Analysis of Habitat Factors Influencing Invertebrate Communities in Streams of the Bay of Plenty Region

October 2019

Prepared By:

Ton Snelder[1]

Katie Dey[1]

Alastair Suren[2]

[1] LWP Ltd

[2] Bay of Plenty Regional Council, Whakatane

For any information regarding this report please contact:

Ton Snelder

Phone: 03 377 3755

Email: ton@lwp.nz

LWP Ltd
PO Box 70
Lyttelton 8092
New Zealand

LWP Client Report Number: LWP Client Report 2018-14

Report Date: August 2019

LWP Project: 2018-14

Quality Assurance Statement

[Click here and type text]

Version	Reviewed By	
1	Caroline Fraser	

Table of Contents

Executive Summary

1. The Bay of Plenty Regional Council (BoPRC) undertakes annual state of environment (SoE) monitoring of stream invertebrate communities. Stream ecological health at the monitoring sites is subsequently quantified using biotic metrics calculated from the monitoring data. Information about habitat conditions are also routinely collected as part of this work, encompassing both semi-quantitative assessments of habitat condition using measures such as the Rapid Habitat Assessment (RHA) Protocol, and quantitative measurements of factors describing aspects of the streambed, hydraulic conditions, and riparian conditions.

2. Although BoPRC has routinely collected this habitat data for many years, linkages between invertebrate communities and both macro-scale (climate, geology and land cover) and micro-scale (habitat) explanatory variables have not previously been studied. We analysed 5 years of SoE data to determine what macroscale and microscale explanatory variables were associated with the between-site variation in invertebrate communities.

3. Our study was based on fitting linear models to both the biotic metric (univariate) response data and to the multivariate community matrices. We used forward and backward stepwise model building to fit parsimonious models and interpreted the model parameters to determine the direction of the relationships between the response and explanatory variables and their relative importance.

4. We showed strong relationships (i.e., ~60% of the variation explained) between biotic metrics (e.g., MCI, QMCI, EPT_richness and % EPT) and both macroscale and microscale explanatory variables.

5. Important micro-scale variables included measurements of fine sediment deposition (negative relationships with biotic metrics indicating ecological health) and bank stability, Rapid Habitat Assessment scores, and the amount of overhanging vegetation (all positive relationships with biotic metrics).

6. These findings have important implications for BoPRC's ongoing riparian protection work, highlighting the importance of riparian shade and sediment deposition on the ecological health of streams. It may be possible to develop criteria for some micro-scale habitat variables to support ecological health objectives. This would result in more focussed habitat management and surveys in the future.

1 Introduction

Under the Resource Management Act (RMA) 1991, regional councils must promote the sustainable management of their region's natural and physical resources. In particular, section 35 (2)(a) of the RMA requires regional councils to monitor the state of the environment and to report on state and trends of freshwater environments throughout their regions and report on the effectiveness of regional plans. State of Environment (SoE) monitoring and reporting is therefore designed, amongst other things, to detect changes in environmental conditions, to provide early warning of environmental problems, and to illustrate where environmental management has been effective. It provides councils and communities with information on the state or condition of the environment and key environmental pressures, and to assess possible and actual ecological responses to these pressures. Ideally, SoE monitoring informs decision-making by helping determine the need for further action, and by indicating where policies and actions can be improved.

Within the Bay of Plenty Regional Council (BoPRC), the Natural Environment Regional Monitoring (NERM) programme has been established to meet the requirements of s35 (2)(a) of the RMA. A key part of the programme is annual SoE river invertebrate monitoring, which has been operating since 1991. This monitoring programme provides a representative summary of stream health throughout the region and can be used to assess the direction of trends in stream health. Stream invertebrates are used worldwide to assess the ecological condition of waterways (Resh and Rosenberg, 1993; Stark, 1985; Stark and Maxted, 2007) and all regional councils in New Zealand monitor river invertebrate communities as part of statutory responsibilities for SoE monitoring.

Recent analysis of the current SoE invertebrate monitoring data by Suren *et al.* (2017) showed that characteristics of invertebrate communities in Bay of Plenty rivers are associated with catchment land cover and other factors such as location within a catchment, stream size (and flow), substrate size and habitat diversity. Catchment land cover in particular had a large influence on invertebrate communities, with sites draining catchments dominated by native bush or exotic plantation forest having the highest ecological condition, and sites representing streams draining agricultural catchments had poorer condition. Sites in urban streams generally had the poorest ecological condition. This pattern is consistent with results of studies worldwide (e.g., Hall *et al.*, 2001; Harding and Winterbourn, 1995; Lenat and Crawford, 1994; Paul and Meyer, 2001; Quinn *et al.*, 1997; Suren and Elliot, 2004).

In response to degradation of freshwater ecosystems in New Zealand, the government has promulgated the National Policy Statement for Freshwater Management (NPS-FM: Ministry for the Environment, 2017). The NPS-FM recognises two compulsory values of freshwater ecosystems: ecosystem health and human health. Objective A1 is to safeguard the life-supporting capacity of fresh water, including ecosystem processes indigenous species. The NPS-FM directs councils to undertake integrated and sustainable management of fresh water by setting freshwater objectives and defining the associated water quality and quantity limits and other actions to achieve these.

The ability of councils to set relevant objectives and define the actions to achieve these depends on knowledge of the attributes that support “life-supporting capacity”, ecosystem processes and indigenous species. The National Objective Framework (NOF) of the NPS-FM defines several compulsory attributes that regional councils must use to set numeric freshwater objectives. The attributes that are relevant to ecosystem health in rivers are:

algae (periphyton); nitrate and ammonia (for toxicity); dissolved oxygen below point source discharges). NOF attributes define four discrete state bands (A, B, C, and D) to assist communities to set objectives and also define the C/D band as a minimum acceptable state or 'national bottom line'.

The compulsory NOF attributes are primarily measures of water quality, with the exception of periphyton biomass. Periphyton was included as an attribute because it has effects on a number of values, including ecological, recreational, aesthetic and cultural (Biggs, 1985, 2000), and is responsive to resource uses that cause nutrient enrichment (e.g., point and diffuse discharges of nutrients), or alterations to flow regime (e.g., by water abstractions). Another ecological metric, the Macroinvertebrate Community Index (MCI) is also included in Policy CB3 of the NPS-FM such that every Council shall monitor the MCI in rivers and streams and investigate the cause of declining trends or when the MCI score is less than 80.

The NOF emphasises water quality parameters (and periphyton) as attributes for setting numeric objectives. However, water quality is not the only requirement for a healthy ecosystem. Policy CA2 of the NPS-FM recognises this and requires regional councils to identify attributes in addition to those prescribed by the NOF and to use these to define freshwater objectives that are relevant to the compulsory values (i.e., ecosystem health). Physical habitat is an important driver of the invertebrate communities in streams and rivers and is impacted by human activities. However, a significant challenge to formulating additional attributes is identifying the most effective measures of physical habitat and their desirable state so that the outcome of management on invertebrate communities can be predicted.

Elucidating relationships between measures of physical habitat and ecological health is challenging because invertebrate communities are structured by a wide range of factors. In addition, these factors can represent different scales including small-scale factors such as hydraulics, substrate, water chemistry and riparian vegetation (e.g., Biggs *et al.*, 2001; Minshall, 1984; Parkyn, 2004; Scarsbrook *et al.*, 2000; Statzner and Higler, 1986), and landscape level factors such as land use, geology, and climate (Death, 1995; Greenwood and Booker, 2015; Richards *et al.*, 1996). Analysis is complicated since factors that are measured at smaller scales are generally influenced by larger scale factors and disentangling these relationships is complex. Different studies have therefore shown different results, illustrating the difficulty of identifying the most effective measures of physical habitat and their desirable state.

Conceptual frameworks have been developed in freshwater ecology where ecological communities are described as a product of a series of filters operating at different scales, for example, regional, basin, reach, channel unit, and microhabitat. The filters are considered to be hierarchical such that microscale patterns are constrained by meso-scale patterns, which in turn are constrained by macro scale patterns (Frissell *et al.*, 1986). Several studies have used this conceptualisation to analyse relationships between river community assemblages and environmental factors. For example, the importance of microscale factors in explaining between site variation of invertebrate communities has been emphasised by studies in Finland (Mykrä *et al.*, 2007), Sweden (Sandin and Johnson, 2004) and Brazil (Ferreira *et al.*, 2014). Other studies have found that environmental factors characterising a range of scales contribute equally significantly to explaining variation in invertebrate communities (e.g., Marzin *et al.*, 2013). Leps *et al.* (2015) found clear differences in linkages between invertebrate community metrics and different scales of land use and found that riparian land

use was less important in large rivers than small rivers in determining community composition and structure.

As part of the NERM monitoring protocol, a suite of small-scale physical habitat measurements (representing “micro-scale” factors) have been made at monitoring sites in the Bay of Plenty region once a year in summer (mid November – mid February) for the past five years (2012-13 to 2016-17 inclusive). A mixture of categorical variables have also been collected using protocols developed by Clapcott (2015), as well quantitative variables such as bank undercutting, substrate size, flow type and shade. Benthic invertebrates were also sampled allowing community composition to be described in terms of either the taxonomy or biotic metrics. Apart from Townsend et al (2003) in Otago, we were not aware of any other study in New Zealand that investigated relationships between habitat quality and invertebrate communities, and in particular, are not aware of any attempt to understand the relative importance of microscale factors and larger meso-scale factors (i.e., the characteristics of the stream reach on which the site is located) or macro-scale variables (e.g., land cover, geology, climate).

This study combined the NERMN monitoring data with data describing larger scale factors such as climate, geology, and landcover. Analyses were performed to examine the proportion of variation in invertebrate communities that could be explained by micro, meso and macro-scale factors. The statistical methods of variance partitioning and step-wise regression models were used to examine the relationships between of the large set of explanatory variables and invertebrate community composition. These analyses also considered the stability of these relationships through time and differences in the ability to explain variation in the invertebrate community composition using the environmental factors over all time compared to in specific years.

The results of these analyses are relevant to at least two areas of BoPRC’s current work programmes. First, if micro-scale habitat variables are uniquely important in explaining site variation in invertebrate community composition (i.e., the information provided by micro-scale variables cannot be replaced by the meso or macro-scale variables), then these variables are likely to be important determinants of community composition at individual sites. This finding would indicate that micro-scale variables are attributes that are relevant to ecosystem health and therefore need to be measured and managed. Second, the relative importance and direction of the relationships between community composition and the variables provides information that can help to guide decisions concerning actions to improve ecological health.

2 Data

2.1 Invertebrate community composition data

Invertebrates were sampled at 126 sites between mid-November and mid-February each austral summer between 2012 and 2017 inclusive. Invertebrate samples were collected using one of three sampling protocols. First, five replicate quantitative samples were collected from 17 cobble-bottomed streams using a Surber sampler (0.5 mm mesh, area = 0.096 m2) as per Protocol C3 (Stark *et al.*, 2001), and pooled. These sites were mainly in large cobble-bed rivers in the eastern part of the region. Second, in the lowland mid and western parts of the region, invertebrates were sampled from all habitats using a kick-net, where invertebrates and organic matter were dislodged from the streambed material upstream of the net and collected in the downstream net. Although this method was based

on Protocol C1 of Stark et al. (2001), all habitats other than riffles were sampled, as riffles were not found in all hard-bottomed streams. Many of the streams in the region were also soft-bottomed, making the standard kick-sampling Protocol C1 problematic. In these streams we used the third sampling method, Protocol C2 of Stark *et al.* (2001). Here, woody debris, submerged logs, aquatic macrophytes and other potential invertebrate habitat was sampled in proportion to its percentage occurrence. Only a single pooled sample was collected from each site for all kick sampling of hard and soft-bottomed streams, so that approximately 1 m² of stream bed or organic material was sampled. This is equivalent to approximately 10 Surber samples.

All collected material was preserved with iso-propyl alcohol prior to processing. All samples collected using the Surber sampler were processed using methodology based on Protocol P3 (Stark et al., 2001) while samples collected using the kick net were processed by a modification of Protocol P2, where a fixed count of 200 invertebrates was used.

We used all invertebrate data collected between 2012 and 2017 inclusive. A taxa by site matrix describing the percent relative abundance of invertebrates belonging to 130 taxa in each year was developed. For all analyses that operated on the taxonomic matrix we removed taxa with less than 5% occupancy (i.e., that occurred at fewer than 5% of sites) to reduce the weight given to rare taxa. The resultant taxonomic matrix was used to calculate six biological indices including taxonomic richness (Richness), the macroinvertebrate community index (MCI), its quantitative variant (QMCI), the richness of Ephemeroptera (mayfly), Plecoptera (stonefly) and Trichoptera (caddisfly) (EPTr), the percentage EPT richness (pEPTr), and percentage EPT to total abundance (pEPTn). For the MCI and QMCI, we used the applicable soft-bottomed or hard-bottomed tolerance values, depending on the dominant substrate of the stream (e.g., pumice sand vs gravels and cobbles).

