stark environmental

Freshwater ecology specialists

Trends in river health of the Manawatu-Wanganui region 2008 with comments on the SoE biomonitoring programme

Stark Environmental Report No. 2008-07

August 2008

Cover Photograph: Whanganui River at Te Maire (Ohinepa Reserve). J.D. Stark, 11 December 1988.

stark environmental

Freshwater ecology specialists

Trends in river health of the Manawatu-Wanganui region 2008 with comments on the SoE biomonitoring programme

John D. Stark

Prepared for

Horizons Regional Council

Stark Environmental Limited
P.O. Box 1831
Nelson 7040, New Zealand.
Ph. +64 3 545 1766

StarkEL@paradise.net.nz

Recommended citation:
Stark JD 2008. Trends in river health of the Manawatu-Wanganui region 2008 with comments on the SoE biomonitoring programme. Prepared for Horizons Regional Council. Stark Environmental Report No.2008-07. 61p.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

EXECUTIVE SUMMARY

This report presents the results of trends testing for macroinvertebrate State of the Environment (SoE) sites in the Manawatu-Wanganui region. Although 78 different SoE monitoring sites have been sampled since monitoring began in 1999, and 35 in 2007, this report only considers data from 21 sites that have been sampled on 6 – 9 occasions. Trends testing is not considered worthwhile for time series comprising fewer than six occasions.

The non-parametric Mann-Kendall test detected three statistically significant trends in the Macroinvertebrate Community Index (MCI) (two +ve and one -ve) and six significant trends in the Quantitative Macroinvertebrate Community Index (QMCI) (three +ve and three -ve). Positive trends in MCI were detected for the Whanganui River (Te Maire) and the Makakahi River (Konini), and a negative trend for the Manawatu River (Maxwell's Line / Teachers' College). Positive trends in QMCI were detected for the Whanganui River (Te Maire, Pipiriki, & Estuary). Two sites in the Oroua River showed negative trends in QMCI, as did the Manawatu River at Maxwell's Line / Teachers' College. However, all trends were comparatively weak and were eliminated by the application of the Benjamini-Hochberg False Discovery Rate (FDR) procedure. The FDR procedure is required when multiple correlations are undertaken all at one time, because some of them could have arisen by chance. It has the effect of eliminating the weakest correlations. In future, as more sampling is undertaken and timer series lengthen, the power of trends testing is expected to increase and some of these trends are likely to remain significant after the FDR analysis.

In addition to testing for trends, SoE biomonitoring data also permit an evaluation of the current state of river health. The simplest method is to use a biotic index and the MCI is recommended. Of the 21 sites examined in this report, the MCI suggested that one site had 'excellent' stream health, five were 'good', 12 were 'fair', and three were 'poor'. Only one site (Mangawhero River @ DOC Headquarters) remained in the same quality class ('excellent') for the entire sampling period. All other sites have fluctuated between quality classes over the years (some over three classes). The 'poor' sites were located on the lower Manawatu River (Karere Road / 42 Mile Hydro Station & Whirokino) and the Oroua River (Awahuri Bridge). Conditions at these sites may be able to be improved once the causes of the degradation(and practical means to deal with them) are identified, provided there is sufficient political will to initiate and undertake remediation. Between one and three 'poor' MCI values (i.e., < 80) have been recorded at nine other sites over the 6-9 years of SoE monitoring. These sites may also warrant attention to determine the cause/s of these low values.

It is unrealistic to expect the lower reaches of long rivers to yield the same 'excellent' MCI values as their headwater streams - some decrease in MCI with distance downstream in a river is inevitable. Any sites that are classified as 'poor' or 'fair' or any sites showing negative trends in stream health are likely to have room for improvement. However, it is also possible that some 'good' sites have been degraded and require remediation to return them to 'excellent' condition. Despite the existence of numerical methods for assessing river health (such as quality classes based on the MCI), inevitably the decision concerning whether river health is acceptable or not is one for water managers to make. In the end it is a value judgment. Data from water quality monitoring, consent compliance monitoring programmes, and knowledge of land-use, industrial activity, and urbanisation within the catchment must be evaluated when deciding whether river health at a particular location meets the desired standard.

This report also reviews aspects of Horizons' macroinvertebrate SoE monitoring programme and makes the following recommendations for future monitoring.

1. Sampling should comprise collection of single hand-net samples from each monitoring site.

2. The coded-abundance sample processing protocol should be used, although a 200 fixed count (with scan for rare taxa) would be an acceptable alternative.

3. The above methods will provide data suitable for calculating MCI, taxon richness, % EPT richness, and SQMCI / QMCI.

4. However, the MCI should be the primary index for reporting SoE monitoring results. The QMCI is more suitable for compliance monitoring and should not be used. The use of multiple indices can be confusing for laypersons and should be avoided.

5. The suggested change of sampling method and processing protocol should not have significant implications for the integrity of the existing database and data time series. All indices from historical data may need to be re-calculated based on pooled data from the five replicate Surber samples. This will ensure comparability with data from single hand-net samples (which are equivalent in effort to five Surber samples).

6. The SoE programme should continue with sampling once per year, but it is suggested that the number of sites should be increased from 35 (2007) to 50 - 60. Sites that have already been sampled several times should have priority for inclusion. Sites should encompass a range of land-uses (from reference condition to highly impact urban or rural), river types, and be distributed throughout the region.

7. The change in sampling and sample processing methods should reduce the overall costs of field work and sample processing dramatically, so that the overall programme can be expanded, and costs should remain well within existing monitoring budgets.

INTRODUCTION

Regional Councils are responsible for undertaking long term biological monitoring of freshwater systems in order to assess the "health" or state of the environment within their regions. The purposes of State of the Environment (SoE) monitoring are to:-

- Obtain representative data for freshwater macroinvertebrates biodiversity in the region;
- Detect the presence and direction of trends;
- Provide baseline information to policy, planning, consents and compliance departments within the council;
- Identify the effects of activities – particularly land use changes and point source discharges – on water quality;
- Determine the effectiveness of management initiatives directed at enhancing water quality.

One question that cannot be addressed until SoE monitoring has been undertaken for several years is whether or not conditions have gotten worse, improved, or stayed the same. This is a typical application for a class of statistics called time series or trends analysis. This report examines trends in biotic indices now that Horizons has undertaken SoE monitoring for nine years.

There are a variety of different techniques for time series analyses that vary in their data requirements and complexity. Linear regression-based parametric methods should only be applied when the trend is expected to be linear (which it often is not) unless data transformations are used. Furthermore, there are problems with parametric methods like 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 found favour 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 for water quality data by Bill Vant[1] (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 - 60 data points). This seldom is the case for biotic data where sampling may be seasonal at best, and is more often undertaken only once or twice per 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 (Bill Vant, Environment Waikato, *pers. comm.*).
Statistical testing of trends in biotic indices is a recent development – especially in New Zealand – because few consistent data sets have existed. Collier & Kelly (2006) (see also Collier 2006) employed a non-parametric method based on Spearman rank correlations to

[1] http://www.ew.govt.nz/enviroinfo/water/healthyrivers/waikato/documents/SeasonalKendallTrendAnalysis.xls

detect trends in the health of rivers and streams in the Waikato region. At the same time Stark & Fowles (2006) developed an approach based on the Mann-Kendall test for examining trends in the health of Taranaki Rivers and streams. The method can be applied to any time series including biotic indices such as the Macroinvertebrate Community Index (MCI) and Quantitative Macroinvertebrate Community Index (QMCI) (Stark 1985, 1993). The report by Stark & Fowles (2006) was reviewed and the approach used was validated by Graham McBride (NIWA, Hamilton). Both approaches gave similar results (Stark & Fowles 2006).

Horizons Regional Council has undertaken macroinvertebrate SoE monitoring at selected sites throughout the region since 1999, with some sites sampled on up to nine occasions. This report uses the trends testing approach developed by Stark & Fowles (2006) to assess aquatic ecosystem state and trends based on the MCI and QMCI biotic indices for sites in the Manawatu-Wanganui region that have been sampled on six or more occasions.

In addition to detecting trends that may be statistically significant, sites that are in 'good' or 'poor' condition are identified.

Horizons Regional Council has not sampled all SoE monitoring site annually. For example, some sites have been sampled every three years. The pros and cons of alternative SoE monitoring strategies are discussed and recommendations are made for an ongoing monitoring strategy following 10 years of data collection.

METHODS

Timeline of SoE monitoring in Manawatu-Wanganui

Horizons Regional Council commenced SoE monitoring in 1999. Dr Russell Death (Massey University) advised on the initial design of the monitoring programme, and prepared the first annual SoE monitoring report (Death 1999). In subsequent years until 2007 the macroinvertebrate sample processing and preparation of the annual biomonitoring reports have been contracted to Massey University.

The timeline below (supplied by Carol Nicholson, Horizons) summarises the initial setup (1999) and notes (for 2000 – 2007) the changes that have occurred over the period of SoE monitoring.

1999 Biomonitoring programme initiated – (Dr Russell Death)
Invertebrate sampling - 5 replicate 0.1 m^2 Surber Samples (250µm mesh) taken at each site in riffles during normal flows (May – November). Full counts.
Biotic Indices - Number of taxa, Number of animals, MCI.
Habitat variables– Conductivity, temperature, dissolved oxygen, pH, depth, width, velocity, % substrate composition (Wolman Walk), channel cover, riparian vegetation, embededness, habitat stability (Pfankuch – 15 attributes) hydrological data included

Periphyton – (4 stones, 3 scalpel scrapings) Foil used to measure stone surface area. Biomass and algal community composition measured.

2000 (Massey University - Dr Russell Death, Sjaan Charteris, Kirsty Francis, Stephen Minchin, Rachel Boisen, Ashley Vosper)
Sampling carried out in low flow conditions (February – June)
Hydrological data not included
Biotic indices - QMCI, %EPT taxa, % EPT animals added
Three dimensions of stone used to calculate surface area.

2001 (Massey University - Dr Russell Death, Tanya Cook, Kirsty Francis, Mark Hamer, Carol Nicholson)
Sampling carried out in low flow conditions (February – April)
Stream habitat assessment form added

2002 (Massey University - Dr Russell Death, Kate McArthur, Richard Pedley, Ian Johnston, Zoe Dewson)
Sampling period same as previous year (February – April)
5 stones collected for periphyton analysis
Biotic Index - O/E taxon richness added.

2003 (Massey University - Russell Death, Troy Makan, Kiryn Weaver, Erna Zimmerman)
Sampling carried out in low flow conditions (March – May)
Biotic index - O/E taxon richness ratio added.

2004 (Massey University - Russell Death, Fiona Death, Rebecca Lewis)
Sampling carried out in low flow conditions (March - May)
Corrections made assuming only ½ of stone exposed to light and therefore suitable for periphyton growth.
Stream habitat assessment form discontinued.

2005 (Massey University - Russell Death, Fiona Death)
Sampling carried out in low flow conditions (February – July)
Pfankuch (1975) index – only the bottom section of channel stability index (last 5 attributes) used to assess stability.
Algal community composition assessment ceased after 2005 and only biomass and/or visual assessment continued.

2006 (Massey University - Russell Death, Fiona Death)
Sampling carried out in low flow conditions (February – March)
Dissolved oxygen measurements discontinued.
Added SHMAK visual assessment methodology for periphyton cover.
Periphyton scalpel scrapings of rocks discontinued (replaced by SHMAK).

2007 (Massey University – Zoe Dewson, Fiona Death, Russell Death)

Sampling carried out in low flow conditions (January – March)

Macroinvertebrate sampling methods

Since monitoring began in 1999 macroinvertebrate sampling has comprised collection of five replicate Surber samples (area 0.1 m^2, mesh 250 µm) from riffle habitat at each sampling site. Samples were preserved in 10% formalin[2].

Sampling sites

Horizons Regional Council has undertaken SoE monitoring of macroinvertebrates in the Manawatu- Wanganui region since 1999 with some sites sampled annually and others every third year. Figure 1 is a map of site locations with the size and colour of the symbols indicating the number of occasions each site has been sampled. Most key sites have remained the same over the past nine years apart from two of the permanent sites (Manawatu @ Maxwells Line and Makakahi @ Konini) which were moved to better align with Horizons hydrology monitoring sites in 2006. Data for the Manawatu River at Maxwells Line and the Manawatu River at Teachers' College have been combined for analyses, and since 2006 the latter site has been the preferred sampling location. These sites are only 4 km apart with no apparent impacts occurring between them (Carol Nicholson, Horizons, *pers. comm.*)

A total of 78 different sites have been sampled over the years with 14 sites sampled on all nine occasions (1999-2007). Twenty-one sites sampled on six or more occasions were selected by Horizons Regional Council for trends testing (Table 1).

A number of biological indices have been calculated from the macroinvertebrate samples to reflect water quality or stream health. These are:-

- the MCI, where each macroinvertebrate taxon present at the site is assigned a tolerance value according to its tolerance to nutrient enrichment (Stark 1985, 1993, Stark & Maxted 2007a);
- %EPT Taxa which measures the proportions of mayflies, stoneflies and caddisflies, (three groups of insects that are sensitive to pollution);
- %EPT individuals which measures the proportion of individuals collected that are mayflies, stoneflies and caddisflies;
- Number of Taxa which measures the variety of different taxa recorded.
- Observed over expected (O/E) taxa richness (Joy & Death 2003) which is a measure of how many invertebrates are at a site that would be expected to occur according to a predictive model based on catchment-scale habitat variables.

[2] Note that formalin is very hazardous substance as is out of favour as a fixative for invertebrate samples Instead samples should be preserved with ethyl alcohol at a final concentration of 70-90% in the sample pottle. A cost-effective form is 'ethanol solution' (e.g., Mobil SDA-3A) (Stark et al. 2001). Glyoxal (1-4%) can be added to the preservative as a fixative to improve retention of colour and shape of invertebrates. Glyoxal is a non-vaporable (odourless) form of formaldehyde that is much safer and more pleasant to use (but equally effective at fixing tissue it comes in contact with).

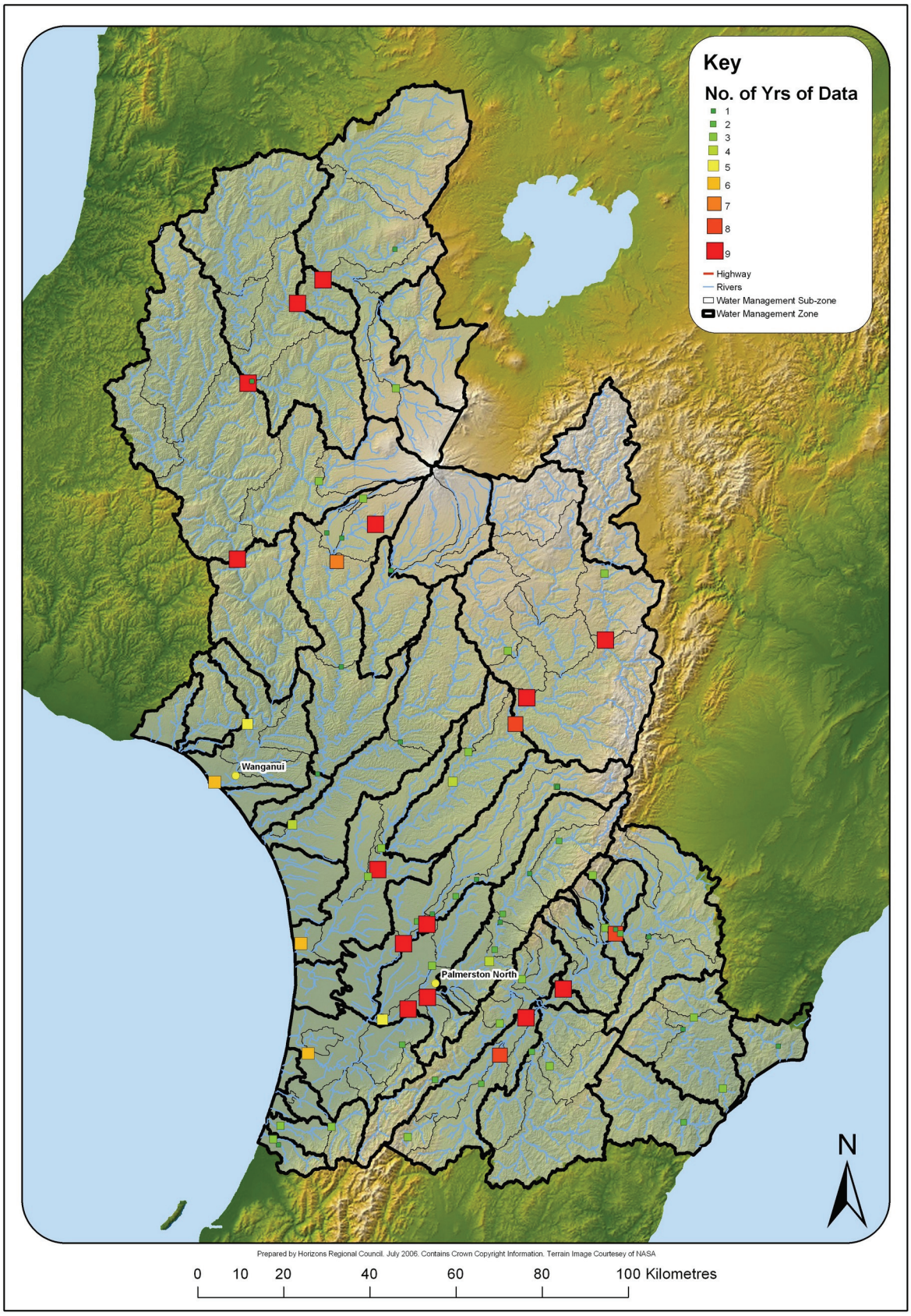


Figure 1 Macroinvertebrate biomonitoring sites in the Manawatu-Wanganui region with an indication of the number of sampling occasions.

- the QMCI, which weights the tolerance values according to percentage community composition and is more sensitive to subtle changes in community composition (Stark 1985, 1993, Stark & Maxted 2007a);

Table 1 SoE monitoring sites in the Manawatu-Wanganui region that have been sampled on six or more occasions since 1999.

River	Site	N	Easting	Northing
Manawatu catchment				
MANAWATU RIVER	Hopelands Reserve	9	2761500	6089800
MANAWATU RIVER	Maxwells Line / Teachers' College	9	2729937	6087844
MANAWATU RIVER	Karere Rd / 42 Mile Hydro Station	9	2725462	6085118
MANAWATU RIVER	Whirokino	6	2702200	6074700
OROUA RIVER	Upstream MBP discharge @ Nelson St	9	2729800	6104800
OROUA RIVER	Awahuri bridge	9	2724400	6100300
MAKAKAHI RIVER	Konini	8	2746700	6074300
MANGATAINOKA RIVER	SH2 Bridge Mangatainoka	9	2752800	6083100
MANGATERA STREAM	Confluence @ Timber Bay	8	2773600	6102600
Rangitikei catchment				
HAUTAPU RIVER	Upstream Rangitikei River	9	2753000	6157400
RANGITIKEI RIVER	Pukeokahu	9	2771300	6170800
RANGITIKEI RIVER	Mangaweka	8	2750300	6151300
RANGITIKEI RIVER	Kakariki	9	2718400	6117500
RANGITIKEI RIVER	Estuary	6	2700500	6100300
Whanganui catchment				
WHANGANUI RIVER	Cherry Grove	9	2705700	6254500
WHANGANUI RIVER	Te Maire	9	2699800	6249000
WHANGANUI RIVER	Downstream Retaruke confluence	9	2688300	6230500
WHANGANUI RIVER	Pipiriki	9	2685800	6189600
WHANGANUI RIVER	Estuary	6	2680500	6137800
Whangaehu catchment				
MANGAWHERO RIVER	DOC Headquarters	9	2717900	6197700
MANGAWHERO RIVER	Downstream Makotuku	7	2708894	6189024

SoE monitoring results have been reported annually (Death 1999; Charteris *et al.* 2000; Cook *et al.* 2001; Death *et al.* 2002, 2003, 2004, 2005, 2006; Dewson *et al.* 2007).

For this review, Horizons Regional Council decided to focus only on the MCI and QMCI. Calculated MCI and QMCI values were provided for the sites listed in Table 1. Raw macroinvertebrate data were not provided for analysis. PDF copies of previous monitoring reports were provided also.

In addition, six-replicate Surber sample data were available from surveys I undertook in 1988-

89 for the Wanganui Low Flow Appeal hearing. Three Whanganui River sites, sampled in May 1989, were in common to SoE monitoring sites that had been sampled six or more times (viz., Cherry Grove, Te Maire, and Pipiriki). Time series analyses for these sites were conducted with and without these additional data.

Trends testing methodology

The following procedure, developed by Stark & Fowles (2006) was used to examine trends in river health (MCI & QMCI) for 21 sites in the Manawatu-Wanganui region.

1. **Visualise the trend**: Scatterplots of biological index vs time with LOWESS fit. Stark & Fowles (2006) recommended a tension of 0.4 (which was suitable for their data sets comprising (mostly) 15 or more data points). With 6 – 9 data points, a tension of 0.7 provides a more appealing smooth and is used in this report.
2. **Test for significance using the Mann-Kendall test**: When multiple tests are required we recommend 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. If the number of data values (n) for each site is similar, then the p-value resulting from the test will provide a reliable indication of the strength of trends. Given the fact that statistical significance is not the same as ecological significance, explanations should be sought for these clear trends starting with the one with the lowest p-value and working down the list.

Benjamini & Hochberg FDR

Some explanation of the Benjamini & Hochberg (1995) FDR analysis is warranted here. A statistical problem arises when undertaking multiple comparisons. Put simply, when multiple correlations are undertaken there is a chance that some will be found to be significant purely by chance. Put another way, testing a hypothesis for each correlation at a level α (say 0.05) will inflate the overall Type I error[3] rate (McBride 2005). A method for dealing with this problem was advocated by Benjamini & Hochberg (1995). The FDR is defined as the expected proportion of true hypotheses rejected out of the total number of rejections. It considers how many of the α-level rejections may be in error. McBride (2005) describes the method and it can be applied using a computer programme written by Ian Jowett (NIWA, Hamilton) or on an Excel Spreadsheet available from the first author of this report or the following URL.

http://www2.tltc.ttu.edu/Westfall/images/6348/p-vals.xls [Accessed 24 September 2008]

The overall effect of applying the Benjamini-Hochberg FDR method to a table of multiple correlations is that FDR is controlled at the α level. This means that the expected proportion of rejections that are in error is less than α. Thus, the number of correlations deemed to be statistically significant is reduced.

[3] A Type I error is made if we reject a hypothesis when it is true and the risk of this happening is the probability level (normally 0.05 or 0.01).

When testing for significance of a trend from a single site (as opposed to testing multiple sites) FDR analysis is not required but in such cases Stark & Fowles (2006) recommended that statistical significance should be assessed at the 1% (rather than 5%) level.

RESULTS

Trends in river health based on MCI and QMCI

Appendix 1 presents the detailed results of trends testing. These are summarised in Table 2. Statistically significant trends in MCI were detected for three sites. A positive trend was detected in the Whanganui River at Te Maire when data from May 1989 were included (but no trend for the period 1999 – 2007). A positive trend in MCI was detected for the Makakahi River (Konini), and a negative trend for the Manawatu River at Maxwells Line / Teachers' College.

Six statistically significant trends in QMCI were detected – three positive and three negative (Table 2). A positive trend was detected for the Whanganui River at Pipiriki (1999 – 2007) but no trend when data from May 1989 were included. QMCI trended up at the Te Maire site on the Whanganui River but (as for the MCI) only with data from May 1989 included. The remaining positive trend in QMCI was for the Whanganui River Estuary. Both sites in the Oroua River showed negative trends in QMCI, as did the Manawatu River at Maxwells Line / Teachers' College.

Application of the Benjamini & Hochberg (1995) FDR to the results in Table 2 effectively eliminates all of the statistically significant results for the MCI (Appendix 2) and the QMCI (Appendix 3). FDR analyses produced a critical p-value of 0.002 for both the MCI and QMCI (based upon the 24 correlations undertaken concurrently). Since the p-values for the best correlations for the MCI (Whanganui River Te Maire p=0.004) and the QMCI (Whanganui River Te Maire p=0.019) are greater than this critical value we cannot reject the null hypothesis. In other words, we conclude that no statistically significant trends in MCI and QMCI were detected.

