrichment and altitude, the algorithm can result in an index of altitude when an index of enrichment was desired (Bruce Chessman, pers. comm.), but the starting point remains unimportant. In our experience with Auckland Regional Council's State of the Environment (SoE) data, and data from soft-bottomed streams from other regions, running the Chessman process on sample data, site-averaged data and various subsets of the data all produced tolerance values that were similar. This suggested that these data embodied a strong environmental gradient, and gave confidence that the process was likely to produce a useful result. Subsequent testing of the new indices (by rank correlations with environmental variables) confirmed that the Chessman process does produce biotic indices that perform well.

3.2.2. Deriving new tolerance values for existing biotic indices

Once a biotic index has been developed, it is inevitable that new taxa (*i.e.* not previously scored) will be encountered. How should tolerance values for these new taxa be derived? There are several options here.

1. *Adopt tolerance values from another biotic index*. For example, MCI tolerance values are likely to be a reasonable substitute if a particular taxon has not yet been assigned a tolerance value for the MCI-sb. It would be better to substitute tolerance values in this way than to exclude unscored taxa from the index calculations. When developing the MCI, Stark (1985) used family scores from the BMWP Score System as a guide for assigning tolerance values when no better information was available.

2. *Professional judgement.* Most of the additional tolerance values for the MCI (*i.e.* those added to the list provided by Stark [1985]) were assigned by the professional judgement of one or more experienced freshwater macroinvertebrate ecologists (Stark 1993b, 1998; Stark *et al.* 2001; Winterbourn *et al.* 2006). These tolerance values are not necessarily unreliable or incorrect, but this process is subjective rather than objective, and has been criticised for that reason (Hickey & Clements 1998; Joy & Death 2003).

3. The Chessman process was designed to assign tolerance values objectively when developing new biotic indices (Chessman 2003), but can be used to derive additional tolerance values for previously unscored taxa. To be practical, any procedure that sets tolerance values needs to be quick and conservative (*i.e.* cause little or no change to existing tolerance values), because it would be very undesirable if the index was re-invented each time a new tolerance value was required.

Of these options, the last is the most objective, but there remains the issue of how best to carry it out. The initial development of the MCI-sb was based on 2000–2004 data (117 taxa x 179 samples) from soft-bottomed streams in Auckland. Auckland Regional Council's 2005 SoE monitoring data set (45 samples) included seven new taxa that did not have MCI-sb tolerance values. We added the seven samples containing the new taxa to create a 124 taxa x 224 sample data matrix and re-ran the Chessman process. We then adopted the seven new tolerance values from this analysis, while retaining existing tolerance values (Stark & Maxted 2004). The fact that 86% of existing tolerance values were unchanged and 99% changed by less than ±1 justified this approach.

For the final version of the MCI-sb, however, Stark & Maxted (2007) adopted a different approach. The entire Auckland soft-bottomed data set (2000–2005, 224 samples) was used to derive tolerance values for 124 taxa using the Chessman process. These tolerance values were calculated to the nearest 0.1 rather than to the nearest integer (*e.g.* the MCI-sb tolerance value for the mayfly *Acanthophlebia* is 9.6, *cf.* 7 for the MCI), because this improved the performance of the resulting indices (*i.e.* it gave higher correlations with environmental variables). Tolerance values for an additional 35 taxa were derived by running the Chessman process on all of the soft-bottomed data available to us – a total of 1,159 samples from Northland, Auckland, Waikato, Bay of Plenty, Hawke's Bay, Taranaki, and Otago. These data contained 35 new taxa. Comparison of the 124 existing tolerance values (derived from the Auckland analyses) with those produced by this analysis showed that 21% were the same, with 57% within ±1, 82% within ±2, and over 93% within ±3 of the Auckland-derived tolerance values. This agreement is good enough, in our view, for us to retain the existing tolerance values and adopt the 35 new ones.[5]

3.3. Calculating the MCI, QMCI and SQMCI

The MCI is calculated from presence-absence data as follows.

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

where S = the total number of taxa in the sample, and a_i is the tolerance value for the ith taxon (see Table 1).

The QMCI is calculated from count data as follows.

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

where S = the total number of taxa in the sample, n_i is the abundance for the ith scoring taxon, a_i is the tolerance value for the ith taxon (see Table 1) and N is the total of the coded abundances for the entire sample.

The SQMCI is calculated in a similar way to the QMCI, except that coded abundances (assigned to the R, C, A, VA and VVA[6] abundance classes) are substituted for actual counts:

[5] Chessman (2003) derived tolerance values (grades) from 24 regional data sets and expressed scores as means with a standard error (SE) provided as a measure of confidence in the averaged national SIGNAL2 grade. Most SEs were less than one unit; the higher SEs (up to 3.2) were usually for rarer taxa. We could not adopt a similar approach because there were insufficient regional data sets available. However, the variability in scores that we encountered based on analyses of various data sets (and combinations of samples) corresponded to SEs between 0 and 2.5, with SEs for over 77% of taxa ≤ 1. Thus, we believe that the approach we used provided tolerance values that should be fairly reliable.

[6] R = Rare; C = Common; A = Abundant; VA = Very Abundant; VVA = Very Very Abundant.

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

where S = the total number of taxa in the sample, n_i is the coded abundance for the ith scoring taxon (*i.e.* R = 1, C = 5, A = 20, VA = 100, VVA = 500), a_i is the tolerance value for the ith taxon (see Table 1), and N is the total of the coded abundances for the entire sample.

Versions of the MCI developed specifically for soft-bottomed (SB) streams are calculated in exactly the same way, except that a different set of taxon tolerance values is used (see column SB in Table 1). Most taxa commonly encountered in soft-bottomed streams have been assigned tolerance values. If a taxon that has not been scored is encountered, the hard-bottomed tolerance value can be used. Alternatively, if data containing the unscored taxa are available, the first author of this report (John Stark) could derive new tolerance values using the Chessman process.

QMCI and SQMCI values range from 0 to 10 and are directly comparable with each other (Stark 1998). MCI values range from 0 to 200 (Stark 1985). Only when no taxa are present are these indices zero. In practice it is rare to find MCI values greater than 150 (or SQMCI and QMCI >7.5) and only extremely enriched stony riffle sites score less than 50 (QMCI and SQMCI <2.5). The soft-bottomed versions are analogous (Stark & Maxted 2004, 2007). The different scales for the indices were chosen deliberately to avoid inappropriate comparisons.

Table 1. Tolerance values for MCI-based biotic indices in hard-bottomed (HB) (Stark *et al.* 2001) and soft-bottomed (SB) (Stark & Maxted 2007) streams

Taxon	HB	SB	Taxon	HB	SB	Taxon	HB	SB
COELENTERATA			**Odonata (continued)**			**Diptera** (continued)		
Hydra	3	1.6*	*Procordulia*	6	3.8*	Sciomyzidae	3	3.0
PLATYHELMINTHES	3	0.9	*Uropetala*	5	0.4	Stratiomyidae	5	4.2
RHABDOCOELA	-	0.9*	*Xanthocnemis*	5	1.2	Syrphidae	1	1.6*
BRYOZOA	-	4.0*	**Hemiptera**			Tabanidae	3	6.8
NEMATODA	3	3.1	*Anisops*	5	2.2	Tanypodinae	5	6.5
NEMATOMORPHA	3	4.3	*Diaprepocoris*	5	4.7*	Tanytarsini	3	4.5
NEMERTEA	3	1.8	*Microvelia*	5	4.6	*Tanytarsus*	3	-
OLIGOCHAETA	1	3.8	Saldidae	5	3.9	Thaumaleidae	9	8.8
POLYCHAETA	-	6.7*	*Sigara*	5	2.4	Tipulidae	5	3.4
HIRUDINEA	3	1.2	**Coleoptera**			*Zelandotipula*	6	3.6
TARDIGRADA	-	4.5*	*Antiporus*	5	3.5	**Trichoptera**		
CRUSTACEA			*Berosus*	5	-	*Alloecentrella*	9	-
Amphipoda	5	5.5	*Copelatus*	5	3.7	*Aoteapsyche*	4	6.0
Cladocera	5	0.7*	Dytiscidae	5	0.4*	*Beraeoptera*	8	7.0*
Copepoda	5	2.4*	Elmidae	6	7.2	*Confluens*	5	7.2*
Halicarcinus	-	5.1*	*Enochrus*	5	2.6	*Conuxia*	8	-
Helice	-	6.6*	Hydraenidae	8	6.7	*Costachorema*	7	7.2*
Isopoda	5	4.5	Hydrophilidae	5	8.0	*Cryptobiosella*	9	-
Mysidae	-	6.4*	*Liodessus*	5	4.9*	*Diplectrona*	9	-
Ostracoda	3	1.9	*Onychohydrus*	5	-	*Ecnomina*	8	9.6
Paracalliope	5	-	*Podaena*	8	-	*Edpercivalia*	9	6.3*
Paraleptamphopus	5	-	Ptilodactylidae	8	7.1	Ecnominidae	8	-
Paranephrops	5	8.4	*Rhantus*	5	1.0	*Helicopsyche*	10	8.6
Paranthura	-	4.9*	Scirtidae	8	6.4	*Hudsonema*	6	6.5
Paratya	5	3.6	Staphylinidae	5	6.2	*Hydrobiosella*	9	7.6*
Tanaidacea	4	6.8*	**Neuroptera**			*Hydrobiosis*	5	6.7
INSECTA			*Kempynus*	5	-	*Hydrochorema*	9	-
Ephemeroptera			**Diptera**			*Kokiria*	9	-
Acanthophlebia	7	9.6	Anthomyiidae	3	6.0	*Neurochorema*	6	6.0
Ameletopsis	10	10.0	*Aphrophila*	5	5.6	*Oecetis*	6	6.8
Arachnocolus	8	8.1	*Austrosimulium*	3	3.9	Oeconesidae	9	6.4
Atalophlebioides	9	4.4*	*Calopsectra*	4	-	*Olinga*	9	7.9
Austroclima	9	6.5	Ceratopogonidae	3	6.2	*Orthopsyche*	9	7.5
Austronella	7	4.7	Chironomidae	2	3.8	*Oxyethira*	2	1.2
Coloburiscus	9	8.1	*Chironomus*	1	3.4	*Paroxyethira*	2	3.7
Deleatidium	8	5.6	*Corynoneura*	2	1.7*	*Philorheithrus*	8	5.3*
Ichthybotus	8	9.2	*Cryptochironomus*	3	-	*Plectrocnemia*	8	6.6*
Isothraulus	8	7.1	*Culex*	3	-	*Polyplectropus*	8	8.1
Mauiulus	5	4.1	Culicidae	3	1.2	*Psilochorema*	8	7.8
Neozephlebia	7	7.6	Diptera indet.	3	2.9	*Pycnocentrella*	9	-
Nesameletus	9	8.6	Dixidae	4	7.1	*Pycnocentria*	7	6.8
Oniscigaster	10	5.1*	Dolichopodidae	3	8.6	*Pycnocentrodes*	5	3.8
Rallidens	9	3.9	Empididae	3	5.4	*Rakiura*	10	-
Siphlaenigma	9	-	Ephydridae	4	1.4*	*Synchorema*	9	-
Tepakia	8	7.6	Eriopterini	9	7.5	*Tiphobiosis*	6	9.3
Zephlebia	7	8.8	*Harrisius*	6	4.7	*Triplectides*	5	5.7
Plecoptera			Hexatomini	5	6.7	*Triplectidina*	5	-
Acroperla	5	5.1	*Limnophora*	3	4.5	*Zelandoptila*	8	7.0
Austroperla	9	8.4	*Limonia*	6	6.3	*Zelolessica*	10	6.5*
Cristaperla	8	-	*Lobodiamesa*	5	7.7	**Lepidoptera**		
Halticoperla	8	-	*Maoridiamesa*	3	4.9	*Hygraula*	4	1.3
Megaleptoperla	9	7.3	*Mischoderus*	4	5.9	**Collembola**	6	5.3
Nesoperla	5	5.7	*Molophilus*	5	6.3	**ACARINA**	5	5.2
Spaniocerca	8	8.8	Muscidae	3	1.6	**ARACHNIDA**		
Spaniocercoides	8	-	*Nannochorista*	7	-	*Dolomedes*	5	6.2
Stenoperla	10	9.1	*Neocurupira*	7	-	**MOLLUSCA**		
Taraperla	7	8.3*	*Neolimnia*	3	5.1	*Gundlachia = Ferrissia*	3	2.4
Zelandobius	5	7.4	*Nothodixa*	4	9.3	*Glyptophysa = Physastra*	5	0.3*
Zelandoperla	10	8.9	Orthocladiinae	2	3.2	*Gyraulus*	3	1.7
Megaloptera			*Parochlus*	8	-	*Hyridella*	3	6.7
Archichauliodes	7	7.3	*Paradixa*	4	8.5	*Latia*	3	6.1
Odonata			*Paralimnophila*	6	7.4	Lymnaeidae	3	1.2
Aeshna	5	1.4*	*Paucispinigera*	6	7.7	*Melanopsis*	3	1.9
Anisoptera	5	6.0	Pelecorhyncidae	9	-	*Physa = Physella*	3	0.1
Antipodochlora	6	6.3	*Peritheates*	7	-	*Potamopyrgus*	4	2.1
Austrolestes	6	0.7	Podonominae	8	6.4*	Sphaeriidae	3	2.9
Hemianax	-	1.1*	*Polypedilum*	3	8.0			
Hemicordulia	5	0.4	Psychodidae	1	6.1			
Ischnura	-	3.1*	*Scatella*	7	-			

Notes: '-' indicates tolerance value not yet assigned. All tolerance values were derived from Auckland data only, except for those marked '*', where data from other regions were used.

3.4. Interpreting the MCI

The interpretation of index values when applied to stony (MCI, SQMCI, QMCI) or soft-bottomed (MCI-sb, SQMCI-sb, QMCI-sb) streams throughout New Zealand is given in Table 2. The quality thresholds are the same for hard-bottomed and soft-bottomed streams, making the new indices easy to implement. These thresholds do not work well when the hard-bottomed indices are applied to soft-bottomed streams, however. For example, soft-bottomed reference sites (which should be high quality) had MCI scores <119, indicating possible mild pollution (Maxted *et al.* 2003; Stark & Maxted 2004, 2007). This provided the motivation for developing a separate set of tolerance values for taxa found in soft-bottomed streams.

Although Stark (1998) provided interpretive descriptions based on enrichment or pollution, we now prefer to use the quality classification used by Stark & Maxted (2004, 2007) (see Table 2). This recognises that the MCI (and its variants) respond to an interacting complex of environmental variables including (but not limited to) enrichment.