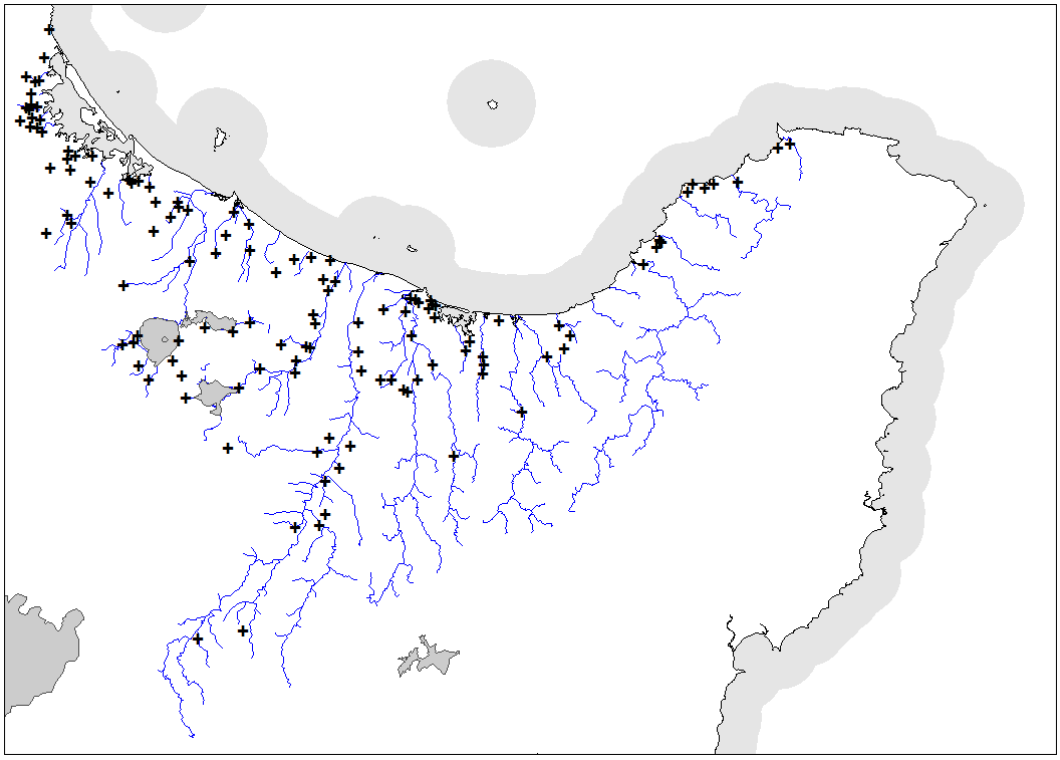

Figure 1. Map showing the location of the invertebrate monitoring sites in the Bay of Plenty region.

2.2 Explanatory variables

We grouped the environmental explanatory variables into three distinct scales, based on the hierarchical classification of Frissell *et al.* (1986): macroscale, mesoscale and microscale. These characterise not only the spatial scale, but also the timescale of persistence. Thus, macro scale variables (or "Stream System" as per Frissell *et al.*,1986) characterise spatial scales of 10^3 m, and persist over large timescales (10^5 – 10^6 years) and best correspond to overall catchment conditions. Mesoscale variables ("Segment System") characterise smaller spatial scales (10^2 m) and persist over shorter time periods (10^3 – 10^4 years) and are best represented by individual segments of the waterway network. The microscale variables characterise the spatial scales of between 10^0 m to 10^1 m, which mostly persist over relatively short time frames (10^0 – 10^2 years). These microscale variables are most representative of the local conditions characterised at the scale of the stream reach to pool/riffle system sampled in the monitoring programme.

2.2.1 Micro-scale (habitat) variables

Micro-scale habitat variables are variables that describe physical conditions in close proximity (i.e., 1 to 10 meters) from the location(s) at which the biological sample is taken and were collected at each site in most years. In total there were 38 micro-scale variables that were assessed using either categorical or quantitative measures, from either within the stream, or on the left and right banks (Table 1).

At each site, microscale habitat assessments were made using a mixture of quantitative and categorical methods. For quantitative measurements, five transects were selected at equally spaced locations along the study reach (defined as 40 times the stream width). At each transect, measurements were made of stream width, water depth and depth of the fine bed sediment at ¼, ½ and ¾ across the width of the channel. Velocity was also recorded at these locations at each transect. Velocity readings were made using the ruler technique (Harding, 2009) in the 2012-2013 and 2013-2014 summers, and by using a depth integrating velocity meter in the other three summers (2014-15, 2015-16 and 2016-17). Measurements were also made of the degree of bank undercutting, and of the distance into the stream of overhanging vegetation. Instream flow diversity was assessed by calculating the number and percentage of riffles, runs or pools along the study reach. Technical problems with a field data collection device meant that velocity measurements were not collected in the 2016-17 summer.

Substrate size was assessed using the (Wolman, 1954) technique, and the resultant percentage cover of the different substrate classes was converted to a substrate index (Jowett, 1993), which ranged from 0.1 (sand or silt dominated) to 0.8 (bedrock dominated).

Assessments were made of ten categorical habitat parameters including: stream shade; bank stability; the width, intactness and vegetation composition of bankside and riparian buffers; stock access and stock damage, and overall stream habitat diversity (Table 1). Most of these assessments were based on assigning each parameter to one of either 4 or 5 categoryseach of which were assigned a specific score (1, 5, 10, and 20) or (1, 5, 10, 15 and 20. Low scores were assigned to factors that were detrimental to stream health.

For example, "Buffer intactness" was assessed as being:

- Completely intact (score = 20);
- Occasional breaks i.e. 1-20% gaps in reach (score = 15);

- Breaks common i.e. 20 - 50% gaps in reach (score = 10);
- Breaks frequent i.e. 50-99% gaps in reach (score = 5);
- Buffer absent (score = 1)

“Stock damage” was assessed as:

- None (score = 20);
- Low (score = 10);
- Modest (score = 5);
- High (score = 1).

Where relevant (e.g., for bank stability and bankside vegetation), assessments were made of these parameters on both left and right banks.). All parameter scores were summed to create an overall stream habitat score (HABSCORE), with a theoretical range of 19 to 380. To reduce the number of explanatory variables used in the analysis, only the HABSCORE variable was used instead of the individual categorical variables.

Both quantitative and qualitative measures were possible in shallow, streams (<0.5 m deep) where it was possible to safely wade across. In larger rivers, where deep water (> 0.5m deep) and fast flows made it impractical to measure any of the quantitative factors, only categorical measurements were made.

Habitat quality was also assessed by a Rapid Habitat Assessment protocol (Clapcott, 2015). Two RHA protocols were used: an initial developmental phase protocol that assessed 9 instream habitat variables and a final phase protocol that assessed 10 habitat parameters (Table 2). The main differences in these protocols were

1. The “invertebrate habitat” and “fish cover” variables used in the initial RHA had been divided into invertebrate and fish habitat diversity and abundance

2. Bank Vegetation was scored separately for left and right banks, and averaged in the initial survey, whereas in the final survey the score was for both banks.

3. The channel alteration variables in the initial RHA had been dropped in the final RHA

4. The original RHA protocol was based on a 1 – 20 score, while the final protocol was based on a 1 – 10 score.

Table 1. List of micro-scale habitat variables measured at each of the 126 streams. Quantitative factors were measured at five transects placed across the stream, or were an assessment of the whole stream, while categorical factors were measured along the whole length of the stream, or its riparian area along the stream's left or right banks. All categorical variables were summed to derive HABSCORE.

Variable type	Measured variables	Measured where
Quantitative	Stream width (mean and standard deviation)	5 transects
	Stream depth (mean and standard deviation)	5 transects (at 3 locations across each transect)
	Degree of bank undercutting (mean[1])	Left and right banks at 5 transects
	Overhanging vegetation (mean[1])	Left and right banks at 5 transects
	Fine sediment depth (mean and standard deviation)	5 transects (at 3 locations across each transect)
	Velocity (mean and standard deviation, CV)	5 transects (at 3 locations across each transect)
	Flow diversity	Whole stream
	% backwaters, rapids, riffles, runs, pools	Whole stream
	Substrate index and diversity	Whole stream
Categorical	Stream shading	Left and right banks
	Width of bankside buffer vegetation	Left and right banks
	Buffer intactness	Left and right banks
	Composition of buffer vegetation	Left and right banks
	Composition of adjacent vegetation	Left and right banks
	Bank stability	Left and right banks
	Stock access	Left and right banks
	Composition of buffer groundcover	Left and right banks
	Composition of adjacent groundcover	Left and right banks
	Instream diversity	Whole stream

[1]The standard deviation for these variables was not calculated as invertebrate communities were thought to respond only to the overall quantity of these variables, not their variability

Table 2 List of categorical habitat variables. These variables were measured in the initial and final RHA protocols as developed by Clapcott (2013, 2015) showing which variables were collected in each period.

Initial RHA	Final RHA	2013-14	2014-15	2015-16	2016-17
Sediment deposition	Sediment deposition	Y	Y	Y	Y
Invertebrate habitat		Y	Y		
	Invertebrate habitat diversity			Y	Y
	Invertebrate habitat abundance			Y	Y
Fish cover		Y	Y		
	Fish cover diversity			Y	Y
	Fish cover abundance			Y	Y
Hydraulic heterogeneity	Hydaulic heterogeneity	Y	Y	Y	Y
Bank Stability_LB	Bank erosion_LB	Y	Y	Y	Y
Bank Stability_RB	Bank erosion_Rb	Y	Y	Y	Y
Bank_Vegetation_LB	Bank vegetation	Y	Y	Y	Y
Bank_Vegetation_RB					
Riparian buffer (width)_LB	Riparian width_LB	Y	Y	Y	Y
Riparian buffer (width)_RB	Riparian width_RB	Y	Y	Y	Y
Riparian shade	Riparian shade	Y	Y	Y	Y
Channel alteration		Y	Y		

Although the invertebrate monitoring programme collected samples from 126 sites throughout the region, habitat was not consistently sampled at all sites. The most consistently measured data included assessments of HABSCORE as well as quantitative physical measurements. RHA assessments were not made in 2012-13, the preliminary RHA1 assessments were made during the 2013-14 and 2014-15 summers, and more recent RHA2 assessments made in 2015-2016 and 2016-17 (Table 2).

The average and standard deviation of all quantitative micro-scale variables were calculated, as well as the CV of velocity, with the exception of measurements of the substrate index, substrate diversity, and shuffle index, and of the % of backwaters, rapids, riffles, runs and pools, where just the measured values were used. These variables, plus assessments of HABSCORE and the RHA1 and RHA2 resulted in a total of 24 microscale variables used in the analysis (Table 3). The predicted response of invertebrate metrics to these variables was also determined a priori, based on expert opinion.

Table 3. Micro-scale explanatory variables included in this study.The variable types are either quantitative (Q) or categorical (C). The hypothesised response describes the expected relationship with ecological health. Note that HABSCORE and both RHA measurements were made up of individual Categorical variables assessed in the field.

Explanatory variable	Description	Type	Predicted response
Av_Bank_Under	Average bank undercutting	Q	negative
Av_Depth	Average_Depth	Q	negative
Av_Sed_Depth	Average _Sediment_Depth	Q	negative
Av_Veg_over	Average _Vegetation_overhang	Q	positive
Av_Vel	Average _Velocity	Q	positive
Av_Width	Average width measured at 4 cross-sections	Q	neutral
CV_Vel	Coefficient_Vairation_Velocity	Q	positive
Flow_Hetero	Measured flow heterogeneity (the number of hydraulic types)	Q	positive
HAB_SCORE	Sum of individual components	C	positive
LENGTH	Segment_Length	Q	
RHA1_Score	RHA score Version 1	C	positive
RHA2_Score	RHA score Version 2	C	positive
Shuffle_Index	Deposited sediment_Shuffle Index	Q	negative
Std_Depth	Standard_deviation of Depth	Q	positive
Std_Sed_Depth	Standard_Deviation_Sediment_Depth	Q	neutral
Std_Width	Standard_Deviation of average_width	Q	neutral
Stdev_Vel	Standard_Deviation_Velocity	Q	positive
Sub_Divers	Substrate_diversity (the number of substrate classes)	Q	positive
Sub_Index	Substrate_Index (from Wolman sampling)	Q	positive
perc_Back	% backwaters	Q	neutral
perc_Pool	% pools	Q	neutral
perc_Rapid	% rapids	Q	positive
perc_Riffle	% riffles	Q	positive
perc_Run	% run	Q	neutral

2.2.2 Meso-scale (segment) variables

Several meso-scale (segment-scale) and macro-scale (catchment-scale) explanatory variables were obtained from the Freshwaters of New Zealand database (FWENZ; Snelder *et al.*, 2006). FWENZ is based on a digital representation of New Zealand's river and stream network, comprising segments and catchments (hereafter called the digital stream network) (Snelder and Biggs, 2002). The digital stream network represents New Zealand's rivers as 560,000 segments (delineated by upstream and downstream confluences) and their catchments. The digital river network was combined with spatial data layers describing the climate, topography, geology, and land cover of New Zealand to describe many catchment and segment-scale characteristics for each network segment (Booker and Snelder, 2012; Leathwick *et al.*, 2011). FWENZ is contained in a geographic information system (GIS) and data pertaining to the monitoring sites were obtained by linking the site location to a specific segment of the digital stream network based on geographic coordinates. Individual segments in the digital stream network are identified by a unique identifier known as the NZReach.

Meso-scale variables are geomorphic characteristics of the stream segment that the site is located on. Because segments have a typical length of 700 metres (Snelder and Biggs, 2002), meso-scale variables describe conditions at a scale significantly larger than micro-scale variables. Mesoscale variables describe the site's immediate physical context and include distance to sea, average elevation, and average air temperature (Table 4). Most meso-scale variables were derived from FWENZ, however, two categorical variables (GRND_ADJ and VEG_STR_ADJ) were measured in the field (Table 4).

Table 4. Meso-scale explanatory variables included in this study. The variable types are either quantitative (Q) or categorical (C). The hypothesised response describes the expected relationship with ecological health.