The same **overall** result was obtained from analyses undertaken only with data supplied by Horizons (*i.e.*, not including the May 1989 data). The only difference was that without the May 1989 data trends in MCI and QMCI for the Whanganui River at Te Maire were not detected at all (pre FDR analysis) (Table 2).

Although the Benjamini-Hochberg FDR analysis has ruled out all of the trends that the Mann-Kendall tests regarded as significant at the 5% level (as shown in Table 2), it is worth examining LOWESS plots for the nine trends that would be regarded as significant if they had been tested alone.

Table 2 Summary of Mann-Kendall p-values for MCI and QMCI for 21 rivers in the Manawatu-Wanganui region. Shaded cells include data from May 1989. See Appendix 1 for detailed results. Statistically significant results are shown in red text.

	River	Site	MCI		QMCI	
H16	MANAWATU RIVER	Hopelands Reserve	0.532	NS	0.144	NS
H20	MANAWATU RIVER	Maxwells Line / Teachers' College	0.026	-ve	0.037	-ve
H19	MANAWATU RIVER	Karere Rd / 42 Mile Hydro Station	0.211	NS	0.211	NS
H21	MANAWATU RIVER	Whirokino	0.851	NS	0.348	NS
H11	OROUA RIVER	Upstream MBP discharge @ Nelson St	0.297	NS	0.037	-ve
H10	OROUA RIVER	Awahuri bridge	0.677	NS	0.037	-ve
H12	MAKAKAHI RIVER	Konini	0.013	+ve	0.621	NS
H14	MANGATAINOKA RIVER	SH2 Bridge Mangatainoka	0.211	NS	0.404	NS
H18	MANGATERA STREAM	Confluence @ Timber Bay	0.322	NS	0.239	NS
H15	HAUTAPU RIVER	Upstream Rangitikei River	0.211	NS	0.404	NS
H17	RANGITIKEI RIVER	Pukeokahu	0.113	NS	0.404	NS
H13	RANGITIKEI RIVER	Mangaweka.	0.896	NS	0.138	NS
H9	RANGITIKEI RIVER	Kakariki	0.916	NS	0.061	NS
H23	RANGITIKEI RIVER	Estuary	0.851	NS	0.573	NS
H6	WHANGANUI RIVER	Cherry Grove	0.677	NS	0.095	NS
H7	WHANGANUI RIVER	Cherry Grove	0.109	NS	0.069	NS
H4	WHANGANUI RIVER	Te Maire	1.000	NS	0.835	NS
H5	WHANGANUI RIVER	Te Maire	0.004	+ve	0.019	+ve
H3	WHANGANUI RIVER	Downstream Retaruke confluence	0.404	NS	0.532	NS
H1	WHANGANUI RIVER	Pipiriki	0.061	NS	0.022	+ve
H2	WHANGANUI RIVER	Pipiriki	0.068	NS	0.336	NS
H24	WHANGANUI RIVER	Estuary	0.188	NS	0.039	+ve
H8	MANGAWHERO RIVER	DOC Headquarters	0.916	NS	0.677	NS
H22	MANGAWHERO RIVER	Downstream Makotuku	0.652	NS	0.881	NS

Whanganui River

Three sites in the lower Whanganui River (Te Maire, Pipiriki, Estuary) showed significant positive trends in QMCI with the first also showing a significant increase in MCI (Figures 2-5). The trends at the Te Maire site were significant only when data collected in May 1989 were included. The strongest trend was that for the MCI at Te Maire ($p = 0.004$), with the remaining trends much weaker ($p = 0.019 – 0.039$) (Table 2).

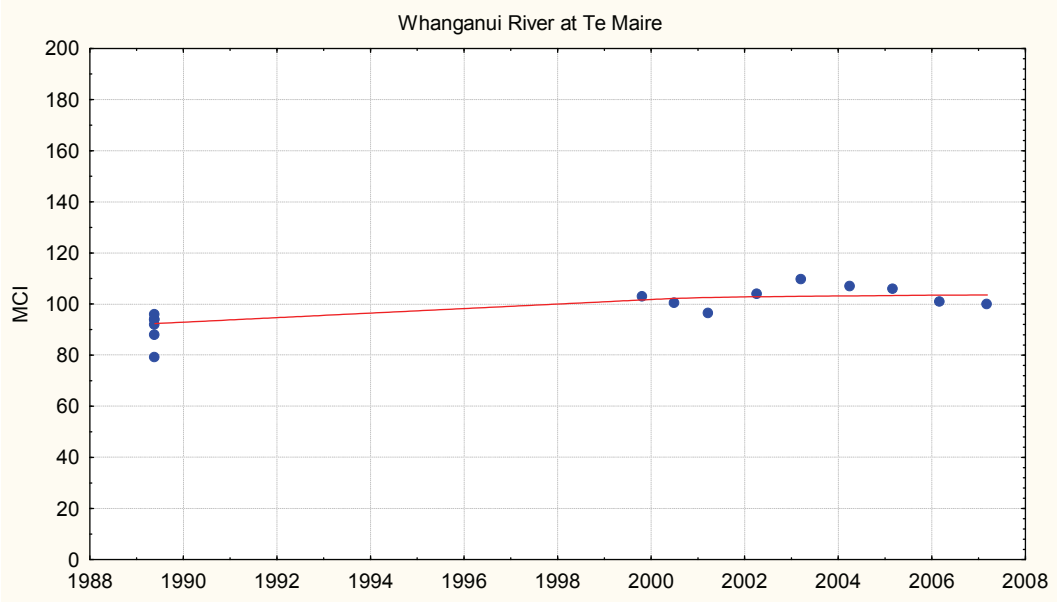


Figure 2 Significant positive trend in MCI for the Whanganui River at Te Maire.

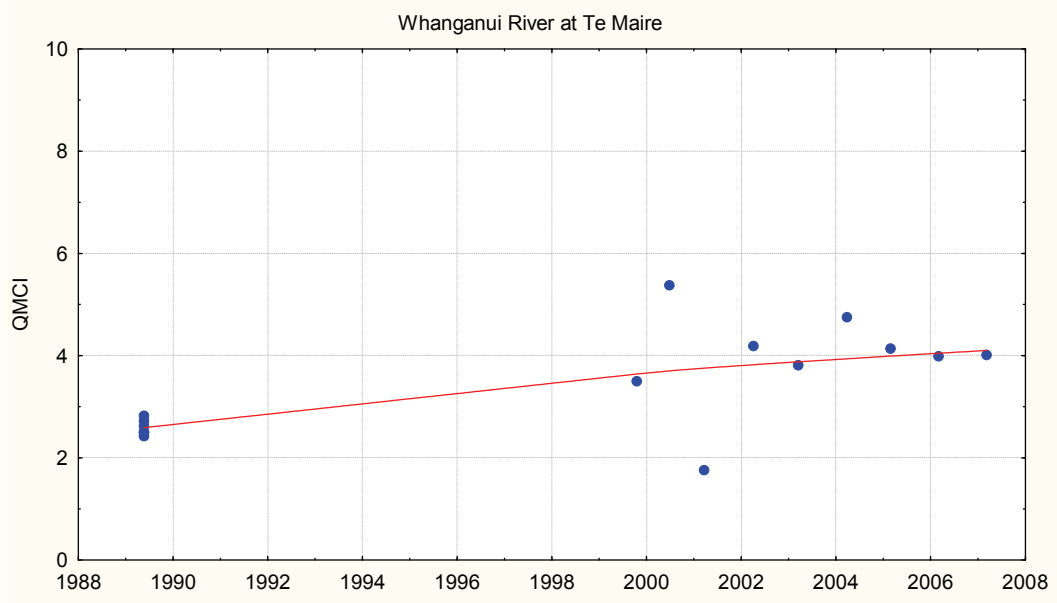


Figure 3 Significant positive trend in QMCI for the Whanganui River at Te Maire.

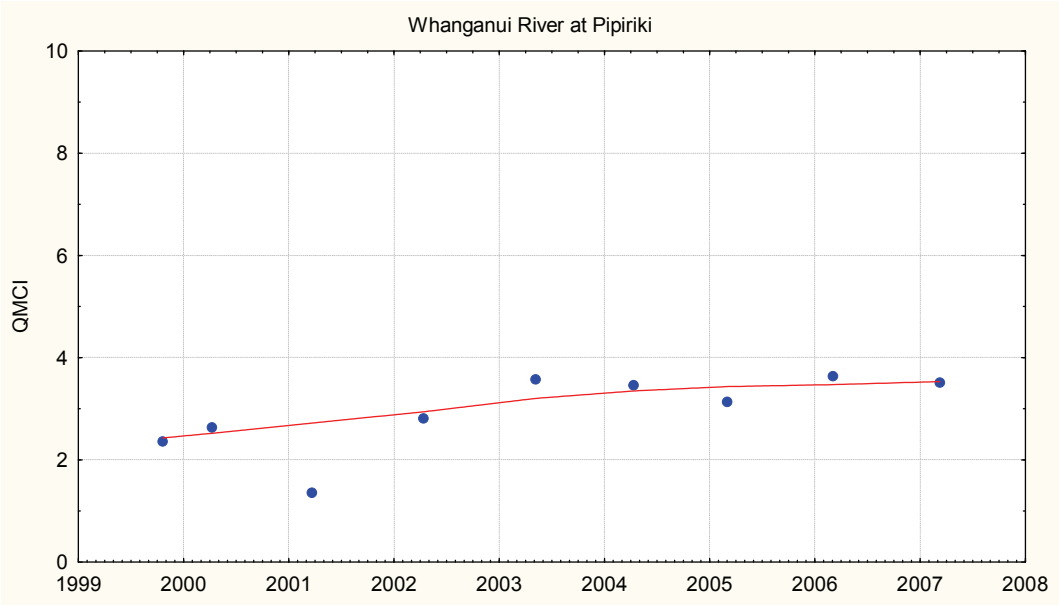


Figure 4 Significant positive trend in QMCI for the Whanganui River at Pipiriki.

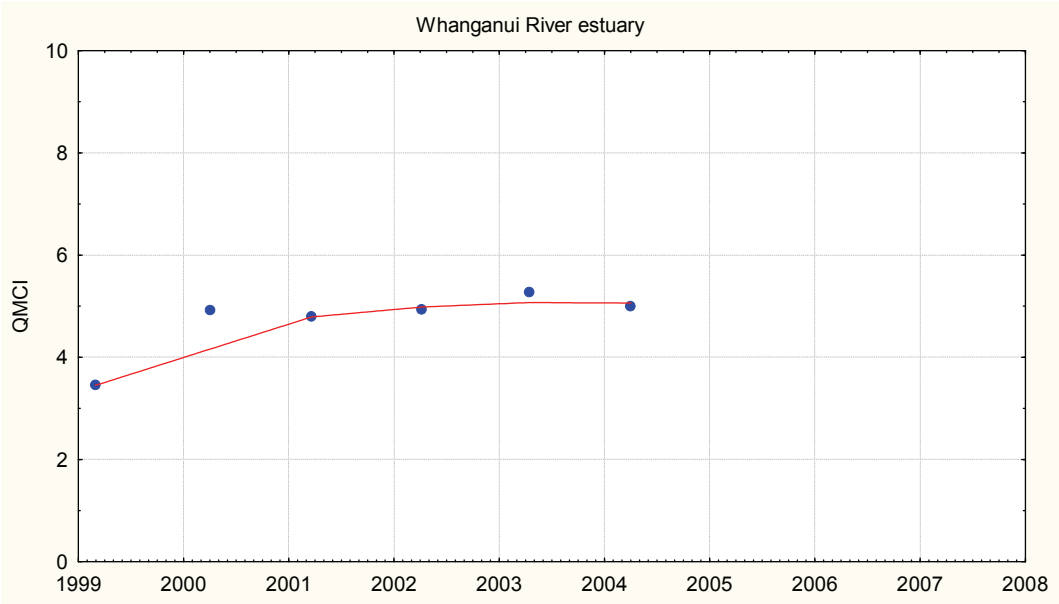


Figure 5 Significant positive trend in MCI for the Whanganui River estuary.

Oroua River

Two sites in the Oroua River showed significant decreases in QMCI (Figures 6 & 7), although neither trend was especially strong ($p = 0.037$) (Table 2). Although both sites appear to have deteriorated in "health" by about 30% over the past nine years, QMCI values at the more upstream site (Nelson St) remain about 2 units higher than for the site further downstream (Awahuri Bridge).

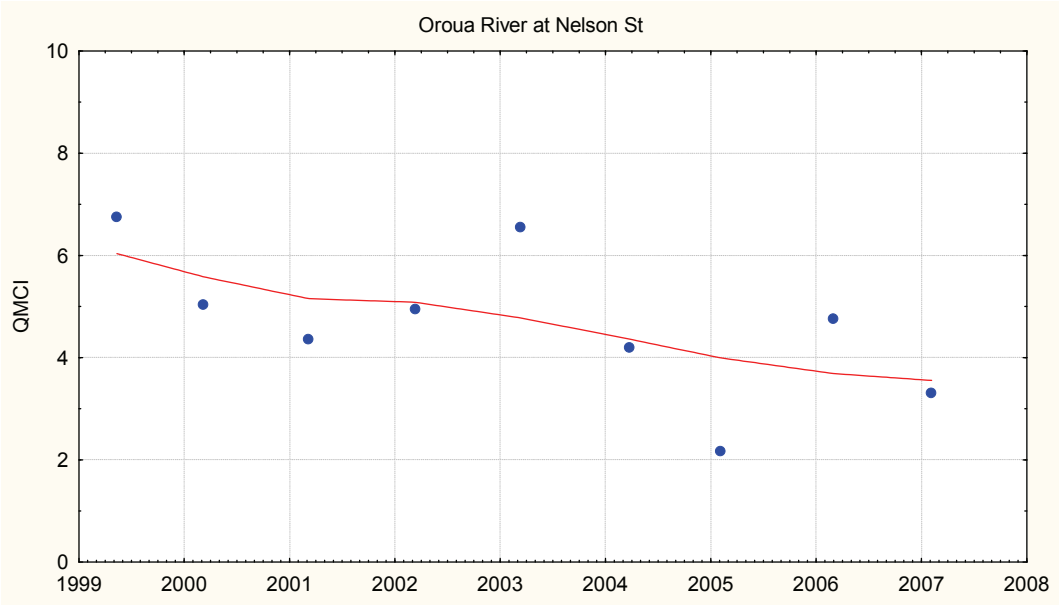


Figure 6 Significant negative trend in QMCI for the Oroua River at Nelson St upstream of the MBP discharge.

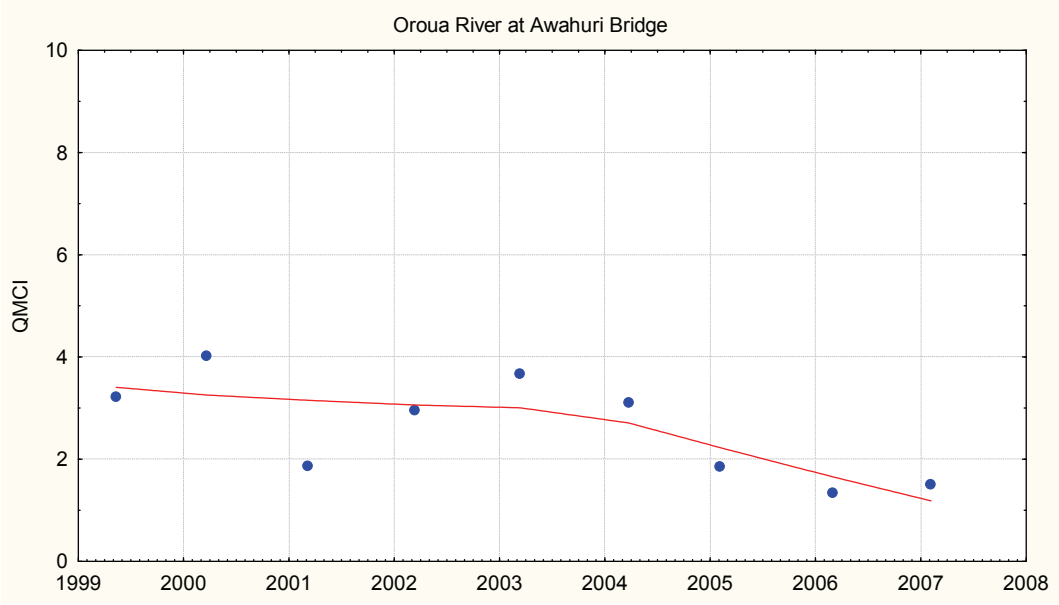


Figure 7 Significant negative trend in QMCI for the Oroua River at Awahuri (SH3 bridge).

Manawatu River

The Teachers' College sites is approximately 4 km upstream of the Maxwells Line site on the Manawatu River (Kate McArthur, Horizons *pers. comm*.). The Maxwells Line site was sampled in 1999, 2001, 2002, 2004, and 2005, and the Teachers' College site in 2000, 2003, 2006, and 2007. Although MCI and QMCI values at the upstream site tend to be higher on average (by 2 and 0.17 units respectively) than at the downstream site in neither case are the differences statistically significant (MCI: t = -0.273, p = 0.793; QMCI: t = -0.159, p = 0.878). The decision to combine data from these sites is, therefore, justifiable. In future, only the

Teachers' College site will be sampled because it is also a water quality SoE and flow monitoring site. MCI and QMCI both showed a statistically significant decrease over time in the Manawatu River at this point (Table 2, Figures 8 & 9).

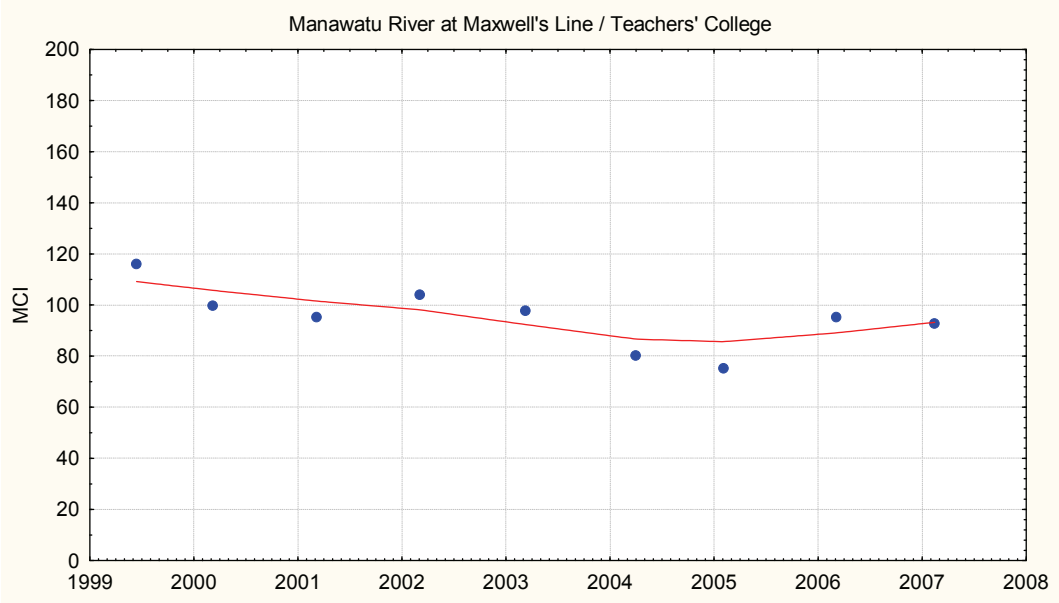


Figure 8 Significant negative trend in MCI for the Manawatu River at Maxwells Line / Teachers' College.

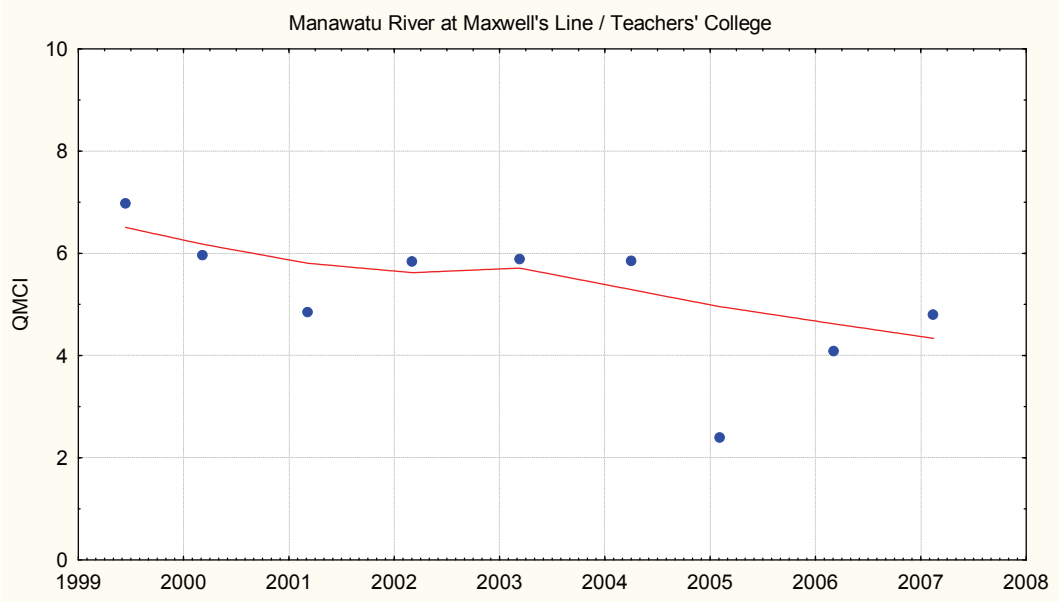


Figure 9 Significant negative trend in QMCI for the Manawatu River at Maxwells Line / Teachers' College.

A statistically significant increase in MCI has occurred in the Makakahi River at Konini (Table 2, Figure 10).

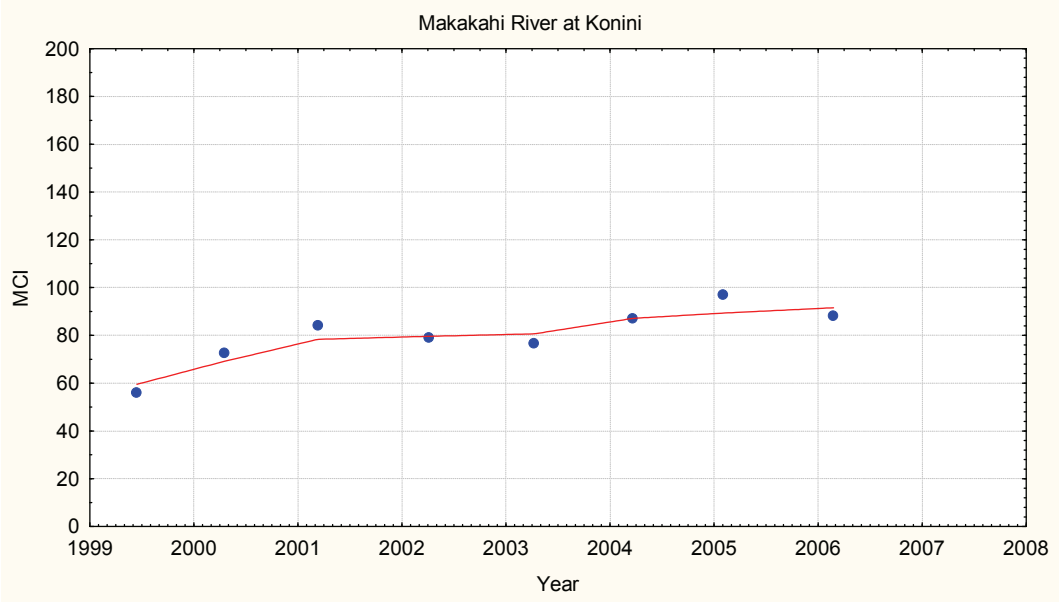


Figure 10 Significant positive trend in MCI for the Makakahi River at Konini.

Are these trends in biotic indices significant or not?

The concept that a statistically significant trend based on a single test, might no longer be regarded as significant if that test was done along with several (or many) other tests may seem difficult to grasp. However, it is now generally accepted that when multiple tests are undertaken some correction is required, and the effect of this will be to reduce the number of correlations that remain significant.