Table 2. Interpretation of MCI-type biotic indices

Stark & Maxted (2004, 2007) quality class	Stark (1998) descriptions	MCI MCI-sb	SQMCI & QMCI SQMCI-sb & QMCI-sb
Excellent	Clean water	> 119	> 5.99
Good	Doubtful quality or possible mild pollution	100–119	5.00–5.90
Fair	Probable moderate pollution	80–99	4.00–4.99
Poor	Probable severe pollution	< 80	< 4.00

The index values corresponding to divisions between the four quality classes were selected initially by Stark (1993b) based on professional judgement. However, Stark & Maxted (2004, 2007) used an objective procedure based on the statistical distribution of biotic index values at references sites, together with an estimation of the lowest practical index value, to determine divisions between quality classes. A similar procedure had been used previously in the United States by Maxted *et al.* (2000). In brief, the "excellent" quality class was set at the 25th percentile of the reference site biotic index distribution. This means that 75% of all reference samples have higher index values than this threshold and are assigned to the "excellent" quality class. The midpoint between "excellent" and the "lowest practicable" index value was set as the threshold between "fair" and "poor". The range between the "excellent" and "fair/poor" thresholds was then bisected to set the threshold between the "good" and "fair" classes. When this procedure was applied to the MCI-sb scores for Auckland soft-bottomed streams, it resulted in the same thresholds that Stark (1998) had provided for the MCI.[7]

We believe that you should be flexible when interpreting the divisions between quality classes (see Table 2), and that it is best to regard the boundaries between them as fuzzy. This concept is not new: see Figures 1 to 3 in Stark (1985), where it was suggested that the divisions

[7] The "worst site" (an unnamed tributary of Wairau Creek, off Goldfield Road, North Shore City) had an MCI-sb of 40, and the 25th percentile of the reference site biotic distribution was 126.1 (which was rounded to 120).

between three site groups (which were, in effect, pollution classes) should be 120 ±5 units, 100 ±5 MCI units, and so on.

The same suggestion was made by Wright-Stow & Winterbourn (2003) following their examination of the correspondence between the MCI and QMCI using fixed-count data from 230 stream and river sites in Canterbury. The two indices ranked sites similarly (r_s = 0.86), but the MCI placed most sites in the "good" and "fair" pollution classes, whereas most sites were assigned to the "excellent" and "poor" classes by the QMCI. Wright-Stow & Winterbourn concluded that either the MCI was a more conservative index, or that the boundaries between pollution classes were not equivalent. The latter reason was considered more likely, and given the difficulties inherent in defining classes based on continuous distributions and the fact that there is no way of knowing which index gives the "right" answer, Wright-Stow & Winterbourn suggested a return to fuzzy boundaries between classes (MCI: Excellent 125–200, Good 105–115, Fair 85–95, Poor <75; QMCI: Excellent 6.2–10, Good 5.2–5.7, Fair 4.2–4.7, Poor <3.7).

Alternatively, when comparing large numbers of sites, as in SoE monitoring, Wright-Stow & Winterbourn (2003) suggested that the percentile within which the site of interest falls could be stated. A site with an MCI of 130, for example, could be described as being within the top 10% of sites in the region.

Fuzzy boundaries are desirable because there is always error when estimating biotic indices. Stark (1998) has shown that the MCI from a single hand-net sample has a precision of approximately ±10%. For example, an MCI of 117 taken at face value would assign a site to the "good" quality class, but given the ±10% error inherent in the index estimate, the true MCI could have been anywhere from 105 to 129. The balance of probability would still place that site in the "good" class, but it could possibly be classified as "excellent". We quantified fuzzy boundaries for soft-bottomed streams and found the error twice as high for the QMCI-sb (±12–22%) compared to the MCI-sb (±4–9%) (Stark & Maxted 2004, 2007). This error in QMCI-sb estimates is a major reason why the MCI-sb is recommended for assessing soft-bottomed streams.

In such cases, how should you decide which quality class to assign the site to? Consider, for example, SoE reporting that is based on coloured dots on maps – green dots denote "excellent" stream condition, yellow "good", orange "fair", and red "poor". If a site has an MCI of 126 in year one, 119 in year two and 124 in year three – values that are unlikely to be significantly different – the site would be regarded as "excellent" in years one and three but only "good" in year two if these values were interpreted strictly in terms of the guidelines in Table 2. If there was no reason why there should have been a decrease in stream health in year two, then we believe that the site could remain classified as "excellent". Alternatively, it could be described as "good-excellent" with a symbol that was 50% green and 50% yellow.

Thus, for borderline biotic index values (*i.e.* threshold ±5 MCI units or 1 SQMCI or QMCI unit), we suggest that the ecologist should be able to choose the more appropriate pollution

class to assign the site to, based on other information (such as knowledge of water quality, catchment land use, or the existence of point or diffuse sources of enrichment). A borderline site alternating between two quality classes from year to year is undesirable when annual SoE reports are prepared because it is more likely to reflect sampling error (combined with the quality class threshold effect) than indicate any real change in stream condition.

This is not an issue with more sophisticated analyses of biotic indices (such as time series analyses) because the assignment of sites to pollution classes based on single estimates of index values is not required.

3.5. Strengths and weaknesses of the MCI and variants

The MCI, QMCI, and SQMCI were developed to assess organic enrichment in stony streams by sampling in stony riffles, where the greatest variety of the most sensitive macroinvertebrates may be expected (Stark 1985, 1998). The MCI-sb, SQMCI-sb, and QMCI-sb have been developed for assessing the condition of soft-bottomed streams (Stark & Maxted 2004, 2007). These indices are designed to be used with samples collected according to the national protocols (Stark *et al.* 2001).

The MCI and MCI-sb respond to any perturbation that alters the list of taxa (*i.e.* taxonomic composition) present at a site. The QMCI and SQMCI, and their soft-bottomed variants, respond to changes in taxonomic and numerical composition or relative abundances.

Because the MCI reflects changes in community taxonomic composition, not numerical composition, it is less sensitive to subtle changes in community composition than the QMCI or SQMCI. Sometimes this is an advantage (*e.g.* for SoE monitoring, see Section 3.2, Part 2), but for compliance monitoring, where subtle changes in community composition need to be assessed, the QMCI (or SQMCI) would be more appropriate. Overall, it is important to use a version of the MCI that meets the aims of the investigation.

High MCI values can be derived from taxonomically poor communities (*e.g.* situations where a few individual mayflies and stoneflies are found). High MCI values when taxa richness is low (say <5 taxa per sample) may be an indication of impairment and should be interpreted with caution.

Note that the MCI-type biotic indices have been developed to assess nutrient enrichment/sedimentation in stony- and soft-bottomed streams: they have not been evaluated for other habitats (*e.g.* lakes, ponds, wetlands, large non-wadeable rivers, hot springs), or for other types of disturbance (*e.g.* toxic discharges, flow variation). It is possible that the MCI (or similar indices) might work in these other situations, but they should be used and interpreted with extreme caution. For example, Maxted *et al.* (2003) found that the MCI and SQMCI performed acceptably in soft-bottomed streams, but that the interpretation differed (*e.g.* MCI >100 indicated a soft-bottomed stream in reference condition, *cf.* >120 in a stony stream). However, the MCI-sb, which was developed specifically for soft-bottomed streams,

performs much better with an interpretation that is consistent with the stony-stream MCI, an expanded range (permitting greater discrimination in stream health both between and within land-use classes), and higher correlations with environmental factors and land uses that are known to affect stream communities (Stark & Maxted 2007).

The main criticism of indices (including biotic indices) is the inevitable loss of information compared with the raw data from which index values are derived. Such criticisms generally are made by biologists who can make sense of raw data, but who may not always appreciate the needs of water managers who cannot. We maintain that it is better to convey 40% of the information so that all of it is understood, than all of the information in a form in which only 10% is understood.

Unfortunately, the attractiveness of biotic indices to water managers can lead to their misuse and misinterpretation. For example, if the objective is to assess between-site differences in water quality, and one site is a stony riffle while the other is silty or sandy, then the difference in MCI will not be entirely due to the quality of the water. Invariably, biotic indices respond to a complex of factors (primarily water quality, substrate, and disturbance), so interpretation can be difficult and should be made only by those with suitable training and experience.

Biotic indices should not be the sole means of analysing or depicting biological data if a comprehensive assessment is required. In addition, we recommend looking at EPT richness (either the number or percentage of taxa richness comprising mayflies, stoneflies and caddisflies[8]) and macroinvertebrate densities (if available), and discussing or tabulating the dominant (say top five) taxa at monitoring sites. Total taxa richness is not, in our view, a particularly useful indicator of stream health. Multivariate analyses (clustering and/or ordination) may also be useful to "let the data tell their own story".

The major problem with the use of biotic indices is establishing that they actually measure features of the environment that are of interest, and that they reflect environmental change in some ecologically meaningful way (hopefully linearly) (Norris & Norris 1995). One measure of the performance of a biotic index is to determine whether interpretations based on indices are consistent with those produced by other methods. For example, Stark (1985) found that the MCI produced interpretations consistent with those based on the more traditional quantitative and descriptive analyses used prior to its introduction, by workers such as Hirsch (1958), Winterbourn *et al.* (1971), Winterbourn & Stark (1978), and Marshall & Winterbourn (1979). The MCI (and variants), however, do also have a track record of proven performance in the scientific literature (see Appendix 2).

The tolerance values for most biotic indices usually assume a particular type of pollution (frequently organic enrichment), a particular habitat type (*e.g.* stony riffles), and a particular geographic area. Be careful when using indices to assess different kinds of pollution (*e.g.* sedimentation, inorganic chemicals, metal toxicity), or in different habitats (*e.g.* weed beds, swamps, lakes and estuaries). Applying biotic indices to different regions will require the

[8] EPT = Ephemeroptera, Plecoptera and Trichoptera.

derivation of tolerance values for the taxa encountered, but as yet there is no convincing evidence to suggest that different indices would be required for different parts of New Zealand (although it is almost certain that better-performing indices for use in different eco-regions could be derived from suitable regional data sets).

A strong correlation of biotic indices with chemical pollution measures may seem desirable. Indeed, some workers have suggested that "subjective tolerance estimates" should be replaced by "quantitative tolerance determinations" by extensively examining the correlations between species presence and water quality (Herricks & Cairns 1982). Others have questioned the validity of this approach, however. Washington (1984) noted that biotic indices have been developed primarily to assess organic pollution, so they should have high correlations with biological oxygen demand (BOD), dissolved oxygen (DO) and total organic carbon (TOC), but not necessarily with other chemical parameters. In fact, there is no *a priori* reason why a biotic index should correlate only, or primarily, with chemical data, because chemical changes are not mirrored uniformly by biological organisms or communities (Washington 1984).

Benthic macroinvertebrate communities respond to changes in water quality and bed sediments (Katoh 1992). When using biotic indices to assess water quality, it is essential to minimise or eliminate between-site variation in other factors (particularly substrata) otherwise it will be difficult to determine the causes of any changes in biotic indices. Artificial substrates can be used for this purpose (De Pauw *et al.* 1986). Often, however, it is the overall quality of the habitat (*i.e.* both the substrate and water quality) that is of interest, and biotic indices are suitable for this purpose.

3.6. Alternative or complementary approaches

3.6.1. Predictive models

Although this report is about the use of the MCI, there are alternative or complementary ways of undertaking biological assessments. Indeed Stark (1985) concluded by noting:

> *Finally, I must stress that a biotic index (such as the MCI) must not become the be-all-and-end-all of biological monitoring programmes. A biotic index can be a useful management tool but if progress is to be made, especially in the understanding of habitat requirements and tolerances of macroinvertebrate species, then it is essential that detailed quantitative and taxonomic studies continue to be undertaken whenever possible.*

The MCI has stood the test of time and has been the most often used measure of stream health in New Zealand since it was developed in the mid-1980s (see Appendix 2). However, it has not been without its critics. The Ministry for the Environment (1997), for example, in a discussion document outlining the proposed Environmental Performance Indicators Programme, noted:

> *In New Zealand scientists have developed the Macroinvertebrate Community Index (MCI), but this was developed explicitly to assess nutrient enrichment for Taranaki streams. At the time it was developed the MCI was considered "state of the art", but techniques overseas have now moved well beyond the MCI.*

This statement is somewhat misleading, because although the MCI was developed initially using a Taranaki ringplain macroinvertebrate data set, it was tested on data from Manawatu, Canterbury and Southland, leading Stark (1985) to conclude that it showed potential for application throughout New Zealand – an assertion that was validated subsequently by Quinn & Hickey (1990) (see Appendix 2). Furthermore, biotic indices are far from obsolete, are widely used around the world, and are still being developed, not only for freshwater (*e.g.* Artemiadou & Lazaridou 2005; Davy-Bowker *et al.* 2005; Jiang 2006), but also for marine ecosystems (*e.g.* Borja & Muxika 2005).

AUSRIVAS was trialled in the Waikato region of New Zealand (Coysh & Norris 1999), but it has not been adopted nationwide. Joy & Death (2003) are strong advocates of undertaking biological assessments using predictive models derived from the British RIVPACS (Clarke *et al.* 2003) or Australian AUSRIVAS (Davies 1997). In our view, predictive models and biotic indices are complementary, and both may be part of stream health assessment programmes. In addition, multivariate data analyses using canonical correspondence analysis or non-metric multi-dimensional scaling can be used to analyse raw macroinvertebrate data. These methods can provide further insight into the data and the summary measures (*e.g.* MCI) derived from them.

3.6.2. Multi-metric indices

A multi-metric index comprises several metrics that incorporate biological components that are sensitive to a broad range of human activities (*e.g.* sedimentation, organic enrichment, toxic chemicals, or flow alteration). The QMCI or MCI is included as one of seven metrics in NIWA's adaptation of the Index of Biotic Integrity (IBI) of the US EPA Rapid Bioassessment Protocol III (Plafkin *et al.* 1989) (*e.g.* Quinn *et al.* 1997a).

PART TWO: GUIDELINES FOR USING THE MCI, QMCI AND SQMCI

1. APPLYING MCI INDICES IN DIFFERENT FRESHWATER ENVIRONMENTS

1.1. Hard- and soft-bottomed streams

Traditionally, freshwater ecologists have favoured wadeable, hard-bottomed or stony streams for biological monitoring programmes. Such streams often are more visually appealing and support communities dominated by mayflies, stoneflies and caddisflies, which are not only more sensitive to pollution but also are more exciting or attractive to many ecologists than the snails, worms and chironomids that dominate soft-bottomed stream habitats. Furthermore, sampling macroinvertebrates from stony streams is easier, with well-known and well-proven sampling methodologies (even before the publication of standard methods – Stark *et al.* 2001).

For these reasons, soft-bottomed streams have been a neglected habitat, despite (or perhaps because of) their proximity to centres of population and their consequential pollution. The macroinvertebrates that inhabit soft-bottomed streams generally are more tolerant of enrichment and (especially) sedimentation effects, and so are less sensitive indicators for monitoring disturbance. This could also explain why ecologists have avoided undertaking biomonitoring programmes in soft-bottomed streams.

Lack of standard methods is no longer a reason to ignore biomonitoring in soft-bottomed streams. Stark *et al.* (2001) have provided standard sampling, sample processing, and quality control procedures for macroinvertebrate communities in hard- and (for the first time in New Zealand) soft-bottomed streams, and Stark & Maxted (2004, 2007) have developed versions of the MCI specifically for soft-bottomed streams. Auckland's SoE monitoring network is dominated by soft-bottomed streams, with 45 of 62 sites sampled in 2005, and there are likely to be soft-bottomed streams in most, if not all, other regions of New Zealand.