Variable	Description	Type	Hypothesised response
GRND_ADJ	Field measurement of structure of riparian ground cover of adjacent land (4 categories)	C	positive
VEG_STR_ADJ	Field measurement of structure of riparian vegetation of adjacent land (4 categories)	C	positive
Avg_Seg_Elev	Average segment elevation from FWENZ	Q	neutral
Order2	Strahler stream order from FWENZ	C	neutral
ReachHab	Segment modelled habitat from FWENZ	Q	positive
ReachSed	Segment sediment grain size estimate from FWENZ	Q	positive
SegJanAirT	Segment January average air temperature from FWENZ	Q	negative
SegSlope	Segment slope from FWENZ	Q	positive
Shade_Cat	Segment shade category from FWENZ	C	positive

2.2.3 Macro-scale (catchment) variables

Macro-scale variables are characteristics of the catchment upstream of a site (Table 5). Catchment characteristics represent the average value of a variable such as slope, climate or geological conditions over the whole upstream catchment and therefore describe the site's broad-scale physical context. All macro-scale variables were obtained from FWENZ except CONDUCTIVITY, which was measured in the field on the day of sampling (Table 5).

Table 5. Macro-scale explanatory variables included in this study. The variable types are either quantitative (Q) or categorical (C). The hypothesised response describes the expected relationship with ecological health.

Explanatory variable	Description	Type	Hypothesised response
area_sqkm	Upstream catchment area	Q	neutral
AV_US_Slope	Average upstream catchment slope	Q	neutral
Bare_Ground	Bare_Ground from LCDB4	C	negative
CONDUCTIVITY	Field measurement of conductivity	Q	negative
Exotic_Bush	Exotic_Bush from LCDB4	Q	positive
DISTSEA	Distance to Sea	Q	neutral
Exotic_Scrub	Exotic_Scrub from LCDB4	Q	positive
FRE3	Frequency of flow > 3 x median	Q	negative
Hort	Hort from LCDB4	Q	negative
Native_Bush	Native_Bush from LCDB4	Q	positive
Pasture	Pasture from LCDB4	Q	negative
SegCluesN	SegCluesN from FWENZ	Q	negative
SegFlow	SegFlow from FWENZ	Q	positive
SegLowFlow	SegLowFlow from FWENZ	Q	negative
Urban	Urban from LCDB4	Q	negative
USCalcium	USCalcium from FWENZ	Q	neutral
USDaysRain	USDaysRain from FWENZ	Q	positive
USHardness	USHardness from FWENZ	Q	neutral
USPhosporu	USPhosporu from FWENZ	Q	negative
Wetlands	Wetlands from LCDB4	Q	positive

2.2.4 Geographic coordinate data

The physical location (geographic coordinates) of each site can be used to describe spatial patterns in the biological data. If geographic coordinates explain significant amounts of variability to the ecological data that is independent of the other explanatory variables, it suggests that either unmeasured physical variables or inherent biogeographic patterns are associated with the geographic distribution of the invertebrate communities. The spatial patterns can vary at differing and any characteristic scale and the detection of these patterns is facilitated by converting the geographic coordinates to the terms of a cubic trend surface regression of the form (Legendre and Legendre, 1998);

$$Z = X + Y + XY + X^2 + Y^2 + X^2Y + XY^2 + X^3 + Y^3$$

Where Z represents the invertebrate community at each site and X and Y are the eastings and northings (i.e., geographic coordinates) of the sites obtained from their physical location data. The X and Y terms of the trend surface describe any simple linear spatial patterns in the data, while the higher order terms model more complex landscape features such as patches and gaps. All the X and Y terms were included as explanatory variables in the statistical analyses that follow.

3 Statistical analyses

3.1 Temporal variation of measured biological and habitat variables

The biological and habitat variable data was collected at most sites every year over a five-year period. All these variables were subject to variability due to both (1) temporal variability in the invertebrate community composition and small-scale habitat factors and (2) imprecision in the measurements (particularly associated with inter-operator variability). This study and the data collection were not designed to discriminate between these two sources of variation. However, we tested the impact of variation on relationships between the biotic variables and the measured habitat variables by performing statistical analyses (described below) on data pertaining to individual years and to data that represented the mean value of the individual variables over all years.

Our first step was to describe temporal variation in the biological and habitat variable data. This variation was described for all micro-scale field measured explanatory variables (other meso-scale and macro-scale explanatory variables derived from GIS did not change over time). We characterised the variation in the invertebrate community data based on the six biological indices. Temporal variation was quantified by first calculating the coefficient of variation (standard deviation divided by the mean) value of each variable at each site. The overall variability of each variable was then characterised by the median and ranges of the site coefficients of variation.

3.2 Variance partitioning

Describing the relationship between physical factors and biological characteristics of rivers is complicated because a wide range of environmental factors are involved in structuring biological communities in rivers. These include biogeography, climate (e.g., temperature; Guégan *et al.*, 1998), position along the river network (e.g., altitude, distance from the source, mean flow; Horwitz, 1978), geomorphic characteristics of the segments (e.g., slope, mean water depth and velocity) and micro-scale habitat characteristics, (e.g., size of bed substrates; Angermeier and Winston, 1999; Lamouroux *et al.*, 1999). In addition, many of the

variables that represent these factors are correlated because they share strong hierarchical relationships and they tend to vary as a monotonic function of position in the river network (e.g., Poff, 1997; Vannote *et al.*, 1980). Correlation between these environmental factors may lead to overestimation of the strength of the relationship between biotic responses and any physical variables if covariance is not taken into account (Borcard *et al.*, 1992; Fortin and Dale, 2005; Legendre and Troussellier, 1988).

In this study we examined co-variation of the invertebrate community data and our explanatory variables provided as four 'tables' that represented three scales of influence (macro, meso and micro) and space (i.e., the geographic terms). Our objective was to quantify the strengths of relationships between these four tables and invertebrate community structure and determine the extent to which there is redundancy in these explanatory factors. We used all the available explanatory variables representing each of the four tables (i.e., macro, meso and micro and space). Covariation of variables within the tables was not an issue with this analysis because we were not attempting to interpret the fitted model parameters.

Variance partitioning analyses were performed using 1) redundancy analysis (essentially linear regression) when the biological responses were biotic indices, and 2) canonical correspondence analysis (CCA) when the biological response was represented by the original taxonomic matrix. CCA simultaneously analyses the assemblage and explanatory data by combining an ordination method (correspondence analysis; CA) and multiple linear regression (ter Braak, 1986; Legendre and Legendre, 1998). We first performed these analyses on the data pertaining to each individual summer sampling period, and then on the data that represented the average of both biological and field measured micro-habitat data over all summer sampling periods.

The variance partitioning analysis used the procedure of Borcard *et al.* (1992) to partition the total explained variation in the invertebrate data (i.e., each biotic index and the community matrix) into 16 components that included the individual, shared and unique contributions of the four sets of explanatory factors; macro, meso, micro and spatial. See section 4.2 for an explanation of individual, shared and unique contributions to explained variation.

The significance of all components was tested using permutation tests that randomly permuted the invertebrate assemblage data, while holding the explanatory variables constant. The significance was determined from the number of random permutations in which the total inertia of the constrained axes (i.e. the explained variation) exceeded that of the original CCA. The significance of the unique fractions were tested by permutation of partial CCAs on the sets of variables under examination with the other set of variables included as co-variables (i.e., their effect was removed; Legendre and Legendre, 1998).

Estimates of explained variation derived from samples are generally biased (Zar, 1999). This bias is influenced by the number of independent variables in the model and sample size. We used the method of Peres-Neto *et al.* (2006) to adjust the estimate of variation explained by each component to make valid comparisons between sets of variables of differing size. All CCA analyses and variance partitioning was performed in R using the 'vegan' package (R Development Core Team, 2004).

3.3 Stepwise model building

We used stepwise model building to fit the individual explanatory variables to the biological responses. Models were built for each of the individual biotic indices and for the taxonomic

matrix. Stepwise model building was used to select and fit the most parsimonious model starting with all the explanatory variables. We evaluated the relative importance of the explanatory variables that were included in each model. In addition, we interpreted the relationship between each explanatory variable and the response for the models describing the biological indices.

Prior to model building, we applied appropriate transformation of the response variables to approximate a normal distribution. Thus, the following variables were log transformed before analysis: Distsea, area_sqkm, avg_seg_elev, bare_ground, exotic_bush, exotic_scrub, urban, hort, wetlands, segslope, segflow, seglowflow, segcluesn, av_bank_under, av_width, std_width, std_depth, av_sed_depth, std_sed_depth, length, perc_pool, perc_rapid, perc_back, perc_riffle.

We also examined the explanatory variables for collinearity. Because in this analysis we were interpreting the parameters of the fitted stepwise models, collinearity, or excessive correlation among explanatory variables, was an issue. True relationships among the response and explanatory variables will be masked if explanatory variables are collinear. Therefore, we used variance inflation factors (VIF) to identify collinearity (Kutner *et al.*, 2004). A VIF for a single explanatory variable (*j*) is obtained using the r-squared value of the regression of that variable against all other explanatory variables:

$$VIF_j = \frac{1}{1 - R_j^2}$$

A VIF value of one indicates that the explanatory variable is orthogonal to the other explanatory variables (i.e., the variable represents unique information). High values of VIF indicate multicollinearity that increases the uncertainty of the coefficient for the explanatory variable when fitted in the regression. VIF was calculated for each explanatory variable and the explanatory variable with the highest VIF value was removed. This was performed sequentially until all variables with 'high' VIF values were removed. . The definition of 'high' is arbitrary but values in the range of 5-10 are commonly used. We used VIF > 5 as our threshold for which to remove colinear variables.

3.3.1 Biological metrics

For the biological metrics, we applied standard forward and backward stepwise linear regression to the saturated models (i.e., models that had access to all explanatory variables). The Akaike information criterion (AIC; Akaike, 1973) was used to apply a penalised log-likelihood method to evaluate the trade-off between the degrees of freedom and fit of the model as explanatory variables were added or removed (Crawley, 2002). AIC is an estimator of the relative quality of a statistical model for a given set of data. Given a collection of models for the data, AIC provides a measure of the quality of each model, relative to each of the other models. Given a set of candidate models, the most parsimonious model is the one with the minimum AIC value. The procedure identified the preferred model as that with the lowest AIC value.

We interpreted fitted relationships between explanatory variables and biological responses associated with the preferred model. First, we used the sign of each explanatory variable's coefficient (proportional [positive] vs. inverse [negative]) as an objective measure of the direction of its relationship with the water quality variable. Second, we used a measure of relative importance to rank the explanatory variables that were included in the models. The importance metric decomposed the model R^2 into non-negative contributions associated with each explanatory variable that sum to the total R^2. An issue with decomposition of R^2 for

regression models is that each order of regressors yields a different decomposition of the model sum of squares. The importance metric is therefore based on the average value for each explanatory variable derived from calculating each variable's R^2 value for all possible orderings (Grömping, 2006).

3.3.2 Community matrix

For the community matrix, we used an automated stepwise model selection routine to fit a canonical correspondence analysis (CCA) model. Because constrained ordination models are not implemented in an AIC framework, permutation p-values were used to determine whether variables should be included in the model. The routine iteratively tests whether variables should be added and then dropped from the model using permutation p-values.

A measure of the importance of the explanatory variables that were included in the CCA was calculated for each variable using the correlation of that variable with each significant axis of the CCA, weighted by the variation explained by each axis. The weighted correlations were then combined into a single value by using a generalisation of the Pythagorean theorem. For example, if a CCA model has 3 significant axes for which explanatory variable E has correlations C1, C2, C3 with the axes, and the axes explain V1, V2 V3 amount of the variation in the community matrix then the variable importance of E was given by:

$$Importance_E = \sqrt{(C1 \times V1)^2 + (C2 \times V2)^2 + (C3 \times V3)^2}$$

4 Results

4.1 Temporal variation of measured variables

Box and whisker plots summarise the distribution of the coefficient of variation (CV) of the measurements made in the field (including the biological data and the habitat variables) at each site over the sampling occasions (summers) (Figure 2 and Figure 3). Note that only 23 of the 24 micro-scale variables are shown because Shuffle_Index was only sampled during one season. The CV values for the habitat variables differed considerably between sites with values ranging from zero % to 223%. Some variables exhibited consistently higher variability than others. For example measurements of the site dimensions (LENGTH, Av_Width) had site CV of less than 100% but the median CV of the site measures of perc_Rapid and perc_Back were 173% and some sites had CV values up to 200% (Figure 2). Note that the CV values for each site and variable are provided as supplementary data.