In the present case all nine trends, which would have each been regarded as statistically significant (at the 0.05 level) if undertaken singly, have been rendered non-significant by the Benjamini-Hochberg FDR procedure. There were no exceptionally strong trends (*e.g.*, all p values were 0.004 or greater and all but one was 0.013 or greater). The interpretation here is not that these trends do not exist, but rather that they are not yet strong enough to remain after the FDR procedure has been applied. In future, when a longer time series has been established (and assuming that these trends continue) at least some of these trends are likely to remain significant following application of the FDR procedure.

Stark & Fowles (2006) suggested that the Benjamini-Hochberg FDR procedure applied at the 0.05 level yielded a similar number of statistically significant trends as the Mann-Kendall test when applied at the 0.01 level. When testing for trends in biotic indices, it was suggested, therefore that a single test should be interpreted at the 0.01 rather than 0.05 level.

Indeed, if we interpret the Mann-Kendall results at the 0.01 level, we find only one trend remains statistically significant (Table 2, Figure 2): the Whanganui River at Te Maire ($p = 0.004$). This trend is significant only if the May 1989 data are included, and not if based entirely on Horizons' SoE monitoring data collected from 1999 – 2007 (Table 2).

Evaluation of river health in the Manawatu-Wanganui region

We probably should conclude that no statistically significant trends in MCI and QMCI were detected at the 21 SoE monitoring sites in the region that had been sampled on six or more occasions. However, it would be prudent to take some regard of the results and to note the *suggestion* of an improvement of river health in the Whanganui River (@ Te Maire – Pipiriki – Estuary) and the Makakahi River (@ Konini) and a deterioration in river health in the Oroua River (@ Nelson St – Awahuri bridge) and the Manawatu River (@ Maxwells' Line / Teachers' College).

Table 3 (columns labelled (a)) provides guidelines for interpreting the MCI and QMCI (Stark & Maxted 2007a). Wright-Stow & Winterbourn (2003) suggested that the boundaries between quality classes (or degradation categories) should be fuzzy – a concept put forward originally by Stark (1985 – see Figures 1-3) where a band of uncertainty equivalent to 10 MCI units (*i.e.*, ±5) was proposed between quality classes. Wright-Stow & Winterbourn (2003) retained the criteria for defining quality classes but suggested that fuzzy boundaries of ±5 and ±0.2 should be applied to the MCI and QMCI (Table 3 columns labelled (b)).

Table 3 Interpretation of MCI and QMCI according to (a) Stark & Maxted (2007a), and (b) Wright-Stow & Winterbourn (2003).

Quality class	MCI		QMCI		Degradation category
	(a)	(b)	(a)	(b)	
Excellent	> 119	> 124	> 5.99	> 6.1	Clean
Good	100-119	105-115	5.00-5.90	5.2-5.7	Mild
Fair	80-99	85-95	4.00-4.99	4.2-4.7	Moderate
Poor	<80	<75	<4.00	<3.8	Severe

Table 4 presents an evaluation of river health for 21 sites in the Manawatu-Wanganui region based on the interpretive guidelines provided by Stark & Maxted (2007a) (column (a)) and Wright-Stow & Winterbourn (2003) (column (b)).

The MCI and the QMCI may classify sites into different quality classes. For example, when interpreted according to Stark & Maxted's (2007a) guidelines, the MCI classified one site as 'excellent', five sites as 'good', 12 sites as 'fair', and three sites as 'poor'. On the other hand, the QMCI classified one site as 'excellent', five sites as 'good', six sites as 'fair', and nine sites as 'poor'. The MCI and QMCI produced the same result for 11 of the 21 sites (1E, 3G, 5F, & 2P) (Table 4).

The effect of the fuzzy boundaries between quality classes is best visualised by re-ordering sites based on their MCI (Table 5) or QMCI (Table 6) values.

Table 4 Evaluation of river health based on average MCI and QMCI values (1999 – 2007) for 21 sites in the Manawatu-Wanganui region according to the interpretive guidelines in Table 3.

River	Site	MCI (a)	MCI (b)	QMCI (a)	QMCI (b)
MANAWATU RIVER	Hopelands Reserve	93	93	4.66	4.7
MANAWATU RIVER	Maxwells Line / Teachers' College	95	95	5.18	5.2
MANAWATU RIVER	Karere Rd / 42 Mile Hydro Station	76	76	2.61	2.6
MANAWATU RIVER	Whirokino	73	73	4.24	4.2
OROUA RIVER	@ Nelson St	98	98	4.67	4.7
OROUA RIVER	Awahuri bridge	74	74	2.61	2.6
MAKAKAHI RIVER	Konini	80	80	3.75	3.8
MANGATAINOKA RIVER	SH2 Bridge Mangatainoka	86	86	4.57	4.6
MANGATERA STREAM	Confluence @ Timber Bay	90	90	3.43	3.4
HAUTAPU RIVER	Upstream Rangitikei River	85	85	2.71	2.7
RANGITIKEI RIVER	Pukeokahu	117	117	5.79	5.8
RANGITIKEI RIVER	Mangaweka	105	105	5.67	5.7
RANGITIKEI RIVER	Kakariki	96	96	5.18	5.2
RANGITIKEI RIVER	Estuary	83	83	4.01	4.0
WHANGANUI RIVER	Cherry Grove	107	107	5.37	5.4
WHANGANUI RIVER	Te Maire	103	103	3.94	3.9
WHANGANUI RIVER	Downstream Retaruke conflunce	101	101	3.64	3.6
WHANGANUI RIVER	Pipiriki	84	84	2.94	2.9
WHANGANUI RIVER	Estuary	87	87	4.73	4.7
MANGAWHERO RIVER	DOC Headquarters	138	138	7.83	7.8
MANGAWHERO RIVER	Downstream Makotuku	87	87	2.61	2.6

Key to Quality Classes

Excellent
Good - Excellent
Good
Fair – Good
Fair
Poor – fair
Poor

Table 5 Effect of fuzzy boundaries on the assignment to quality classes based on the MCI using to the interpretive guidelines in Table 3. See Table 4 for key to quality classes.

River	Site	MCI (a)	MCI (b)	Rank Order
MANGAWHERO RIVER	DOC Headquarters	138	138	1
RANGITIKEI RIVER	Pukeokahu	117	117	2
WHANGANUI RIVER	Cherry Grove	107	107	3
RANGITIKEI RIVER	Mangaweka	105	105	4
WHANGANUI RIVER	Te Maire	103	103	5
WHANGANUI RIVER	Downstream Retaruke conflunce	101	101	6
OROUA RIVER	@ Nelson St	98	98	7
RANGITIKEI RIVER	Kakariki	96	96	8
MANAWATU RIVER	Maxwells Line / Teachers' College	95	95	9
MANAWATU RIVER	Hopelands Reserve	93	93	10
MANGATERA STREAM	Confluence @ Timber Bay	90	90	11
MANGAWHERO RIVER	Downstream Makotuku	87	87	12=
WHANGANUI RIVER	Estuary	87	87	12=
MANGATAINOKA RIVER	SH2 Bridge Mangatainoka	86	86	14
HAUTAPU RIVER	Upstream Rangitikei River	85	85	15
WHANGANUI RIVER	Pipiriki	84	84	16
RANGITIKEI RIVER	Estuary	83	83	17
MAKAKAHI RIVER	Konini	80	80	18
MANAWATU RIVER	Karere Rd / 42 Mile Hydro Station	76	76	19
OROUA RIVER	Awahuri bridge	74	74	20
MANAWATU RIVER	Whirokino	73	73	21

Table 5 reveals that 12 sites were classified unambiguously into the 'excellent', 'good', 'fair', or 'poor' quality classes using the MCI irrespective of whether fuzzy or fixed boundaries were used. One site in the 'good' class could have been 'excellent' and two could have been 'fair'. Two sites in the 'fair' class could have been 'good', and three could have been 'poor' (3). One site in the 'poor' class may have been 'fair' if fuzzy boundaries had been used.

Table 6 shows that the QMCI assigned 17 sites to the same quality classes irrespective of whether or not fuzzy boundaries were used. One site in the 'good' class could have been 'excellent', one site in the 'fair' class could have been 'good', and two sites in the 'poor' class may have been 'fair' if fuzzy boundaries had been used.

The use of fuzzy boundaries does not, of course, alter the rank order of sites because this is based directly on the MCI or QMCI values themselves.

Table 6 Effect of fuzzy boundaries on the assignment to quality classes based on the QMCI. See Table 4for key to quality classes.

River	Site	QMCI (a)	QMCI (b)	Rank Order
MANGAWHERO RIVER	DOC Headquarters	7.83	7.8	1
RANGITIKEI RIVER	Pukeokahu	5.79	5.8	2
RANGITIKEI RIVER	Mangaweka	5.67	5.7	3
WHANGANUI RIVER	Cherry Grove	5.37	5.4	4
MANAWATU RIVER	Maxwells Line / Teachers' College	5.18	5.2	5=
RANGITIKEI RIVER	Kakariki	5.18	5.2	5=
WHANGANUI RIVER	Estuary	4.73	4.7	7
OROUA RIVER	@ Nelson St	4.67	4.7	8
MANAWATU RIVER	Hopelands Reserve	4.66	4.7	9
MANGATAINOKA RIVER	SH2 Bridge Mangatainoka	4.57	4.6	10
MANAWATU RIVER	Whirokino	4.24	4.2	11
RANGITIKEI RIVER	Estuary	4.01	4.0	12
WHANGANUI RIVER	Te Maire	3.94	3.9	13
MAKAKAHI RIVER	Konini	3.75	3.8	14
WHANGANUI RIVER	Downstream Retaruke conflunce	3.64	3.6	15
MANGATERA STREAM	Confluence @ Timber Bay	3.43	3.4	16
WHANGANUI RIVER	Pipiriki	2.94	2.9	17
HAUTAPU RIVER	Upstream Rangitikei River	2.71	2.7	18
OROUA RIVER	Awahuri bridge	2.61	2.6	19=
MANAWATU RIVER	Karere Rd / 42 Mile Hydro Station	2.61	2.6	19=
MANGAWHERO RIVER	Downstream Makotuku	2.61	2.6	19=

The MCI and QMCI do not provide exactly the same classification of sites into quality classes. Overall, the QMCI presents a more pessimistic picture with three times as many sites in the 'poor' class compared with the MCI. This is not unexpected, and normally is a result of extreme numerical dominance of taxa such as chironomids, worms, and hydroptilid caddisflies (and the exclusion of higher scoring taxa such as mayflies, stoneflies and (most) caddisflies). It was for this reason that Scarsbrook *et al.* (2000) recommended the MCI rather than the QMCI for reporting the results of macroinvertebrate SoE monitoring programmes, a recommendation that was echoed by Stark & Maxted (2007a). In statistical terms the MCI and QMCI are equally good at ranking sites from best to worst (Spearman R = 0.709, p = 0.0003).

DISCUSSION

PRESENT STATE AND TRENDS IN RIVER HEALTH

This report has analysed time series of MCI and QMCI values for 21 sites on rivers in the Manawatu – Wanganui region. Only sites that had been sampled six or more times were included because time series analysis is problematical if there are too few data points.

SoE biomonitoring data have two primary uses (1) assessing the state or condition of the environment, and (2) detecting changes in the state or condition of the environment (i.e., detecting trends).

When the state of the environment is assessed the fundamental question is "What is the state (or present condition) of the environment?" The next question that follows is "Is the present state of river health acceptable?" If the answer is "No.", then that should lead to other questions such as "Can river health be improved?" "How can it be improved?", "What will it cost?", "How long will it take?", and "Does the political will exist to make the resources available to effect some improvement?" Answering all of these questions is beyond the scope of a report such as this. However, freshwater ecologists need to be able to answer the first question – "What is the present state of river health?"

Before one can determine whether the present state is acceptable or not, one must define what the present state is. There are many ways that this can be achieved based on freshwater macroinvertebrate biomonitoring data and a long history of the practice overseas (Rosenberg & Resh 1993) and in New Zealand (Boothroyd & Stark 2000). Perhaps the simplest is to use a biological index to summarise complex biotic data down to single numbers that are more easily interpreted by water managers and others who are not specialist freshwater ecologists (Stark 1985). Stark & Maxted (2007a) discussed the merits of various biotic indices and recommended that the MCI (for hard-bottomed streams and rivers) and the MCI-sb (for soft-bottomed streams) were the indices that should be used for SoE reporting.

Dewson *et al.* (2007), in the 2007 SoE biomonitoring report, assessed the state of river health in the Manawatu – Wanganui region using mean (from five replicate Surber samples) QMCI and MCI values assigned to quality classes according to the fuzzy criteria given by Wright-Stow & Winterbourn (2003). They also reported O/E taxon richness – a measure of how many invertebrate taxa were observed at a site compared with how many were expected - calculated using a predictive model based on the physical and chemical characteristics of the site (Joy & Death 2003). An O/E ratio of 0.8 of more (the 90th percentile of all references sites used to build the model) was considered unimpacted. O/E ratios between 0.51 and 0.79 were considered impacted and O/E ratios of 0.50 or less were regarded as severely impacted.

I have some concerns about the use of O/E taxon richness as a measure of whether or not river health is a good as expected. In theory, O/E ratios are a good way of determining whether a site (or river) is as healthy as it might be expected to be. AURIVAS embodies O/E taxon

richness and O/E SIGNAL (the Australian equivalent of the MCI) (Simpson & Norris 1997). However taxon richness is not the best indicator of river health. Highest values often are associated with slight nutrient enrichment (such as might occur when a native forested stream exits into farmland), and lowest values may be recorded in torrential mountain streams of pristine water quality. The O/E ratio overcomes those issues somewhat (because it compares observed values with those calculated from a reference dataset – although the composition of the reference dataset can have a marked effect on the results) but does not deal with the main problem with estimates of taxon richness, which is that the primary cause of variability in taxon richness is the flow regime (and especially the effects of large floods, which can cause severe reductions in taxon richness). Thus, sampling too soon after a servere flood could result in a low O/E taxon richness ratio and lead to a conclusion that stream health was severely impacted. In my view, Joy & Death (2003) and, subsequently, Dewson et al. (2007) should have determined O/E MCI values instead of (or as well as) O/E taxon richness since this is likely to provide a more reliable assessment. MCI (being an average tolerance value per taxon) is affected less by sample size and flow variability than taxon richness.

Following Stark &Maxted's (2007a) recommendation, I propose that the MCI should be the primary biotic index for interpreting SoE biomonitoring results. [Note that Stark & Maxted (2007a) recommend that the SQMC and QMCI should not be used for reporting SoE biomonitoring – they are more suited to compliance monitoring.] The first thing that the MCI enables is for sites to be ranked in order of stream health (from highest to lowest MCI values) (e.g., Table 5). That is the easy part. The next step is to assign the sites to quality classes or to determine which sites are considered of acceptable health and which require improvement.

Table 3 in this report presents the published criteria for assigning quality classes based on MCI values. Here, again we can choose whether to interpret these criteria with fixed (Stark & Maxted 2007a) or fuzzy (Wright-Stow & Winterbourn 2003) boundaries. It makes no difference, of course, to the rank order. The fuzzy boundaries are of most use when deciding each year which quality class to assign a site to. For example with a time series of MCI values such as 125, 118, 120, 117, 122 it might be considered undesirable to have the site oscillating between the 'excellent' and 'good' quality classes. None of those MCI values is likely to be statistically significantly different from any of the others (all within 10.83 MCI units: Stark (1998)). In this case it would make sense to assign the site to either the 'excellent' or the 'good' quality class. In this case I would choose 'excellent' because in the first year it was 'excellent' (125) and subsequent values have not been low enough to warrant a change. The average of this series is also (just) over 120 (which is, of course, only known in hindsight), but justifies the retention of 'excellent' status.

Table 7 shows MCI values recorded from 21 SoE biomonitoring sites in the Manawatu – Wanganui Region classified into quality classes according to Stark & Maxted's (2007a) criteria (Table 3). On average, one site had ''excellent'' stream health, five were 'good', 12 were 'fair', and three were 'poor' (Table 7). When using fixed boundaries between quality classes, only one site (Mangawhero River @ DOC Headquarters) has remained in the same quality class for the entire sampling period. Only two sites showed statistically significant trends in MCI (even though they were eliminated subsequently by Benjamini-Hochberg FDR analysis): These data suggest that the Manawatu River at Maxwells Line / Teachers' College

has deteriorated from 'good' to 'fair' and the Makakahi River at Konini has improved from 'poor' to 'fair'. All other sites have fluctuated between quality classes over the years (some over three classes), but there is no suggestion of overall trends (+ve or –ve) at these sites.

It would be tempting to suggest that all sites within a regions should be classified as ''excellent'' or 'good' and that any sites that are classified as 'fair' or 'poor' should be subjected to remediation aimed at raising stream health. However, this is an overly simplistic approach.

Firstly, some decrease in MCI with distance downstream in a river is inevitable. It would be unrealistic to expect the lower reaches of long rivers to yield the same 'excellent' MCI values recorded from their headwater streams. The downstream decline in MCI values occurs not only because of increasing nutrient enrichment (which can be from natural sources), but also because rivers process their sediments into finer size fractions as they progress downstream. Highest MCI values tend to be associated with boulder- and cobble-dominated substrates and lower values with sandy or muddy substrates. Furthermore, highest MCI values are also associated with riffle habitats. In large rivers such habitats may not exist and deep runs and pools may be the only habitats present. On the other hand, just because a site is rated 'good', or 'excellent' does not necessarily mean that is as healthy as it could be – there may still be habitat degradation or water pollution occurring from anthropogenic sources.

Despite the existence of numerical methods for assessing river health (whether they be stream quality classes like those used for the MCI, or O/E ratios like those embodied in AUSRIVAS and other predictive modelling approaches), inevitably the decision concerning whether river health is acceptable or not is one for water managers to make. In the end it is a value judgment. Certainly, any sites that are classified as 'poor' or 'fair' or any sites showing negative trends in stream health are likely to have room for improvement. Detection of a significant decline in river health is evidence that, in the past, the river has been in better condition. If that's the case, then it raises the possibility that the downward trend may be able to be reversed if the factor/s causing the degradation or negative trends can be identified. Obviously, whether remediation is possible does depend on the cause. If the decline in river health was due to the presence of a recently constructed large hydro-electric dam, then the effects may not easily be reversed; a polluting discharge may be much more easily controlled.

Identifying cause/s is the first step towards remedial action. Point source discharges from municipal oxidation ponds, industries, dairy sheds, stormwater runoff (especially in urban areas), and poor sediment management from logging or land developments may be the first to receive attention, but diffuse-source enrichment from farmland, direct stock access to waterways, and poor condition of riparian margins may also have significant impacts on stream health. For most, if not all, of these there are practical ways to ameliorate or minimise adverse environmental consequences.

Table 7 Time series of MCI values at 21 SoE biomonitoring sites in the Manawatu-Wanganui region. Sites are listed in decreasing order of stream health based on overall average MCI values. See Table 3 for key to quality classes. Sites have been assigned to quality classes based on the fixed criteria of Stark & Maxted (2007a). Trend: -ve = negative, +ve = positive, NS =- Not Significant.

River	Site	1999	2000	2001	2002	2003	2004	2005	2006	2007	Average	Trend
MANGAWHERO RIVER	DOC Headquarters	148	133	139	138	134	138	130	141	143	138	NS
RANGITIKEI RIVER	Pukeokahu	122	130	113	121	107	117	114	117	113	117	NS
WHANGANUI RIVER	Cherry Grove	112	115	97	120	97	116	99	103	102	107	NS
RANGITIKEI RIVER	Mangaweka	131	-	94	102	94	102	102	117	97	105	NS
WHANGANUI RIVER	Te Maire	103	100	97	104	110	107	106	101	100	103	NS
WHANGANUI RIVER	Downstream Retaruke conflunce	108	97	94	105	104	111	104	100	88	101	NS
OROUA RIVER	@ Nelson St	120	91	98	107	113	82	80	100	92	98	NS
RANGITIKEI RIVER	Kakariki	91	99	97	102	103	91	106	96	77	96	NS
MANAWATU RIVER	Maxwells Line / Teachers' College	116	100	95	104	98	80	75	95	93	95	-ve
MANAWATU RIVER	Hopelands Reserve	101	95	87	100	91	88	90	93	93	93	NS
MANGATERA STREAM	Confluence @ Timber Bay	105	92	99	94	60	78	91	98	-	90	NS
MANGAWHERO RIVER	Downstream Makotuku	-	101	90	76	76	94	84	89	-	87	NS
WHANGANUI RIVER	Estuary	74	84	82	80	101	99	-	-	-	87	NS
MANGATAINOKA RIVER	SH2 Bridge Mangatainoka	78	83	88	86	80	90	101	74	91	86	NS
HAUTAPU RIVER	Upstream Rangitikei River	86	92	95	87	88	81	92	70	72	85	NS
WHANGANUI RIVER	Pipiriki	68	78	61	89	89	97	96	90	90	84	NS
RANGITIKEI RIVER	Estuary	78	86	80	95	89	72	-	-	-	83	NS
MAKAKAHI RIVER	Konini	56	73	84	79	77	87	97	88	-	80	+ve
MANAWATU RIVER	Karere Rd / 42 Mile Hydro Station	105	71	66	73	87	90	71	66	59	76	NS
OROUA RIVER	Awahuri bridge	66	76	71	101	77	87	77	54	55	74	NS
MANAWATU RIVER	Whirokino	68	77	74	84	64	73	-	-	-	73	NS
Number classified 'EXCELLENT'		4	2	1	3	1	1	1	1	1	1	
Number classified 'GOOD'		7	4	1	7	6	4	5	6	3	5	
Number classified 'FAIR'		2	9	15	9	10	13	10	7	7	12	
Number classified 'POOR'		7	5	4	2	4	3	3	4	4	3	

In the context of the sites listed on Table 7, the three sites with an overall classification of 'poor', and another nine that have ever recorded a 'poor' MCI value (i.e., < 80) could be first to be investigated in order to determine the reasons for the poorer than average river health. Often, however, determining the cause/s of the problem is far less difficult than taking remedial action because that usually involves changing attitudes, politics, and finances.

The Manawatu River downstream of Palmerston North does appear to be in comparatively 'poor' condition. More often than not MCI values are less than 80 (Table 7). Perusal of a topographical map indicates that there are a number of industrial facilities, farming activities, and the city itself, that all have the potential to affect the river adversely. Determining whether they do or not would require access to more information than I have available to me. Similarly, MCI values for the Oroua River at Awahuri Bridge have also been indicative of 'poor' river health on most occasions (with the exception of 2002 and 2004 when 'good' and 'fair' condition, respectively, was indicated) (Table 7).

Concern about pollution in the Manawatu and Oroua Rivers is not new (Hirsch 1985, Suckling 1982). In March 1957, Hirsch (1958) sampled the Manawatu River and data he collected enabled Stark (1985) to calculate MCI values. MCI values of 92 - 98 were recorded in the vicinity of Horizons' Teachers' College /Maxwells Line sites, which is comparable with the mean of 95 (range 75 – 116) recorded more recently (Table 7). Further downstream in the vicinity of the Karere Road / 42 Mile Hydro Station site Hirsch's (1958) data yielded an MCI of 78, which is also comparable with the mean of 76 (range 59 – 105) of recent data (Table 7). It is difficult to determine from these data whether the Manawatu River is in better or worse condition now than it was in 1957 because Hirsch (1958) sampled only once and more recent sampling over a period of years has revealed occasions when MCI values have been significant higher and significantly lower than he recorded. It is, however, difficult to conclude that there has been any improvement given the statistically significant decline in MCI from 1999 to 2007 (Table 2, Table 7, Figure 8).

Unfortunately, Suckling (1982) did not provide the data necessary for MCI calculation in his paper, but he noted that invertebrate species diversity in the Manawatu River between Maxwells Line and Karere Road generally was "good (allowing for the effects of agricultural runoff), although midsummer organic loading induced marked changes in benthic species composition, including the disappearance of *Deleatidium* sp. from several sites."