Clearly, a soft-bottomed stream is not simply a consequence of underlying geology, but will also depend on stream slope, land use and other factors. Interrogation of the River Environment Classification for the various factors that might determine the nature of stream substrates could enable the prevalence of soft-bottomed streams to be estimated, but this is beyond the scope of this review. The bottom line is that soft-bottomed streams are likely to comprise a significant proportion (perhaps 20–40%) of New Zealand's streams and rivers, so the continuing development of methods for their bioassessment is worthwhile. Ultimately,

whether or not you have a soft-bottomed or a hard-bottomed stream is a decision that requires local knowledge, and is best made when standing on the stream bank!

Hard- or soft-bottomed MCI?

Now that there are hard- and soft-bottomed versions of the MCI available, which indices should you use? In most cases, the MCI tolerance values should be used on samples collected using the hard-bottomed sampling protocols (C1 or C3: Stark *et al.* 2001) and the MCI-sb tolerance values used on samples collected using the soft-bottomed protocols (C2 or C4: Stark *et al.* 2001). However, there may be exceptions due to the aims of the investigation. For example, if the stream of interest is a hard-bottomed stream inundated with fine sediment and there are no riffles to sample using the hard-bottomed protocols, then the soft-bottomed sampling protocol C2 would minimise the filling of the net with fine sediment, which would otherwise cause a processing nightmare. If the objective is to assess the degree of disturbance relative to its potential as a hard-bottomed stream, then the data collected using the soft-bottomed protocol might be more accurately assessed using the hard-bottomed tolerance values for MCI or QMCI calculations. The key is to consider the project objectives when selecting the index to use, and to recognise that rules should not take the place of common sense.

1.2. Wadeable versus non-wadeable waterways

All variants of the MCI have been developed using data from wadeable streams. Data are limited, due to sampling difficulties, so there has been no formal evaluation of the performance of these indices for large, non-wadeable rivers. There is no reason why biotic indices such as the MCI cannot be applied to, or developed for, non-wadeable rivers. The major difficulty is obtaining representative samples and then calibrating the interpretation of the index values.

An analysis of River Environment Classification data indicates that nearly 89% of New Zealand's mapped streams are 1st to 3rd order. Most of these are likely to be wadeable. Higher-order (*i.e.* 4–8) streams and rivers are not necessarily unable to be sampled using methods developed for wadeable streams (Stark *et al.* 2001). Large braided rivers in Canterbury, for example, such as the lower Waitaki, have smaller braids or shallow margins along major braids that are accessible. Other large rivers, such as the lower Waikato River, may be of similar stream order but certainly are not wadeable.

1.3. Other freshwater habitats

Although macroinvertebrate biotic indices have not been developed in New Zealand for assessing wetland or lake health, given suitable data sets there is no reason why they could not be developed, just as they have been for wetlands in Western Australia (Chessman *et al.* 2002) or lakes in France (Verneaux *et al.* 2004).

Use of the MCI for other freshwater habitats

The existing versions of the MCI and MCI-sb should not be used to assess the environmental health of wetlands or lakes because they have not been calibrated or evaluated for these habitats. This may seem like an unnecessary caution given that the existing indices were developed for stony- and soft-bottomed streams, but we have seen the MCI used to assess the health of lake margins, which is not recommended. There is, of course, no reason why macroinvertebrate biotic indices could not be developed for other freshwater habitat types.

1.4. Incorporating the MCI-sb into existing biotic monitoring programmes

The MCI-sb is calculated in exactly the same way as the hard-bottomed MCI except for the different list of tolerance values that are used (see Table 1). Consequently, the MCI-sb can easily be integrated into existing monitoring programmes once it has been determined that the sites in question are soft-bottomed sites. It is a simple matter to recalculate MCI-sb values for existing data. Use of the MCI-sb would have no effect on the integrity of existing time series data, although if trends testing has already been undertaken these would need to be re-calculated using the MCI-sb values. The quality thresholds (*e.g.* excellent, good, fair, poor) developed for the hard-bottomed indices (Stark 1998) were found to be applicable to the soft-bottomed indices (Stark & Maxted 2007), making it easy to incorporate the new index scores into existing monitoring programmes.

2. TAXONOMY

The top priority when processing samples to enable the calculation of biotic indices is to have good taxonomy. After all, the tolerance values that reflect environmental health vary among taxa, so if taxa identifications are not correct, then you may end up with an incorrect assessment. Stark *et al.* (2001) have provided quality control (QC) methods that can be used to ensure samples are processed to a high standard.

Identifying aquatic macroinvertebrates is not easy despite the existence of good keys such as Winterbourn *et al.* (2006), and there are plenty of traps for the inexperienced. The increasing adoption of QC procedures in sample processing since they were provided by Stark *et al.* (2001) has highlighted that accurate taxonomy in sample processing cannot always be taken for granted. There has also been some confusion over exactly how best to undertake the sample processing QC required by the protocols, so we provide further clarification here.

2.1. Taxa identification and data quality control

It is important to remember that the overall objective of QC is to ensure data quality, and because QC is an overhead cost we believe that this objective should be achieved as cost-effectively as possible. This entails reconciling any differences between the identifications made by the original sample-processing laboratory and the laboratory chosen to undertake QC (which should be separate agencies).

We recommend the following procedure.

1. The processing laboratory should provide the client with a spreadsheet containing the data, at least one vial for each sample containing representatives of all taxa that have been identified from the sample, and the sample material re-potted and preserved in the original sample containers.
2. The client will then choose at random 10% of the samples to be subjected to QC by another laboratory.
3. The QC laboratory should be provided with the spreadsheet of all data, and the vials and sample residue for these 10% of samples.
4. The second laboratory will then work through the vials and the sample residue according to the procedures described by Stark *et al.* (2001), aiming to check the identifications, find any taxa whose identifications are disputed, detect any taxa in the sample residue that may have been overlooked, and check the counts or relative abundances (depending on the type of sample).

It may seem more objective to require the QC laboratory to undertake QC without having a copy of the data generated by the processing laboratory. It is possible, for example, that an inexperienced person could undertake QC simply by agreeing with the identifications provided on the data sheets, which would not be a QC check at all. However, from our experience, both here and in the United States of America, when QC is undertaken blind (*i.e.* without the data provided by the sample processors) it is extremely difficult to reconcile any differences in identifications (without someone having another look at the specimens, which adds extra time and cost). Furthermore, given the very small size of some specimens, it is easy for the QC laboratory to overlook taxa that should be in the vial, leaving doubt as to whether they were there and were missed, or whether the sample processors forgot to put specimens into the vial. When the QC is undertaken blind like this, the result can be two slightly different lists of taxa for each sample and an additional step of reconciliation is required. Since a reconciled data set is the aim of QC, we believe that the reconciliation should be part of a one-step QC procedure undertaken by the second laboratory, leaving a third stage only if there is disagreement over the identification of specific taxa. In such cases, these can be provided to an agreed independent expert, as stated in the protocols (Stark *et al.* 2001).

QC is not required on every batch of samples processed, especially when the processing laboratory has a proven track record of excellent performance. Even then, however, QC should be undertaken every now and then to ensure that high-quality work is being maintained.

The value of reference collections (*i.e.* a set of vials containing clearly labelled identified examples of different macroinvertebrate taxa) in QC should not be overlooked, especially if reporting is to be based on the MCI or MCI-sb (which require only presence-absence data). Although examining a reference collection containing all taxa identified from a particular batch of samples does not evaluate the complete processing performance of the processing laboratory, it is the most cost-effective way to confirm the taxonomy and resolve any disputes over identifications.

2.2. Should we now do better than MCI-level taxonomy?

The MCI was developed initially in the early 1980s (Stark 1985), shortly after the publication of Winterbourn & Gregson's (1981) landmark first edition of the "Guide to aquatic insects of New Zealand". Although the Guide is a huge aid to better identifications, it does require some experience and training to use reliably. This was one of the reasons why it was decided to develop the MCI based on generic (at best) taxonomy. This appears to have been a wise decision given that the two most recent editions of the Guide have reverted to generic keys (except for the Simuliidae) "because so many described species are unknown as larvae" (Winterbourn *et al.* 2006). Stark (1985) also found that an MCI based on family-level data was not very sensitive and could distinguish only gross pollution from everything else. Generic-level taxonomy was adopted because there was sufficient sensitivity for assessing stream health at this level of taxonomy, and also because it was more cost-effective and practical than species-level taxonomy.

There is no doubt that MCI-level taxonomy has become the norm in New Zealand, and, in general, this has proven suitable for bioassessments and SoE reporting based on the MCI and its variants. The fact that most data sets are identified to the MCI level does mean, however, that we don't have as many data identified to the species level as would have been the case if the MCI had not constrained the identifications. It could well be that an MCI based on species-level identifications (where practical) may perform even better than the existing generic-level indices. However, to develop a species-level MCI, species-level data are required, which, in general, are not being collected.

Wright-Stow & Winterbourn (2003) evaluated the effect of taxonomic resolution on biotic index performance by comparing MCI and QMCI values determined using ordinal-level taxonomy with the conventional MCI and QMCI. Ordinal-level tolerance values were obtained by averaging MCI scores for 10 insect orders and five other higher taxonomic groups. They found that biotic indices based on coarse-level taxonomy ranked the health of streams in Canterbury in a comparable way to the MCI and QMCI.

3. MONITORING AND REPORTING

Biological systems are complex and unstable in space and time. As a result, biologists often feel compelled to study all of their components, but one need not sample everything. For monitoring it is more important to focus on biological attributes that respond reliably to human activities, are minimally affected by natural variation, are cost effective to measure and can be presented in a way that conveys useful information to water managers or to the general public. Macroinvertebrate biomonitoring using biotic indices has a long history of meeting these objectives (Karr & Chu 1999).

This section looks at different types of bioassessments and biomonitoring programmes and provides some guidance on how the MCI and its variant indices should be used within these programmes.

3.1. Compliance monitoring and environmental impact assessments

3.1.1. Compliance monitoring

Compliance monitoring is routine monitoring to ensure that an activity that is allowed under a resource consent is not having any significant adverse effects.

We recognise that regional councils have the responsibility for determining the scope of consent monitoring on a case-by-case basis. However, we have seen examples of monitoring conditions that seem to us to be overly burdensome for consent holders. There really is no need to monitor everything. Put simply, if a stream or river has healthy macroinvertebrate communities then it is almost certain that other ecosystem components will be in good shape too. There is often no need to monitor water quality, bed sediments, periphyton biomass or fish populations to obtain reasonable assurance that a consented activity is not having significant adverse effects. If macroinvertebrate biomonitoring does reveal disturbing trends, or if a problem is observed when sampling for macroinvertebrates, then that is when additional investigations should be undertaken.

There is some justification for more frequent or more intensive monitoring programmes for new consents, but if no adverse effects have been detected by the first review (after perhaps five years), then we believe that monitoring requirements should be reduced. In our view, annual macroinvertebrate sampling represents the minimum desirable level for compliance monitoring.

The first step in choosing which MCI variant to use is to ensure that the chosen index is sensitive to the type of impact or disturbance that the consented activity is expected to have (or might have if something untoward happens – bearing in mind that under the Resource Management Act 1991 (RMA) impacts from consented activities are not supposed to have *any*

significant adverse effects). One decision that will have a significant effect on costs is whether or not quantitative data are required.

Using the MCI for consent compliance monitoring

All variants of the MCI and MCI-sb are suitable for use in consent compliance monitoring programmes. The choice of which variants to use will depend on the objectives and available budget. The SQMCI and QMCI (and their soft-bottomed variants) are more suited to compliance monitoring and synoptic surveys (where all samples are collected on the same day under similar conditions) than to SoE monitoring (where samples may be collected over a month or more and yet need to be compared on a common basis). The SQMCI and QMCI (and their soft-bottomed variants) are best used where changes in stream community composition might be an anticipated consequence of the consented activity – an enriching discharge is a prime example.

The design of compliance monitoring programmes is discussed further in Section 4.

3.1.2. Assessments of environmental effects (AEEs)

Assessments of environmental effects (AEE) are undertaken when a new activity that is expected to have effects on the environment is proposed. The AEE will form part of the application for consent to undertake this activity under the provisions of the RMA. An AEE is also likely to be required when the term of an existing consent expires and permission is required for the activity to continue.

Biotic indices are a measure of stream health and can be used as part of an AEE to show the effect of an activity, provided the activity is capable of causing the kinds of changes to macroinvertebrate communities that biotic indices can detect. For example, it is entirely appropriate to use the MCI (or one of its variants) to assess an activity that has the potential to cause nutrient enrichment or sedimentation. However, if there was a proposed discharge containing chemicals that had no effect on macroinvertebrate communities (but did have other adverse effects on the environment), then the MCI would not be an appropriate tool to use. For example, Hickey & Clements (1998) found that the QMCI did not detect the impacts of heavy metal pollution in streams on the Coromandel Peninsula. They noted that this was because the QMCI had "incorrect tolerance scores for some taxa to heavy metals." This is not really a valid criticism of the QMCI, which was developed to detect organic pollution and nutrient enrichment, but rather a warning about using indices for assessing impacts (*e.g.* metal toxicity) for which they were not designed.

Although research is still in progress, there is evidence to suggest that the MCI has limitations when assessing the effects of extremely low flows. As flows reduce (whether during natural droughts or as the result of abstraction), macroinvertebrate communities in many stony streams

change from being dominated by mayflies, stoneflies, and caddisflies, to being dominated by chironomids, worms, snails and hydroptilid caddisflies. This occurs because periphyton on the stone surfaces changes from a thin diatom film to thick algal mats or even filamentous algae. These changes are reflected in a sharp decrease in the MCI. However, once the entire riverbed is covered with thick periphyton, the MCI stabilizes (perhaps around 80) even though flow, wetted perimeter, and space for aquatic communities continue to decrease. It follows that MCI values alone should not be used to support arguments that extreme abstractions do not have significant adverse effects on aquatic communities.

3.2. State of the Environment monitoring and reporting

Effective management of water resources (or environmental quality) requires knowing about changes that occur in the environment and having an understanding of the underlying cause(s) of any changes that might be predicted or observed. It is also desirable to be able to distinguish anthropogenic (human-caused) changes from natural ones. This kind of information is gathered mainly by State of the Environment (SoE) monitoring and reporting, which in New Zealand usually is undertaken by regional and unitary councils.

Specifically, SoE monitoring and reporting programmes aim to:

- Obtain representative data for each of the resources or resource compartments;
- Detect the presence and direction of trends;
- Identify the effects of activities – particularly land-use change – on resource quality;
- Determine the effectiveness of management initiatives directed at enhancing degraded resources.

Long-term data sets, including biomonitoring data sets, are vital (Likens 1998). They can address scientific and environmental questions at a scale that is realistic and applicable to environmental management, and can document the responses to disturbance by natural and anthropogenic events and activities.

3.2.1. Which biotic index should be used?

Recommended indices for SoE monitoring

We believe that the MCI and MCI-sb are the best biotic indices for state of the environment monitoring and reporting, and that the SQMCI and QMCI (and their soft-bottomed stream versions) should not be used for SoE reporting.