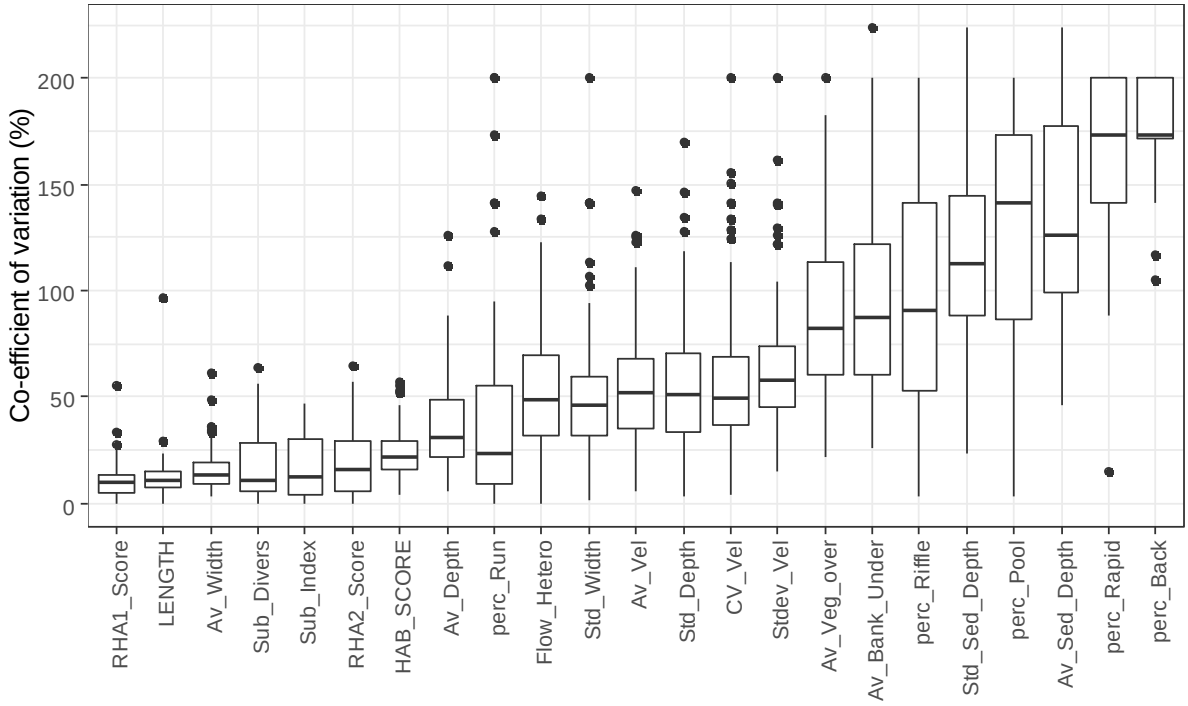

Figure 2. Temporal variability of the measured habitat variables. The plots show the distributions of the site coefficients of variation for each of the biological metrics. The box indicates the inter-quartile range and the horizontal bar within the box indicates the median. The whiskers indicate the lowest datum still within 1.5 IQR of the lower quartile, and the highest datum still within 1.5 IQR of the upper quartile. Outliers are indicated by black dots.

The coefficients of variation for the biological indices also exhibited considerable variation between sites with values ranging from close to zero % to 200%. Some indices exhibited consistently higher variability than others. For example, site MCI had site coefficients of variation of less than 25% and the median CV of all sites was 10%, whereas CV of pEPTn had a value of 200% at one site and the median CV for all sites was 49% (Figure 3).

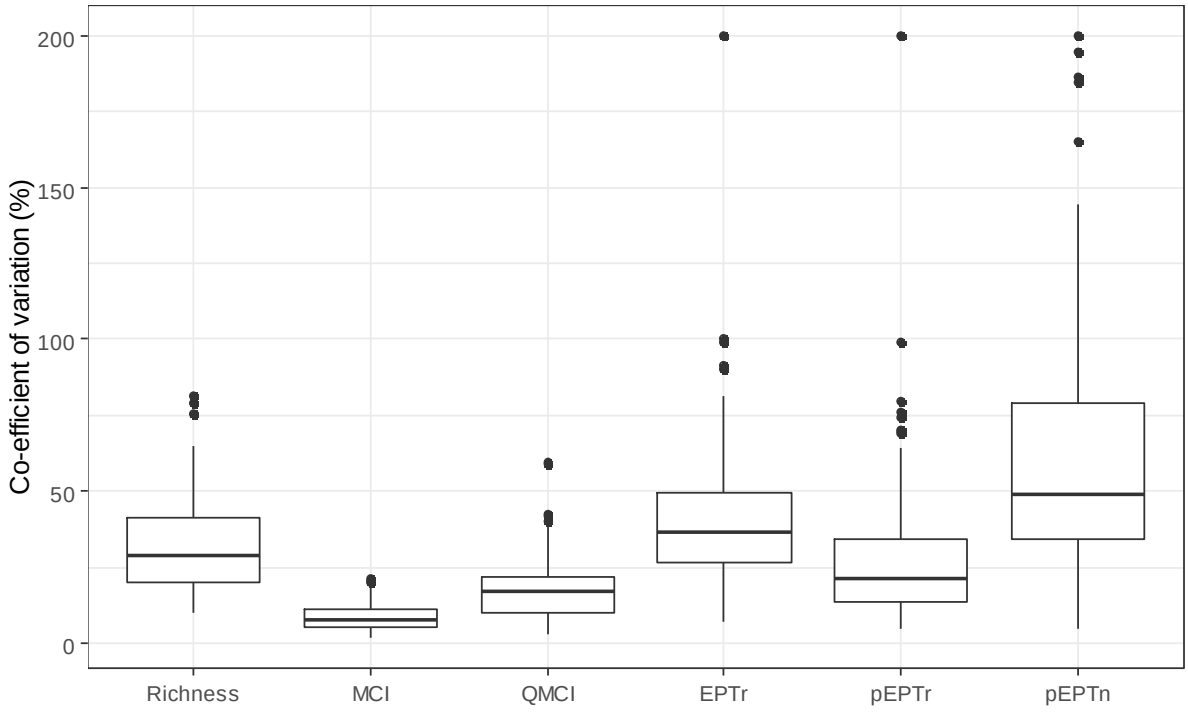

Figure 3. Temporal variability of the measured biological metrics. The plots show the distributions of the site coefficients of variation for each of the biological metrics. The box indicates the inter-quartile range and the horizontal bar within the box indicates the median. The whiskers indicate the lowest datum still within 1.5 IQR of the lower quartile, and the highest datum still within 1.5 IQR of the upper quartile. Outliers are indicated by black dots.

4.2 Variance partitioning

4.2.1 Biological metrics

For individual summer sampling periods, the total variation in all six biological metrics that was explained (adjusted R^2 values) by the four tables of explanatory variables (macro, meso, micro and space) was uniformly low (Figure 4). For example, the largest R^2 value was 35%, which occurred for EPTr in the 2016-17 summer (Figure 4). Of the 30 models that represented the biological metrics for individual summers, only nine had adjusted R^2 value greater than 20% (Figure 4). Some of these models had adjusted R^2 values that were negative. It is noted that while regression R^2 values are always positive real numbers, adjusted R^2 values can be negative. In this study, some models had negative R^2 because the explained variation was low (close to zero) and the subsequent penalty associated with the number of explanatory variables reduced the adjusted value below zero. Negative adjusted R^2 values can be interpreted as zero, but they also suggest that some of the explanatory variables that were included in the model are redundant and a simpler model would be preferable.

In contrast to the models representing the individual summer sampling periods, the total variation explained by the models representing the average of all microscale factors and biotic indices over all summer sampling periods was high (Figure 4). The lowest R^2 for the average models was 27% for Richness and all other models had R^2 in excess of 55%. The average models for MCI, QMCI and pEPTr had adjusted R^2 greater than 70%. The variance

partitioning results presented below have therefore concentrated on the biological indices that represent the average over all years (and referred to hereafter as the biological indices).

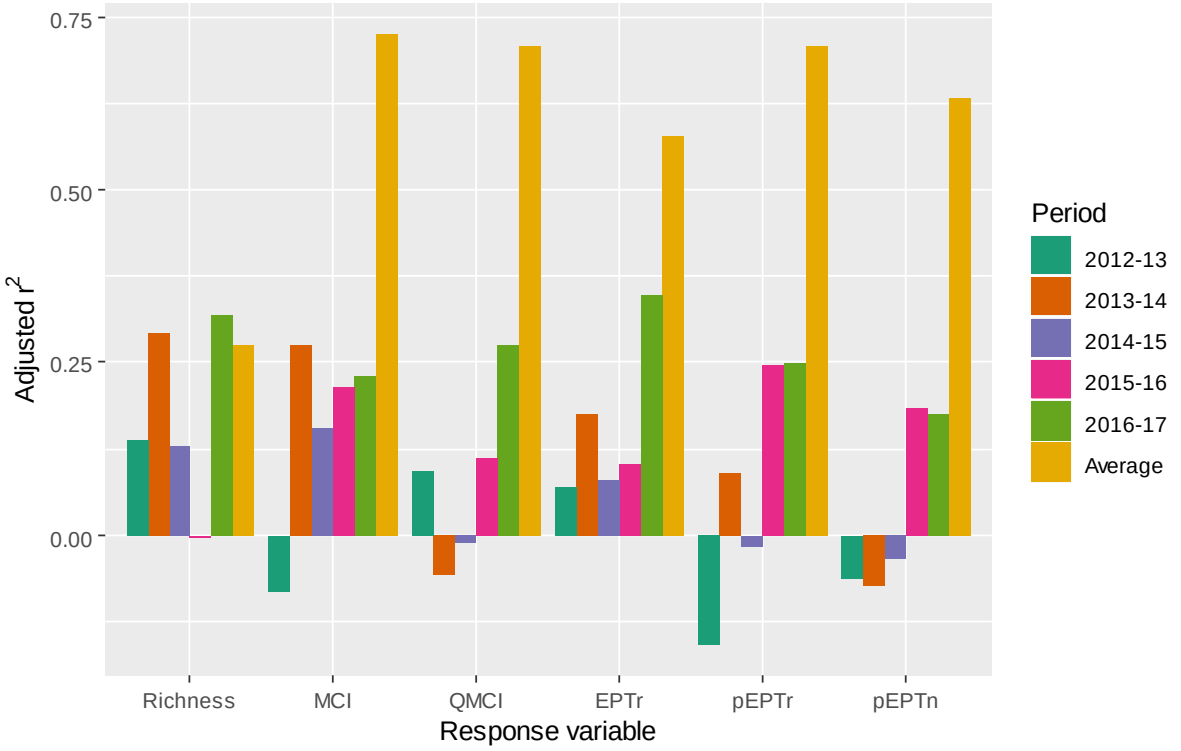

Figure 4. The total variation associated with the six biological metrics that was explained (adjusted R^2) by the micro, meso, macro and spatial variables. The plot shows adjusted R^2 for each individual summer period and on average over all summer sampling periods.

All components of the variance decomposition are indicated schematically on Figure 5 and for the analyses corresponding to the six biological metrics in Figure 6. The total variation in the response is represented by the area bounded by the outer rectangle. The total variation in the response that can be explained using all the explanatory variables (i.e. the four tables) is represented by the Venn diagram. The individual variation explained by the four tables is indicated by the four primary ellipses (e.g., the lower left ellipse represents the variation explained by the micro-scale variables). The unique variation explained by each table is represented by the portion of each ellipse that does not intersect any other ellipse (e.g., the portion of the lower left ellipse labelled [a] represents the unique variation explained by the micro-scale variables). All shared components of variation are represented by the portions of the Venn diagram that represent the intersection of the ellipses (e.g., the variation explained that is shared by the micro and meso scale variables is indicated by the portion of the Venn diagram labelled [e]).

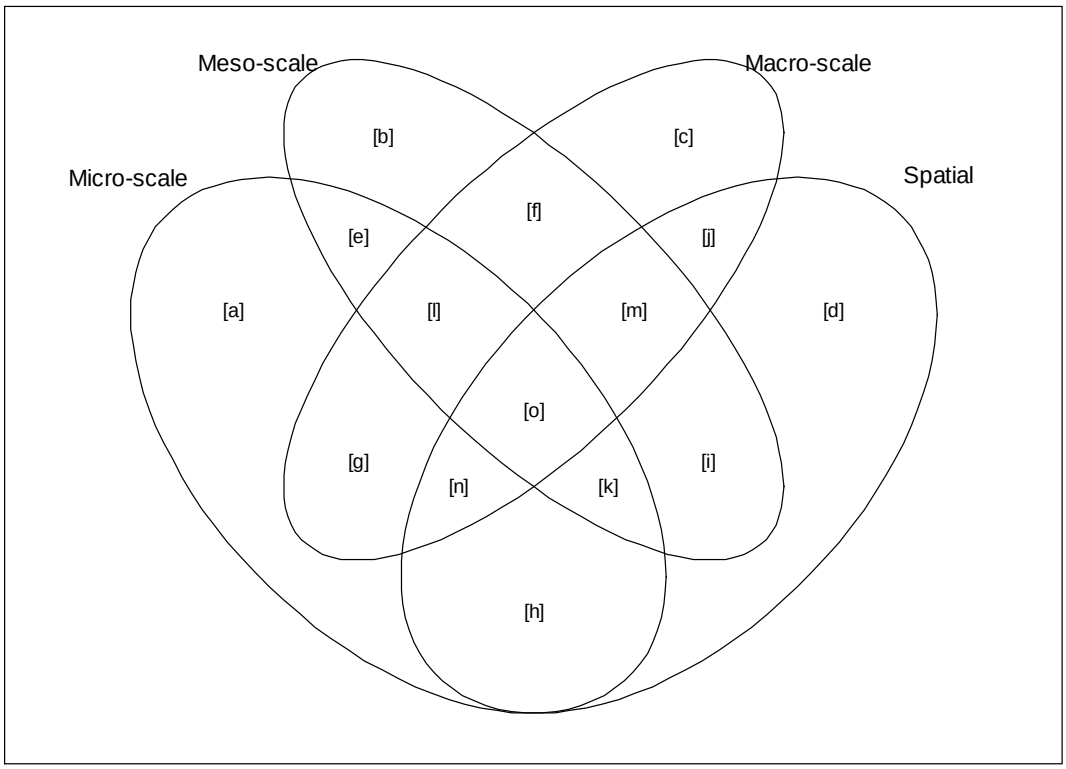

Figure 5. Schematic diagram of all components of variation provided by the variance partitioning of the four environmental data tables. Letters denote the variability explained by variables either uniquely (e.g., a, b, c, d) or in combination with variables from other data tables.