Hirsch (1958) also sampled the Oroua River and data he collected returned an MCI value of 35 for the Awahuri Bridge site. He noted effluent discharges exceeded the assimilative capacity of the river in this vicinity but that there was recovery downstream prior to the confluence with the Manawatu River (MCI 62 -65). Recent data suggest that the condition of the Oroua River at Awahuri Bridge has improved markedly with a mean MCI of 74 (range 54 – 101) (Table 7). However, it is likely that further improvement would be possible.

Detecting trends in river health addresses the question "Is river health improving or getting worse?" Trends testing can answer that question but SoE monitoring data alone seldom answers the "Why?" question. Other information, such as data on land-use, water quality, the

hydrological regime, or knowledge concerning point-source or diffuse-source discharges is required to do that.

It should be emphasised also that differences between data or trends that are described as 'statistically significant' are not necessarily ecologically significant or meaningful to management. Managers, with advice from experienced freshwater ecologists. need to confirm the scale of effect that will trigger a management response.

If we are statistically rigorous, we should conclude that no statistically significant trends in MCI and QMCI were detected at the 21 SoE monitoring sites in the region that had been sampled on six or more occasions. This was because all trends were comparatively weak and were eliminated by the Benjamini-Hochberg FDR procedure. As the time series lengthens over coming years, it is likely that some of these trends will become significant. This is because the power of the trends tests will increase as the number of data points in the time series increases.

However, it might be prudent note the *suggestion* of an improvement of river health in the Whanganui River (@ Te Maire – Pipiriki – Estuary) and the Makakahi River (@ Konini) and a deterioration in river health in the Oroua River (@ Nelson St – Awahuri bridge) and the Manawatu River (@ Maxwells' Line / Teachers' College). Council staff may like to consider likely causes of these "trends".

GENERAL COMMENTS ON SOE MONITORING

Design criteria

Monitoring networks must be designed with management objectives in mind. While it is beyond the scope of this review to describe State of the Environment (SoE) monitoring programme design in any detail, the following matters must be considered:-

- Site locations.
- Reference sites and impacted sites – samples should be representative in space.
- Number of 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 data be analysed?
- How much data do the statistical analyses require?
- How will data be translated into information that water managers can use?
- How much will it cost?
- Is there a long-term political commitment to funding?

In the following sections I discuss some critical aspects of SoE monitoring programme design, including those that could be affected if the integrity of the monitoring programme is

disrupted.

Collect data that provide information for water managers

When designing long-term SoE monitoring programmes the specific management uses for information must be defined at the outset. Likens (1983) lamented the collection and storage of vast amounts of data at great cost that have subsequently been ignored because there was no experimental design or management objective specified initially. He observed that "it is exceedingly difficult to determine what information might be important in the future, and particularly, what error limits will be required so that accumulating these data would be significant for answering problems in the future."

Routine SoE monitoring should not be regarded as an opportunity to collect data "in case it might be useful in future". Only data that assist interpretation of monitoring results or are of proven value for water managers should be included. Special investigations or research projects (which may be undertaken by, or in collaboration with, external research providers) are an appropriate means to determine those environmental variables that provide useful information for water managers (and might then be added to routine SoE monitoring programmes).

Recommendations: *Avoid the temptation to sample everything. Concentrate on maintaining cost-effectiveness and collecting data of proven usefulness for water managers.*

Statistical requirements

Statistical power

Data requirements vary according to the variables being measured. Byl & Smith (1994) commented that it is difficult to detect any significant trends in water quality over time in large areas because of the variability due to natural events. Thus, many years of data must be gathered before a statistically significant improvement in water quality is demonstrated. The same is true of SoE biomonitoring data, because sampling might occur only once per year, several years of data are required in order to undertake time series analyses.

Smith *et al.* (1996) analysed the first five years (1989-93) water quality data from New Zealand's NRWQN in order to detect any trends that may exist. They considered that data from at least five years of monthly sampling (*i.e.*, 60+ data values per determinand) were required.

Stansfield (2001) examined the effect of sampling frequency on the ability to detect trends in water quality. He found that if water quality sampling frequency was reduced from monthly to quarterly, many trends were no longer detected. The reason was that smaller data sets have larger standard errors in the estimation of statistics, and sometimes these are large enough to discount a trend that was evident from monthly data. Furthermore, the trends that were detected from seasonal data often were slightly different from those detected using monthly data.

Unlike water quality sampling, macroinvertebrate sampling is seldom (if ever) undertaken on a monthly basis. Macroinvertebrate SoE monitoring is most often annual (occasionally twice per year). Scarsbrook *et al.* (2000) examined data from annual sampling of macroinvertebrates from 66 of the NRWQN sites from 1989 to 1996. They noted that analysis of long-term trends in invertebrate community composition is rare – probably because there have been few consistent, long-term data sets. A major issue with analysis of long-term trends is the need to maintain consistent sample collection and processing methods (Smith *et al.* 1996). This is assisted by the use of standard methods (Stark *et al.* 2001) and by doing macroinvertebrate SoE monitoring every year. Having a break of several years between monitoring occasions would make it much harder for personnel to maintain the necessary degree of consistency (and could cause problems for trends analysis due to missing values also).

Due to the lack of experience with trend analysis of macroinvertebrate data (because few data sets exist), the minimum number of sampling times required before meaningful trends can be detected is uncertain. It will depend also on the strength of the trend in relation to natural or random variability, and is likely to vary between different stream types and regions too. Scarsbrook *et al.* (2000) had a time series of eight data points corresponding to the years 1989 to 1996 inclusive. These data were sufficient to detect trends in various macroinvertebrate community measures (at the 0.10 level), 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 the national scale. However, no causal links were established, and it is possible that some of the trends were not real because they were based on only eight data points each one year apart. There are many environmental factors (other than water quality, which often is the factor of interest) that can affect macroinvertebrate community composition. Although annual SoE monitoring using invertebrates is undertaken at the same time of year and under similar flow conditions, in order to minimise the influence of season and flow, some variation from year-to-year is inevitable even if water quality remains consistent.

More recently Collier & Kelly (2006) and Stark & Fowles (2006) have examined trends in macroinvertebrate community indices from the Waikato and Taranaki regions respectively. Collier & Kelly (2006) examined trends based on 8-10 years of annual sampling whereas most sites examined by Stark & Fowles (2006) had been sampled at least twice per year for ten years (range 6 – 38 sampling occasions).

Recommendations: *In my view trends testing should not be undertaken unless a time series comprising at least six data points has been compiled. Data series with 10 or more data points are preferable.*

Coping with variability

When evaluating long-term physico-chemical (or biological) data, it is important to take short-term temporal variability into account (Lowell & Culp 2002). If the concentration of a particular environmental variable can vary within one day or from day-to-day, then single measurements once per month are not likely to be indicative of average conditions during the

month. River flow is perhaps the best example of this and automated river level recording at 15 minute intervals is a cost-effective means of obtaining reliable flow data. A major cost of water level recording is the equipment required and its installation, so while the number of water level recording sites may be limited, it is cost-effective to maintain existing recorders in perpetuity (providing that they are still producing data of use to water managers). Furthermore, flow statistics (such as the median or mean annual low flow) are critically important in water management, and become more and more reliable as the years of continuous record increase.

Water quality variables such as temperature, and dissolved oxygen show significant diurnal variation, but provided the time of day is recorded, spot measurements made on a monthly basis can be interpreted in relation to the diurnal and seasonal cycles that are now well-understood. Similarly, concentrations of many water quality variables are flow- or temperature-related but, because of previous research focused on understanding such relationships, techniques for flow-correction and de-seasonalisation can be used to make sense of data from monthly sampling.

Biological communities are products of their environments – the fundamental basis for their value in biomonitoring. Poor quality environments support poor quality communities, and it follows that if habitats are improved, then an improvement in community "health" will follow. Biotic data generally show temporal variability over a longer term, so need not be sampled as frequently as water quality or flow.

Recommendations: *Annual sampling is the minimum recommendation for routine SoE monitoring. Sampling in two seasons, especially during the first few years, can hasten establishment of a time series of data. It is difficult to imagine why sampling more than four times per year would ever be required for SoE monitoring. If there is a perceived need for that, consider a separate "research" project.*

How much precision do you really need?

When designing long-term SoE monitoring programmes the specific management uses for information must be defined at the outset so that adequate statistical resolution can be achieved in those areas that are of importance to managers, and to avoid the unnecessary expense of over-defining trivial changes that may be well below any practical thresholds of concern. If water managers only want to detect a 20% change, for example, then why collect five replicate macroinvertebrate samples to detect a < 5% change in MCI, when a single sample can detect a change of around ± 11 % (Stark 1998)? For similar resources, a greater geographical coverage could be achieved, or supplementary synoptic surveys could be undertaken to address the research needs of specific management objectives.

Note, however, that in the first few years of SoE monitoring within a region it may be appropriate to sample a greater number of sites or collect replicate samples in order to define the variability within the region and the precision of estimates obtained from each sampling site. Ongoing SoE monitoring may, however, involve a carefully selected subset of these sites

and no sample replication in order to improve cost-effectiveness and to release resources for special investigations. The precision of estimates from single samples can be assumed based on previously collected replicate data, or, for the MCI family of indices, from the table of detectable differences provided by Stark (1998).

Recommendations: *Avoid the temptation to be too ambitious with SoE macroinvertebrate monitoring programmes. Keep it simple and maintain cost-effectiveness by collecting single hand-net samples from each monitoring site and use a coded abundance sample processing protocol. Although different biotic indices may reveal a slightly different picture (which can be confusing for laypersons and non-specialist water managers) most biotic indices are highly correlated with one another. Consequently, don't use too many different indices - the MCI is the preferred index for reporting macroinvertebrate SoE monitoring results.*

Methods consistency

The integrity of SoE monitoring programmes is dependent upon consistency. Changes to methods, varying the sites sampled, and missing sampling occasions can all have adverse impacts on the ability of SoE programmes to provide quality information to water managers. The recent development and widespread adoption of standard methods for macroinvertebrate (Stark *et al.* 2001) monitoring assists in maintaining methods consistency, but personnel also need to be suitably trained and experienced and to be collecting samples or processing them routinely in order to maintain a high and consistent standard.

Recommendations: *Data consistency and integrity are fundamental to defensible and robust SoE monitoring programmes. Consistency of sampling methods, sample processing methods, and using trained or experienced personnel can contribute greatly to achieving this. Although,* ***some*** *changes can be made without destroying the integrity of the time series of data, seek expert advice first!*

Sampling strategies

The ideal strategy for SoE monitoring is to select a number of sites within the region and sample them all at regular intervals (usually once per year at the same time each year) on an ongoing basis. However, other sampling strategies have been employed. Indeed, Horizons has sampled some sites on a rolling basis (*i.e.*, every three years) rather than each year.

The US EPA's Environmental Monitoring and Assessment Program (EMAP) is a research program that aims to develop the tools necessary for monitoring and assessing the status and trends of national ecological resources (http://www.epa.gov/emap/).

EMAP is based on a probabilistic sampling design that invariably results in a very large number of sampling sites in order to satisfy statistical requirements. The probabilistic design involves placing a grid over the study area and selecting sampling sites within each grid systematically from a random start location. This is a form of stratified random sampling that ensures geographic coverage, while ensuring that all possible sampling locations have equal

probabilities of being selected. It assumes no prior knowledge about the nature of the region or the sampling sites. This can prove expensive.

Perhaps because the probabilistic sampling designs result in large numbers of sites and given the large geographical areas of basins or study units in the USA, EMAP has developed several strategies for the allocation of sample sites in space and time.

Rotating panels: Under this scheme, a subset of the total number of SoE monitoring sites is sampled on each occasion. For example, say there were 40 SoE monitoring sites in a region, but only sufficient resources to sample 10 of them at a time. Each block of 10 sites would be spread at random throughout the entire region. Sites 1-10 would be sampled in Year 1, sites 11-20 in Year 2, sites 21-30 in Year 3 and sites 31-40 in Year 4. In Year 5, sites 1-10 would be revisited etc. This is an example of a four-year rotating panel. There is some coverage of the entire region each year, but it takes four years (in this case) for all sites in the entire region to be sampled (and the statistical requirements of the probabilistic sampling design to be met).

Rotating basins: The rotating basin design is similar to the above except that the sites in each block are confined to a particular basin or catchment. Under this scheme, all sites in a particular basin are sampled in one year, but it takes several years for the entire region to be sampled (and the statistical requirements of the probabilistic sampling design to be met).

The rotating panel and rotating basin strategies promoted by EMAP, which might appear to be a means of interrupting SoE monitoring, are, in fact, a compromise that enables finite resources (*i.e.*, personnel and funds) to cover large numbers of sites within large areas by spreading the effort out over several (say 4 - 5) years. However, there is a price to pay. It takes 4 - 5 years to cover the region and to meet statistical requirements. At any one site, because sampling times are 4-5 years apart it takes a long time to establish a reasonable time series of data, so trends analysis – a key objective of SoE monitoring - is compromised. In other words, it also takes 4-5 times as long to compile a time series suitable for trends testing. In reality, with rotating sapling strategy on a 4-5 year cycle, it will take 24 – 30 years of data collection before any trends testing can be contemplated!

A purely random strategy may be acceptable for selecting sampling sites in a region that is completely uncharacterised, but that is seldom the case in New Zealand. We have information about land-use, geology, stream type, source of flow, and River Environment Classification categories that can all provide a foundation for a stratified sampling design that would prove much more cost-effective.

Trends analysis becomes increasingly reliable when the time series lengthens or sampling is more frequent. By suspending or disrupting sampling it will take longer to assemble the time series data necessary for detection of trends.

Recommendations: *Decide on a basis for designing your SoE monitoring programme, select the sampling sites to be representative of the region and stick with it. Sites can be added and subtracted, but avoid chopping and changing too much. Ideally all SoE monitoring sites*

should be sampled at the same frequency (no less frequently than annually), otherwise it takes too long to establish a time series of data suitable for trends testing. Avoid the use of sampling strategies like rotating basins and rotating panels.

How many SoE monitoring sites should there be?

There are many factors to consider when deciding how many sites should be included in a macroinvertebrate SoE monitoring programme. The size of the region and the variety of river types should be considered. SoE monitoring sites should be located throughout the region in various river types and sizes (perhaps selected based on the River Environment Classification), with sites ranging from reference (near-pristine native vegetation) to heavily impacted by farming, forestry, or urban land-use.

Cost is an important consideration, as is the long-term political will to continue with the programme in perpetuity (or until a better alternative approach is developed). Clearly, the methods employed and the sampling effort (which can have a marked effect on costs) are also important factors.

A brief informal survey of other regional councils suggests some consistency in the number of sites surveyed. Greater Wellington, for example, surveyed 48 sites in 2002, 42 sites in 2003, and 56 sites in 2004. Hawkes Bay Regional Council increased their SoE programme from 38 sites in 2003 to 50 sites in 2005. Auckland Regional Council included 18 sites in 2001, 52 sites in 2002 - 2006 and 65 sites in 2007, however, prior to 2006 triplicate samples were collected from each sampling site, but between 2000 and 2001 processing changed from fixed count to coded abundance, and between 2005 and 2006 collection of triplicate hand-net samples was abandoned in favour of single hand-net samples. These changes were based on scientific evaluation of existing data, and enabled more sites to be added to the SoE monitoring programme and, at the same time, reduced the overall costs. Taranaki Regional Council has the longest-running SoE macroinvertebrate monitoring programme in New Zealand comprising 51 sites for most of the first 12 years with an increase to 56 sites in 2007 (Chris Fowles, Taranaki Regional Council, *pers. comm.*).

Recommendation: *Between 50 and 70 macroinvertebrate SoE monitoring sites would seem to be an appropriate number with more or less depending upon the size of the region, the variety of different river or land-use types, the information needs of water managers, and resource constraints (budget, time, personnel, equipment).*

Scope of SoE monitoring

Effective management of water resources (or environmental quality) requires knowledge of changes that occur in the environment and an understanding of the underlying cause/s of any changes that might be predicted or observed. The ability to distinguish anthropogenic changes from natural ones is also desirable. However, SoE monitoring does not have to meet all of these demands alone. Many of these issues are better addressed by undertaking special investigations or experimental research.

SoE monitoring involves ongoing routine collection of data from established sampling sites. These sites normally are selected to represent both reference and impacted conditions and common land-uses within the region. It is appropriate to monitor different variables at different time intervals (*e.g.*, water level – every 15 minutes, water quality – monthly, macroinvertebrates – annual). In my view, an essential characteristic of SoE monitoring programmes is that they collect data using standard methods, from established monitoring sites at the stated sampling frequency without interruption. The integrity of SoE monitoring programmes is dependent upon consistency. Changes to methods, varying the sites sampled, and missing sampling occasions can all have adverse impacts on the ability of SoE programmes to provide quality information to water managers. Note however, that it is data consistency that is important, not methods consistency *per se* (although the use of the same methods is a good was to minimise data inconsistency). For example, a single hand-net sample may be equivalent to collection of fiver replicate Surber samples. Although the methods may appear inconsistent, comparable estimates of taxon richness, MCI, QMCI (or SQMCI), and %EPT richness are likely to be obtained from either sampling method.

SPECIFIC COMMENTS ON HORIZONS' SOE MONITORING

Recommendations for future SoE monitoring

Number of sites

Seventy-eight different sites have been sampled as part of Horizon's macroinvertebrate SoE monitoring programme since it began in 1999 (Figure 1, Appendix 4). A total of 21 sites have been sampled on six or more occasions (Table 1). At the other extreme, 19 sites have been sampled only once, 12 sites on two occasions, and 21 sites only three times (Appendix 4). Twenty-two sites are regarded as permanent, four as reference sites and sampling at one site has been discontinued. Fifty-one sites have been sampled on a rolling basis (*i.e.*, not every year), with 87% of these sampled between one and three times over the last nine years.

The Manawatu-Wanganui region covers a large area compared with regions like Taranaki and Auckland and yet only 35 sites were sampled in 2007 (although five replicate Surber samples (area 0.1m^2, mesh 250µm) were collected from riffle habitat at each site). In Taranaki, 56 sites are sampled twice per year and in Auckland 65 sites are sampled annually. In both cases sampling is semi-quantitative (D-net 0.5 mm mesh) (Protocol C1 Stark *et al.* 2001) with samples processed to yield coded-abundance data (Protocol P1 Stark *et al.* 2001). I estimate that the annual cost of SoE macroinvertebrate sample processing for Taranaki and Auckland could be 30 - 60% of the cost for Manawatu-Wanganui (assuming the same charge-out rates), and yet their programmes include 60 – 85% more sites.

Recommendation: *Horizons considers increasing the number of SoE macroinvertebrate monitoring sites to between 50 and 60.*

Site selection

SoE monitoring sites should cover the region, represent different land-use types, and cover a range of river health from reference (near-pristine) sites through to sites that are subjected to disturbance or pollution.

All 21 sites that have already been sampled at least six times should be included because that retains the value of the investment already made in their data series. Five additional sites have been sampled four or five times, 21 sites three times, and 12 sites twice, and 19 sites on one occasion. When considering which 39 (or so) sites to add to the SoE programme, priority should be given to sites that have already been sampled several times unless there are good reasons not to do so.

Recommendation: *When deciding which sites to add to the SoE programme priority should be given to sites that have already been sampled several times unless there are good reasons not to do so.*

SoE sampling strategies

Although there are techniques like rotating panels and rotating basins that enable a greater number of sites to be included in a monitoring programme than resources can cater for on any one occasion, these methods should be avoided because they complicate data analyses and compromise the ability to undertake trends testing (by delaying the compilation of long time series).

Monitoring should be undertaken each year at the same time of year, preferably under similar flow conditions. Sampling should be confined to a period of consecutive days or a few weeks (at most) to minimise the effects of temporal variability. Avoid the temptation to collect too much extraneous information (*e.g.*, complex and time consuming habitat assessments that record data that seldom change from year to year), if it extends the sampling time unduly.

Recommendation: *Decide on the SoE monitoring sites and sample them within as short a time-frame as possible under similar river flow conditions at the same time of year each year..*

SoE sampling methods and indices for reporting

The first decision to make is whether quantitative macroinvertebrate data are required for SoE biomonitoring. Some scientists have a strong preference for quantitative data. With quantitative data you can calculate macroinvertebrate densities – but these can vary by several orders of magnitude from time to time at any given site for reasons unrelated to river health, so what value are they for SoE biomonitoring? Quantitative data enable %EPT abundance to be calculated. This is a well-performing index of stream health, but tends to be highly correlated with other well-performing indices such as the MCI. Quantitative data allow the percentage community composition to be determined, but isn't it enough to know which taxa dominate community composition? Estimates of taxon richness may be more reliable when determined from samples collected from a strictly-defined area of streambed, but taxon richness is not a

good indicator of stream health. Highest values tend to be associated with slightly enriched habitats, with lowest taxon richness often associated with swift, steep rivers with pristine water quality (Stark & Maxted 2007a). Incidentally, despite the quantitative sampling, the 2007 SoE monitoring report used the quantitative data only for calculating percentage community composition, QMCI, and %EPT abundance (Dewson *et al.* 2007). There was no discussion of invertebrate densities in the report.

Coded-abundance (semi-quantitative) data are unsuitable for calculating invertebrate densities, and %EPT abundance. However, indices like the MCI, SQMCI (which is numerically similar to QMCI), % EPT richness, and taxon richness can all be calculated.

When the issue of cost or cost-effectiveness is introduced into the discussion, I believe that the choice of sampling and sample processing methods becomes very clear indeed.

Stark (1998) provided a table of Detectable Differences (DD) that highlights the precision of estimates of the MCI and SQMCI or QMCI that can be achieved by semi-quantitative (hand-net) or quantitative (Surber) sampling. Put simply, two MCI values calculated from single hand-net samples would need to differ by 10.83 in order to be considered significantly different. This is equivalent to a difference of 11% compared with an average MCI of 100. To achieve such precision in the MCI by Surber sampling would require at least four replicate samples. Similarly, a single hand-net sample provides an estimate of the SQMCI that would take three Surber samples to achieve.

Although there are a variety of different indices that can be used to report macroinvertebrate SoE results, most are highly correlated with one another, although there will be differences on a site by site basis. These differences may cause confusion for lay persons if several different indices are used in SoE reporting. The MCI (and the equivalent MCI-sb for soft-bottomed streams) is widely used, with various aspects of its performance well investigated (Stark 1985, 1993, 1998, Stark & Maxted 2007a, b), and is the biotic index of choice for macroinvertebrate SoE reporting.

Recommendations: *Single hand-net samples per site are recommended for SoE macroinvertebrate monitoring programme, and provide estimates of MCI similar in precision to those obtained from quantitative sampling costing more than four times as much. The MCI (and MCI-sb) is the preferred biotic index for SoE reporting.*

Compatibility with existing methods and existing data

Horizons' SoE programme is based upon collection of five replicate Surber samples. For the MCI this provides comparable performance to the single hand-net sample (see Appendix 5). For the QMCI, five Surber samples combined perform better than a single hand-net sample for the numerically equivalent SQMCI (albeit at much greater cost).

The collection of five replicate Surber samples is, for all practical purposes, equivalent in sampling effort to collection of a single D-net sample (Stark *et al.* 2001: cf. Protocols C1 & C3). Estimates of taxon richness, and biotic indices calculated from data from five pooled Surber samples and a single D-net sample should be directly comparable. Horizon's could, therefore, change to collection of single D-net samples for future sampling without compromising the existing time series of taxon richness, MCI, and QMCI/SQMCI. It is most unlikely that there would be a noticeable change in the MCI, taxon richness, or %EPT richness as a consequence of this change in sampling method because overall sampling volumes would be similar. As Figure 2 in Stark (1993) indicates, because they are average scores per taxon, the MCI and QMCI are relatively independent of sample size – reliable estimates can be obtained from a single hand-net sample or a composite sample of 5 – 7 Surber samples. Only with samples smaller than this is there any possibility that unreliable MCI values could be obtained

Another advantage of the MCI is that it relies on presence-absence data for its calculation. It does not matter whether all animals have been counted, relative- (or coded-) abundances have been determined, or if a fixed count with scan for rare taxa has been done. In each case, the taxon list should be identical resulting in the same estimates of the MCI, %EPT richness, or taxon richness. Thus, any council that collects single hand-net samples according to Protocol C1, or at least five Surber samples (Protocol C3) (with data combined), will be able to calculate MCI values that are comparable, provided that all taxa present in each sample are identified correctly to the level (mostly generic) required for MCI calculation. The key point here is that there can be data comparability without necessarily using identical methods.