This view probably is quite contentious, given that most regional councils have reported the SQMCI (from coded-abundance data) or QMCI (from fixed-counts) along with the MCI in their SoE reports. So why do we recommend using only the MCI?

Most regional councils take several weeks (on each monitoring occasion) to collect their SoE samples. Most have rules about when to sample in relation to the last significant flood. However, when sampling is spread over several weeks (and even up to a month or two), there will be differences in macroinvertebrate community composition relating to when the sample was collected. For example, consider two sites that are identical in condition or "health." If they are sampled on the same day, their MCI, SQMCI or QMCI values will be about the same. However, if these two sites are sampled 30 days apart, it is likely that the biotic indices will be different. If the river has been in recession the entire time, the site sampled second will probably have lower index values because there will be increasing development of periphyton-associated communities with increasing dominance by low-scoring taxa such as chironomids, worms, snails and hydroptilid caddisflies. If there has been a significant fresh[9] between sampling these sites, the site sampled second could have a higher index value, because low-scoring algal-associated taxa are displaced by mayflies, stoneflies and caddisflies that are characteristic of comparatively clean stone surfaces. The difference between index values is a consequence of when the samples were collected rather than a measure of the health of the sites. This problem affects the MCI to a lesser extent than the SQMCI or QMCI, because the list of species present at a site is affected less when samples are collected than the densities or relative abundances of taxa.[10]

For SoE reporting to the public or to other laypeople, the KIS (Keep It Simple) principle should apply. SoE reports that consider taxa richness, MCI and SQMCI may be confusing. These three indices do not measure the same thing, so assessments based on them are not necessarily in complete agreement, often requiring an experienced biologist to explain the differences.

Taxa richness is not strictly a measure of stream condition or stream health. In fact, highest taxa richness often is associated with slightly enriched streams (*e.g.* those experiencing diffuse-source nutrient enrichment from farmland), rather than pristine streams in reference condition. Low taxa richness can be associated with quite "sterile" environments with extremely pure water, perhaps where torrential water velocities and lack of nutrients result in low productivity. Such places are naturally unproductive. In other words, there is no valid basis to assume that high taxa richness is good and low taxa richness is bad. Furthermore, estimates of taxa richness are highly dependent on sample size, which, in turn, can be influenced by sampling or processing effort (which can vary markedly with different personnel). For these reasons, use of taxa richness for SoE reporting is not recommended.

[9] A fresh is a sudden increase in stream or river flow due to rainfall or snow/ice-melt.

[10] Research on the effects of floods and droughts on biotic index values in stony streams is currently being undertaken by John Stark under NIWA's Water Allocation FRST programme. One aim is to determine correction factors that would enable biotic index values to be standardised to factor out the influence of floods, droughts, season or sampling time. Results are expected by June 2007.

We do not recommend the SQMCI (or QMCI) for SoE reporting either, because community percentage composition (more so than taxonomic composition) can change during the sampling period (which can be several weeks) in response to small freshes or as a flow recession lengthens. Consequently, differences between sites arise as a result of when samples were collected, and are therefore an artefact of the sampling regime rather than a true measure of stream health. We also found substantially higher variances in SQMCI-sb scores compared to MCI-sb scores from replicate samples collected in soft-bottomed streams in the Auckland regions (Stark & Maxted 2004, 2007). The high variance impaired the ability of the index to discriminate between sites of different qualities.

As we have seen, the MCI is a well-proven and reliable index for assessing stream health. Reliable MCI values can be derived from samples collected according to the national protocols, including fixed counts as low as 100 (although a minimum fixed count of 200 is recommended) (Stark *et al.* 2001; Duggan *et al.* 2003). The MCI is not affected by changes in percentage composition, and, consequently, is affected much less than the SQMCI and QMCI by flow-related or seasonal factors. The MCI is essentially a scaled average score per taxon, and therefore (being an average) is relatively unaffected by sample size (unlike taxa richness) provided that samples are collected according to the standard protocols (Stark *et al.* 2001). The MCI normally is highly correlated with the SQMCI or QMCI (Stark 1993, 1998), so using indices that tell much the same story seems somewhat superfluous, and could be confusing for laypeople. As a result, we recommend the MCI as the index of choice for SoE reporting.

In theory, the MCI can be affected by samples containing taxa that have drifted into the sampling area from upstream habitat or tributaries. However, this also affects other indices such as taxa richness and EPT richness, and we do not believe this is a significant problem for bioassessments, although the effect of this on stream health assessments has not been evaluated. We suspect it would be more of an issue when sampling is undertaken soon after a significant fresh. Drift is a major mechanism for re-colonisation of downstream reaches, and who is to say whether one or two higher-scoring taxa that may have drifted in from upstream really belong there or not. All macroinvertebrate communities have rare taxa present.

3.2.2. SoE monitoring versus research

In our view, SoE monitoring is not the best way to undertake basic research into the relationships between biological communities, land use, or other human activities that may affect water quality. This is not to say that data from SoE monitoring cannot be used for this purpose (as was done in the Auckland region; see Stark & Maxted 2004). Research objectives can be achieved by undertaking additional investigations in association with SoE monitoring, but it is important to keep the fundamental aims of SoE monitoring in focus, and not compromise the integrity of the SoE monitoring programme.

3.2.3. Planning an SoE monitoring programme

The design of monitoring programmes is discussed in Section 4 in more detail, but it is worth making some preliminary comments here. SoE monitoring involves sampling at defined (fixed) locations, at predefined intervals, according to the parameters being measured. Sites should be representative of least-disturbed conditions (reference) and impacted areas, and of a range of common land uses within the region. Monitoring methods should be robust, standardised, and applied consistently over time.

SoE monitoring networks have to be designed with management objectives in mind, and we consider the following matters should be at the forefront of design planning.

- What are the site locations?
- Use both reference sites and impacted sites, and samples should be representative in space (*e.g.* across different land uses and different River Environment Classification classes, and spread throughout the region).
- How many sites?
- What indicators will be measured?
- What degree of change do you want to detect (and how will it be distinguished from natural variability)?
- How often will monitoring be undertaken? (Samples should be representative in time.)
- What information is required?
- How will the data be analysed?
- How many data do the statistical analyses require?
- How much replication is required?
- How will data be translated into information that water managers can use?
- How much will it cost?
- Is there a long-term commitment to funding?

3.3. Biodiversity monitoring

Whereas the MCI focuses on cost-effective stream health assessment, there is increasing interest these days on freshwater macroinvertebrate biodiversity. This is a different issue, and here MCI-level identifications fall a little short of the ideal.

How best to undertake surveys to assess biodiversity is a topic that warrants further investigation. A large number of samples are almost certain to be required to obtain a complete list of the macroinvertebrate taxa in a particular stream habitat. Stark (1993b: Figure 1) showed, for example, that a single hand-net sample collected from riffle habitat contained

about 57% (range 34%–78%) of the taxa collected in 12 replicates combined, and even 12 samples were insufficient to collect all the taxa that were likely to be present. Furthermore, other taxa found in non-riffle habitats (*e.g.* runs, pools, under banks) would require additional sampling effort to detect them. In our view, routine sampling for biodiversity is likely to be prohibitively expensive.

So how can we continue to undertake cost-effective SoE monitoring and collect data suitable for assessing biodiversity at the same time? Perhaps the solution to the bioassessment–biodiversity issue is to take single hand-net samples at each SoE site according to protocols C1 (hard-bottomed) or C2 (soft-bottomed) (Stark *et al.* 2001), identify all taxa to the species level (where possible), and record only presence/absence. These data would enable calculation of MCI and MCI-sb values that would be consistent with those calculated from data collected previously. Additional samples could be collected from other habitats and scanned for taxa that were not found in the sample collected from the habitats targeted for SoE monitoring using protocols C1 or C2. The additional effort previously put into relative abundance counts or abundance coding could be put into the specific identification of more samples and habitats. A reference collection should also be assembled.

It is unlikely that sufficient sampling could be undertaken on each SoE monitoring occasion to compile complete macroinvertebrate biodiversity inventories for each sampling site. However, over time, the list of taxa recorded from each site will increase. Alternatively, every few years (say five or ten) extra sampling effort focusing on biodiversity assessment could be added to the SoE programme.

Although collecting presence-absence data for SoE monitoring may seem like a backward step, we believe that the positives outweigh the negatives. The positives include:

1. Cost-effective biomonitoring is retained (or even improved).
2. More sites can be monitored within a given budget because sample processing costs are reduced.
3. Species-level data suitable for biodiversity assessments are collected.
4. These data are suitable for developing a species-level MCI using the rank correlation iterative method developed by Bruce Chessman (as used by Stark & Maxted 2004, 2007 for the MCI-sb).
5. Assessments for soft-bottomed streams would be improved and much more cost-effective – the MCI-sb performs much better than its semi-quantitative and quantitative variants in soft-bottomed streams, and obtaining coded abundances or counts of invertebrates from samples from soft-bottomed streams is very time consuming.
6. Data quality control would be simplified – checking the reference collection and/or checking identifications in vials and for missed taxa in the sample residue.
7. The confounding influences of floods, low flows, season and extended sampling periods are minimised because they affect taxa richness and MCI much less than relative abundances, densities, SQMCI or QMCI.

On the other hand, if only presence-absence data are collected:

1. There are some well-performing indices that cannot be calculated (*e.g.* %EPT abundance), but they are often strongly correlated with MCI anyway, and so may be redundant.
2. There will be a loss of information, because knowing which taxa are dominant is useful information for an ecologist but understanding the influences of the confounding effects of flow and season is required to make full use of this information.

4. DESIGN OF MONITORING PROGRAMMES

To make sure you achieve what you set out to do, it is essential that monitoring programmes are well-designed. In particular, the methods used for monitoring should be linked to the study or management aims so that the programme will:

- Provide useful information for water managers;
- Be scientifically robust, to satisfy the ecologists and statisticians;
- Be cost-effective, to enhance the security of ongoing funding.

Data that do not help interpretation, or have no proven value for environmental managers, should not be collected (Likens 1983). The statistical analyses used should be envisaged as part of the study design – not an after-thought once data are collected.

4.1. Sample site selection

Study aims (*i.e.* data requirements) and budgetary considerations dictate the selection of the sample sites. If an upstream versus downstream (or control versus impact) design is required, then sample site selection can have a major impact on the results. If you require an assessment of differences in water quality (say, as a result of a discharge), then it is essential that the sites are as similar as possible in all other physical features, because habitat differences between sites can lead to differences in MCI values that are unrelated to the health, enrichment or pollution status of the stream ecosystem. It is crucial that the stream substrates to be sampled are similar, but the degree of exposure of a site to sunlight can also be important. One site in the shade for most of the day (where the stones are clean), and the other in full sunlight (where algae proliferate) will invariably have different macroinvertebrate communities. In such cases a false impression of the effects of an impact are highly likely.

The highest MCI values generally are recorded from good stony riffle habitat, with lower values associated with deeper water, slower current velocities, and/or finer bed sediments (in runs and pools). Increasing levels of periphytic algae on the streambed will also reduce MCI values. This is because conditions become less favourable for taxa that require clean

conditions (such as mayflies, stoneflies, and most caddisflies) and more favourable for taxa that prefer living in algal mats (such as worms, snails, and chironomids). In soft-bottomed streams, large woody debris appears to be the habitat that is equivalent ecologically to stony riffles, where the most pollution-sensitive taxa exist and the highest MCI values will be recorded (Maxted *et al.* 2003).

It is almost inevitable that there will be differences in substrate, current velocity, water depth, and aspect (*i.e.* the orientation of the reach in relation to the sun) between sites that may affect monitoring results. However, one approach that largely overcomes the influence of between-site physical habitat differences on biotic index values is to monitor changes over time by sampling at the same place one to four times a year. This approach is particularly useful for SoE monitoring or consent compliance monitoring, where one wants to know whether conditions are being maintained, improved, or are deteriorating. The long-running biomonitoring programme at Kapuni (Taranaki region, North Island) uses this approach, with time-series analysis to examine trends in MCI (and species richness) at monitoring sites (Stark 1993a). Monitoring at Kapuni began in September 1981 and remains ongoing, with sampling four times per year.

Although Stark (1993b) found that water depths, current velocities and substrata commonly encountered in stony riffles appear to have little impact on index values, we recommend limiting sampling in stony streams to water depths of 0.1–0.4 m, current velocities of 0.2–1.2 ms^{-1}, and substratum median rock diameters of 60–140 mm, where possible.

There are a number of different ways of selecting sample sites for SoE monitoring. Perhaps the most intensive approach is the US EPA's EMAP, which involves placing a grid over the study area and selecting sampling sites within each grid systematically from a random start location (Herlihy *et al.* 2000). This probabilistic sampling design invariably results in a very large number of sampling sites in order to satisfy statistical requirements, assumes no prior knowledge about the nature of the sampling sites, and is very expensive.

In our view, the EMAP approach is not an efficient use of resources when you already have some information about the streams and rivers within the region. A stratified sampling design with sites classified into categories (*e.g.* based on stream type, source of flow, and land use) will prove much more cost-effective. The River Environment Classification system could provide a more efficient basis for site selection, with sites replicated within the dominant classes within the region (Snelder & Biggs 2002; Snelder *et al.* 2004). Alternatively, a stratified sampling design based on representative stream/land-use types across the full disturbance gradient could be adopted, as it was for Auckland Regional Council's SoE monitoring programme (which provided a superb data set for developing the MCI-sb and for SOE reporting).

4.2. Sampling frequency

Recommended sampling frequency

Biotic data generally change over a longer period, so need not be collected as frequently as data on water quality (often monthly) or flow (often every 15 minutes). For routine SoE monitoring, annual sampling is common and it is difficult to imagine why sampling more than four times per year would ever be required (except for specific research projects).

Expanding on the general recommendation given above, sampling frequency is determined by the study objectives, modified by cost considerations (Resh & McElravy 1993). New Zealand stream faunas show less seasonal variation in species presence than is seen in the Northern Hemisphere (Towns 1985), and absence of key species or segments of the fauna is related more to environmental disturbance than life-history patterns. Towns (1985) suggested that one summer and one winter sampling should be sufficient to provide a close approximation of potential species richness at a site. For most biomonitoring programmes, sampling one or two times per year is likely to provide a reasonable balance between cost and ensuring that in-stream conditions are acceptable. If time-series analyses are required, then seasonal sampling and replication may be advisable for the first few years to build up a picture of biotic index variability, or to establish the reference condition if sampling begins before the exercise of a discharge consent. After this, sampling frequency can be reduced to once or twice per year, particularly for long-term or SoE monitoring, where year-to-year trends are of interest.

Sampling macroinvertebrate communities more frequently than seasonally is unlikely to be appropriate for SoE monitoring, due to the cost and because macroinvertebrate communities are a product of their environmental experience over the past weeks and months. However, more frequent sampling might be appropriate for more intensive special investigations.

Frequency versus number of sites

We suggest to those councils that sample more than once per year, or collect replicate samples, that they might consider devoting the same resources to sampling more sites once per year. If resources are concentrated on annual sampling, more sites can be included in SoE monitoring programmes for a given investment, and better regional (and national) coverage can be achieved.