The micro-scale variables individually explained (i.e., the sum of each region contained in each ellipse on Figure 5), between 23% and 57% of the variation (adjusted R^2 values) in the six biological indices (Table 6). The meso-scale variables individually explained between 18% and 44% of the variation in the six biological indices (Table 6). The macro-scale variables individually explained between 3% and 20% of the variation in the six biological indices (Table 6). The spatial variables individually explained between 0% and 10% of the variation in the six biological indices (Table 6).

The micro-scale variables uniquely explained between 4% and 19% of the variation (i.e., the region labelled [a] on Figure 5), depending on the biological index (Table 6). The meso-scale variables uniquely explained between 0% and 6% of the variation (i.e., the region labelled [b] on Figure 5), depending on the biological index. The macro-scale variables uniquely explained between 3% and 8% of the variation (i.e., the region labelled [c] on Figure 5), depending on the biological index. The spatial variables uniquely explained between 2% and 6% of the variation in the six biological indices (Table 6).

The individual contributions of all four tables of explanatory variables were significant at the 1% level (Table 7). The unique contributions from the micro-scale and macro-scale variables were significant at between the 1% and 4% level. The unique contributions of the meso-scale variables were significant at the between the 1% and 2% level. No unique contributions of the spatial variables were significant at the 5% level.

Figure 6. All components of the variance decomposition for all six of the biological variables. The individual components are as shown on the schematic diagram on Figure 5.

Table 6. The variation explained (adjusted R^2) by the variance partitioning analysis for all six biological indices by the micro, meso, macro and spatial variables. The individual and unique components of variation explained attributable to each table of explanatory variables are the result of adding the associated individual components shown in Figure 6.

Explanatory variables	Component of variation	Richness	MCI	QMCI	EPTr	pEPTr	pEPTn
Micro	Individual	20	55	57	50	58	54
Meso	Individual	15	49	45	37	44	43
Macro	Individual	11	16	22	11	13	18
Spatial	Individual	-4	10	6	-2	6	6
Micro	Unique	1	12	12	13	19	9
Meso	Unique	-3	10	3	3	7	0
Macro	Unique	17	5	3	13	6	3
Spatial	Unique	-1	3	1	1	0	4

Table 7. The p-values associated with the variance partitioning analysis of all six biological indices. P-values are shown for the individual and unique components of variation explained that are attributable to the micro, meso, macro and spatial variables.

Explanatory variables	Component of variation	Richness	MCI	QMCI	EPTr	pEPTr	pEPTn
Micro	Individual	0.001	0.001	0.001	0.001	0.001	0.001
Meso	Individual	0.001	0.001	0.001	0.001	0.001	0.001
Macro	Individual	0.001	0.001	0.001	0.001	0.001	0.001
Spatial	Individual	0.001	0.001	0.001	0.001	0.001	0.001
Micro	Unique	0.001	0.001	0.001	0.001	0.001	0.001
Meso	Unique	0.016	0.022	0.019	0.013	0.015	0.014
Macro	Unique	0.005	0.003	0.002	0.002	0.002	0.002
Spatial	Unique	0.148	0.169	0.155	0.174	0.156	0.150

4.2.2 Community matrix

For individual summer sampling periods, the total variance in the community matrix that was explained (adjusted R^2 values) by the four tables of explanatory variables (macro, meso, micro and space) was uniformly low (Figure 7). For example, the largest R^2 value was 9%, which occurred for the 2016-17 summer (Figure 7). In contrast to the models representing the individual summer sampling periods, the total variation explained by the model representing the average of all environmental variables and biotic indices over all summer sampling periods was 43% (Figure 7). The variance partitioning results presented below has therefore concentrated on the community matrix that represents the average over all years (and referred to hereafter as the community matrix).

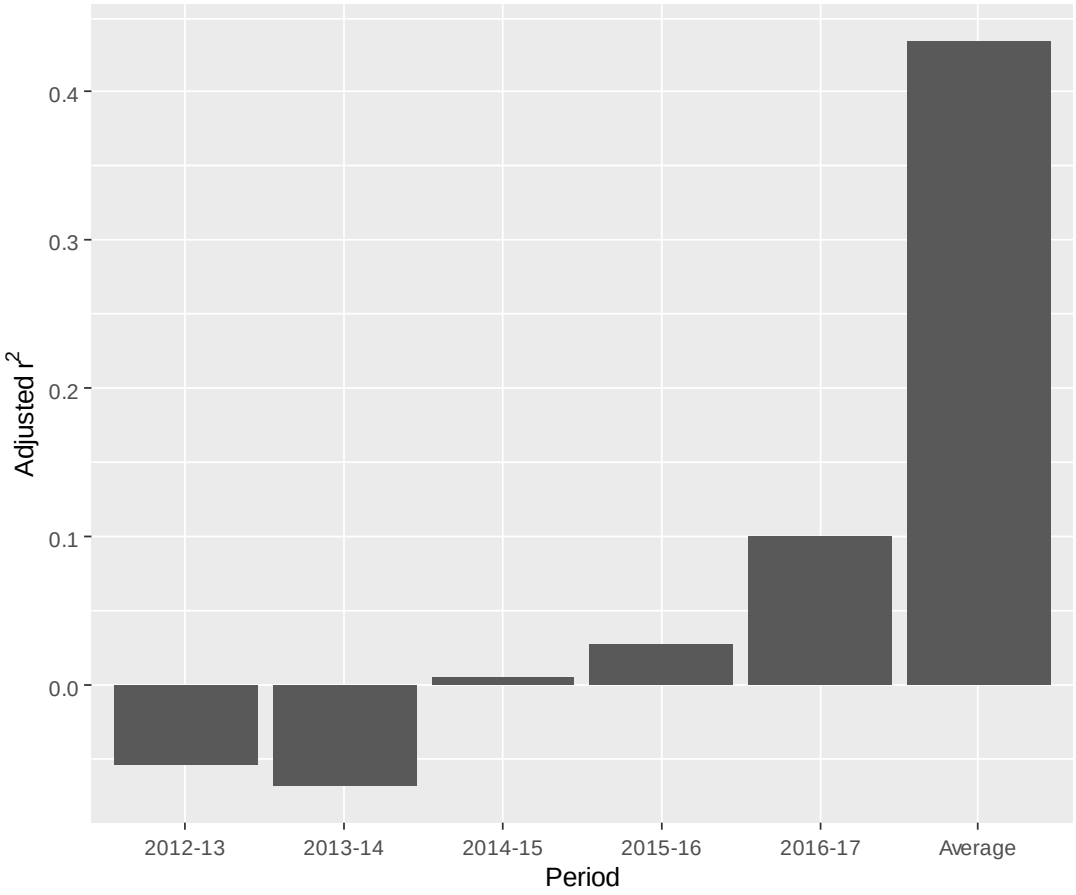

Figure 7. The total explained variation in the taxonomic community matrices (adjusted R^2). The plot shows associated with the summer sampling periods and the average over all period that was explained by the four tables of explanatory variables (macro, meso, micro and space).

All components of the variance decomposition are indicated schematically on Figure 5 and for the analyses corresponding to the community matrix in

Figure *8*. The micro, meso and macro-scale variables individually (e.g., for the micro scale the individual variance includes components [a], [e], [g], [l], [n], [o], [h] and [k] on Figure 5) explained 31%, 18% and 18% of the variation (adjusted R^2 values) in the community matrix, respectively (Table 8). The spatial variables individually (i.e., variance components [d], [j], [m], [i], [o], [k], [n] and [m] on Figure 5) explained 10% of the variation in the community matrix (Table 8). All individual models were significant (p=0.001).

The micro, meso and macro-scale variables uniquely (i.e., variance components [a], [b] and [c] respectively on Figure 5) explained 11%, 2% and 5% of the variation, respectively (Table 8). The contributions for the micro, meso, and macro scale were significant at the 1% level. The spatial variables uniquely (i.e., variance component [d] on Figure 5) explained 1% of the variation in the community matrix (Table 8) but this contribution was not significant.

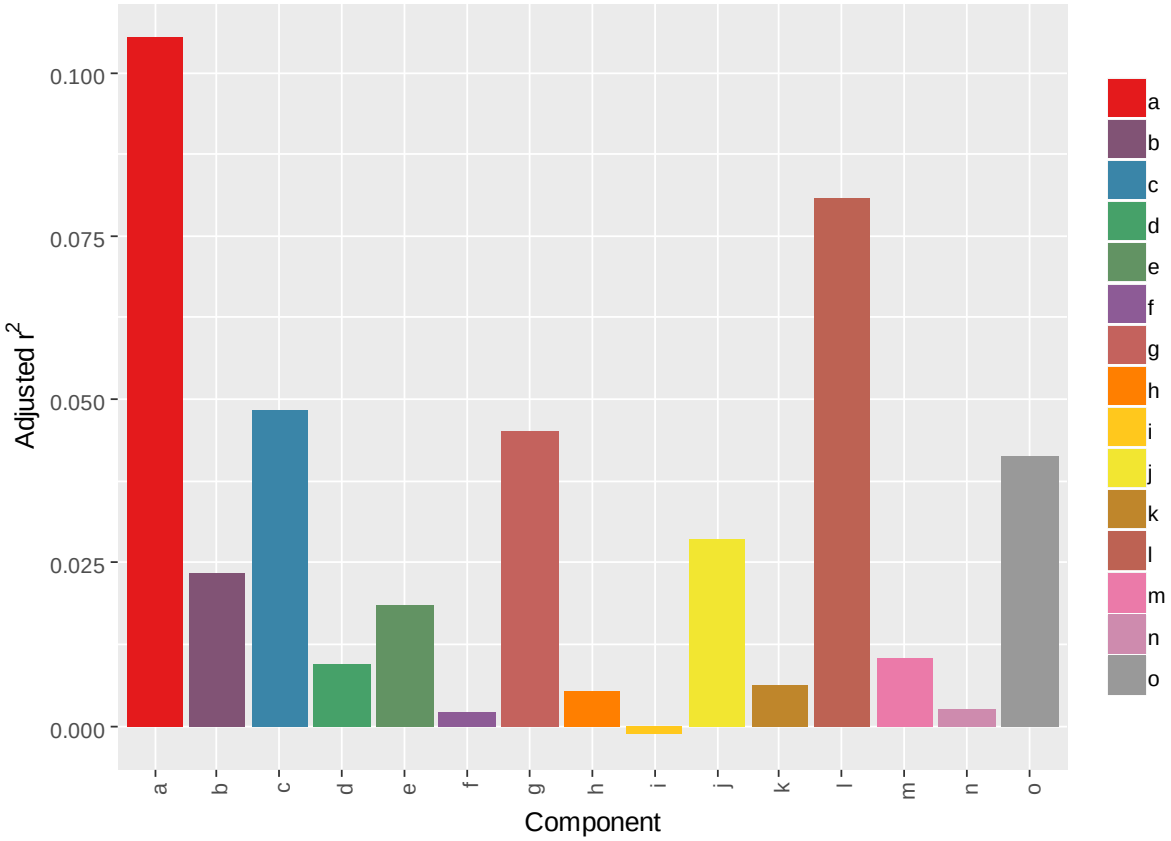

Figure 8. All components of the variance decomposition for the community matrix. The individual components are as shown on the schematic diagram on Figure 5.

Table 8. The variation explained (adjusted R^2) by the variance partitioning analysis of the average community matrix by the micro, meso, macro and spatial variables. The individual and unique components of variation explained attributable to each table of explanatory variables are the result of adding the associated individual components shown in Figure 8..

Explanatory variables	Component of variation	Adjusted R^2	*p*-value
Micro	Individual	31	0.001
Meso	Individual	18	0.001
Macro	Individual	18	0.001
Spatial	Individual	10	0.001
Micro	Unique	11	0.001
Meso	Unique	2	0.016
Macro	Unique	5	0.005
Spatial	Unique	1	0.148

4.3 Stepwise models

There were 14 explanatory variables with VIF values > 5. This resulted in the following variables being excluded from the stepwise model building: SegFlow, perc_Run, area_sqkm, Av_Width, Av_Sed_Depth, HAB_SCORE, USCalcium, Avg_Seg_Elev, Native_Bush, Av_Vel, FRE3, Order2, Pasture, ReachSed. This left 39 environmental variables available for the stepwise models.

4.3.1 Biological metrics

Of the 39 available explanatory variables (after exclusion of those with VIF > 5) 28 were retained in at least one of the biological metrics models following stepwise elimination. Note that these models represent the average over all summer sampling periods. The following eleven variables were not selected by any of the biological metric models: Stdev_Vel, CONDUCTIVITY, perc_Pool, perc_Rapid, Sub_Divers, AV_US_Slope, Exotic_Bush, Exotic_Scrub, SegCluesN, USHardness, USPhosporus.

The adjusted R2 values for the stepwise linear regression models fitted to the biological indices ranged between 41% and 70% (Figure 9). The importance and directions of the relationships between the explanatory variables and the biological indices are summarised in Figure 10. Plots showing these results for the individual biological metrics are provided in Appendix A. The plot represents the mean of each variable's importance over all six biological metric models. The length of the bar can be interpreted as the mean variation explained by the explanatory variable. The bars in Figure 10 are colour coded to indicate whether the variable was positively or negatively related to the biological indices. Some variables differed in the direction of their relationships with the biological indices between biological indices. Bars with both colours indicate the proportion of the biological indices for which the relationships were in each direction.