Recommendations: *The MCI calculated from single hand-net samples per site is the most cost-effective biotic index for reporting SoE monitoring results. Each sample should be equivalent in sampling effort to six Surber samples (i.e., 0.6m^2). This is similar to the sampling effort used previously by Horizons and should not cause any significant issues for the integrity of the existing data series.*

Responses to additional questions

In this section of this report I address additional questions posed by Horizons. In some cases the responses will be clear from previous discussion, but here serve as a summary of key points.

The One Plan uses QMCI as a measure of water quality for compliance monitoring - is it best to keep the sampling method (*i.e.*, Surber samplers – 5 replicates) aligned to this for reference purposes?

No, the QMCI is an appropriate biotic index for compliance monitoring when, for example, all samples are collected on a single day from upstream and downstream of the discharge point (see Stark & Maxted 2007a Section 3.2.1). In my view, SoE monitoring is much more cost-effective if the MCI is used. The lower cost of MCI-based SoE monitoring can (ideally) permit more sites to be included in the SoE monitoring programme or can (alternatively) free resources that can be devoted to

special investigations.

Would it be better to sample fewer sites more often to reduce the effects of seasonal/flow variation?

No, research in Taranaki streams suggests that seasonal variation in the MCI is within 3% of the annual mean for a hard-bottomed stream, and for soft-bottomed streams (MCI-sb) it is within 4.7%. Most other indices showed greater seasonal variability in hard-bottomed streams (SQMCI ±4.3%, %EPT richness ±7.4%, taxon richness ±7.7%) and soft-bottomed streams (SQMCI-sb ±3.6%, taxon richness ±4.7%, %EPT richness ±11.2%). Variability in the MCI associated with season is within the sampling error and can safely be ignored.

In New Zealand streams flow variability (especially large floods and prolonged periods of low flow) can affect biotic index values significantly. The solution is not to sample more often, but to sample at a similar time of year under similar flow conditions (if possible), and to have flow data available to inform when significant floods (or droughts) occurred.

In my view, detecting trends in macroinvertebrate SoE monitoring data (i.e.,, addressing the question "Are conditions getting better or worse or staying the same?") is more important than assigning sites to quality classes based on the latest results. Over the long term, season and flow variability should not have a major influence on trends if sampling is undertaken at the same time each year when river flow conditions permit it. In my view, it is better to put additional resources into sampling more sites.

The Surber sampler is a quantitative method and suitable for the calculation of QMCI. Can kick net samples give a robust sample to be used for calculating QMCI?

Yes. A 200 Fixed Count from a kick net sample will provide a robust estimate of the QMCI. However, the SQMCI (calculated from coded-abundance data) is essentially the same index and was designed to be calculated from kick net samples. QMCI and SQMCI values can be compared directly.

If so should Horizons be adopting a kick net sampling technique in line with other Regional Councils particularly in light of the proposed national invertebrate database and the fact that it is more cost effective?

Yes. The cost-saving could be put towards increasing to number of SoE monitoring sites from 35 to 50 - 60.

What will be compromised if this is adopted and what are the risks involved?

A single kick sample per site collected according to Protocol C1 (Stark et al. 2001) with sampling effort equivalent to sampling 0.6m^2 (the minimum recommended in the protocol) will provide data that are compatible with existing data held by Horizons (from five replicate Surber samples combined). Estimates of MCI, %EPT richness,

and SQMCI should be directly comparable with MCI, %EPT richness, and QMCI values calculated from the five pooled Surber samples collected to date. Macroinvertebrate densities or %EPT abundance will not be able to be calculated (although despite the collection of quantitative samples the former were not calculated in the 2007 SoE monitoring report (Dewson et al. 2007)). Densities and taxonomic richness are not good indicators of river health. Densities can vary by several orders of magnitude from time to time due to the influence of flows and prolonged recessions, and %EPT abundance tends to be highly correlated with other indices (such as the MCI and QMCI – see Figure 7 in Dewson et al. 2007). Taxonomic richness, although unlikely to be affected greatly by the change in sampling method, shows a non-linear response to the enrichment gradient with highest values at slightly enriched sites.

Is a pooled Surber sample feasible (*i.e.*, is a QMCI still able to be calculated) for bio-monitoring purposes as a cheaper alternative? Pooled Surber samples were used for QMCI calculations for the New Zealand's National River Water Quality Network (NRWQN) surveys (Scarsbrook et al, 2000).

In my view pooled Surber samples is a not a good approach. When the NRWQN was being designed in 1984 I was seconded to the Water Quality Centre (Hamilton) preparing Stark (1985) and I was asked to comment on the proposed methods. I wrote at the time that I did not agree that collecting seven Surber samples and combining them in the field was a sensible thing to do and I have not changed my opinion. It does not save much time – there is still the same volume of sample to process. The samples are large, extensive sub-sampling is required, and they are much more tedious to process than the same number of separate samples. By combining the samples, one cannot use statistics to determine standard errors for estimates of density, and various indices. The only advantage of this approach over single-hand-net samples is that densities can be calculated because a defined area of 0.7m^2 has been sampled.

Horizons use whole counts to calculate biotic indices at the moment – should we adopt a fixed count method for improved cost effectiveness?

I would strongly suggest that the present sampling methods (five replicate Surber samples) should be abandoned. A 200 Fixed Count (Protocol P2) applied to single hand-net samples (Protocol C1) is a step in the right direction and would reduce sample processing costs to 20-25% of present. There would be a saving of time in the field too. It is much quicker to collect a single hand-net sample than five replicate Surber samples.

Additional cost saving could be achieved by adopting the coded abundance processing protocol (P1). In my experience the time required for P1 is approximately 10% less than for P2.

How will this compromise our previous data and what are the risks involved?

A single hand-net sample equivalent in sampling effort to five replicate Surber samples can be collected in a manner consistent with Protocol C1. If care is taken to ensure that this is the case, there should be a high level of compatibility with the

existing data series. There would be no problem at all for the MCI – once samples reach a minimum threshold (approximately equivalent to three Surber samples), estimates of MCI are similar with increasing sample size. Note however, that MCI values for existing data need to be calculated from five pooled Surber samples, not as a mean of MCI values for the five separate samples. Estimates of taxonomic richness (which I do not regard as a good indicator of stream health) are at greatest risk because more taxa tend to be collected as sample size increases. The key to consistency is ensuring that each hand-net sample is equivalent in effort to five Surber samples.

Should the protocol for sampling soft bottom streams be adopted (Stark *et al*, 2001) for the small number of soft bottomed streams within the region?

Not necessarily. The objective of the soft-bottomed stream sampling protocols (C2 or C4) is to ensure that the sampling net does not get filled with fine sediments. The aim of all macroinvertebrate sampling is to fill the net with animals and minimise the amount of other material. If you are able to collect manageable samples from soft-bottomed streams using the hard-bottomed stream protocol (C1), then there is no reason to change. You should, however, use the MCI-sb for reporting on the state of soft-bottomed streams.

Presently six separate biotic indices, MCI, QMCI, %EPT taxa, %EPT individuals, Number of Taxa and O/E ratio, are measured. Is it necessary to have so many

The short answer is "No". However, it depends on the target audience for your SoE reports. Technical reports might well calculate, compare and contrast various biotic indices. However, for the general public and laypersons (and that may include many regional councillors) I believe that SoE monitoring reports should be kept as simple as possible. Such reports should be about communicating results not impressing people with science and technical terms (although some of these usually are necessary). Most of the above-named indices will be highly correlated with one another. In that respect any one of them can do the job. On the other hand, there will be some differences between the indices. For example, the MCI and QMCI may rank sites slightly differently. These differences can confuse people. At a technical level the explanations are relatively easy, but it's detail that people don't really need or want to know. I would opt for the MCI (and MCI-sb) for reporting SoE monitoring results. I do not recommend the QMCI for SoE monitoring – it is best-suited to compliance monitoring (Stark & Maxted 2007a).

Will a change in methods affect the O/E or Bayesian Belief Network programmes currently underway?

The O/E taxonomic richness ratio should be unaffected by the change from five Surber samples to a single hand-net sample because sampling effort is similar, provided that taxonomic richness estimates are obtained from five pooled Surber samples per site and are not averages from the five replicates.

The Bayesian Belief Network project funded by Horizons and Hawkes Bay Regional Councils aims to develop a predictive model of invertebrate communities in relation to

environmental variables, with a GIS interface. These variables could then be manipulated within the model to look at invert community effects and would be useful for consents work to predict community changes. The final report on this project is still being completed (Kate McArthur, Horizons pers. comm.). However, since collection of single hand-net samples from each SoE monitoring site will collect data comparable with five Surber samples I would not expect the BBN programme to be compromised (unless it requires invertebrate density data).

If the sampling methods were changed what would be the best programmed method of change over without jeopardising the integrity of the existing database?

I recommend that single hand-net samples should be collected from each monitoring site on the next SoE monitoring occasion. An increase in the number of sites should occur as soon as possible so that data points are added to the time series at the earliest opportunity.

What would be a good timeline of method changes with the current Biomonitoring programme?

The change in sampling method can be implemented immediately. If Horizons has some reservations concerning moving from full counts (with subsampling) to coded-abundance sample processing, then a 200 fixed count (with scan for rare taxa) could be introduced in the interim, with a subsequent review undertaken later to decide whether or not to move to coded abundances. Both methods produce a complete list of the taxa present in each sample so there is complete compatibility as far as indices such as the MCI, %EPT richness, taxonomic richness, and O/E richness are concerned.

ACKNOWLEDGEMENTS

Thanks to Kate McArthur and Carol Nicholson (Horizons Regional Council) for providing the data, details of methods, and history of SoE monitoring undertaken by Horizons Regional Council in the Manawatu-Wanganui region and Maree Clark (Horizons) for the providing the biomonitoring site map. Kate McArthur also provided helpful comments on the first draft of this report.

REFERENCES

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.

Boothyrod IKG, Stark JD. 2000. Use of invertebrates in monitoring. p344- 373 in Collier KJ & Winterbourn MJ (Eds) New Zealand Stream Invertebrates: Ecology and

Implications for Management. New Zealand Limnological Society / NIWA. Caxton Press, Christchurch.

Byl T D, Smith G F 1994. Biomonitoring our streams. What's it all about? USGS Open-File Report 94-368. Nashville, Tennessee. 8p.

Charteris SC, Death RG, Francis K, Minchin S, Boisen R, Vosper A 2000. River health of the Manawatu-Wanganui Region. State of the Environmental Report 2000 - Invertebrate and Periphyton Communities. Massey University. 50p.

Collier KJ 2006. Temporal trends in macroinvertebrate metrics for some Waikato streams. New Zealand Natural Sciences 31:79-91.

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. 28p.

Cook T, Death RG, Francis K, Hamer M, Nicholson C 2001. River health of the Manawatu-Wanganui Region. State of the Environmental Report 2001 - Invertebrate and Periphyton Communities. Massey University. 51p.

Death F, Death RG 2005. River health of the Manawatu-Wanganui Region. State of the Environmental Report 2005 - Invertebrate and Periphyton Communities. Massey University. 73p.

Death F, Death RG 2006. River health of the Manawatu-Wanganui Region. State of the Environmental Report 2006 - Invertebrate and Periphyton Communities. Massey University. 65p.

Death RG 1999. River health of the Manawatu-Wanganui Region. State of the Environmental Report 1999 - Invertebrate and Periphyton Communities. Massey University. 88p.

Death RG, McArthur K, Pedley R, Johnston I, Dewson ZS. 2002. River health of the Manawatu-Wanganui Region. State of the Environmental Report 2002 - Invertebrate and Periphyton Communities. Massey University. 65p.

Death RG, Makan T, Weaver K, Zimmerman EM 2003. River health of the Manawatu-Wanganui Region. State of the Environmental Report 2003 - Invertebrate and Periphyton Communities. Massey University. 67p.

Death RG, Death F 2004. River health of the Manawatu-Wanganui Region. State of the Environmental Report 2004 - Invertebrate and Periphyton Communities. Massey University. 68p.

Dewson ZS, Death F, Death RG 2007. River health of the Manawatu-Wanganui Region. State of the Environmental Report 2007 - Invertebrate and Periphyton Communities. Massey University. 63p.

Hirsch A 1958. Biological evaluation of organic pollution of New Zealand streams. New Zealand Journal of Science 1 (4):500-551.

Joy MK, Death RG 2002. Predictive modelling of freshwater fish as a biomonitoring tool in New Zealand. Freshwater Biology 47(11):2261-2275.

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.

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

Lowell RB, Culp JM 2002. Implications of sampling frequency for detecting temporal patterns during environmental effects monitoring. Water Quality Research Journal of Canada 37 (1). 119-132.

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

Rosenberg DM, Resh VH. 1993. Freshwater biomonitoring and benthic invertebrates. Chapman & Hall, London.

Scarsbrook MR, Boothroyd IKG, Quinn JM 2000. 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.

Simpson JC, Norris RH 1997. Biological assessment of river quality: development of AUSRIVAS models and outputs. p 125-142 in Assessing the biological quality of fresh waters, RIVPACS and other techniques, edited by JF Wright, DW Sutcliffe, MT Furse. Ambleside, Cumbria, Freshwater Biological Association.

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.

Stansfield B 2001. Effects of sampling frequency and laboratory detection limits on the determination of times series water quality trends. New Zealand Journal of Marine and Freshwater Research 35:1071-1075.

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 1993. 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.

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

Stark JD, Maxted JR 2007a: A user guide for the MCI. Ministry for the Environment, Wellington, New Zealand. (http://www.mfe.govt.nz).

Stark JD, Maxted JR 2007b: A biotic index for New Zealand' soft-bottomed streams. New Zealand Journal of Marine and Freshwater Research 41:43-61.

Suckling DM 1982. Organic wastewater effects on benthic invertebrates in the Manawatu River. New Zealand Journal of Marine and Freshwater Research 16:263-270.

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

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 LOWESS (tension = 0.7) plots MCI and QMCI versus time for 21 Manawatu-Wanganui streams and rivers. Results of the Mann-Kendall non-parametric test for assessing the statistical significance of trends are given also.

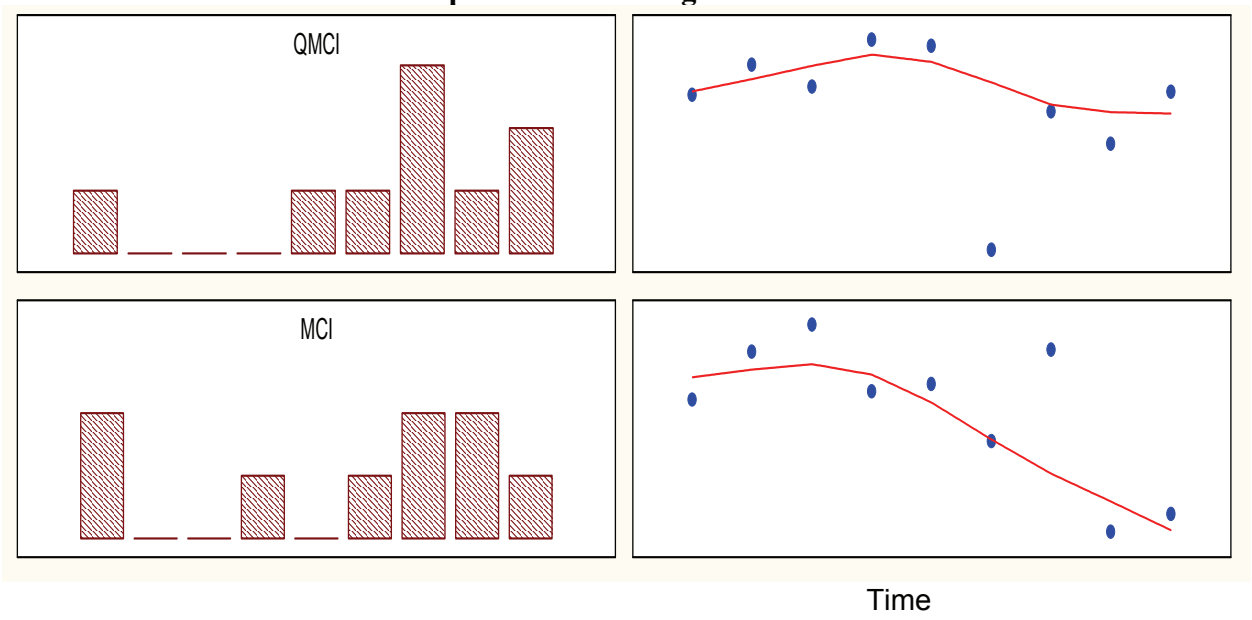
Results in red type are considered statistically significant at the 1% level, and results in orange type at the 5% level. The sign of the Kendall tau statistic (*i.e.*, ±) indicates whether the trend is positive (increasing with time) or negative (decreasing with time).[4] See Tables 1 and 2 for key to site codes and locations.

The histograms to the left of the LOWESS plot show the distribution of data values for each of the biological indices, in nine bins. For the graph on the next page (Hautapu River u/s Rangitikei River) there is one low QMCI value and two low MCI values. These are denoted by the left-most bars in the two bar graphs and can be seen as the lowest one (QMCI) or two (MCI) blue dots on the LOWESS plots.

There is no way to label scales on the axes of the LOWESS plots produced by STATISTICA using the procedure outline in the footnote below when two Y-variables have very different scales (i.e., MCI: 0 – 200, QMCI: 0 – 10). Automatic scaling works fine but as soon as manual axis labels are inserted both graphs must have the same scale. However, it is a very quick procedure for obtaining an overall picture of trends. The scatterplot procedure should be used if fully labelled graphs are required.

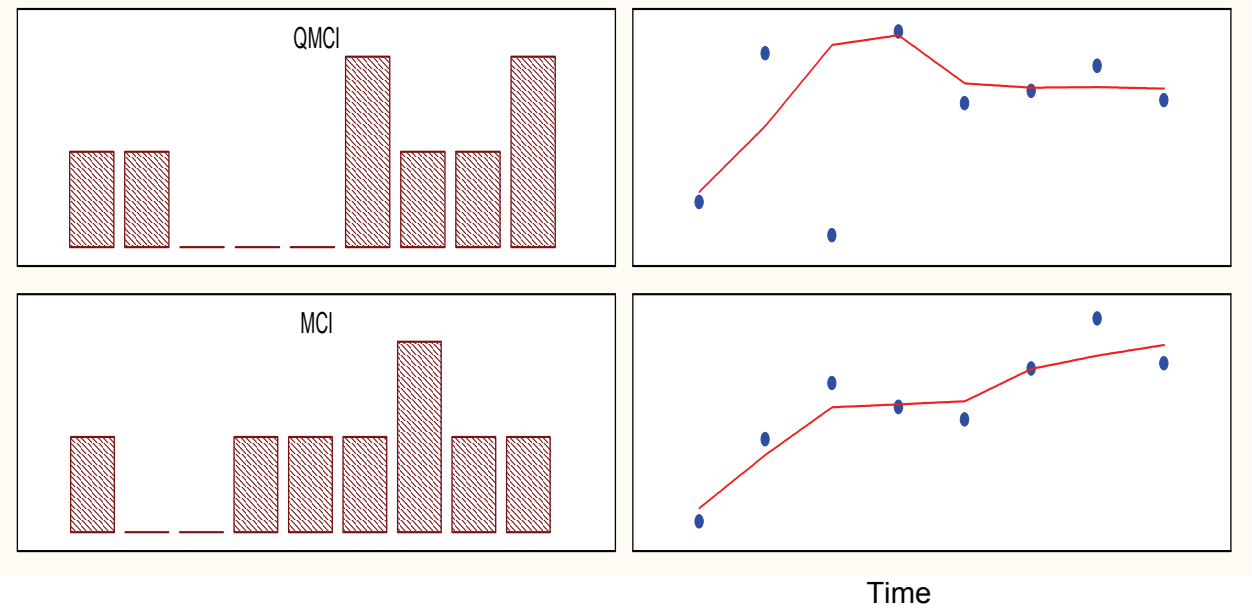
[4] These results were obtained with STATISTICA 8.0 using the Statistics Batch (By-Group) Analyses, non-parametric correlations (Kendall tau) with MCI and QMCI as the first variables and YEAR (expressed as a whole number *e.g.*, 2002) as the second variable. The grouping- or by-variable was SITE$ (*i.e.*, a code for each sampling site). The radio button labelled "create Kendall tau" should be checked and "all results" and "detailed results" should be selected on the General tab of the analysis. Once the graph was displayed double-clicking on the fitted linear regression line enables the fit to be changed to LOWESS with the tension (called stiffness in this module) to be changed from the default of 0.25 to 0.7.