From a sample processing viewpoint, spring sampling has some advantages because most aquatic insects are present as large, later instar[11] larvae or pupae, and are more easily identified. Most keys (*e.g.* Winterbourn *et al.* 2006) are based on final instar larvae, and caddisfly pupae often develop adult characteristics (while retaining larval sclerites), enabling specific identifications to be made. Summer can be the preferred season for SoE monitoring, however, because that is when flows are lowest, water temperatures are highest, and aquatic communities are under most stress. The summer season was selected in the Auckland region because many catchments are small (<1000 ha) and dominated by small streams that dry up in late summer. Sampling during the dry late summer ensures that the streams are perennial and excludes intermittent streams affected by frequent desiccation. However, sample processing can be tedious due to the presence of many early instar larvae, which can be difficult to identify.

So, what is the preferred season for undertaking SoE monitoring? Preliminary investigations into seasonal variability (J. D. Stark, unpublished) suggest that spring and winter MCI values may be higher, on average, than those recorded in summer or autumn. The difference appears to be small (perhaps 4–5%), but often is statistically significant. However, given that there is approximately ±10% error in estimating the MCI from single hand-net samples (Stark 1998), it could be argued that seasonal variability is not a major cause for concern.

4.3. Sampling methods

Standard protocols for macroinvertebrate sampling in wadeable streams have been available since late 2001 (Stark *et al.* 2001). These protocols were developed in association with regional council personnel and aim to promote standardised methods by building on (rather than overturning) current practice. As a result, the standard protocols have experienced wide acceptance.

4.4. Replication

Replicate samples are collected for one or more of the following reasons:

- To obtain sample mean values of indices such as the MCI and a measure of variance (*e.g.* standard deviation, standard error);
- To meet requirements of statistical significance tests for detecting differences between sites or times;
- To increase the sampling effort and collect more taxa from the site.

In general, increasing replication improves the precision of estimates of biological indices, taxa richness (expressed as number of taxa per sample), overall macroinvertebrate densities, or

[11] Instar = a stage in the life of arthropods between two periods of moulting (shedding of the exoskeleton in order to grow).

densities of individual taxa. Note, however, that precise estimates of the densities of rare taxa may require collecting hundreds of replicates (see Elliot 1977).

Whether or not sample replication is necessary (or even possible) depends on the study objectives and budget or time constraints. We believe that sample replication is essential when undertaking robust scientific research programmes and intensive AEEs, but that for routine ongoing compliance monitoring or SoE monitoring, sample replication is optional and, in most cases, unnecessary, not only because routine biomonitoring programmes should be as cost-effective as possible (with additional investigations initiated only if routine monitoring detects a problem), but also because single samples provide reasonably precise estimates of biotic index values. Single hand-net samples provide estimates of MCI and MCI-sb that are less than ±12% of average (100) index values (Stark 1998; Stark & Maxted 2007), which we believe is more than adequate for SoE monitoring. Replication to define the variance of index scores from single samples under local conditions is recommended and should cover a range of land-use and habitat quality conditions. Replication need only be done once when sampling methods are established, and then updated periodically when new site types are added to a sampling network. Alternatively, replicate samples collected at a certain percentage of sites (*e.g.* 5%) would gradually build such a data set over time.

Knowledge of index variability among replicate samples is essential if statistically significant differences between samples are to be detected. The number of replicate samples collected is inevitably a compromise between the cost (mainly of sample processing) and the desired sensitivity of the monitoring programme. If we expect a given site to have an MCI around 100, and hope to detect statistically significant differences in index values within ±10% of this value, then two hand-net samples or five Surber samples are likely to be required (see Table 5 in Stark 1998). If detection of a ±20% change is acceptable, a single hand-net sample and duplicate Surber samples would suffice. Table 5 in Stark (1998) also reveals the cost-effectiveness of the SQMCI compared with the QMCI: only three hand-net samples compared with eight Surber samples are required to achieve a detectable difference of 0.48 index units (which represents <±10% of an average index value of 5.0).

Once replicates have been collected, between-site differences can be assessed using standard statistical procedures such as ANOVA and t-tests. Information in Stark's (1998) Table 5 is useful for determining whether index values from single samples are likely to be significantly different.

4.5. Other factors that can affect monitoring results

When designing monitoring programmes it is important to consider the influence of any factors that may confound the interpretation of monitoring results. We have already considered habitat differences between sample sites in Section 4.1 above. However, flow variability can also have a profound effect on bioassessments.

The influence of flow variability on biotic index values is under investigation as part of FRST-funded research in collaboration with NIWA's Water Allocation Research programme (J. D. Stark, unpublished). Exploratory data analyses have indicated that increased MCI values may be significantly correlated with flood events, and that extended periods of low flow (or periods without significant freshes) are related to lower MCI values. At these extremes, MCI values will be higher (soon after floods) or lower (during prolonged low flows) than the annual mean MCI. It has been suggested that high flows may elevate MCI values by washing higher-scoring taxa from "better" habitat upstream and/or by flushing away algal proliferations (which generally are inhabited by communities of lower-scoring taxa). Floods five or six times the magnitude of the preceding base flow appear to have significant influences on MCI values. This is consistent with the findings of Biggs & Close (1989) for periphyton communities.

Most councils try to factor out the effects of significant floods by delaying sampling until a set time after the last significant flood – two weeks is common. This gives the stream time to recover from the immediate effects of the flood, so that it will be closer to its "average" state. Most councils do not, however, have a similar rule that avoids sampling under extreme low flow conditions. Variability in flow over time does not affect synoptic surveys or the interpretation of compliance biomonitoring data collected on a single day (because all sites normally will have experienced similar flows), but it can affect comparisons between times or SoE monitoring programmes if it takes several weeks to sample all monitoring sites within the region.

4.6. What other environmental data should be recorded?

This section focuses on SoE monitoring, although our recommendations may apply also to other forms of monitoring (*e.g.* compliance, AEE, and biodiversity).

The fundamental aim of SoE monitoring is to collect data that will provide robust assessments of stream health or condition. Macroinvertebrate sampling alone can achieve this aim, but it often is desirable to collect additional information to identify the causal environmental factors. Only when these are understood can water managers implement policies or actions designed to improve stream condition. For example, an ecologist might report an MCI of 90 for a stream, which indicates only "fair" condition or "probable moderate pollution". Clearly, there is room for improvement, but this information alone is not sufficient for any action to be taken. Knowing that chironomids dominate the community and that the stream bed is covered in thick periphyton is helpful, but still not enough to suggest remedial action unless we understand what has caused the problem. If we also know that this river is in a rural setting we might suspect that the periphyton proliferation (with its associated macroinvertebrates) might be the result of nutrient enrichment from point or diffuse sources and may have been exacerbated by the removal of riparian shade. Controlling nutrient inputs (by fencing off a riparian strip to prevent stock access and to intercept nutrient run-off, and treating any point-source discharges), and planting shade trees along the stream banks could bring about some improvement in stream condition.

Put simply, an MCI value does not imply any particular remedial action: it is necessary to understand the causes of degradation in order to remediate it. However, there has been considerable research into the effects of agricultural activity on stream community health (*e.g.* Hickey *et al.* 1989; Quinn *et al.* 1992; Quinn *et al.* 1997a), so in the case where we have a site photograph showing the periphyton on the streambed and the stream in a rural setting, and where we know the MCI, it probably is sufficient to enable a water manager to implement remedial strategies. In other words, the macroinvertebrate sample alone provides an assessment of stream health and knowing the land use in the catchment provides sufficient information, when combined with existing knowledge, for remedial action to be undertaken.

How much additional information should be collected along with macroinvertebrate samples for SoE monitoring? There is a time and cost associated with data collection, so it is prudent to ensure that this expenditure is worthwhile and contributes to the monitoring objectives. We agree with Likens (1983), who notes that just because the information "might be useful one day" is insufficient justification.

When balancing practicalities (especially the ideal of collecting all SoE samples within a few days – see Section 4.7) and cost with information needs, we believe that SoE monitoring should comprise the following essential and optional elements.

Essential elements of SoE monitoring are:

1. The site location (including the map or GPS reference), sampling date, sampling time, and name(s) of personnel undertaking sampling.
2. Site photographs showing:
    a. An overview of the site in its setting (stream banks, riparian cover, catchment land use);
    b. The stream bed (substrate type, periphyton type and cover).
3. A single hand-net sample per site collected according to protocol C1 (hard-bottomed) or C2 (soft-bottomed) (Stark *et al.* 2001); descriptions of substrata sampled; and proportions of soft-bottomed substrata sampled (*e.g.* % submerged wood, macrophytes, bank structure).
4. Field observations – local land use and sources of pollution.
5. Catchment variables – land use, proportion of impervious land cover, elevation, distance from the sea, stream order, pollution sources, barriers.

Optional elements are:

1. Field physicochemical measurements – electronic instruments can log pH, dissolved oxygen, temperature and conductivity quite efficiently.
2. Habitat assessments using a formal protocol.
3. Periphyton percentage cover and/or samples for chlorophyll *a* or ash-free dry-weight (AFDW) determinations.

4. Stream substrata percentage composition.
5. Width, depth, current velocity and discharge.
6. River Environment Classification class, or some other description of the environment/eco-region.

4.7. Minimising the time taken for a full sample round

The integrity of the macroinvertebrate data is maximised if all SoE monitoring sites are sampled within the space of a few days rather than over several weeks. This gives less time for conditions to change so that the risk of confounding influences, such as freshes and recessions, affecting assessments is minimised. However, few councils seem able to collect samples from 40 to 60 sites within a few days. There are several reasons for this. Often it is because of all the other sampling and assessment that is being undertaken at the same time (as described in the previous section); sometimes the time taken to travel between sites is a constraint.

We accept that it may be impractical for most councils to sample all sites within a few days, but there are several strategies that may help to achieve this. First, all sites could be visited, with the only work undertaken being site photographs and macroinvertebrate sampling. If electronic meters capable of automatically recording selected water quality variables are available, these could be used too. In this way perhaps eight to 10 sites could be visited each day. All sites could then be visited a second time after the macroinvertebrate samples have been collected in order to undertake more time-consuming activities such as habitat assessments. Another strategy is to send out two or more teams to sample different sites on the same days (although that could increase sample variability if different people collect them).

Ultimately, sampling over a period of weeks may not be a problem. Current research aims to determine seasonal and flow-related correction factors that would enable MCI values (among others) to be standardised to improve the validity of between-site comparisons and stream health assessments.

5. DETECTING TRENDS

One question that crops up after SoE monitoring has been undertaken for several years is whether or not in-stream conditions have deteriorated, improved or stayed the same. To answer this question, a class of statistics called time-series or trends analysis is used.

Different techniques for time-series analyses vary in their data requirements and complexity. Linear regression-based parametric methods should only be used when the trend is expected to

be linear (often it is not), unless data transformations are used. There are also problems with parametric methods such as linear regression if there is heteroscedasticity in the data (*i.e.* variance differs with time). Non-parametric techniques for trend analysis are much better able to handle non-normal data with censored, tied and missing values, so they have been favoured for analysing trends, particularly in water quality data.

A popular non-parametric trend test for water quality data is the seasonal Kendall trend test – a technique described by McBride (2005) and implemented by Bill Vant of Environment Waikato (see Vant & Smith 2004). However, this technique requires monthly data collected for at least three – and preferably five – years or more (*i.e.* 36 to 60 data points). This is seldom the case for biotic data, where sampling may be seasonal at best, and is more often undertaken only once or twice a year. Although the seasonal Kendall test could be adapted for detecting trends in biotic data collected much less frequently, it is doubtful whether it would be worth the effort compared with less complicated methods that are likely to give a similar result.

Stark & Fowles (2006) examined several simple statistical approaches for detecting significant trends in stream macroinvertebrate biological indices. These included non-parametric tests based on the Mann–Kendall or Spearman rank correlations (Collier & Kelly 2006) and a parametric approach using linear regressions.

Recommended method for monitoring trends

We recommend the following cost-effective method for examining trends in macroinvertebrate biological data:

1. Visualise the trend using scatter plots of biological index versus time with LOWESS (tension = 0.4) fit (*e.g.* see Figure 2).
2. Test for significance using the Mann–Kendall test at the 5% significance level, followed by Benjamini–Hochberg (1995) False Discovery Rate (FDR) analysis. The trends remaining significant following this procedure should be clear trends, although the final decision on ecological significance rests with the best professional judgement of an experienced freshwater macroinvertebrate ecologist.

When trend testing is undertaken at a number of sites, it is probable that some significant results will be obtained by chance. The FDR analysis is undertaken when multiple comparisons are made in order to eliminate significant positive or negative trends that may have arisen by chance. The positive trend in MCI shown on Figure 2 had a probability of 0.011, which would normally be considered significant, but since it was one of a batch of 60 trends tests, the FDR cut-off value of P was 0.0065, so this significant result was deemed non-significant (*i.e.* it could have occurred by chance).

When a trend has been identified there are two vital questions that should be asked (Rutherford 1985):

- What caused the trend?
- Will it continue into the future?

Unfortunately, trend analysis alone cannot provide definite answers to these closely related questions. To find out why a significant trend has occurred you will need additional information, which could include data on stream flows, weather patterns, catastrophic erosion events in the catchment, physicochemical water quality, or changes to land or water management practices that may have resulted from water management initiatives, or industry or farming activities. It is worth emphasising that the final decision on whether or not any trend should be considered ecologically significant is reliant on the best practice judgement of an experienced freshwater ecologist. We caution water managers about interpreting statistically significant trends in stream biological health, particularly if these trends are only marginally significant, if they cannot be explained and/or may be unrelated to initiatives aimed at improving stream condition.

If the cause(s) of the trend can be determined, then one may be able to predict the future with some confidence. Conversely, if one cannot determine why a trend has occurred, then extrapolation into the future may be most unwise because there remains a chance that the trend is simply an artefact of natural variability or sampling error.