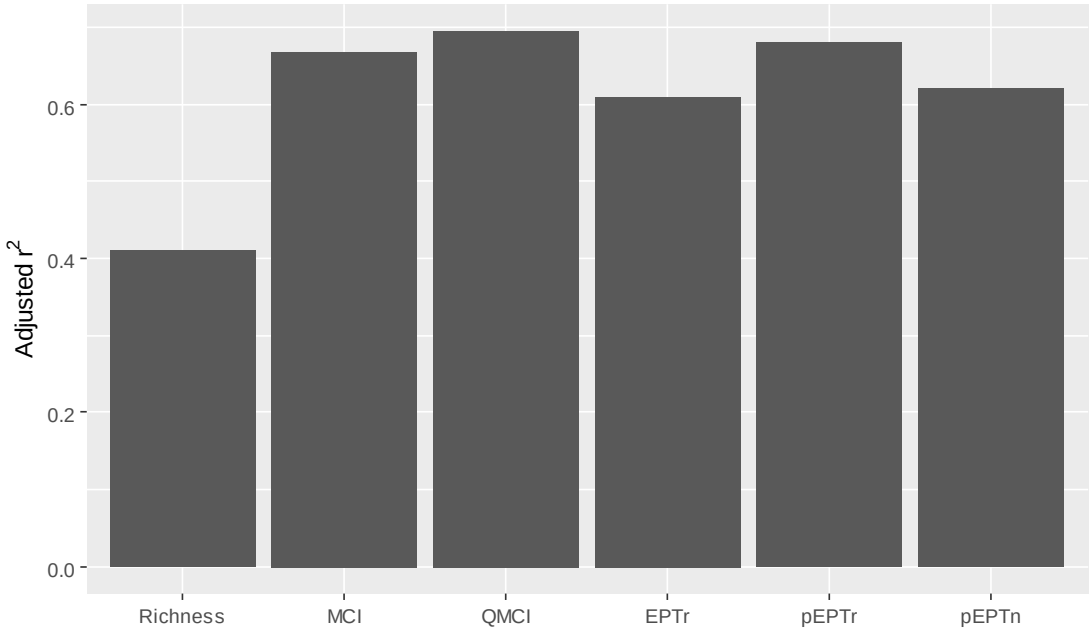

Figure 9. The adjusted R^2 values for the stepwise linear regression models fitted to the biological indices.

The most important variable was RHA1_Score, which was selected in 5 of the 6 models and had a mean importance of 12%. The five biological metrics which used this variable were all positively related to RHA1_Score (i.e., the regression coefficients were positive, indicating the value of the metric increases with increasing values of RHA1_Score). After RHA1_Score there were 10 other variables with importance values greater than 2 (and up to 5.1). These variables were, in decreasing order of importance: Hort and Shuffle_Index (both negatively related to the metrics), ReachHab (positively related), Std_Sed_Depth and Urban (both negatively related to the metrics), VEG_STR_ADJ, Av_Veg_over and GRND_ADJ, Flow_Hetero, perc_Riffle (all positively related).

Micro-scale explanatory variables dominated the variables that were included in the models. Of the 28 explanatory variables included in the models, 15, 6 and 7 were micro, meso and macro-scale variables respectively. Micro-scale variables comprised six of the 11 most important variables. Of the 28 explanatory variables included in the models, half had fitted relationships with the biological indices (Figure 10) that were consistent with our hypothesised responses (Table 3, Table 4, Table 5). However, nine of the 10 most important variables in the biological metrics models had fitted relationships between the explanatory variable and the biological indices that were consistent with our hypothesised responses (Table3, Table 4, Table 5). For example, the following micro-scale variables; RHA1_Score, Flow_Hetero and Av_Veg_over were positively related to the biological indices (Figure 10), which was consistent with the hypothesised responses (Table 3). The micro-scale variable Shuffle_Index was negatively related to the biological indices (Figure 10), which was also consistent with the hypothesised response (Table 3). The only explanatory variable that was in the 10 most important fitted variable for which the fitted relationship did not agree with the hypothesised response was std_sed_depth. This micro-scale variable was predicted to be neutral but its' fitted response was negative (Figure 10). The relationships of the meso and macro-scale explanatory variables with the biological responses were also generally consistent with the hypothesised responses. For example, the meso-scale variables

ReachHab, VEG_STR_ADJ, GRND_ADJ and SegSlope were all positively related to the biological indices and SegJanAirT and Shade_Cat were negatively related, which was consistent with our prediction. The macro-scale variables Hort and Urban were both negatively related to the biological indices, which was also consistent with our hypothesised responses.

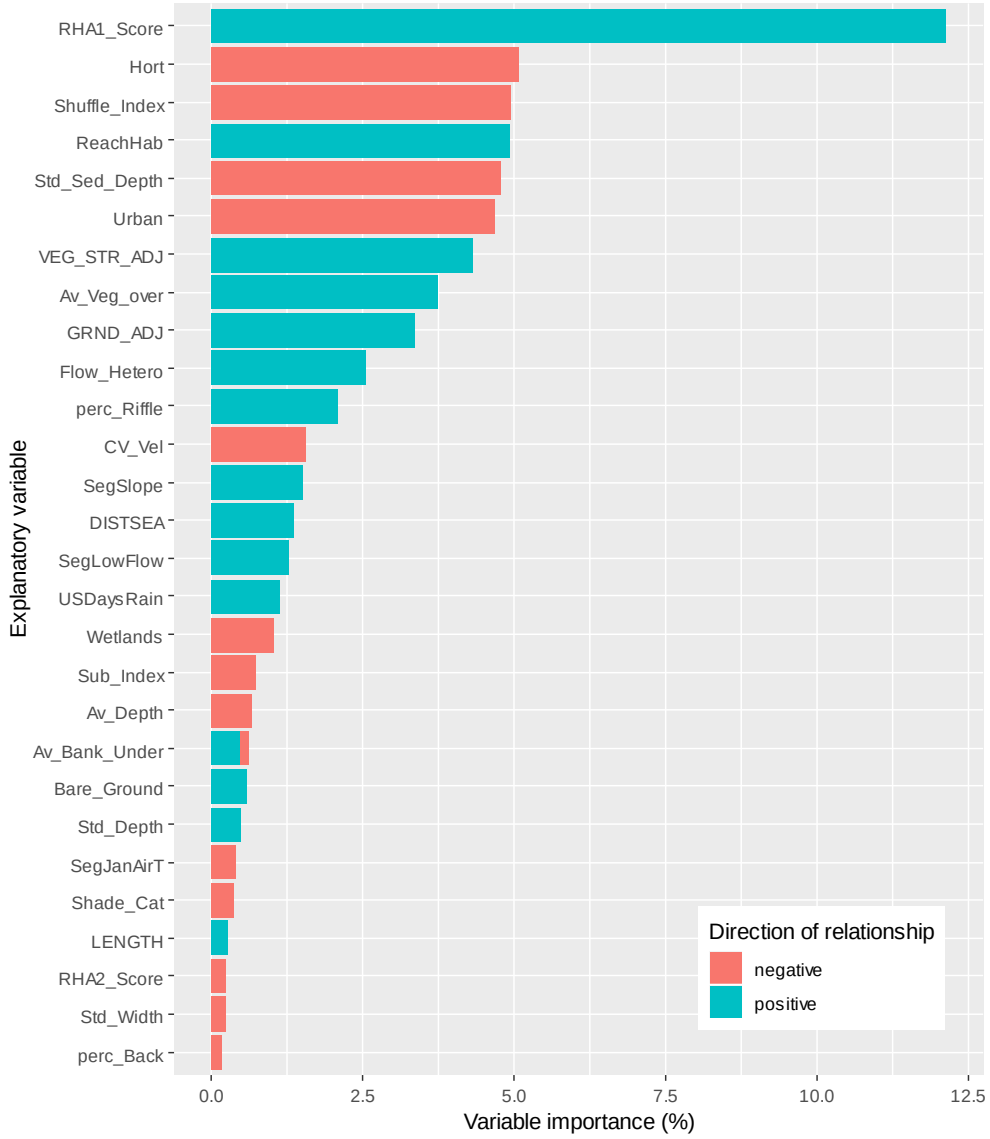

Figure 10. Importance and direction of relationships between the fitted explanatory variables in the stepwise models of the biological indices. The bar indicates the mean of that variable's importance measure over all six biological metric models. The colour of the bar indicates whether the variable was positively or negatively related to the biological indices. Bars with both colours indicate that relationships between that explanatory variable and the indices were both positive and negative and the proportion of the bar in each colour represents the proportion of models having relationships in each direction.

4.3.2 Community matrix

Of the 39 available explanatory variables (after exclusion of those with VIF > 5) 22 were retained in the stepwise CCA model. The following 16 variables were not selected in the CCA model: Std_Depth, Av_Veg_over, CV_Vel, LENGTH, GRND_ADJ, Flow_Hetero,

perc_Back, perc_Rapid, Sub_Divers, Shuffle_Index, RHA2_Score, Exotic_Bush, Exotic_Scrub, Urban, SegLowFlow, SegSlope, SegCluesN. The model explained 38% of the variation in the community matrix. The first two dimensions of the CCA (Figure 10, Figure 12, Figure 13) explained 17% of the variation in the community matrix (or 44% of the total variation explained by the model). The stepwise CCA model included 22 explanatory variables: nine represented micro-scale variables, four meso-scale variables, and nine macro-scale variables, including measured Conductivity. The first axis of the CCA was strongly associated with gradients in microscale habitat, with, for example, Sub_Index, perc_Riffle and RHA1_Score being positively correlated to this axis, and Std_Sed_Depth being negatively correlated. AV_US_Slope, VEG_STR_ADJ and ReachHab were also positively correlated to CCA axis 1. CCA axis 2 was strongly correlated to location (DISTSEA) and shade (Shade_Cat), and negatively correlated to climate (SegJanAirT), and depth (Av_Depth).

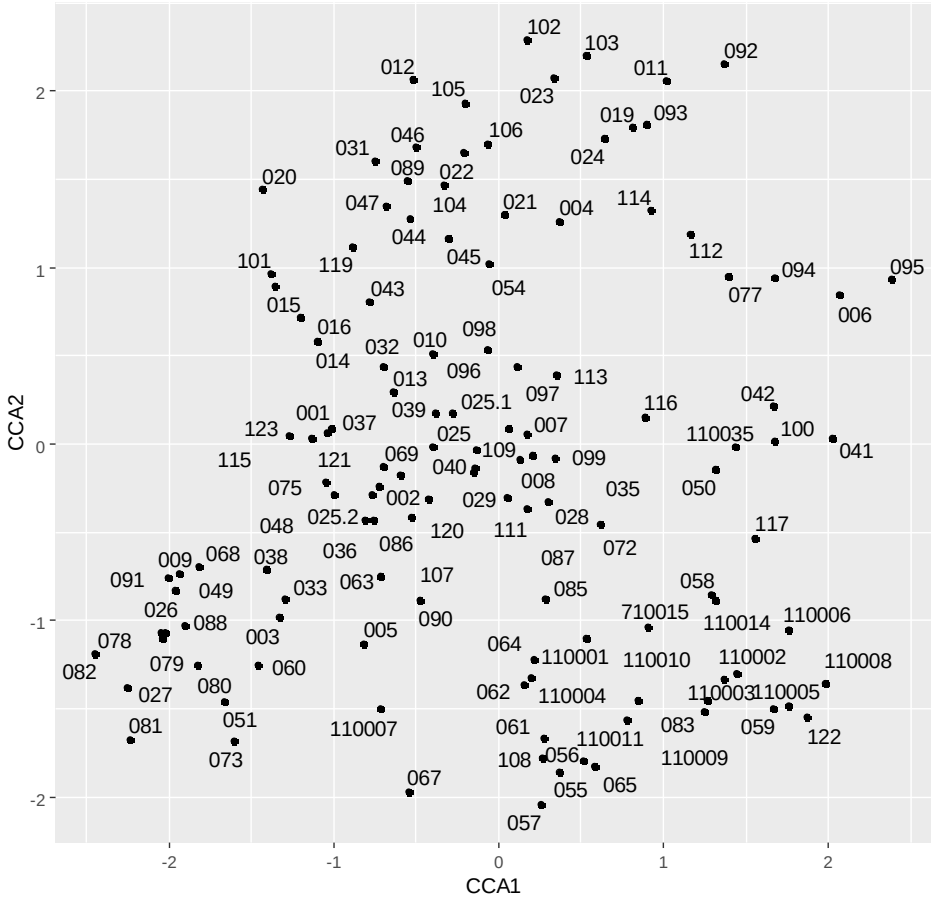

Figure 11. The sites projected onto the first and second components of the stepwise CCA. These components explained 17% of the variation in the taxonomic matrix, which was 44% of the total variation explained.