Hautapu River u/s Rangitikei River



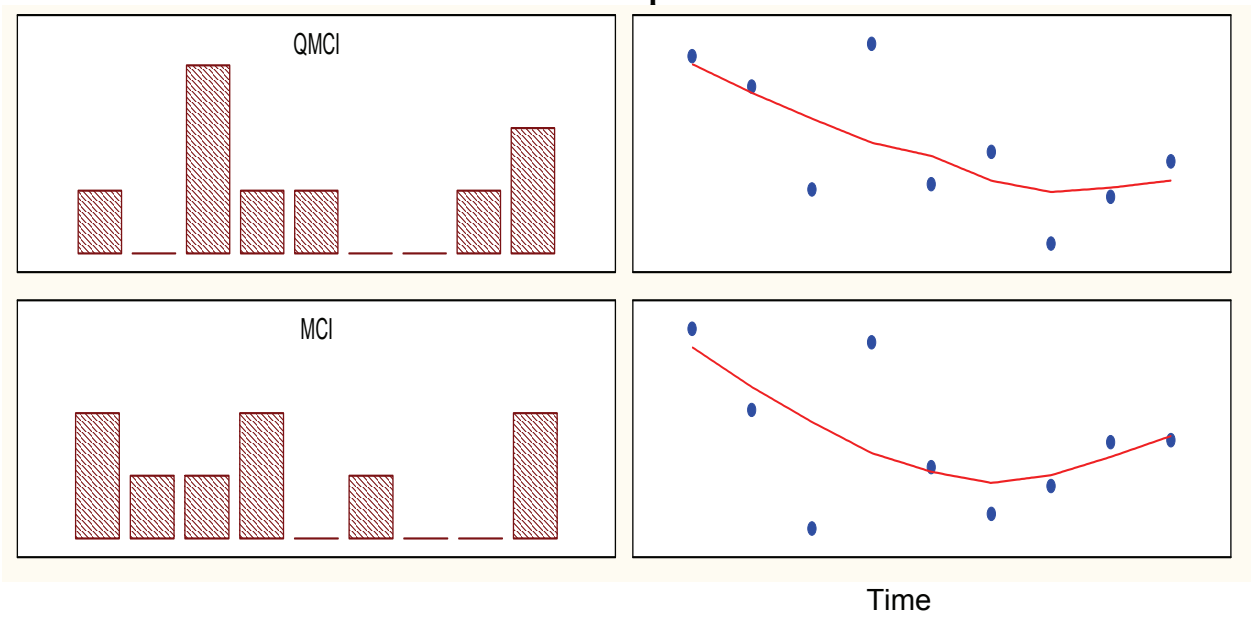
	Valid - N	Kendall – Tau	Z	p-level	p-exact - 1-tailed
QMCI & Year	9	-0.222222	-0.83406	0.404249	.238
MCI & Year	9	-0.333333	-1.25109	0.210903	.130

Makakahi River - Konini



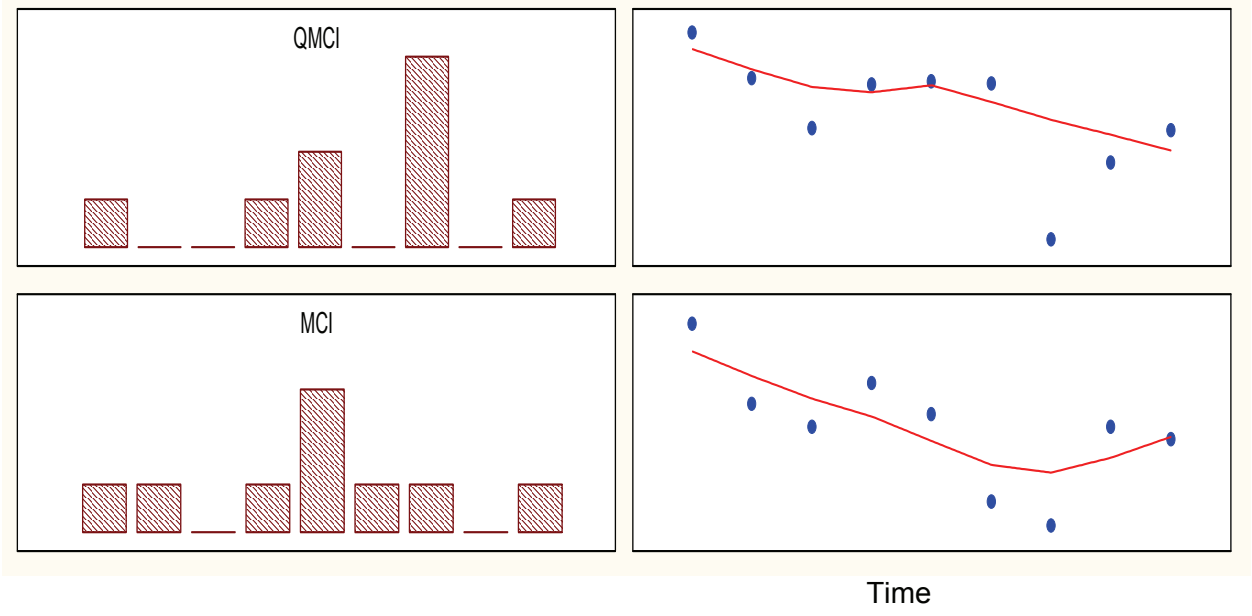
	Valid - N	Kendall – Tau	Z	p-level	p-exact - 1-tailed
QMCI & Year	8	0.142857	0.494872	0.620691	.360
MCI & Year	8	0.714286	2.474358	0.013348	.007

Manawatu River – Hopelands Reserve



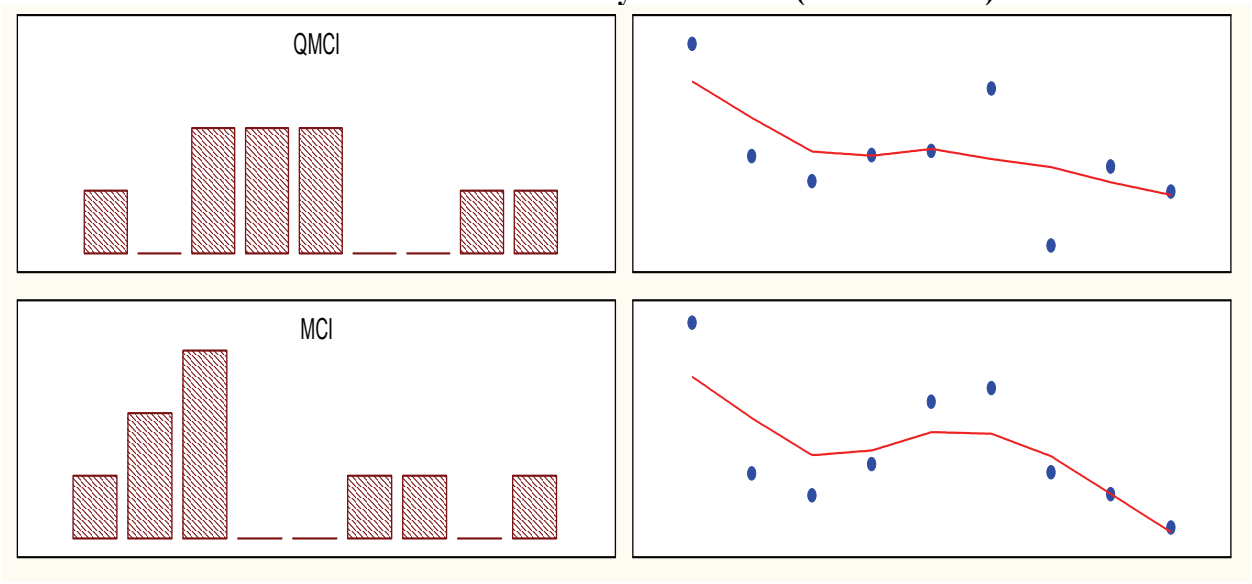
	Valid - N	Kendall - Tau	Z	p-level	p-exact - 1-tailed
QMCI & Year	9	-0.388889	-1.45960	0.144400	.090
MCI & Year	9	-0.166667	-0.62554	0.531615	.306

Manawatu River – Teachers' College + Maxwells Line



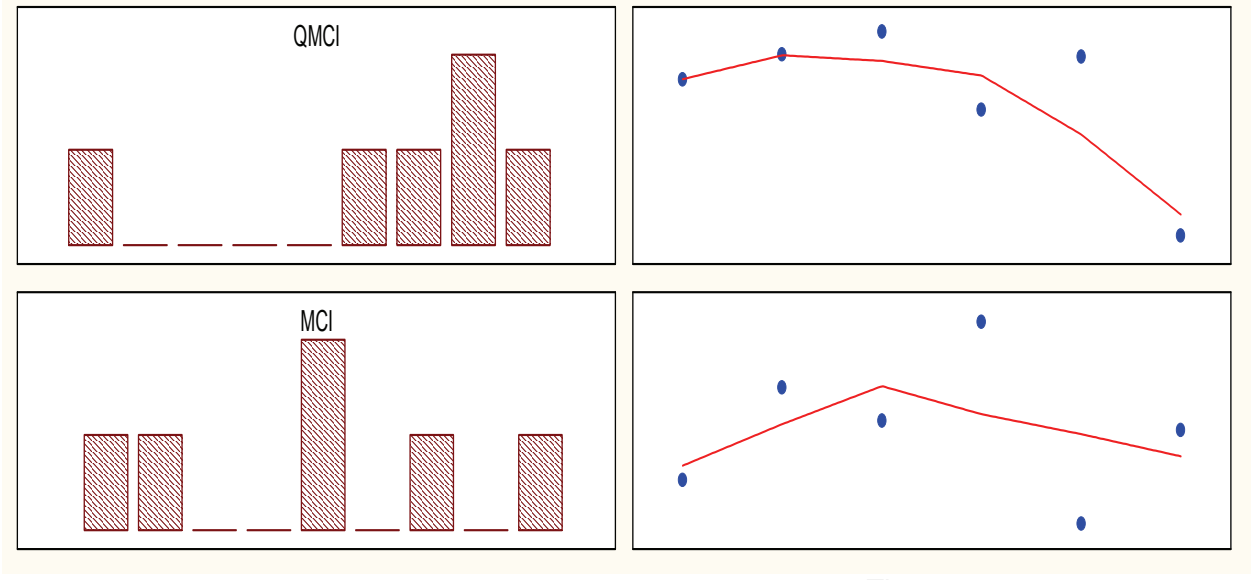
	Valid - N	Kendall - Tau	Z	p-level	p-exact - 1-tailed
QMCI & Year	9	-0.555556	-2.08514	0.037056	.022
MCI & Year	9	-0.591608	-2.22046	0.026388	.022

Manawatu River – 42 Mile Hydro Station (Karere Road)



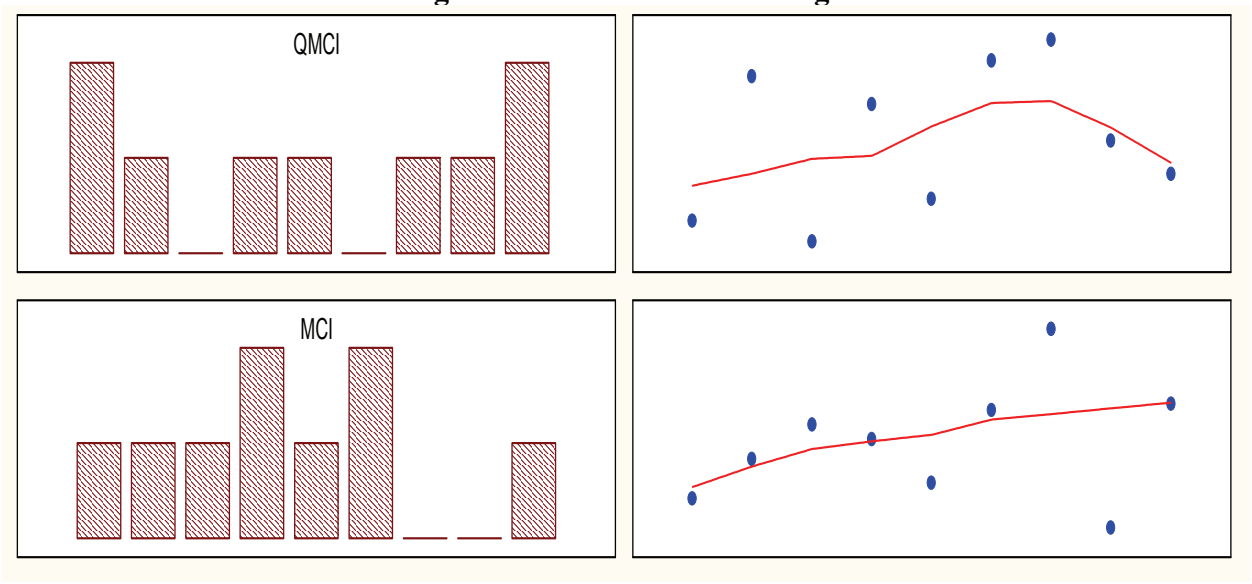
	Valid - N	Kendall - Tau	Z	p-level	p-exact - 1-tailed
QMCI & Year	9	-0.333333	-1.25109	0.210903	.130
MCI & Year	9	-0.333333	-1.25109	0.210903	.130

Manawatu River - Whirokino



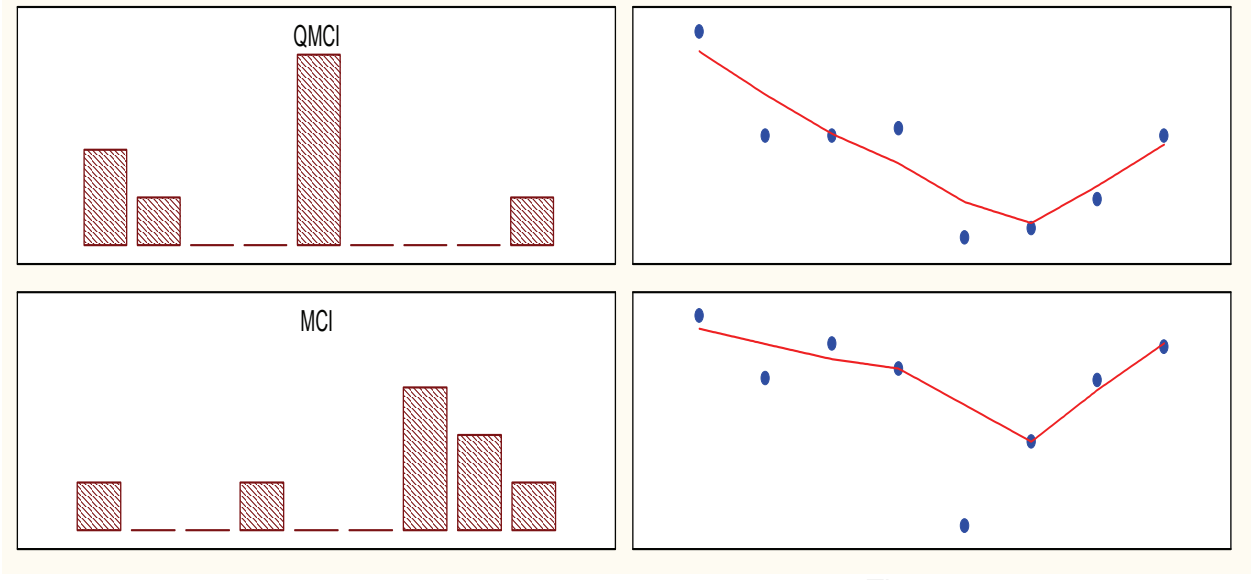
	Valid - N	Kendall - Tau	Z	p-level	p-exact - 1-tailed
QMCI & Year	6	-0.333333	-0.939336	0.347558	.235
MCI & Year	6	-0.066667	-0.187867	0.850981	.500

Mangatainoka River – SH2 Bridge



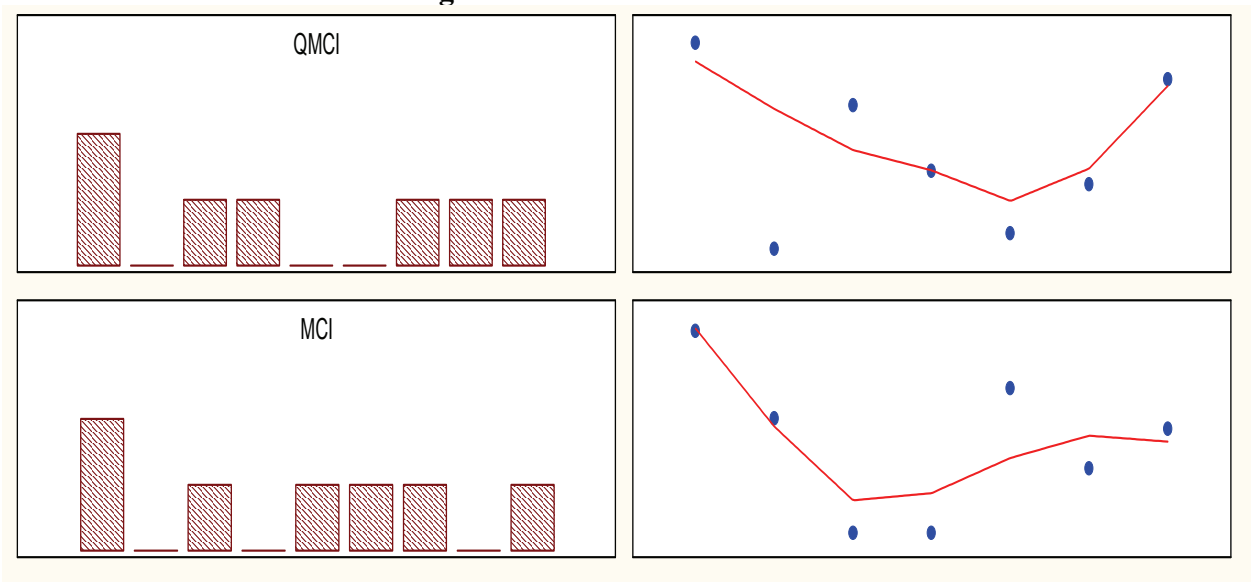
	Valid - N	Kendall - Tau	Z	p-level	p-exact - 1-tailed
QMCI & Year	9	0.222222	0.834058	0.404249	.238
MCI & Year	9	0.333333	1.251086	0.210903	.130

Mangatera Stream – Timber Bay



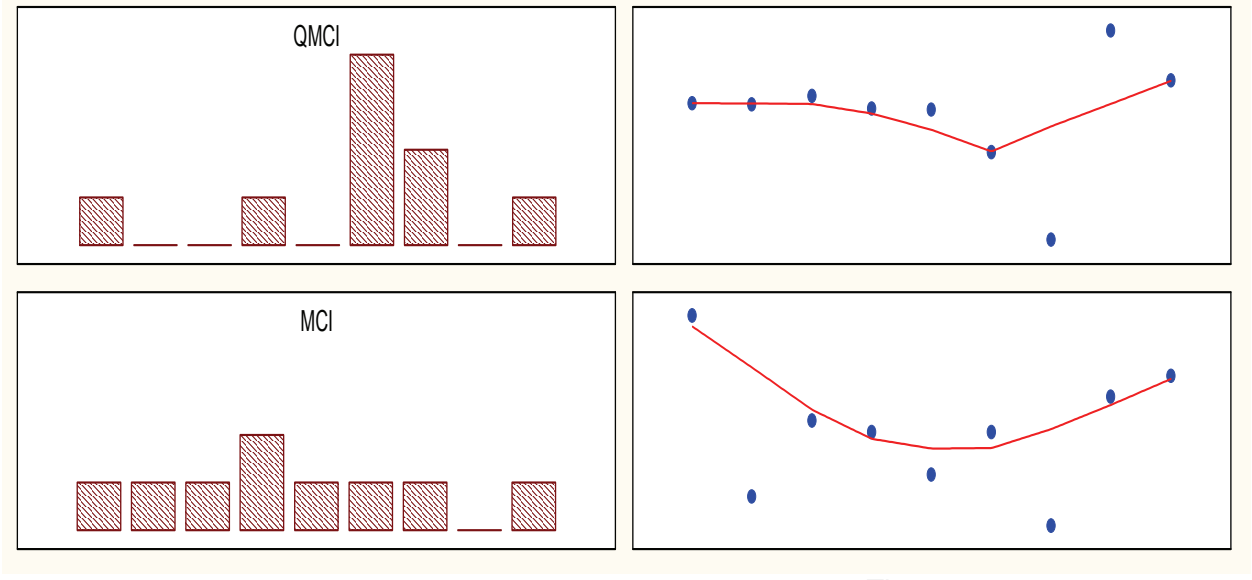
	Valid - N	Kendall - Tau	Z	p-level	p-exact - 1-tailed
QMCI & Year	8	-0.340168	-1.17838	0.238647	.199
MCI & Year	8	-0.285714	-0.98974	0.322300	.199

Mangawhero River d/s Makotuku



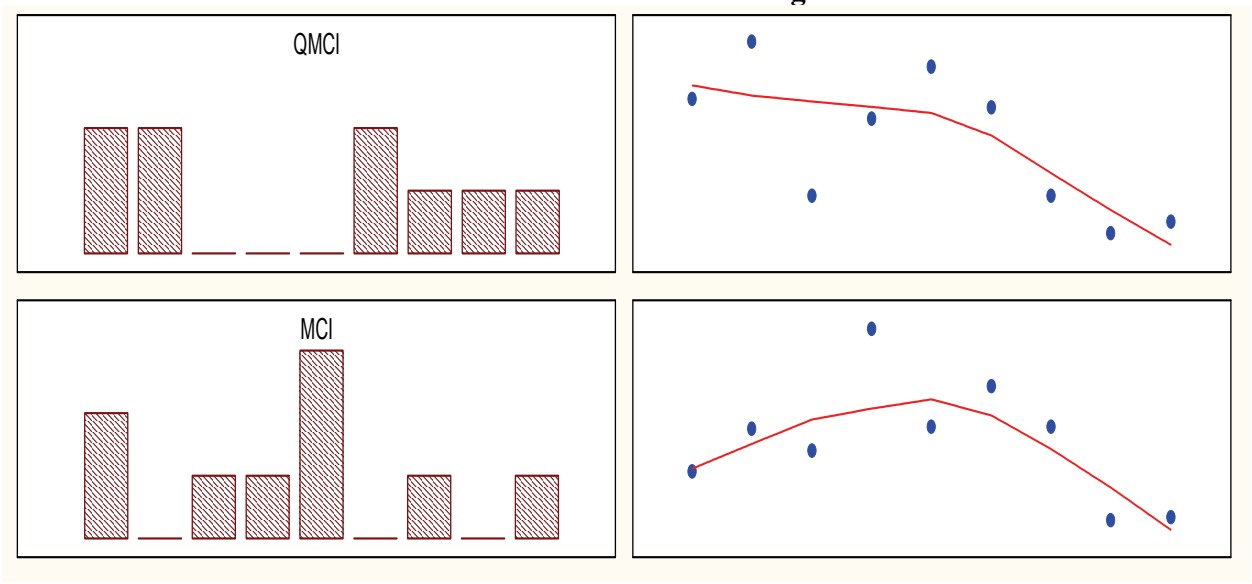
	Valid - N	Kendall - Tau	Z	p-level	p-exact - 1-tailed
QMCI & Year	7	-0.047619	-0.150188	0.880616	.500
MCI & Year	7	-0.142857	-0.450564	0.652304	.386

Mangawhero River – DOC Headquarters



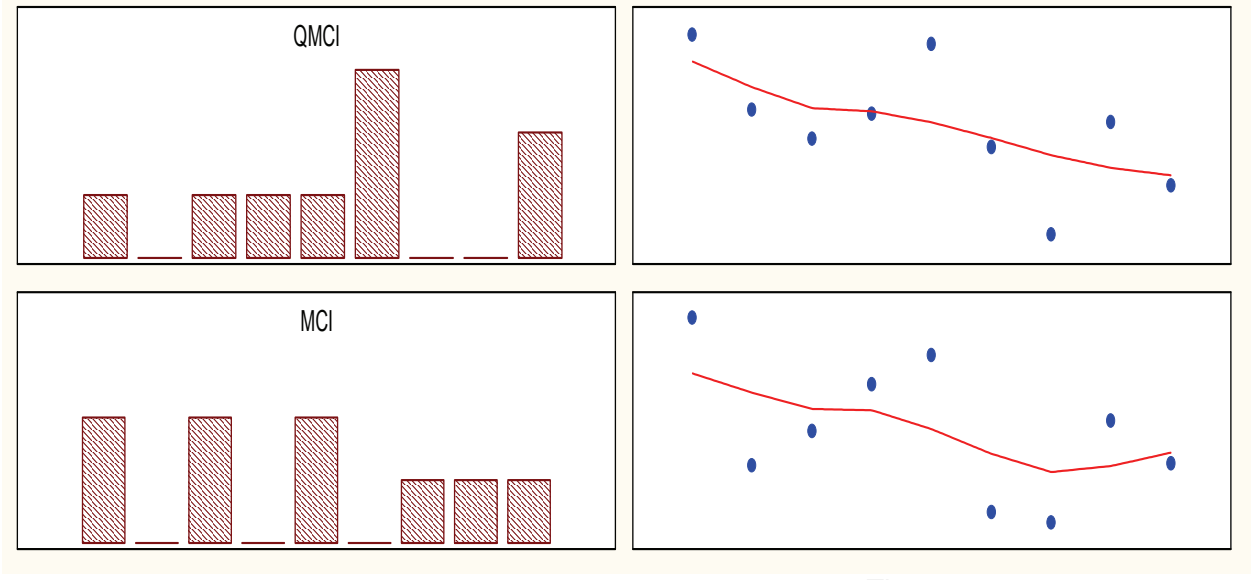
	Valid - N	Kendall - Tau	Z	p-level	p-exact - 1-tailed
QMCI & Year	9	-0.111111	-0.417029	0.676657	.381
MCI & Year	9	0.028172	0.105736	0.915792	.540

Oroua River – Awahuri Bridge



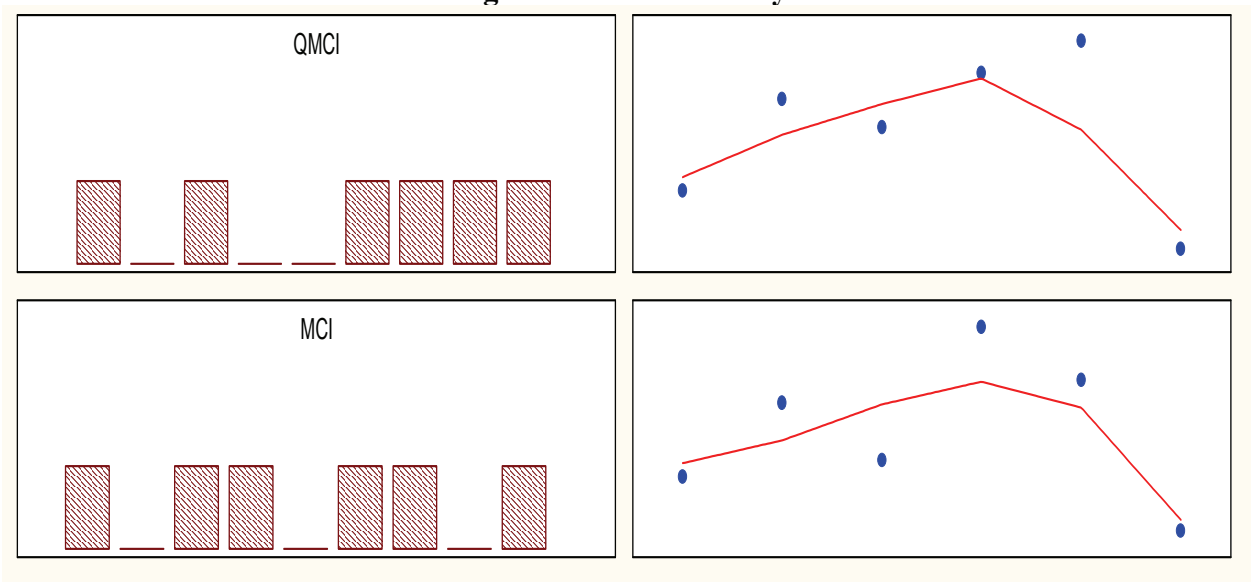
	Valid - N	Kendall - Tau	Z	p-level	p-exact - 1-tailed
QMCI & Year	9	-0.555556	-2.08514	0.037056	.022
MCI & Year	9	-0.111111	-0.41703	0.676657	.381