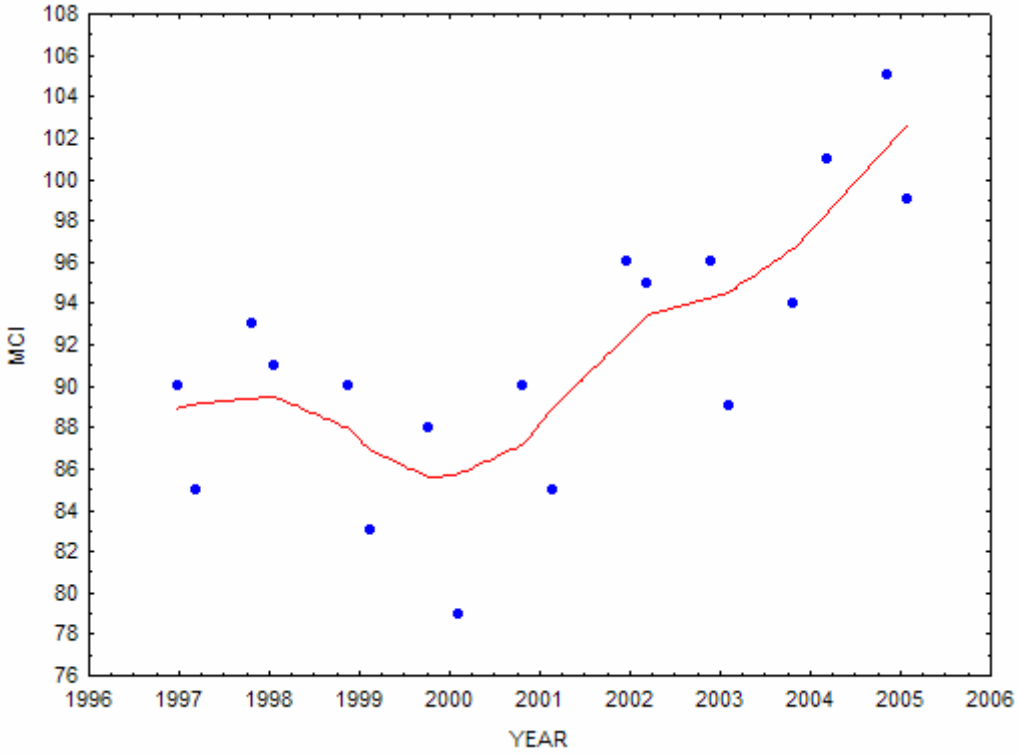

Figure 2. Scatterplot of MCI versus time for the Huatoki Stream at Hadley Drive (Site HTK000350) in Taranaki with a LOWESS (tension = 0.4) fitted line
Source: Stark & Fowles 2006

Experience with trend analysis of macroinvertebrate data is limited (because there are few long-term data sets), so the minimum number of sampling times required before meaningful trends can be detected is uncertain. Collier & Kelly (2006) suggested that a time series of five occasions is the minimum, but Stark & Fowles (2006) consider that trend testing is best undertaken with a time series of 10 or more. Scarsbrook *et al.* (2000a, b) examined data from annual sampling of macroinvertebrates from 66 of the National River Water Quality Network sites from 1989 to 1996 (inclusive). These data were sufficient to detect trends in various macroinvertebrate community measures, many of which coincided with general trends in water quality over the same period (Smith *et al.* 1996), suggesting that at least some of the measured indices (such as MCI and %EPT) were appropriate biological indicators of trends in water quality at a national scale. However, no causal links were established, and it is possible that some of the trends were artefacts of the short time series (*i.e.* only eight data points each, one year apart).

6. REPORTING THE MCI

The different versions of the MCI can be used at all levels, ranging from very specific technical impact assessments or consent compliance assessments, through to high-level SoE reporting at the catchment, regional or national scale. In addition, these indices have found widespread use in the formal scientific literature as a way to assess impacts of various kinds on stream communities, and many of these studies have served to confirm the ability of these indices to assess environmental health in robust and useful ways.

When preparing any report it is essential to consider the target audience, and use tools and language that convey information appropriately. There is no point preparing a highly technical report that only experienced macroinvertebrate ecologists will understand when the target audience comprises regional councillors or the general public. As mentioned earlier, commonly used indices such as taxa richness, MCI and SQMCI do not measure exactly the same thing, so assessments based on them are not necessarily in complete agreement and reporting them all together may present a confusing picture.

We recommend plotting maps similar to Figure 3 to present an overview of macroinvertebrate SoE monitoring on a regional or national basis. In this example, blue dots indicate where MCI values suggest that stream health exceeds expectations, green dots indicate that expectations are met, and red dots indicate sites where some improvement is desirable. We prefer the traffic light analogy, with green indicating stream health that meets expectations, orange indicating that there is some room for improvement, and red indicating that there are real problems that need to be addressed. Classification of index scores (Table 2) provides a consistent basis for colour-coding index scores. These quality classes, coupled with mapping, are an effective way to communicate results to non-scientists.

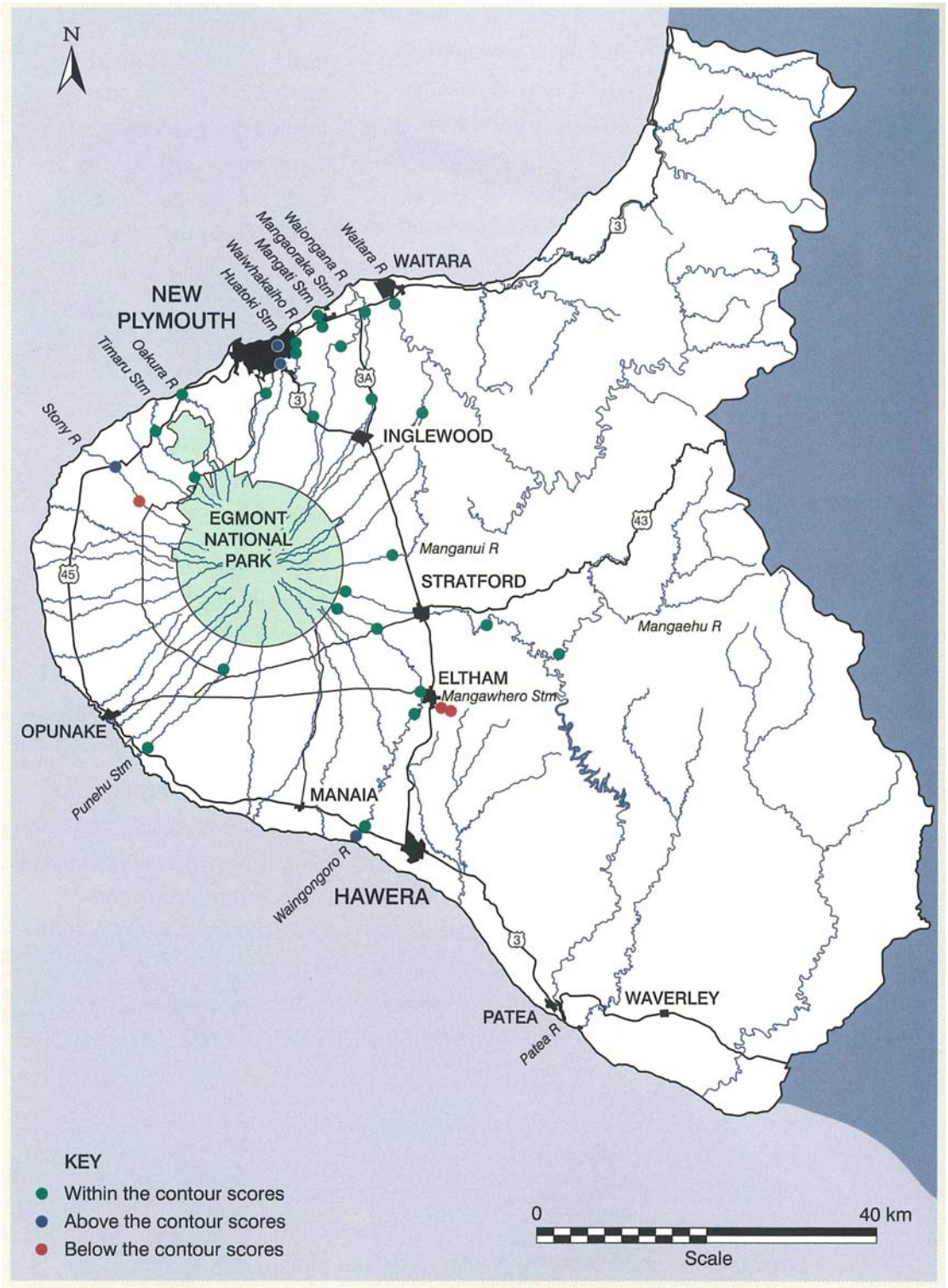

Figure 3. Median MCI scores for the period 1995–2000 in relation to regional MCI reference scores
Source: Taranaki Regional Council 2003

7. REFERENCES

* Includes a bibliography of papers that cite the MCI (and variants), not all of which are cited in this report.

Armitage PD, Moss D, Wright JF, Furse MT 1983. The performance of a new biological water quality score system based on macroinvertebrates over a wide range of unpolluted running water sites. Water Research 17:333-347.

Artemiadou V, Lazaridou M 2005. Evaluation score and interpretation index for the ecological quality of running waters in central and northern Hellas. Environmental Monitoring and Assessment 110(1-3):1-40.

Benjamini Y, Hochberg Y 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society B (57):289-300.

Biggs BJF, Close ME 1989. Periphyton biomass dynamics in gravel bed rivers: the relative effects of flows and nutrients. Freshwater Biology 22:209-231.

BMWP 1978. Final report: assessment and presentation of the biological quality of rivers in Great Britain. Unpublished report, Biological Monitoring Working Party, Department of the Environment, London. 37 p.

*Boothroyd IKG, Stark JD 2000. Use of invertebrates in monitoring. Pp. 344-373 in New Zealand stream invertebrates: ecology and implications for management, edited by Collier KJ, Winterbourn MJ. New Zealand Limnological Society, Christchurch, New Zealand.

Borja A, Muxika H 2005. Guidelines for the use of AMBI (AZTI's Marine Biotic Index) in the assessment of the benthic ecological quality. Marine Pollution Bulletin 50(7):787-789.

Chessman BC 1995. Rapid assessment of rivers using macroinvertebrates: a procedure based on habitat-specific sampling, family-level identification and a biotic index. Australian Journal of Ecology 20:122-129.

Chessman BC 2003. New sensitivity grades for Australian river macroinvertebrates. Marine and Freshwater Research 54:95-103.

Chessman BC, Growns JE, Kotlash AR 1997. Objective derivation of macroinvertebrate family sensitivity grade numbers for the SIGNAL biotic index: application to the Hunter River system, New South Wales. Marine and Freshwater Research 48:159-172.

Chessman BC, Trayler KM, Davis JA 2002. Family- and species-level biotic indices for macroinvertebrates of wetlands on the Swan Coastal Plain, Western Australia. Marine and Freshwater Research 53:919-930.

Chutter FM 1972. An empirical biotic index of the water quality in South African streams and rivers. Water Research 6:19-30.

Clarke RT, Wright JF, Furse MT 2003. RIVPACS models for predicting the expected macroinvertebrate fauna and assessing the ecological quality of rivers. Ecological Modeling 160(3):219-233.

*Collier KJ 1995. Environmental factors affecting the taxonomic composition of aquatic macroinvertebrate communities in lowland waterways of Northland, New Zealand. New Zealand Journal of Marine and Freshwater Research 29:453-465.

*Collier KJ, Kelly J 2006. Patterns and trends in the ecological condition of Waikato streams based on the monitoring of aquatic invertebrates from 1994 to 2005. Environment Waikato Technical Report 2006/04. 28 p.

*Collier KJ, Wilcock RJ, Meredith AS 1998. Influence of substrate type and physico-chemical conditions on macroinvertebrate faunas and biotic indices of some lowland Waikato, New Zealand, streams. New Zealand Journal of Marine and Freshwater Research 32:1-19.

Coysh J, Norris R 1999. Waikato region AusRivAS trial draft report. SMF Contract No. 5091 prepared for the Ministry for the Environment.

Davies PE 1997. Development of a national river bioassessment system (AUSRIVAS) in Australia. Pp. 113-124 in Assessing the biological quality of fresh waters, RIVPACS and other techniques, edited by Wright JW, Sutcliffe DW, Furse MT. Freshwater Biological Association, Ambleside, Cumbria.

Davy-Bowker J, Murphy JF, Rutt GR, Steel JEC, Furse MT 2005. The development and testing of a macroinvertebrate biotic index for detecting the impact of acidity on streams. Archiv fur Hydrobiologie 163(3):383-403.

*Death RG, Baillie B, Fransen P 2003. Effect of Pinus radiata logging on stream invertebrate communities in Hawke's Bay, New Zealand. New Zealand Journal of Marine and Freshwater Research 37:507-520.

De Pauw N, Roels D, Fontoura AP 1986. Use of artificial substrates for standardized sampling of macroinvertebrates in the assessment of water quality by the Belgian Biotic Index. Hydrobiologia 133:237-258.

*Doledec S, Phillips N, Scarsbrook M, Riley RH, Townsend CR 2006. Comparison of structural and functional approaches to determining land-use effects on grassland stream invertebrate communities. Journal of the North American Benthological Society 25(1):44-60.

*Duggan IC, Collier KJ, Lambert PW 2002. Evaluation of invertebrate biometrics and the influence of subsample size using data from some Westland, New Zealand, lowland streams. New Zealand Journal of Marine and Freshwater Research 36:117-128.

*Duggan IC, Scarsbrook MR, Quinn JM 2003. Comparison of coded abundance and fixed count rapid assessment techniques for biomonitoring New Zealand streams. New Zealand Journal of Marine and Freshwater Research 37:23-29.

Elliott JM 1977. Some methods for the statistical analysis of samples of benthic invertebrates. 2nd edition. Scientific Publications of the Freshwater Biological Association No. 25. 156 p.

Graca MAS, Coimbra CN 1998. The elaboration of indices to assess biological water quality: a case study. Water Research 32:380-392.

*Hall MJ, Closs GP, Riley RH 2001. Relationships between land use and stream invertebrate community structure in a South Island, New Zealand, coastal stream catchment. New Zealand Journal of Marine and Freshwater Research 35(3):591-604.

Herlihy AT, Larsen DP, Paulsen SG, Urquhart NS, Rosenbaum BJ 2000. Designing a spatially balanced, randomized site selection process for regional stream surveys: the EMAP Mid-Atlantic Pilot Study. Environmental Monitoring and Assessment 63(1):95-113.

Herricks EE, Cairns J 1982. Biological monitoring Part III: receiving system methodology based on community structure. Water Research 16:141-153.

*Hickey CW, Clements WH 1998. Effects of heavy metals on benthic macroinvertebrate communities in New Zealand streams. Environmental Toxicology and Chemistry 17(11):2338-2346.

Hickey CW, Quinn JM, Davies-Colley RJ 1989. Effluent characteristics of dairy shed oxidation ponds and their potential impacts on rivers. New Zealand Journal of Marine and Freshwater Research 23:569-584.

Hilsenhoff WL 1977. Use of arthropods to evaluate water quality of streams. Department of Natural Resources, Madison, Wisconsin, USA. Technical Bulletin 100. 15 p.

Hilsenhoff WL 1987. An improved biotic index of organic stream pollution. Great Lakes Entomologist 20:31-39.

Hilsenhoff WL 1988. Rapid field assessment of organic pollution with a family-level biotic index. Journal of the North American Benthological Society 7:65-68.

Hirsch A 1958. Biological evaluation of organic pollution of New Zealand streams. New Zealand Journal of Science 1:500-551.

Jiang HG 2006. Development of a new biotic index to assess freshwater pollution. Environmental Pollution 139(2):306-317.

*Joy MK, Death RG 2003. Biological assessment of rivers in the Manawatu-Wanganui region of New Zealand using a predictive macroinvertebrate model. New Zealand Journal of Marine and Freshwater Research 37:367-379.

Karr JR, Chu EW 1999. Restoring life in running waters: better biological monitoring. Island Press, Washington, DC. 206 p.

Katoh K 1992. A comparative study of some ecological methods of evaluation of water pollution. Environmental Science 5:91-98.

Lawrence TM, Harris TL 1979. A quantitative method for ranking the water quality tolerances of benthic species. Hydrobiologia 67:193-196.

Likens GE 1983. A priority for ecological research. Bulletin of the Ecological Society of America 64(4):234-243.