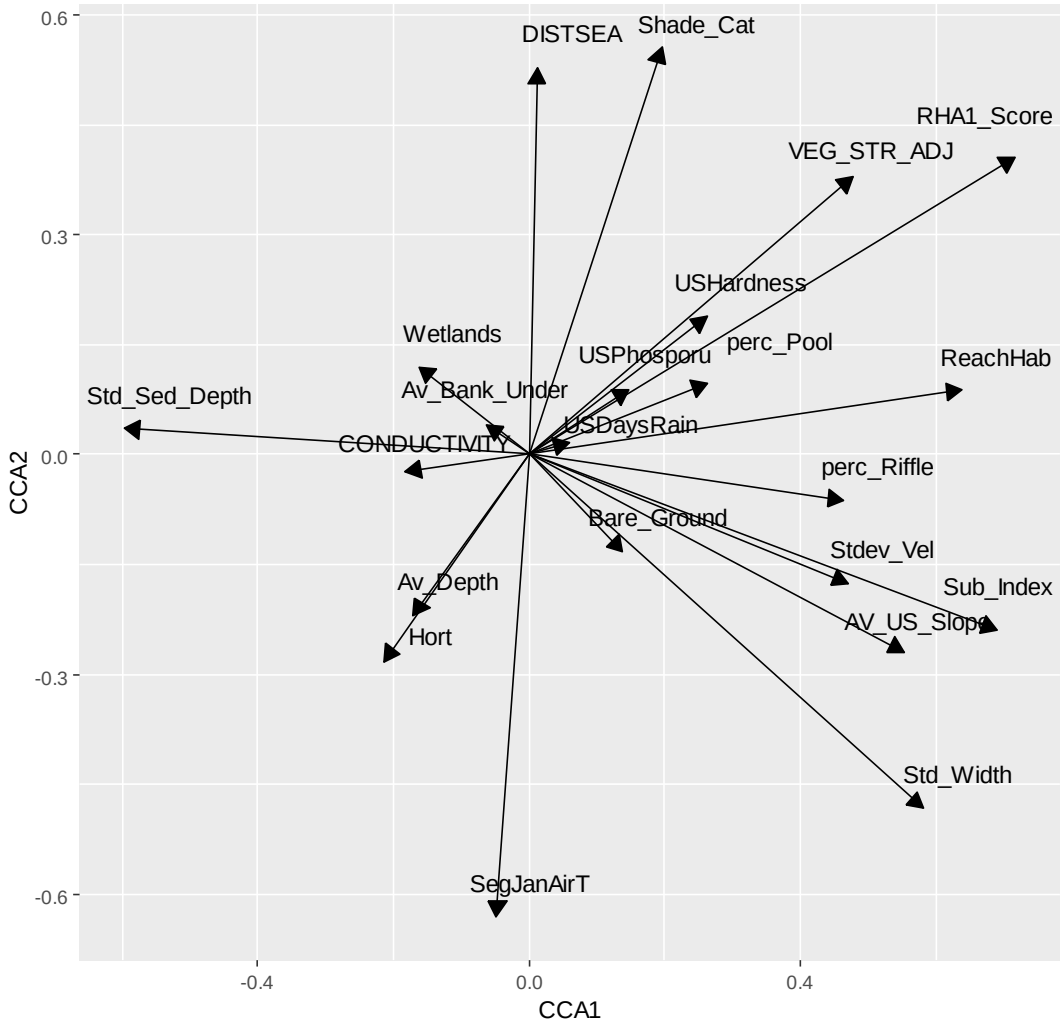

Figure 12. The explanatory variables projected onto the first and second components of the stepwise CCA. These components explained 17% of the variation in the taxonomic matrix, which was 44% of the total variation explained.

The environmental gradients were associated with variation in the abundance of particular taxa. Sites with high CCAaxis 1 and 2 scores were characterised by invertebrates that typically favour fast-flowing water and large substrates, including mayflies (*Acanthophlebia, Ichthybotus*), stoneflies (*Stenoperla, Spaniocerca, Austroperla*), and caddisflies (Alloecentrella, *Zelolessica, Hydrobiosella, Baraeoptera, Helicopsyche, Confluens etc)*). In contrast, sites with low CCA axis 1 and 2 scores were characterised by invertebrates that are associated with slow-flowing water and soft substrates. These included invertebrates such as *Paratya* shrimp, beetles (*Berosus*, Dytiscidae), the snails *Gyraulus* and *Physella*, the backswimmer *Sigara,* ribbon-worms (Nemertea) and leeches (Hirudinea) (Figure 12, Figure 13). A strong gradient thus existed between sites with EPT taxa, and those without.

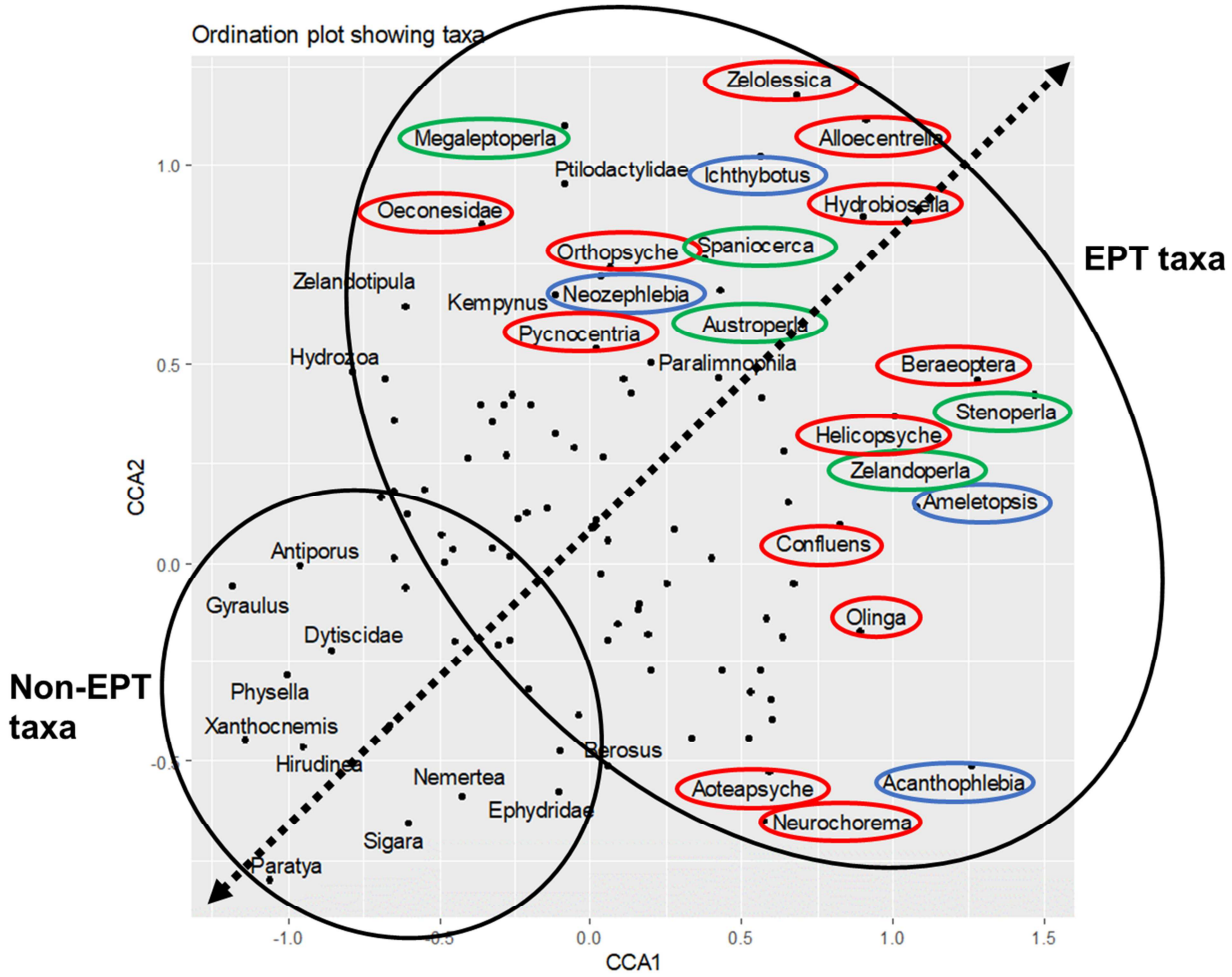

Figure 13. The taxa projected onto the first and second components of the stepwise CCA. These components explained 17% of the variation in the taxonomic matrix, which was 44% of the total variation explained. Note the strong gradient between sites with high and low occurrence of EPT taxa (Blue = Ephemeroptera, Green = Plecoptera, Red = Trichoptera).

The most important explanatory variables were RHA1_Score and Sub_Index, which both had a mean importance of 15%. The next two most important explanatory variables were ReachHab and Std_Width with mean importance of 13%. Micro-scale explanatory variables dominated the variables that were included in the CCA model. Of the 22 explanatory variables included in the models, nine, four and nine were micro, meso and macro-scale variables respectively. Six of the top 10 most important variables and nine of the 22 variables included in the CCA model were micro-scale.

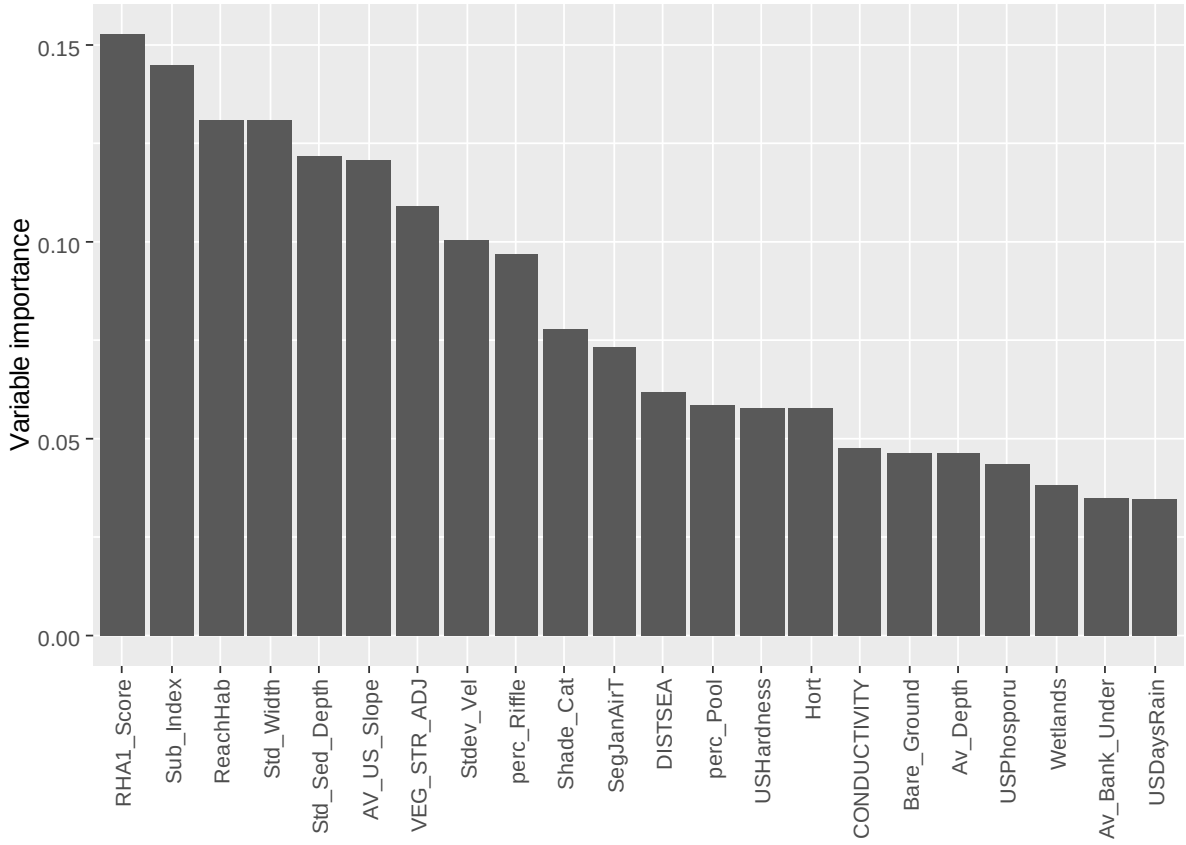

Figure 14. Importance of relationship between the fitted explanatory variables in the stepwise CCA model of the taxonomic matrix. The bar indicates that variable's importance measure.

5 Discussion

5.1 Temporal variation measurements

This study was not designed to discriminate inter-operator variability from within site temporal variability, although temporal variation of measured variables (Figure 2 and Figure 3) is attributable to both of these causes. We found that all models based on data pertaining to individual summer sampling periods performed poorly (i.e., the variation in the responses explained by the models was small). In contrast, the models that represented the average over all summer sampling periods performed well. For example, the annual models for MCI, pEPTr and QMCI had adjusted R^2 less than 35% but the models representing the averaged values had R^2 greater than 65%. In addition, the annual CCA models had R^2 less than 9% but the model representing the averaged values explained 43% of the variation in the community matrix.

We attribute the better performance of the analyses based on average over all summer sampling periods to the smoothing of temporal variability and consequent strengthening of the patterns of spatial variation. Temporal variability is attributable to a mix of inter-operator and within-site temporal variability in the observed micro-scale explanatory variables and variation in the invertebrate community composition (Figure 3). Therefore, we consider that sampling over multiple seasons is required to establish robust measures of ecosystem health and, more importantly, robust relationships between ecosystem health and the environmental variables that influence it.

5.2 Credibility of models

Two aspects of our results suggest our models are a credible basis to make inferences about the relative importance of the environmental determinants of invertebrate communities. First, all the models based on the average of all summer sampling periods performed well, which indicates strong associations between observations of variables and ecological communities at the monitoring sites. Second, nine of the 10 most important variables in the biological metrics models had fitted relationships between the explanatory variable and the biological indices that were consistent with our hypothesised responses. Of the 28 relationships that were observed for the six biotic metrics, 16 (or 50%) agreed with our predictions, while 9 (32%) did not agree. An additional 5 relationships (18%) were inconsistent with our hypotheses in that we predicted neutral but found linear relationships.

The positive associations between the biotic metrics and variables such as RHA1, flow heterogeneity, % riffles, GRND_ADJ, VEG_STR_ADJ reflect the preference for “sensitive” invertebrates such as the EPT taxa (which have high MCI tolerance scores) for sites with silt-free habitats, fast flowing water, and well-vegetated adjacent banks dominated by large trees with good ground cover. In contrast, these taxa were less common in deeper streams with more deposited sediment (associated with high values of Shuffle_index), draining catchments dominated by modified land cover including horticulture or urban development, and with warmer January air temperatures. These results are generally consistent with observations of both hydraulic habitat and substrate preferences for stream invertebrates (e.g., Jowett and Richardson, 1995), effects of deposited sediment on stream invertebrates (Suren and Jowett, 2006; Piggott et al 2012), and on relationships between RHA scores and MCI Scores (Clapcott 2015; Suren and Carter 2018). The results are also in general agreement with observations that stream ecological health is generally highest in un-modified catchments (e.g., native forest), intermediate in streams draining plantation forests, lower in streams draining agricultural areas, and lowest in streams draining urban

catchments (e.g. Quinn and Cooper 1997; Harding et al. 1999; Hall et al. 2001; Walsh et al. 2005; Allan 2007).