Oroua River – Nelson Street



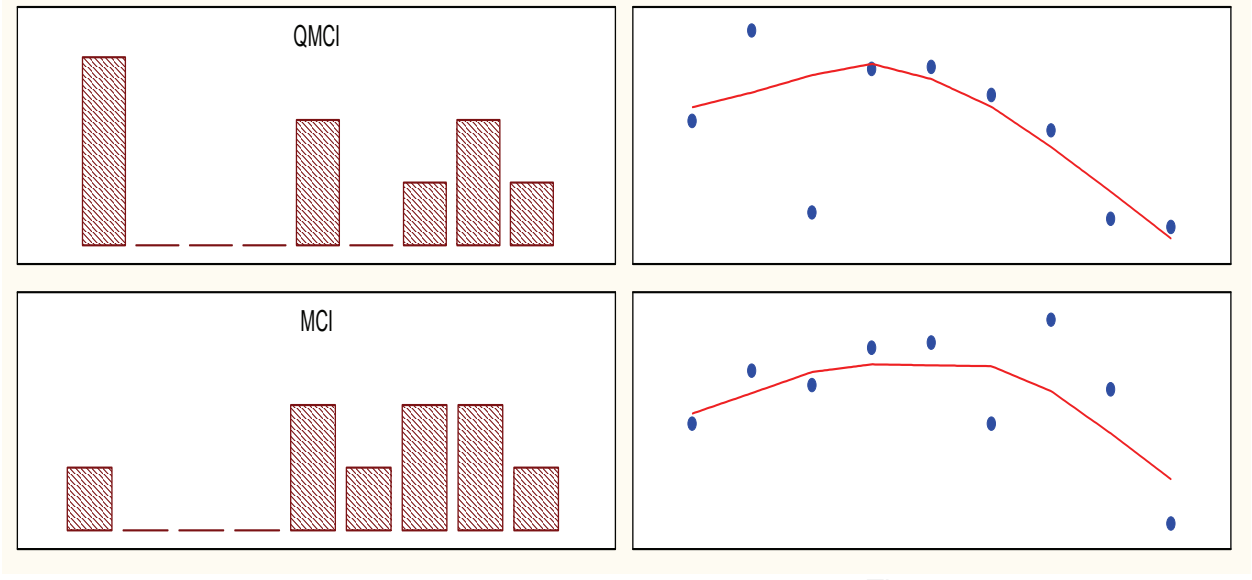
	Valid - N	Kendall - Tau	Z	p-level	p-exact - 1-tailed
QMCI & Year	9	-0.555556	-2.08514	0.037056	.022
MCI & Year	9	-0.277778	-1.04257	0.297147	.179

Rangitikei River - Estuary



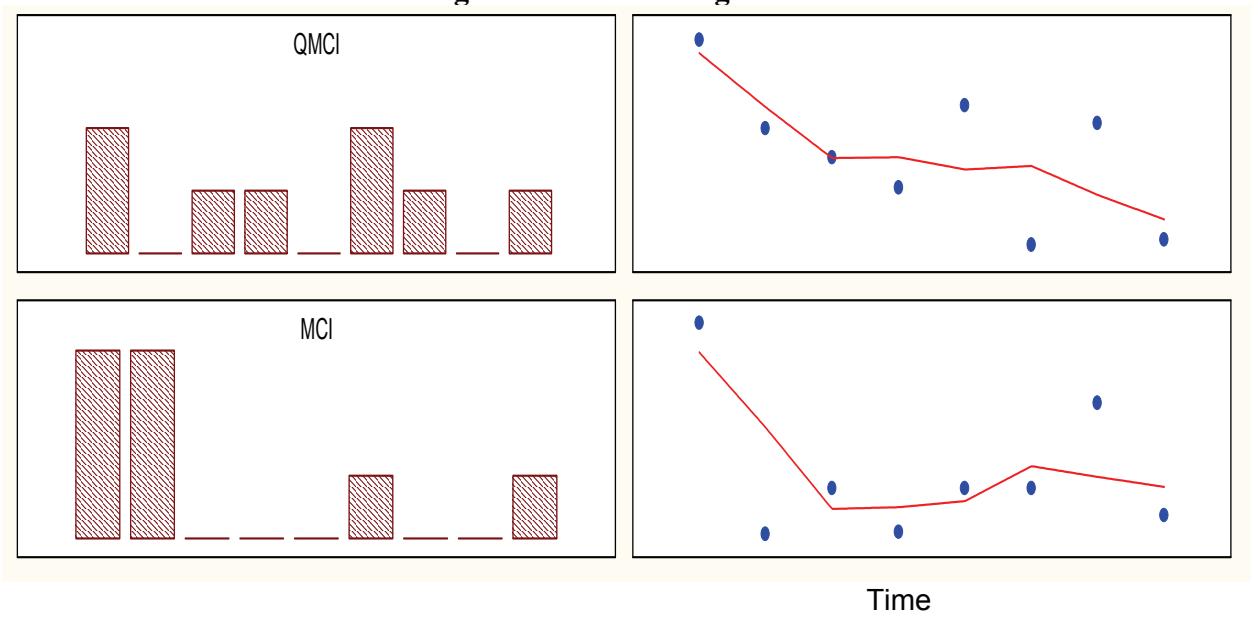
	Valid - N	Kendall - Tau	Z	p-level	p-exact - 1-tailed
QMCI & Year	6	0.200000	0.563602	0.573025	.360
MCI & Year	6	0.066667	0.187867	0.850981	.500

Rangitikei River - Kakariki



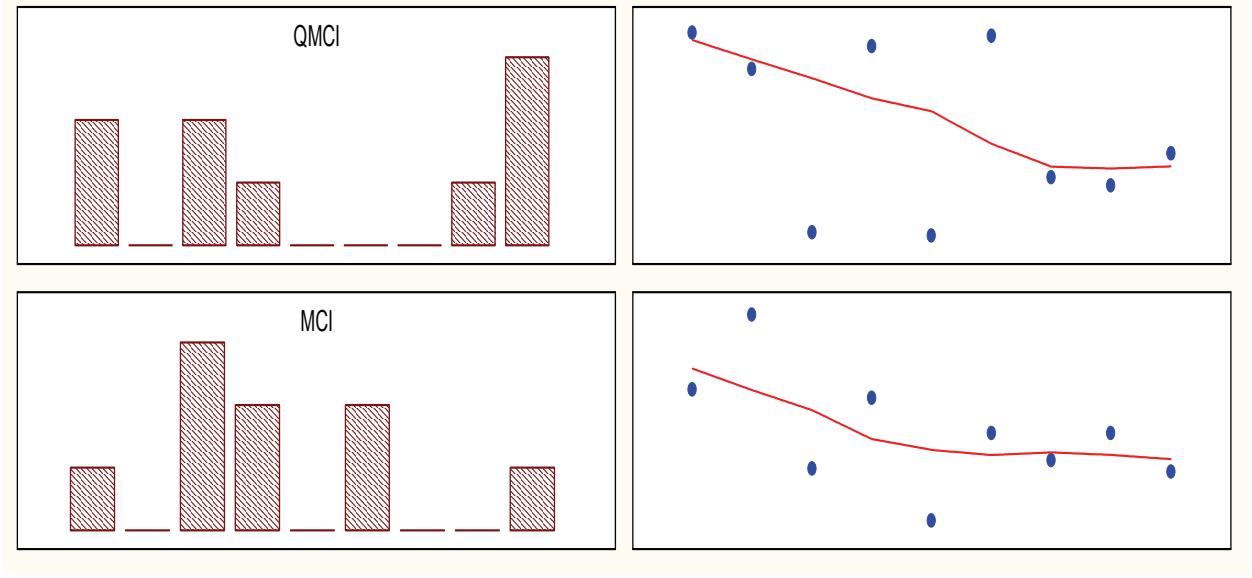
	Valid - N	Kendall - Tau	Z	p-level	p-exact - 1-tailed
QMCI & Year	9	-0.500000	-1.87663	0.060569	.038
MCI & Year	9	-0.028172	-0.10574	0.915792	.540

Rangitikei River - Mangaweka



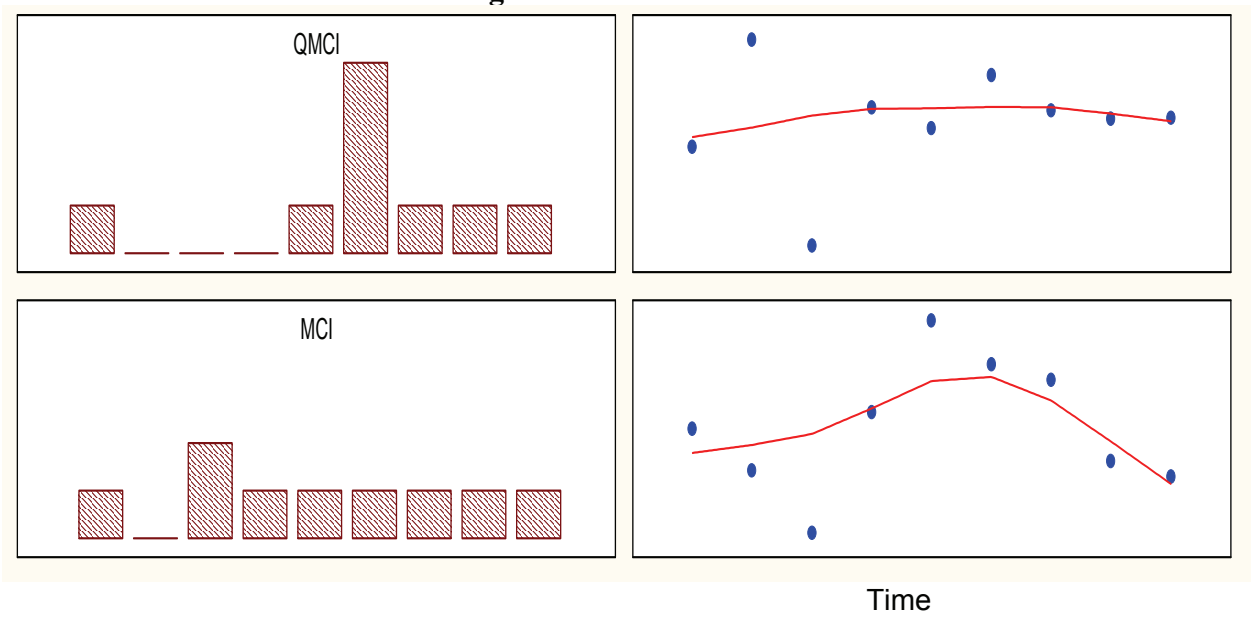
	Valid - N	Kendall - Tau	Z	p-level	p-exact - 1-tailed
QMCI & Year	8	-0.428571	-1.48461	0.137646	.089
MCI & Year	8	0.037796	0.13093	0.895830	.548

Rangitikei River - Pukeokahu



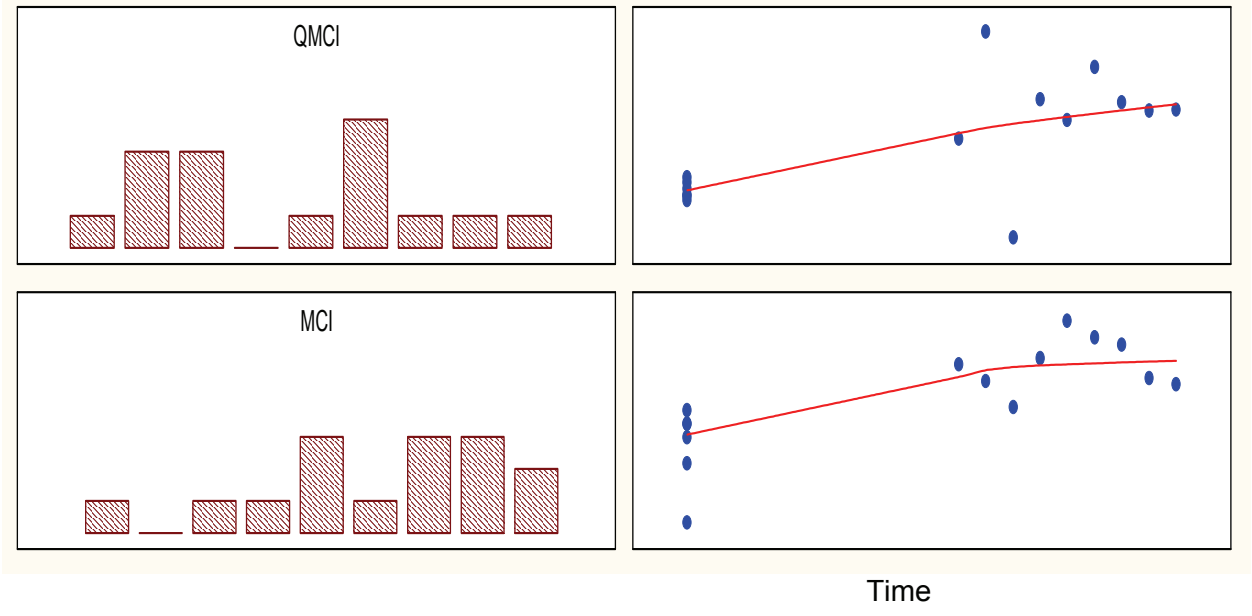
	Valid - N	Kendall - Tau	Z	p-level	p-exact - 1-tailed
QMCI & Year	9	-0.222222	-0.83406	0.404249	.238
MCI & Year	9	-0.422577	-1.58604	0.112730	.090

Whanganui River – Te Maire



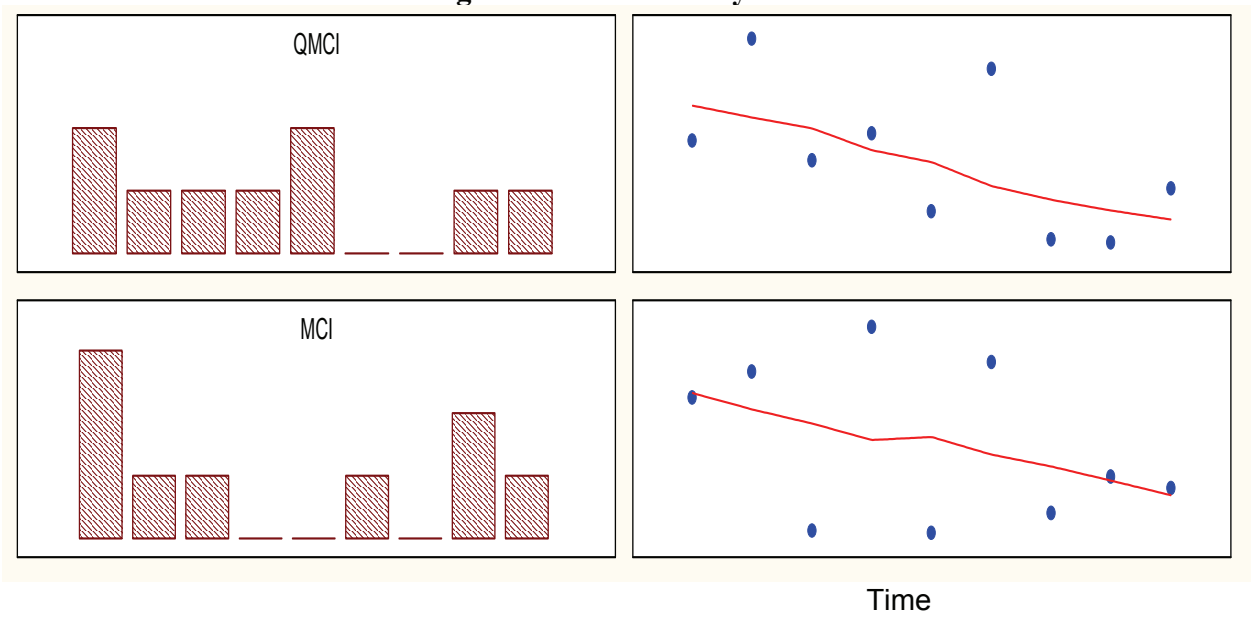
	Valid - N	Kendall - Tau	Z	p-level	p-exact - 1-tailed
QMCI & Year	9	0.055556	0.208514	0.834827	.460
MCI & Year	9	0.000000	0.000000	1.000000	.540

Whanganui River – Te Maire



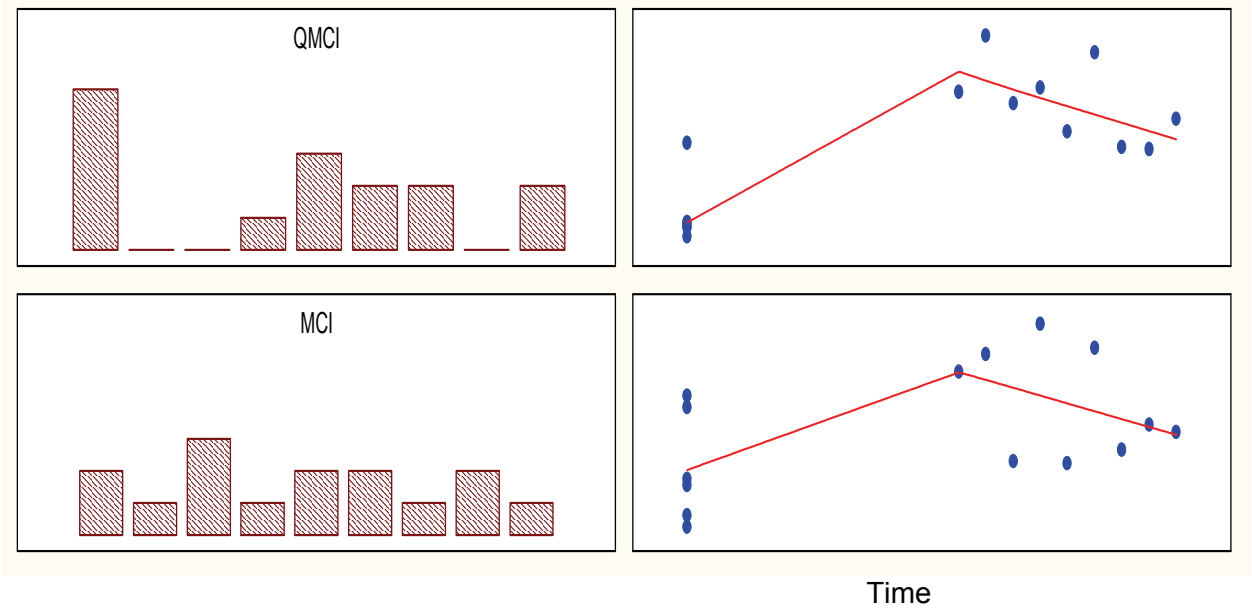
	Valid - N	Kendall - Tau	Z	p-level	p-exact - 1-tailed
QMCI & Year	15	0.452623	2.351899	0.018678	----
MCI & Year	15	0.558156	2.900265	0.003728	----

Whanganui River – Cherry Grove



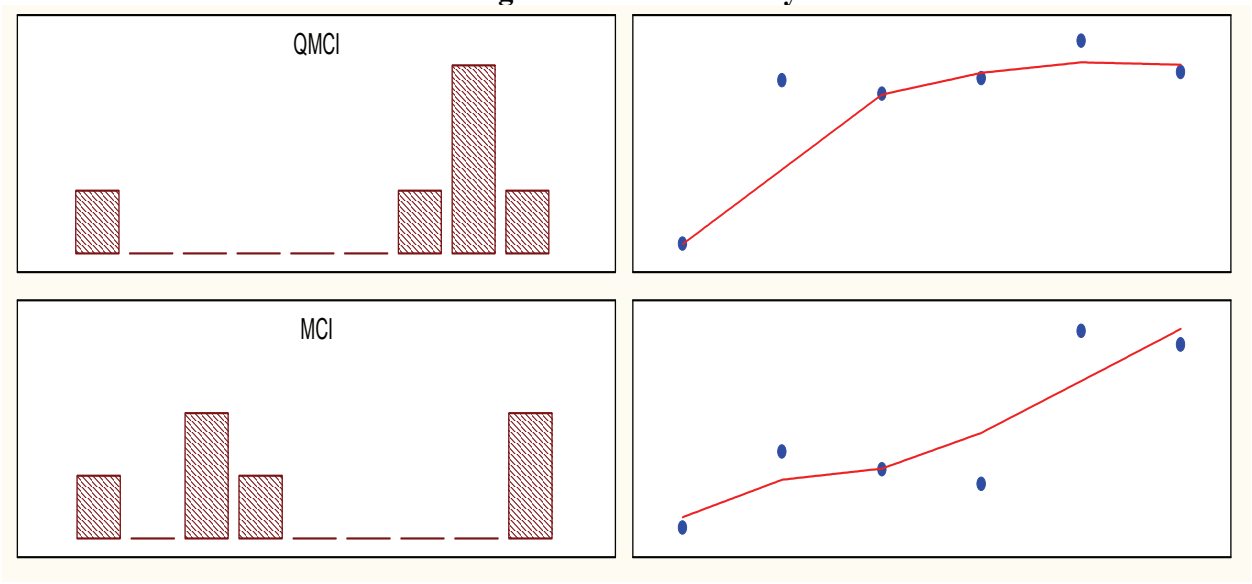
	Valid - N	Kendall - Tau	Z	p-level	p-exact - 1-tailed
QMCI & Year	9	-0.444444	-1.66812	0.095293	.060
MCI & Year	9	-0.111111	-0.41703	0.676657	.381

Whanganui River – Cherry Grove



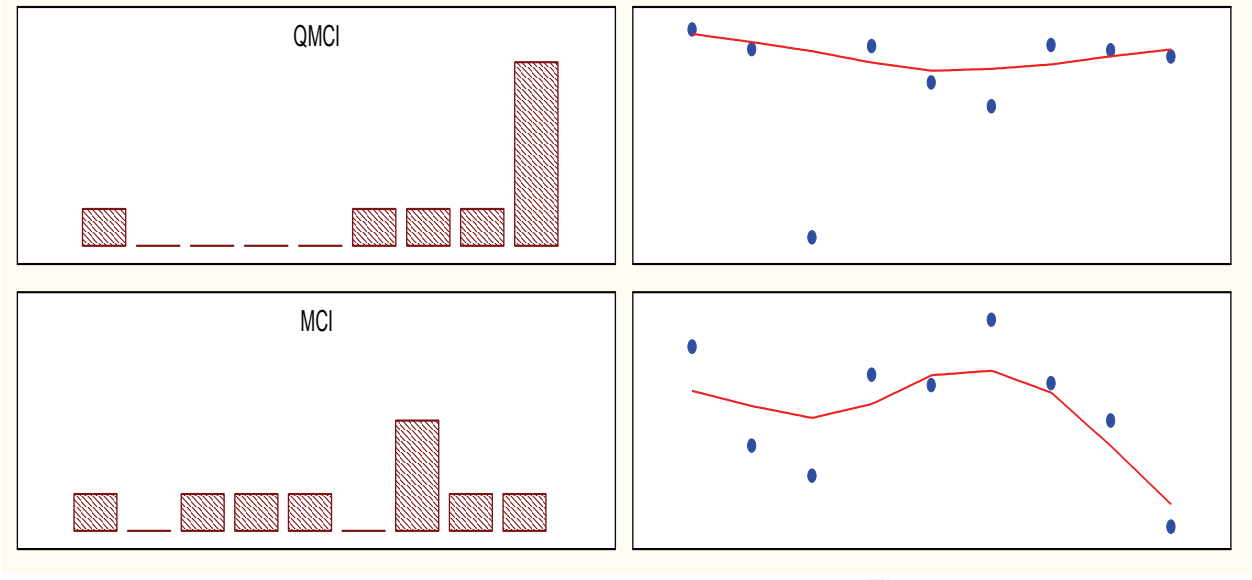
	Valid - N	Kendall - Tau	Z	p-level	p-exact - 1-tailed
QMCI & Year	15	0.349754	1.817376	0.069159	----
MCI & Year	15	0.308607	1.603567	0.108809	----