Likens GE 1998. The value of long-term data for understanding ecosystem structure and function: the Hubbard Brook example. Research in the Taieri Catchment. Occasional paper Number One, Ecology Research Group, University of Otago. Pp. 2-4.

*Maloney RF 1995. Effect of the herbicide triclopyr on the abundance and species composition of benthic aquatic macroinvertebrates in the Ahuriri River, New Zealand. New Zealand Journal of Marine and Freshwater Research 29:505-515.

Marshall JW, Winterbourn MJ 1979. An ecological study of a small New Zealand stream with particular reference to the Oligochaeta. Hydrobiologia 65:199-208.

Maxted JR, Barbour MT, Gerritsen J, Poretti V, Primrose N, Silvia A, Penrose D, Renfrow R 2000. Assessment framework for mid-Atlantic coastal plain streams using benthic macroinvertebrates. Journal of the North American Benthological Society 19: 128-144.

*Maxted JR, Evans BF, Scarsbrook MR 2003. Development of standard protocols for macroinvertebtrate assessment of soft-bottomed streams in New Zealand. New Zealand Journal of Marine and Freshwater Research 37:793-807.

*Maxted JR, McCready CH, Scarsbrook MR 2005. Effects of ponds on stream water quality and macroinvertebrate communities. New Zealand Journal of Marine and Freshwater Research 39:1069-1084.

McBride GB 2005. Using statistical methods for water quality management: issues, problems and solutions. John Wiley & Sons Inc., Hoboken, New Jersey, USA. 313 p.

Ministry for the Environment 1997. Environmental performance indicators: proposals for air, fresh water and land. Signposts for sustainability. Ministry for the Environment, Wellington, New Zealand. 158 p + appendices.

Norris RH, Norris KR 1995. The need for biological assessment of water quality: Australian perspective. Australian Journal of Ecology 20:1-6.

*Parkyn S, Davies-Colley RJ 2003. Riparian management: how well are we doing? Water and Atmosphere 11(4):15-17.

*Parkyn SM, Davies-Colley RJ, Halliday NJ, Costley KJ, Croker GF 2003. Planted riparian buffer zones in New Zealand: do they live up to expectations? Restoration Ecology 11(4):436-447.

Plafkin JL, Barbour MT, Porter KD, Gross SK, Hughes RM 1989. Rapid bioassessment protocols for use in streams and rivers: benthic macroinvertebrates and fish. US Environmental Protection Agency, Office of Water Regulations and Standards, Washington DC.

*Quinn JM, Boothroyd LKG, Smith BJ 2004. Riparian buffers mitigate effects of pine plantation logging on New Zealand streams. 2: invertebrate communities. Forest Ecology and Management 191:129-146.

*Quinn JM, Cooper AB, Davies-Colley RJ, Rutherford JC, Williamson RB 1997a. Land use effects on habitat, water quality, periphyton, and benthic invertebrates in Waikato, New Zealand, hill-country streams. New Zealand Journal of Marine and Freshwater Research 31:579-597.

*Quinn JM, Cooper AB, Stroud MJ, Burrell GP 1997b. Shade effects on stream periphyton and invertebrates: an experiment in streamside channels. New Zealand Journal of Marine and Freshwater Research 31(5):665-683.

*Quinn JM, Hickey CW 1990. Characterisation and classification of benthic invertebrate communities in 88 New Zealand rivers in relation to environmental factors. New Zealand Journal of Marine and Freshwater Research 24:387-409.

*Quinn JM, Williamson RB, Smith RK, Vickers ML 1992. Effects of riparian grazing and channelisation on streams in Southland. 2: benthic invertebrates. New Zealand Journal of Marine and Freshwater Research 26:259-273.

Resh VH, McElravy EP 1993. Contemporary quantitative approaches to biomonitoring using benthic macroinvertebrates. Pp. 159-194 in Freshwater biomonitoring and benthic macroinvertebrates, edited by Rosenberg DM, Resh VH. Chapman & Hall, New York and London.

*Riley RH, Townsend CR, Niyogi DK, Arbuckle CA, Peacock KA 2003. Headwater stream response to grassland agricultural development in New Zealand. New Zealand Journal of Marine and Freshwater Research 37:389-403.

Rutherford JC 1985. Some thoughts on trend analysis. Pp 39-53 in Biological monitoring in freshwaters: proceedings of a seminar, edited by Pridmore RD, Cooper AB. National Water and Soil Conservation Authority, Wellington. 39-53.

*Scarsbrook MR, Boothroyd IKG, Quinn JM 2000a. New Zealand's national river water quality network: long-term trends in macroinvertebrate communities. New Zealand Journal of Marine and Freshwater Research 34(2):289-302.

*Scarsbrook MR, Quinn JM, Boothroyd IKG 2000b. New Zealand's national river quality network – Reply: a balanced approach to interpretation of data. New Zealand Journal of Marine and Freshwater Research 34(4):746-749.

*Scarsbrook MR, Halliday J 1999. Transition from pasture to native forest land-use along stream continua: effects on stream ecosystems and implications for restoration. New Zealand Journal of Marine and Freshwater Research 33(2):293-310.

*Scott D, White JW, Rhodes DS, Koomen A 1994. Invertebrate fauna of three streams in relation to land use in Southland, New Zealand. New Zealand Journal of Marine and Freshwater Research 28:277-290.

Smith DG, McBride GB, Bryers GG, Wisse J, Mink DFJ 1996. Trends in New Zealand's national river water quality network. New Zealand Journal of Marine and Freshwater Research 30(4):485-500.

Snelder TH, Biggs BJF 2002. Multiscale river environment classification for water resources management. Journal of the American Water Resources Association 38(5):1225-1239.

Snelder T, Biggs B, Weatherhead M 2004. New Zealand river environment classification guide. NIWA report prepared for the Ministry for the Environment. 145 p.

*Stark JD 1985. A macroinvertebrate community index of water quality for stony streams. Water & Soil Miscellaneous Publication 87: 53 p. (National Water and Soil Conservation Authority, Wellington, New Zealand).

*Stark JD 1993a. Ammonia-urea plant biomonitoring – is it defensible? Pp. 343-350 in Proceedings of the Regional Resource Futures Conference, Hamilton (27–29 August 1991), edited by Boyce W, Ericksen N, Hingley N. Centre for Environmental and Resource Studies, University of Waikato, New Zealand.

*Stark JD 1993b. Performance of the Macroinvertebrate Community Index: effects of sampling method, sample replication, water depth, current velocity, and substratum on index values. New Zealand Journal of Marine and Freshwater Research 27:463-478.

*Stark JD 1998. SQMCI: a biotic index for freshwater macroinvertebrate coded abundance data. New Zealand Journal of Marine and Freshwater Research 32:55-66.

Stark JD, Boothroyd IKG, Harding JS, Maxted JR, Scarsbrook MR 2001. Protocols for sampling macroinvertebrates in wadeable streams. New Zealand Macroinvertebrate Working Group Report No. 1. Prepared for the Ministry for the Environment. Sustainable Management Fund Project No. 5103. 57 p.
http://www.cawthron.org.nz/coastal-freshwater-resources/downloads.html
http://www.arc.govt.nz/arc/environment/water/publications/publications_home.cfm#ecology

*Stark JD, Fowles CR 2006. An approach to the evaluation of temporal trends in Taranaki state of the environment macroinvertebrate data. Prepared for Taranaki Regional Council. Cawthron Report No. 1135. 88 p.

*Stark JD, Maxted JR 2004. Macroinvertebrate community indices for Auckland's soft-bottomed streams and applications to SOE reporting. Prepared for Auckland Regional Council. Cawthron Report No. 970. Cawthron Institute, Nelson. ARC Technical Publication 303. 59 p.
http://www.arc.govt.nz/arc/environment/water/publications/publications_home.cfm#ecology

*Stark JD, Maxted JR 2007. A biotic index for New Zealand's soft-bottomed streams. New Zealand Journal of Marine and Freshwater Research 41(1).

*Stark JD, Moore SC, Berry A 1999. Biotic indices: their role in water management in New Zealand. Pp. 43-67 in The use of macroinvertebrates in water management, edited by New Zealand Macroinvertebrate Working Group, Ministry for the Environment, Wellington, New Zealand.

*Storey RG, Cowley DR 1997. Recovery of three New Zealand rural streams as they pass through native forest remnants. Hydrobiologia 353:63-76.

*Suren AM, Biggs BJF, Duncan MJ, Bergey EA, Lambert PW 2003. Benthic community dynamics during summer low-flows in two rivers of contrasting enrichment. 2: invertebrates. New Zealand Journal of Marine and Freshwater Research 37:71-83.

Suren A, Snelder T, Scarsbrook M 1998. Urban stream habitat assessment method (USHA). NIWA Client Report No. CHC98/60. 63 p + appendices.

Towns DR 1985. Life history patterns and their influence on monitoring invertebrate communities. Pp. 225-239 in Biological monitoring in freshwaters: proceedings of a seminar, Hamilton, 21-23 November 1984, edited by Pridmore RD, Cooper AB. Water and Soil Miscellaneous Publication 83.

Taranaki Catchment Commission 1984. Freshwater biology: Taranaki ring plain water resources survey. Taranaki Catchment Commission, Stratford, New Zealand. 196 p + appendices.

*Taranaki Regional Council 2003. Taranaki – our place, our future: report on the state of the environment of the Taranaki region, Taranaki Regional Council, Stratford. 206 p.

United States Environmental Protection Agency 2006. Draft wadeable streams assessment: a collaborative survey of the nation's streams. Report No. EPA 841-B-06-002, USEPA Office of Water, Washington DC. 82 p + appendices.

Vant B, Smith P 2004. Trends in river water quality in the Waikato Region, 1987–2002. Environment Waikato Technical Report 2004/02. http://www.ew.govt.nz/publications/technicalreports/documents/TR04-02.pdf

Verneaux V, Verneaux J, Schmitt A, Lovy C, Lambert JC 2004. The Lake Biotic Index (LBI): an applied method for assessing the biological quality of lakes using macrobenthos: the Lake Chalain (French Jura) as an example. Annales de Limnologie – International Journal of Limnology 40(1):1-9.

Washington HG 1984. Diversity, biotic and similarity indices: a review with special reference to aquatic ecosystems. Water Research 18:653-694.

*Winterbourn MJ 1999. Monitoring for sustainable river ecosystem management and the role of macroinvertebrates. Pp. 1-15 in The use of macroinvertebrates in water management: recommendations of the New Zealand Macroinvertebrate Working Group. Ministry for the Environment, Wellington, New Zealand.

Winterbourn MJ, Alderton P, Hunter GG 1971. A biological evaluation of organic pollution in the lower Waimakariri River system 1970-71. New Zealand Marine Department Fisheries Technical Report 67. Pp. 1-69.

Winterbourn MJ, Gregson KLD 1981. Guide to the aquatic insects of New Zealand. Bulletin of the Entomological Society of New Zealand 9. 80 p..

*Winterbourn MJ, Gregson KLD, Dolphin CH 2006. Guide to the aquatic insects of New Zealand. Fourth Edition. Bulletin of the Entomological Society of New Zealand 14. 108 p.

Winterbourn MJ, Stark AW 1978. A biological survey of the Kaiapoi River following fellmongery effluent treatment. Mauri Ora 6: 11-21.

*Wright-Stow AE, Winterbourn MJ 2003. How well do New Zealand's stream-monitoring indicators, the Macroinvertebrate Community Index, and its quantitative variant, correspond? New Zealand Journal of Marine and Freshwater Research 37:461-470.

Appendix 1. Calculating the MCI: Excel macros and user-defined functions

The MCI, SQMCI and QMCI can be calculated manually on spreadsheets, but speed and accuracy are improved if macros or user-defined functions are used. Spreadsheets containing these macros and user-defined functions are available from the first author of this report (John Stark).

The following Excel macros enable MCI, SQMCI and QMCI values to be calculated on a spreadsheet that has coded abundance data entered as R, C, A, VA or VVA. Absences can be blank cells, or indicated by a dash [enter '- (open single quote followed by a dash) in the cell – this prevents Excel from treating the dash as a negative sign or subtraction]. The same macros can be used to calculate the MCI-sb, SQMCI-sb or QMCI-sb if soft-bottomed tolerance values are used.

MCI from coded abundance data

```
=SUMIF(D$8:D$135,">@",$C$8:$C$135)/COUNTIF(D$8:D$135,">@")*20
```

MCI from count data

```
=SUMIF(D$8:D$135,">0",$C$8:$C$135)/COUNTIF(D$8:D$135,">0")*20
```

SQMCI from coded abundance data

```
=(SUMIF(D$8:D$135,"=R",$C$8:$C$135)*1+SUMIF(D$8:D$135,"=C",$C$8
:$C$135)*5+SUMIF(D$8:D$135,"=A",$C$8:$C$135)*20+SUMIF(D$8:D$135
,"=VA",$C$8:$C$135)*100+SUMIF(D$8:D$135,"=VVA",$C$8:$C$135)*500
)/(COUNTIF(D$8:D$135,"R")*1+COUNTIF(D$8:D$135,"C")*5+COUNTIF(D$
8:D$135,"A")*20+COUNTIF(D$8:D$135,"VA")*100+COUNTIF(D$8:D$135,"
VVA")*500)
```

QMCI from count data

```
=(SUMPRODUCT($C$8:$C$135,D$8:D$135))/SUM(D$8:D$135)
```

These macros are entered in a cell at the bottom of column D in an Excel spreadsheet. Tolerance values are in column C (C8:C135) and the first column of data is in column D (D8:D135). The absolute cell references (indicated by '$' signs in the above formulae) enable dragging to the right to calculate MCI, SQMCI and QMCI values for data in columns E, F, G,… etc.

User-defined Excel functions for calculating biotic indices

The following Excel user-defined functions enable the MCI, SQMCI and QMCI values to be calculated on a spreadsheet. Coded abundance data are entered as R, C, A, VA or VVA. Absences can be blank cells, or indicated by a dash (enter '- in the cell). These functions can be used to calculate the MCI-sb, SQMCI-sb, or QMCI-sb if soft-bottomed tolerance values are used. Note that the functions require the macro security level in Excel to be set to medium or low.

These user-defined functions must be in memory (*i.e.* part of the spreadsheet on which the calculations are to be made, or in a separate open spreadsheet) in order to be used. To insert a calculated MCI value, place the cursor in the cell where the result is required (say in cell D136 below the first column of data) and click on <insert.function.user defined>. Double-click on the MCI function and enter the range of tolerance values (*e.g.* C8:C135) into the box labelled 'Scores'. This function argument can also be entered by highlighting the range of scores. Manually make these references absolute by adding $ signs (*e.g.* $C$8:$C$135). Click into the box to the right of the 'Data_column' function argument.

Finally, click anywhere on the spreadsheet in the column of data above the cell containing the function (say cell D20). Making the tolerance value references absolute (both row and column) enables additional MCI values to be calculated by dragging to the right below additional columns of data, and SQMCI or QMCI values (depending on the data type in the spreadsheet) can be calculated by dragging the function in D136 down one cell and manually changing the name from MCI to SQMCI or QMCI.