Whanganui River – Estuary



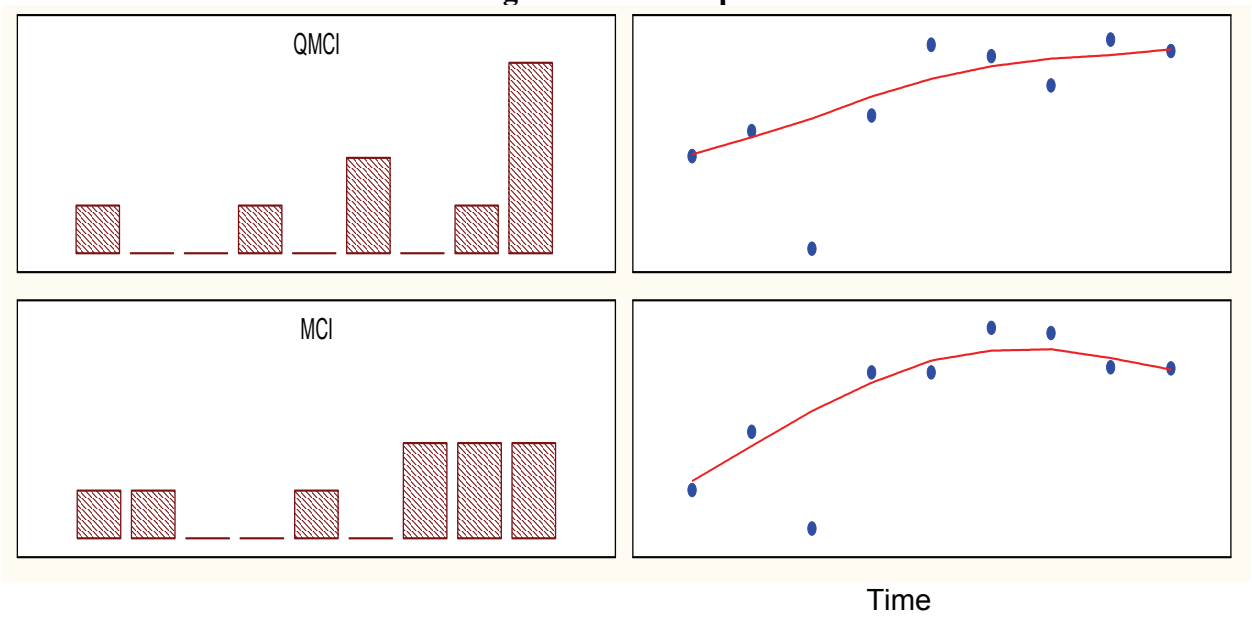
	Valid - N	Kendall - Tau	Z	p-level	p-exact - 1-tailed
QMCI & Year	6	0.733333	2.066540	0.038778	.028
MCI & Year	6	0.466667	1.315071	0.188486	.136

Whanganui River – d/s Retaruke Confluence



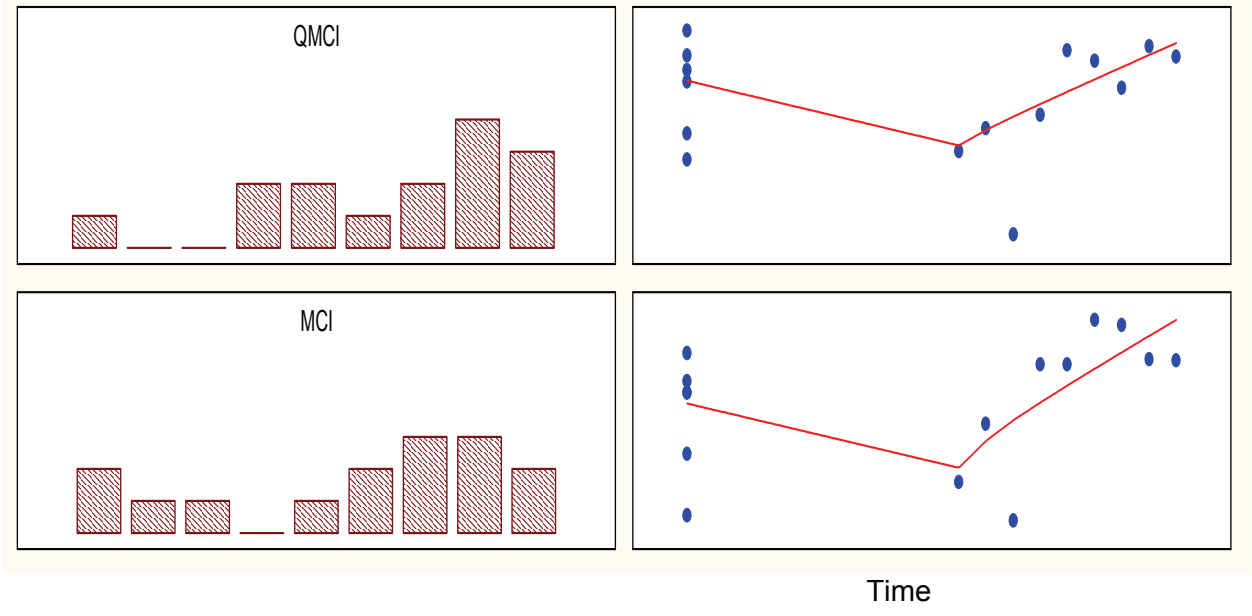
	Valid - N	Kendall - Tau	Z	p-level	p-exact - 1-tailed
QMCI & Year	9	-0.166667	-0.625543	0.531615	.306
MCI & Year	9	-0.222222	-0.834058	0.404249	.238

Whanganui River - Pipiriki



	Valid - N	Kendall - Tau	Z	p-level	p-exact - 1-tailed
QMCI & Year	9	0.611111	2.293659	0.021810	.012
MCI & Year	9	0.500000	1.876630	0.060569	.038

Whanganui River - Pipiriki



	Valid - N	Kendall - Tau	Z	p-level	p-exact - 1-tailed
QMCI & Year	15	0.185164	0.962140	0.335979	----
MCI & Year	15	0.351432	1.826093	0.067836	----

Appendix 2 Benjamini & Hochberg (1995) FDR analysis for trends in MCI values. See Table 2 for key to hypotheses. Shaded rows denote Whanganui River sites including May 1989 data. Hypotheses have been re-ordered from lowest to highest p-values.

Hypothesis	p-value	order (i)	critical value	Adjusted p-value
H5	0.004	1	0.0021	0.0783
H12	0.013	2	0.0042	0.1402
H20	0.026	3	0.0063	0.1847
H1	0.061	4	0.0083	0.2849
H2	0.068	5	0.0104	0.2849
H7	0.109	6	0.0125	0.3382
H17	0.113	7	0.0146	0.3382
H24	0.188	8	0.0167	0.4026
H14	0.211	9	0.0188	0.4026
H15	0.211	10	0.0208	0.4026
H19	0.211	11	0.0229	0.4026
H11	0.297	12	0.0250	0.5200
H18	0.322	13	0.0271	0.5206
H3	0.404	14	0.0292	0.6060
H16	0.532	15	0.0313	0.7443
H22	0.652	16	0.0333	0.7894
H6	0.677	17	0.0354	0.7894
H10	0.677	18	0.0375	0.7894
H21	0.851	19	0.0396	0.8362
H23	0.851	20	0.0417	0.8362
H13	0.896	21	0.0438	0.8362
H8	0.916	22	0.0458	0.8362
H9	0.916	23	0.0479	0.8362
H4	1.000	24	0.0500	1.0000

Appendix 3 Benjamini & Hochberg (1995) FDR analysis for trends in QMCI values. See Table 2 for key to hypotheses. Shaded rows denote Whanganui River sites including May 1989 data. Hypottheses have been re-ordered from lowest to highest p-values.

Hypothesis	p-value	order (i)	critical value	Adjusted p-value
H5	0.019	1	0.0021	0.1358
H1	0.022	2	0.0042	0.1358
H10	0.037	3	0.0063	0.1358
H11	0.037	4	0.0083	0.1358
H20	0.037	5	0.0104	0.1358
H24	0.039	6	0.0125	0.1358
H9	0.061	7	0.0146	0.1815
H7	0.069	8	0.0167	0.1815
H6	0.095	9	0.0188	0.2224
H13	0.138	10	0.0208	0.2757
H16	0.144	11	0.0229	0.2757
H19	0.211	12	0.0250	0.3691
H18	0.239	13	0.0271	0.3855
H2	0.336	14	0.0292	0.4716
H21	0.348	15	0.0313	0.4716
H14	0.404	16	0.0333	0.4716
H15	0.404	17	0.0354	0.4716
H17	0.404	18	0.0375	0.4716
H3	0.532	19	0.0396	0.5876
H23	0.573	20	0.0417	0.6017
H12	0.621	21	0.0438	0.6207
H8	0.677	22	0.0458	0.6459
H4	0.835	23	0.0479	0.7622
H22	0.881	24	0.0500	0.8806

Appendix 4 Complete list in alphabetical order of macroinvertebrate SoE monitoring sites in the Manawatu-Wanganui region (1999 – 2008).

	River	Site	Site_Type	N	Easting	Northing
1	AKITIO RIVER	Weber Road	Rolling	3	2791800	6083200
2	AKITIO RIVER	Estuary	Rolling	3	2798500	6066500
3	HAUTAPU RIVER	NIWA Station (Taihape)	Rolling	3	2748600	6168300
4	HAUTAPU RIVER	u/s Rangitikei River	Permanent	9	2753000	6157400
5	KAHUTERAWA STREAM	Above Confluence	Rolling	1	2730000	6087000
6	KIWITEA STREAM	SH54	Rolling	1	2731000	6107200
7	MAKAKAHI RIVER	Hamua	Permanent	2	2742400	6067600
8	MAKAKAHI RIVER	Konini	Discontinued	8	2746700	6074300
9	MAKINO STREAM	South Street	Rolling	2	2727500	6105400
10	MAKOHINE STREAM	Viaduct	Rolling	3	2739400	6144800
11	MAKOTUKU STREAM	u/s Raetihi	Rolling	1	2706500	6195700
12	MAKOTUKU STREAM	Railway Bridge	Reference	3	2715000	6203700
13	MAKURI RIVER	Tuscan Hills	Rolling	3	2758300	6071700
14	MANAKAU STREAM	flow site	Rolling	1	2695300	6053500
15	MANAWATU RIVER	Opiki	Permanent	5	2719500	6082700
16	MANAWATU RIVER	Upper Gorge (Woodville Domain)	Rolling	3	2749400	6093300
17	MANAWATU RIVER	Hopelands Reserve	Permanent	9	2761500	6089800
18	MANAWATU RIVER	Weber Road	Rolling	2	2774700	6102600
19	MANAWATU RIVER	Ashhurst Domain	Rolling	4	2744268	6096214
20	MANAWATU RIVER	Karere Rd / 42 Mile Hydro Station	Permanent	9	2725462	6085118
21	MANAWATU RIVER	Maxwells Line	Permanent	9	2729937	6087844
22	MANAWATU RIVER	Timber Bay	Rolling	1	2773600	6103600
23	MANAWATU RIVER	Whirokino	Permanent	6	2702200	6074700
24	MANGAHAO RIVER	Ballance	Rolling	3	2746700	6081800
25	MANGAHAO RIVER	Kakariki.	Rolling	2	2731700	6068500
26	MANGANUIOTEAO RIVER	Hoihenga Rd	Rolling	3	2704700	6207700
27	MANGAONE STREAM	Milson Line	Rolling	3	2731000	6095200
28	MANGAPAPA STREAM	Troop Rd Bridge	Rolling	3	2751900	6092100
29	MANGATAINOKA RIVER	Putara	Rolling	3	2725300	6055300
30	MANGATAINOKA RIVER	SH2 Bridge Mangatainoka	Permanent	9	2752800	6083100
31	MANGATERA STREAM	confluence @ Timber Bay	Permanent	8	2773600	6102600
32	MANGATORO STREAM	Mangahei Rd	Rolling	1	2781300	6101900
33	MANGAWHERO RIVER	Raupiu Road	Rolling	1	2709900	6164600
34	MANGAWHERO RIVER	Pakahii Road bridge	Rolling	1	2710000	6194500
35	MANGAWHERO RIVER	DOC Headquarters	Reference	9	2717900	6197700
36	MANGAWHERO RIVER	d/s Makotuku	Permanent	7	2708894	6189024
37	OHAU RIVER	Haines Farm	Rolling	3	2695800	6057900
38	OHAU RIVER	Gladstone Reserve	Rolling	3	2707600	6057700

Appendix 4 continued.

	River	Site	Site_Type	N	Easting	Northing
39	OROUA RIVER	Awahuri bridge	Permanent	9	2724400	6100300
40	OROUA RIVER	u/s MBP discharge @ Nelson St	Permanent	9	2729800	6104800
41	OROUA RIVER	Almadale	Permanent	2	2736500	6111300
42	OROUA RIVER	Apiti Gorge Bridge	Reference	2	2760000	6136700
43	OROUA RIVER	Barrow Rd	Rolling	1	2741300	6115200
44	OROUA RIVER	Main South Rd	Rolling	1	2760060	6136633
45	OWHANGA	Branscombe Bridge	Rolling	2	2789400	6058700
46	POHANGINA RIVER	Saddle Rd Bridge	Rolling	2	2745550	6098900
47	POHANGINA RIVER	Mais Reach	Rolling	1	2746800	6105300
48	POHANGINA RIVER	Piripiri	Rolling	2	2760500	6124100
49	POHANGINA RIVER	Raumai	Rolling	2	2747400	6107200
50	POHANGINA RIVER	Totara Reserve	Rolling	1	2753600	6116600
51	POREWA STREAM	Onepuhi Rd	Rolling	3	2719200	6122500
52	RANGITIKEI RIVER	Kakariki.	Permanent	9	2718400	6117500
53	RANGITIKEI RIVER	Vinegar Hill	Rolling	4	2735800	6137900
54	RANGITIKEI RIVER	Mangaweka.	Permanent	8	2750300	6151300
55	RANGITIKEI RIVER	Old Springvale Suspension Bridge	Reference	3	2771000	6186300
56	RANGITIKEI RIVER	Pukeokahu	Permanent	9	2771300	6170800
57	RANGITIKEI RIVER	Estuary	Permanent	6	2700500	6100300
58	Retaruke River	Retaruke Upper	Rolling	1	2689100	6230900
59	TAMAKI RIVER	Reserve	Rolling	3	2768300	6116200
60	TAMAKI RIVER	SH2	Rolling	3	2771200	6104000
61	TARINGAMOTU STREAM	Oruaiwi Rd	Rolling	1	2722360	6261600
62	TIRAUMEA RIVER	Katiawa Bridge	Rolling	2	2754100	6075200
63	TOKIAHURU STREAM	above confluence	Rolling	1	2721378	6186951
64	TOKOMARU	Horseshoe Bend	Rolling	2	2724100	6076800
65	TURAKINA RIVER	SH3 Bridge.	Rolling	4	2698500	6127900
66	TURAKINA RIVER	Otaire	Rolling	1	2723600	6147100
67	TUTAENUI STREAM	Curls Bridge	Rolling	3	2716100	6115900
68	WAIHI STREAM	SH52	Rolling	1	2789200	6080400
69	WAIKAWA	d/s Manukau	Rolling	3	2694100	6054800
70	WAINUI RIVER	Herbertville Rd	Rolling	1	2811340	6076500
71	WHAKAPAPA RIVER	Below TPD intake	Permanent	3	2722600	6229300
72	WHANGAEHU	Kauangaroa	Rolling	1	2704400	6139700
73	WHANGANUI RIVER	Pipiriki	Permanent	9	2685800	6189600
74	WHANGANUI RIVER	d/s Retaruke confl.	Rolling	9	2688300	6230500
75	WHANGANUI RIVER	Te Maire	Permanent	9	2699800	6249000
76	WHANGANUI RIVER	Cherry Grove	Permanent	9	2705700	6254500
77	WHANGANUI RIVER	Estuary	Permanent	6	2680500	6137800
78	WHANGANUI RIVER	Kaiwhaiki	Rolling	5	2688100	6151300

Appendix 5 Interpreting the MCI and SQMCI / QMCI.

The criteria for interpreting MCI and SQMCI (or QMCI) values have been presented previously (Table 3) including the suggestion by Wright-Stow & Winterbourn (2003) of a return to fuzzy boundaries. I think the suggestion is a good one – after all I suggested it first (Stark 1985) – but I believe Wright-Stow & Winterbourn's (2003) criteria for the QMCI are flawed and are overly optimistic.

The rationale for fuzzy boundaries between quality classes is that there is some error associated with the estimation of biotic index values. With fixed criteria, it's possible that borderline sites could alternate between adjacent quality classes from year to year ('excellent' – 'good' – 'excellent' etc.), simply because of the imprecision of MCI estimation and not because stream health has changed.

Stark (1998) determined that two MCI values calculated from single hand-net samples needed to differ by 10.83 (approximately ±5 about an average MCI value of 100) for them to be considered significantly different. For example, given this imprecision an MCI of 117 could equally have been 104 – 130. At face value an MCI of 117 is in the 'good' quality class, but there is a possibility that the true MCI was 120 or more, and the site would then be classed as 'excellent'.

However, the precision of estimates of the SQMCI (from single hand-net samples) or the QMCI from single Surber samples is not as good. Stark (1998) calculated detectable differences of 0.83 and 1.37 respectively, with would correspond to fuzzy boundaries of ±0.42 and ±0.69 calculated from a single hand-net and Surber samples respectively. If these criteria were used, very few sites would be classified into one of the quality classes (*i.e.*, 'excellent', 'good', 'fair', 'poor') – most would fall between quality classes. This provides further evidence of the unsuitability of Surber (QMCI) sampling or the SQMCI (from a single hand-net sample) for SoE monitoring.

This analysis also indicates that Wright-Stow & Winterbourn's (2003) suggestion that a fuzzy boundary of ±0.2 for the QMCI is overly optimistic, and that ±0.69 is more realistic. This is too imprecise to be useful, so Surber sample replication is essential. Incidentally, by convention QMCI values are reported to two decimal places. Wright-Stow & Winterbourn (2003) proposed fuzzy boundaries for the QMCI to only one decimal place.

Although somewhat repetitious of previous discussion the following graphical analysis may assist in emphasising the advantages of the MCI for SoE monitoring. Figure 11 is based on the Detectable Difference (DD) table provided by Stark (1998: Table 5[5]). For hand-net and Surber sampling, and the MCI and SQMCI ≡ QMCI the relationship between the number of replicates collected and the percentage of samples likely to fall within the fuzzy boundaries between quality classes is presented. For example, Stark (1998) determined that the DD for the MCI calculated from a single hand-net sample was 10.83 units. Thus, the fuzzy boundaries between

[5] Note there is an error in Table 5 of Stark (1985). The DD from the MCI estimated from 9 replicate Surber samples should be 7.22 (not 4.22).

quality classes should be 10.83 units wide. Within the range (0-200) of the MCI there are three fuzzy boundaries between the four quality classes. These are 'excellent' / 'good', 'good' / 'fair', and 'fair' / 'poor'. Therefore, the total range encompassed by the fuzzy boundaries is 3 * 10.83 = 32.49. This represents 16.25% of the range (0 - 200) of the MCI. Assuming, that every possible MCI value within the range 0 - 200 has equal probability of occurring, this means that there is a 16.25% chance of any MCI value falling between quality classes. As sampling effort (replication) increases, this percentage decreases so there is increasing certainty of assignment to one or other of the four quality classes. Clearly, it is desirable that sites can be assigned unambiguously to quality classes, rather than fall between them.

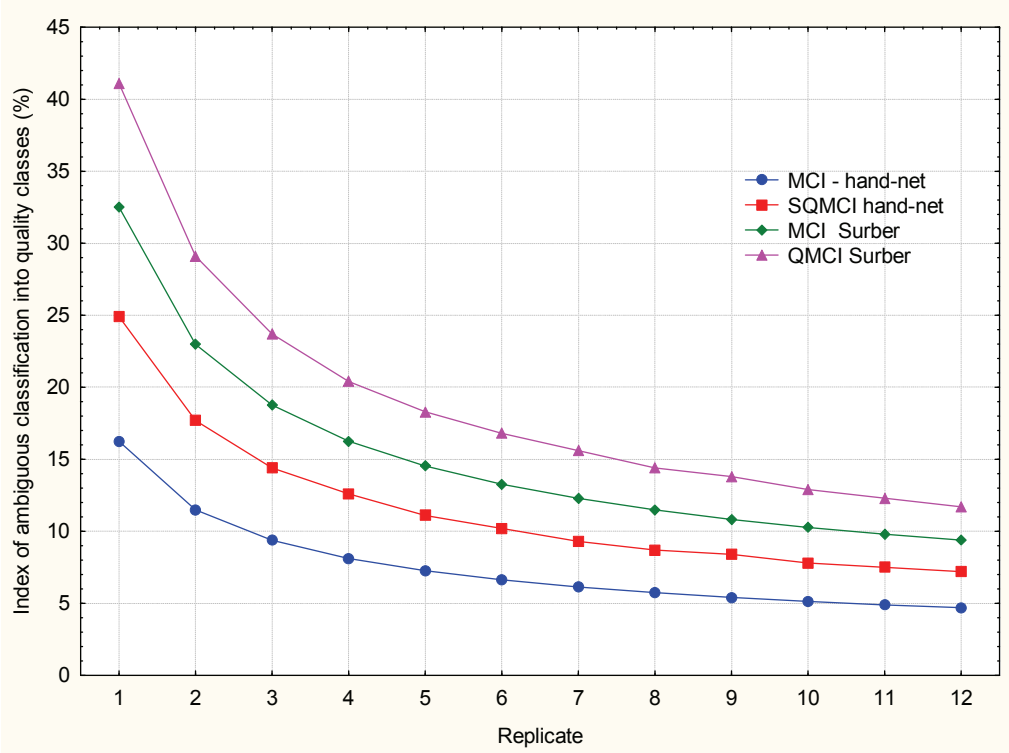


Figure 11 The relationship between sampling method, sample replication, and fuzzy boundaries on the ability to unambiguously assign sites to quality classes.

The assumption that every possible MCI (SQMCI or QMCI) value has equal probability of occurring is unlikely to be correct. In fact, the percentages on Figure 11 are likely to be considerable under-estimates (possibly by around 40%) because in practice, very few MCI values less than 40 or greater than 160 are likely. However, if we don't take too much notice of the actual values on the Y-axis, this analysis does provide a useful means of comparing the relative performance of the various sampling methods and indices.

If the MCI is used for reporting SoE results (as Stark & Maxted 2007a recommend), Figure 11 suggests that four Surber samples provide similar performance to a single hand-net sample. If the SQMCI is to be used, then triplicate hand-net samples will be required in order to perform as well or better than this. If the QMCI is the preferred index, then seven or eight Surber

samples will need to be collected to perform as well as the MCI (one hand-net sample) and SQMCI (triplicate hand-net samples) respectively.

Stark *et al.* (2001) provide the sampling protocols and suggest that the effort required for Protocol C1 (hard-bottomed, semi-quantitative) using a hand-net or D-net was equivalent to sampling 0.6 – 1.0m^2 of streambed. This is equivalent to between six and 10 standard Surber samples. In fact, in my experience, sampling effort equivalent to 0.3 – 0.6m^2 is sufficient to provide reliable estimates of the MCI. However, to remain consistent with the Protocol, and with the sampling effort used previously by Horizons (*i.e.*, five replicate Surber samples), I suggest that each hand-net sample should be equivalent to 0.6 m^2 of habitat.

This analysis provides convincing evidence of the cost-effectiveness of collecting single hand-net samples from each SoE monitoring site and using the MCI for reporting the results. Alternative indices and/or sampling methods could end up costing three to eight times as much (or more!).