For example, assuming that you have a spreadsheet called "indices.xls" containing the user-defined functions for calculating MCI, SQMCI, and QMCI values, and the tolerance values are in column C (from row 8 to row 135). If you want to calculate the MCI for data in column D in cell D136, then in cell D136 you should have the following (the '20' after the 'D' can be any number).

```
=indices.XLS!MCI($C$8:$C$135,D20)
```

Dragging to the right simply changes the 'D' to 'E', 'F', 'G', etc. and calculates MCI values for data in each column. If you drag down from cell D136 into D137 and change the entry to

```
=indices.XLS!SQMCI($C$8:$C$135,D20)
```

or

```
=indices.XLS!QMCI($C$8:$C$135,D20)
```

depending on the data type, you will get SQMCI or QMCI values. These cells can be dragged to the right too, in order to calculate SQMCI or QMCI values for additional columns of data.

Change the number of significant figures – by convention none for MCI and two for SQMCI or QMCI.

Finally, select the rows containing these index calculations and select <copy> and <paste.special.values> to replace the functions with values. If this is not done, the calculated index values will only be displayed when indices.xls is in memory and will be lost otherwise.

A spreadsheet (indices.xls) containing user-defined functions for calculating MCI, SQMCI, and QMCI values is available from the first author of this report.

MCI from presence-absence, coded abundance or count data

```
Option Base 1
Dim score(), data$(), column()

Function MCI(scores, data_column)
'This function calculates MCI from presence-absence, count or coded
abundance data

ReDim score(scores.Count), data$(scores.Count), column(scores.Count)

Row = scores.Row
ScoresCol = scores.column
DataCol = data_column.column
ScoringTaxa = 0

For i = 1 To scores.Count
    ' Read in the column of MCI taxa scores
    score(i) = Cells(Row, ScoresCol)
    ' Read in the column of counts or coded abundance (i.e.
R/C/A/VA/VVA) data
    data$(i) = Cells(Row, DataCol)
    Row = Row + 1
    ' test to see if counts or coded abundances
    ' blank cells and dashes (to indicate absent taxa) are < "0"
    ' and replace data values with 0 (for absent) or 1 (for present)
    If data$(i) > "0" Then column(i) = 1 Else column(i) = 0
    ' only add to scoring taxa if there is an entry in the range 1 -
10 inclusive
    ' in the taxa scores column
    If score(i) > 0 And score(i) < 11 Then ScoringTaxa = ScoringTaxa
+ column(i)

    MCI = MCI + (score(i) * column(i))

Next i

    MCI = MCI / ScoringTaxa * 20

End Function
```

SQMCI from coded-abundance data

```
Option Base 1
Dim scores(), column()
Function SQMCI(scores, data_column)
' This function calculates SQMCI from coded abundance data

' Criteria for assigning coded abundances when sorting samples
' R = 1 - 4 animals
' C = 5 - 19
' A = 20 - 99
' VA = 100 - 499
' VVA = 500+

ReDim score(scores.Count), data$(scores.Count)

Row = scores.Row
ScoresCol = scores.column
DataCol = data_column.column

For i = 1 To scores.Count
    score(i) = Cells(Row, ScoresCol)
    data$(i) = Cells(Row, DataCol)
    Row = Row + 1
    ScoringTaxa = ScoringTaxa + Sgn(Val(data$(i)))

    Select Case data$(i)
        Case "R"
            R = R + score(i) * 1
            TotalR = TotalR + 1
        Case "C"
            C = C + score(i) * 5
            TotalC = TotalC + 5
        Case "A"
            A = A + score(i) * 20
            TotalA = TotalA + 20
        Case "VA"
            VA = VA + score(i) * 100
            TotalVA = TotalVA + 100
        Case "VVA"
            VVA = VVA + score(i) * 500
            TotalVVA = TotalVVA + 500
        Case Else
    End Select
Next i
    TotalCounts = TotalR + TotalC + TotalA + TotalVA + TotalVVA
    SQMCI = (R + C + A + VA + VVA) / TotalCounts

End Function
```

QMCI from count data

```
Option Base 1
Dim scores(), column()

Function QMCI(scores, data_column)

'This function calculates QMCI from count data

ReDim score(scores.Count), data(scores.Count)

Row = scores.Row
ScoresCol = scores.column
DataCol = data_column.column

For i = 1 To scores.Count
    score(i) = Cells(Row, ScoresCol)
    data(i) = Cells(Row, DataCol)
    Row = Row + 1
    If InStr(data(i), "-") > 0 Then data(i) = 0
    If score(i) > 0 And score(i) < 11 Then
        ScoringTaxa = ScoringTaxa + Sgn(data(i))
        TotalCounts = TotalCounts + data(i)
    End If
Next i

For i = 1 To scores.Count
       QMCI = QMCI + (score(i) * data(i) / TotalCounts)
Next i

End Function
```

Appendix 2. Use of MCI, QMCI and SQMCI throughout New Zealand

Use of the MCI and QMCI is well established in New Zealand, possibly because, together with the SQMCI, they are the only comprehensive biotic indices based on tolerances of New Zealand macroinvertebrate taxa for assessing the health of stony streams. All regional councils that undertake SoE monitoring use the MCI and/or SQMCI/QMCI for reporting results. In this appendix we discuss the major findings from the peer-reviewed scientific literature concerning these indices.

A chapter in the "Stream invertebrate book" examines the first 50 years' use of macroinvertebrates in biological monitoring in New Zealand. Boothroyd & Stark (2000) define the term "biomonitoring" and provide a useful discussion on the theory and practice of biomonitoring, including why macroinvertebrate biomonitoring is preferred over other kinds. Some of the historical events that have influenced biological monitoring in New Zealand are detailed. Various methods used for assessing the biological condition of rivers and streams are discussed, and the use of the macroinvertebrate monitoring tool currently favoured in New Zealand, the MCI, is critically examined. Comments on the design of monitoring programmes and data analyses are provided.

Although developed originally in Taranaki for stony streams (Taranaki Catchment Commission 1984; Stark 1985), both the MCI and QMCI were found to be moderately strongly correlated with indicators of enrichment when applied to run/riffle samples from 88 rivers throughout New Zealand (Quinn & Hickey 1990). Quinn & Hickey suggested that the MCI may be a more sensitive index of water enrichment than the QMCI, because it has higher correlations with indicators of enrichment (such as total Kjeldahl nitrogen, periphyton chlorophyll '*a*' and ash-free dry weight) for 88 New Zealand rivers. However, they suggested that the extra effort required (mainly in sample processing) to obtain QMCI values may be warranted where water quality changes are expected over relatively short river reaches (*e.g.* above and below wastewater discharges). In such situations, drift of macroinvertebrates from upstream may introduce taxa (normally in low densities) to polluted downstream sites, where they may not survive in the long term, thereby misrepresenting the "true" character of the site. Quinn & Hickey concluded that these indices were more useful indicators of water quality than species diversity, species richness and the EPT index (*i.e.* the number of ephemeropteran, plecopteran and trichopteran taxa). This was the first validation of these indices nationwide.

In Southland streams, Quinn *et al.* (1992) found that the QMCI was reduced significantly by intensive grazing and channelisation, and Scott *et al.* (1994) found that increasing intensity of pastoral land development decreased the QMCI. At 29 Northland stream sites, Collier (1995) found significant positive correlations between MCI and shade and the percentage area of native forest in the catchment, but significant negative correlations with an index of periphyton biomass and the percentage of the riparian zone in pasture.

Maloney (1995) found that spraying the herbicide triclopyr to control willows and lupins in the Ahuriri River (South Island) did not alter macroinvertebrate community composition or affect

MCI and QMCI values. Quinn *et al.* (1997a) found that pasture streams had significantly lower QMCI than native forest and pine forest catchments at Whatawhata, near Hamilton. Storey & Cowley (1997) found that native forest remnants part-way down second-order streams in pastoral farmland could cause a change from a more enrichment-tolerant macroinvertebrate fauna to one more characteristic of clean water, and that this improvement was detected by the MCI.

Collier *et al.* (1998) evaluated the performance of biotic indices (including the MCI and QMCI) calculated from samples collected from macrophyte, sand, silt, bedrock, and wood substrates at 20 Waikato lowland stream sites. They found that %EPT and MCI were robust under contrasting sampling intensities, and, together with the QMCI, were sensitive to factors relating to water quality and catchment land use, suggesting that these indices are likely to be useful for biomonitoring in lowland stream environments. The MCI was considered particularly suitable for use with rapid bioassessment protocols in lowland streams.

Hickey & Clements (1998) criticised the QMCI because it did not detect the impacts of heavy metal pollution on macroinvertebrate communities in streams on the Coromandel Peninsula. They noted that this was because the QMCI had "incorrect tolerance scores for some taxa to heavy metals". However, this is hardly surprising since the QMCI was developed to detect organic pollution and nutrient enrichment, not metal toxicity.

Hall *et al.* (2001) examined macroinvertebrate communities in streams and rivers dominated by native bush, agricultural or urban land uses within the Water of Leith stream catchment near Dunedin. Both the MCI and QMCI decreased progressively from native bush through agricultural to urban land use, with the QMCI exhibiting the stronger relationship. These results indicate that not only did the representation of pollution tolerant taxa increase along this gradient, but that pollution-tolerant taxa also increased in dominance.

Duggan *et al.* (2002) examined the influence of sample size (100-, 200- and 300-fixed count) on the accuracy and variability of six invertebrate metrics (taxa richness, EPT_{taxa}, %EPT abundance, % dominant taxon, MCI and QMCI). They found that the MCI provided the most consistent results in terms of having low within-site variability and distinguishing differences in stream impacts between sites (although this was affected by seasonal interactions). Although their study reinforced the use of a range of metrics, they suggested that "better performing metrics such as the MCI, could be given higher weighting than some other metrics when interpreting results." Within the range of sample sizes tested, they found that richness measures (including the MCI) were sensitive to sample size, and that one should be cautious when comparing results from different studies. However, within any particular study the effect of sample size on the interpretation of MCI was not considered likely to be significant.

Duggan *et al.* (2003) also compared the performance of 100-, 200- and 300-fixed count subsampling with coded abundance (R, C, A, VA, VVA) subsampling, and full counts for rapid assessment biomonitoring. They found 1:1 relationships between QMCI and %EPT for assessments made with both rapid assessment methods and full counts. However, they

concluded that variability was greater using coded abundance than fixed counts, which they considered could lead to incorrect conclusions "on occasion". These authors used only two data sets in their analyses (from Quinn *et al.* 2002, and Quinn & Hickey 1993), and recommended the 200-fixed count as the preferred rapid assessment sampled processing protocol (which is protocol P2 in Stark *et al.* 2001). We have done calculations using other data sets that do not support the conclusion that fixed count is better than coded abundance, suggesting, perhaps, that the nature of the data may affect the results. The critical issue, however, that Duggan *et al.* (2003) did not consider was the relative time and cost of the different processing protocols. In our experience, coded abundance is usually more cost-effective.

Death *et al.* (2003) examined the effect of exotic forest logging on stream macroinvertebrate communities in Hawke's Bay. They found that the MCI and QMCI reflected the impact of forest harvesting – the main impact was considered to be an increase in fine sediment in the streambed – and that communities had not recovered to pre-harvest condition by 1.5–2.5 years post-harvest. In contrast, recovery from a natural storm event was much more rapid (five months).

Parkyn *et al.* (2003) and Parkyn & Davies-Colley (2003) evaluated the success of stream rehabilitation by comparing agricultural streams where riparian buffers have been restored with unbuffered control reaches upstream or nearby. They found that buffer widths appeared to be closely related to stream health (QMCI), especially those greater than 10 m, although they were unsure of the mechanism for this and suggested that further study was warranted. However, recovery seemed closely related to reducing stream temperatures, suggesting that restoring in-stream communities might take many years and would be achieved only after canopy closure, with long buffer lengths and protection of upstream tributaries.

Wright-Stow & Winterbourn (2003) examined the correspondence between the MCI and QMCI using fixed-count data from 230 stream and river sites in Canterbury. The two indices ranked sites similarly (r_s = 0.86), but the MCI placed most sites in the "good" and "fair" pollution classes, whereas most sites were assigned to the "excellent" or "poor" classes by the QMCI. Wright-Stow & Winterbourn concluded that either the MCI was a more conservative index, or that the boundaries between pollution classes are not equivalent. The latter reason was considered more likely, and given the difficulties inherent in defining classes based on continuous distributions and the fact that there is no way of knowing which index gives the "right" answer, Wright-Stow & Winterbourn suggested a return to fuzzy boundaries between classes as proposed initially by Stark (1985). Alternatively, when comparing large numbers of sites (*e.g.* in SoE monitoring), they suggested that the percentile the site of interest falls within could be stated. A site with an MCI of 130, for example, could be described as being within the top 10% of sites within the region.

Quinn *et al.* (2004) found that riparian buffers mitigate the effects of pine plantation logging in New Zealand streams and that QMCI values were significantly lower following logging if riparian vegetation was removed too.

Maxted *et al.* (2003) studied the effects of sample substrata, sample area, and land use on various metrics, including the MCI and SQMCI. They found that the MCI and SQMCI were sensitive to urban and rural land uses. The differences in metric scores between streams with hard-bottomed and soft-bottomed substrata were used as the basis for developing separate sampling methods for these two stream types (Stark *et al.* 2001).

Riley *et al.* (2003) studied the consequences for stream physico-chemistry and ecology of the agricultural (pastoral) development of tussock grassland. They found no difference in MCI between tussock, grazed tussock, and pasture catchments. However the QMCI was significantly higher in streams in pasture catchments than in tussock or grazed tussock catchments. Despite changes to the physical and chemical nature of the streams due to pastoral development no sensitive taxa had been lost and, in fact, the relative abundances of some sensitive taxa such as the mayfly *Deleatidium* had increased in pastoral streams. All MCI and QMCI values in the study streams were indicative of good or very good stream health.

Maxted *et al.* (2005) studied the effects of in-line ponds on stream water quality and macroinvertebrate communities in the Auckland region. They found that ponds in soft-bottomed geology in rural catchments caused reductions in the EPT richness and SQMCI indices downstream, but in hard-bottomed geology bushed catchments there was no significant difference.

Doledec *et al.* (2006) examined the performance of structural and functional approaches for assessing land-use effects on stream macroinvertebrate communities along a gradient of increasing agricultural development: ungrazed native tussock (UT), grazed tussock (GT), extensively grazed pasture (PA), and intensive deer and dairy farming (DD). Macroinvertebrate densities, EPT_{taxa}, MCI and QMCI differed very little among UT, GT, and PA sites, but densities were somewhat higher and the indices significantly lower at DD sites. The MCI was correlated significantly with the percentage of the streambed covered by sediment particles <1 mm (%FINES), dissolved inorganic nitrogen (DIN) and dissolved reactive phosphorus (DRP), whereas the QMCI was correlated significantly with %FINES and DIN. The authors concluded overall that these traditional structural measures (*i.e.* biotic indices) were just as effective as the species traits approach for differentiating land-use effects on their grassland stream communities. They predicted that the functional species trait approach may be more effective on a larger spatial scale, but this has not yet been tested.