5.3 Importance of microscale variables

Our analyses indicated that up to 35 environmental variables made significant contributions to at least one of the biotic metric or invertebrate community composition models. Of these 35 variables, 18 were micro-scale, 6 meso-scale, and 11 macroscale, highlighting the importance of micro-scale variables in our models. In addition, the micro-scale variables generally explained more variation in macroinvertebrate community variation than the combination of the meso, macro and spatial variables (Table 6; Table 8). Without the micro-scale variables, all contributions to the variance partitioning models' adjusted R^2 value would be available except the component [a] shown on Figure 5. Although a relatively large proportion of variance is explained without the microscale variables, the contribution of microscale variables to the model significantly increases the total variance explained by most of the models (Figure 15).

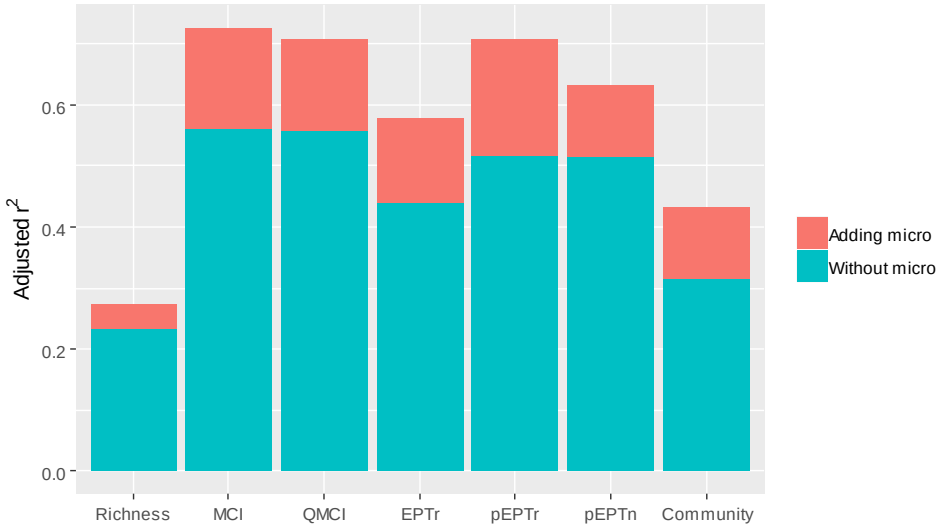

Figure 15. Summary of the results of variance partitioning. The plot shows the total variance explained by the six biotic index model, and the community matrix without microscale variables (i.e., components [b] to [o] of Figure 4: blue bars), and with microscale variables added (i.e., components [a] of Figure 4: red bars).

Similarly, for the stepwise regression analysis of the biotic metrics, the micro-scale variables contributed more to the R^2 of most of the models than the meso-scale and macro-scale variables combined (Figure 16).

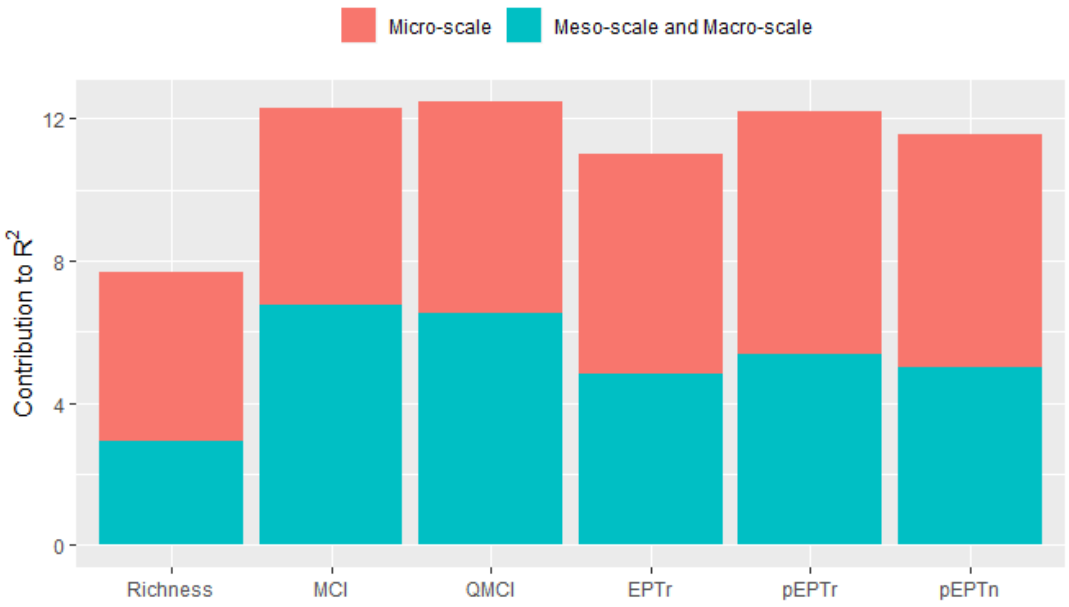

Figure 16. Summary of contribution to overall variance explained (R^2) of the stepwise regression models of the six biological indices. The plot shows the relative importance of microscale variables only (red bars), as well as meso and macro scale variables (blue bars).

The importance of microscale variables was also evident in the results of the CCA model, where the sum of the importance measures ($Importance_E$) over the micro-scale variables was approximately equal to the sum of the importance measures over the meso and macro-scale variables combined (Figure 17).

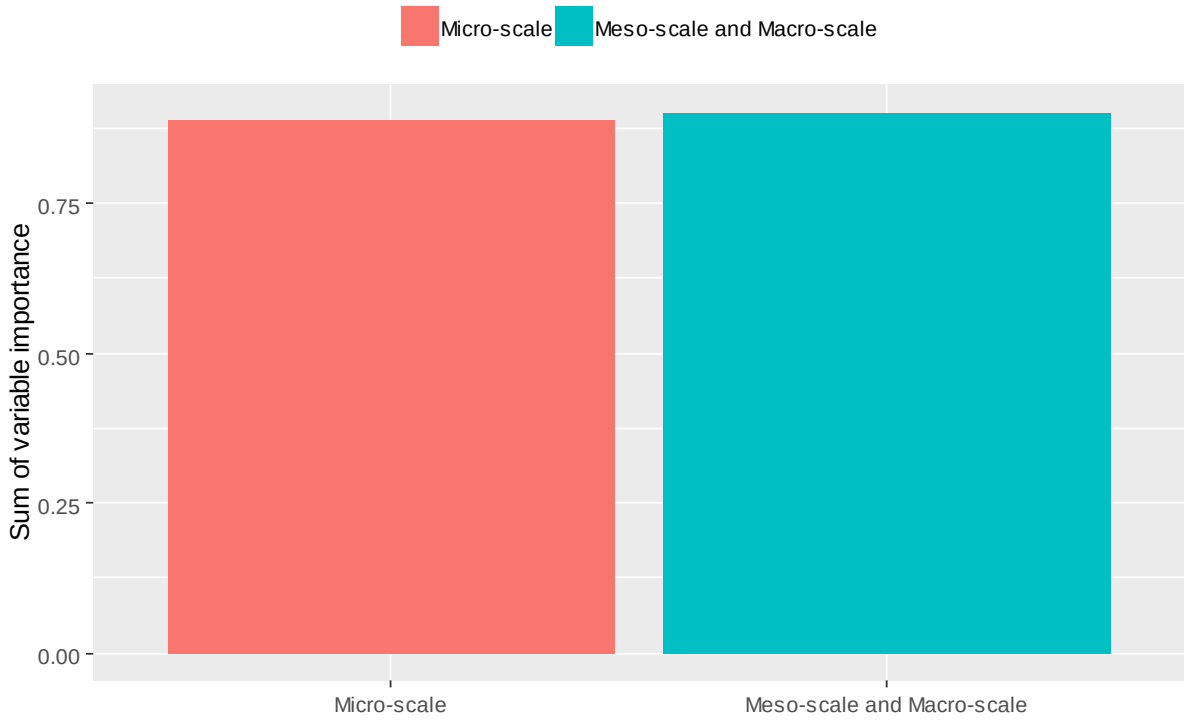

Figure 17. Summary of contribution to variable importance of the stepwise CCA model representing the taxonomic matrix. The plot shows the sum of importance ($Importance_E$) over all microscale variables only (red bars), and the sum of importance ($Importance_E$) over all meso and macro scale variables (blue bars).

Associations between water quality and invertebrate communities and ecological processes in streams were reported by Wagenhoff *et al.* (2017). They used boosted regression trees to model relationships between predictor variables related to land use, and a range of ecological indicators including MCI score, EPT richness and percentage EPT, chlorophyll a, and the rate of decay of cotton strips. They found strong evidence for ecosystem change with small increases in nitrogen concentrations. In particular, they found that the three macroinvertebrate metrics (MCI, EPT richness and percentage EPT) responded negatively to increasing total nitrogen (TN). Wagenhoff *et al.* (2017) suggested the sensitivity of macroinvertebrate communities to TN concentration is due to the eutrophication pathway (i.e., nitrogen enrichment leading to stimulation of growth and changes to periphyton communities and peak biomass). Similar links between periphyton development and invertebrate communities have been observed by Suren *et al.* (2003) in gravel bed streams in North Canterbury, where invertebrate communities changed in proportion to the degree of periphtyon biomass.

Of the 11 macro-scale variables that were included in our models, only two (CONDUCTIVITY and USPhosphorus) represented water quality. Other water quality variables such as SegCluesN were not selected by any model. This result is in contrast to the association between nitrogen and ecological health shown by Death *et al.* (2018). Death

et al.'s (2018) analysis considered associations between MCI scores and nitrate levels at the national scale, and therefore involved a wider range of stream nitrate concentrations. Nitrate concentrations in Bay of Plenty streams draining the predominantly natural conservation estate appear to be higher than in other regions, probably reflecting natural enrichment due to the volcanic geology of much of the area. The relatively high nitrate levels in the Bay of Plenty region may mask relationships between nitrate concentrations and invertebrate communities that are apparent at the national scale where background (i.e., natural state) nitrate concentrations are lower.

The absence of associations between SegCluesN and biotic measures in our study may be because nutrient enrichment in Bay of Plenty streams has a limited effect on periphyton biomass. The beds of just over half the monitoring sites were dominated by fine, unstable pumice sand, where periphyton is unlikely to achieve high biomass. Furthermore, many of the sites were shaded by native or exotic plantation forests: over half our sites had more than 75% shade. As a result of this highly mobile streambed and shading, periphyton biomass and consequent ecological impacts of nutrient enrichment may be supressed. Consequently, strong links between nutrient enrichment and ecosystem health may not be as apparent in the Bay of Plenty streams as they are in more open gravel bed streams. BoPRC is currently monitoring periphyton biomass in 30 gravel-bed streams throughout the region, and this has shown that biomass is consistently low, and in the NPS-FM A-band (A. Suren unpublished data).

Suren and Riis (2010) developed a conceptual model linking change in invertebrate communities to the plant biomass in streams. Their model postulates that changes to invertebrate community composition are linked to the development of plant biomass in streams during times of low, stable flow. Where plant biomass does not increase, the model postulates that communities will be stable, while major shifts to community composition and structure are postulated to occur in streams in response to the development of high algal or macrophyte biomass. Our finding that invertebrate community composition is not strongly linked to water quality in the Bay of Plenty region is consistent with this conceptual model. Our explanation for the lack of response to nutrient enrichment is that the enrichment does not result in a substantial increase in periphyton biomass due to stream shading and the unstable substrate in the generally pumice-bed streams.

Clapcott *et al.* (2017) developed predictive models to estimate reference values for the MCI and tested the effect of spatial scale on predictive accuracy of these models. These models used 22 variables derived from the FENZ database: 16 macro scale variables, and 6 mesoscale variables. They used three statistical techniques (ANCOVA, random forest (RF) models, and boosted regression trees (BRT)) to predict reference conditions at a national scale by modelling MCI as a function of environmental and human impact variables. Their national models explained a relatively large amount of the variability in MCI scores, with cross validated Nash Sutcliff efficiency ranging from 0.55 to 0.63. Clapcott *et al.* (2017) identified four mesoscale environmental predictors for MCI at a national level (segment flow stability, segment habitat, segment substrate, and segment summer temperature), and one macroscale variable (average upstream slope). Our results also highlighted the importance of segment habitat (ReachHab), average upstream slope (AV_US_Slope), and segment summer temperature (SegJanAirT), but modelled substrate size was not identified in our analyses. Instead, measured substrate size, shuffle index scores, and variability in sediment depth of were more important, as was measured assessments of habitat using the RHA1

protocol. The finding that microscale variables measured in our analysis contributed so strongly in explaining variability to invertebrate community composition suggests that the patterns found by Clapcott *et al.* (2017) may have been further improved by addition of these microscale variables. Clapcott *et al.* (2017) suggested that the performance of their models could have been improved by the inclusion of more measures of microscale conditions such as substrate size.

Finally, Clapcott *et al.* (2017) showed that catchment native vegetation cover was the most important predictor of MCI scores. This finding is consistent with numerous studies that have shown catchment land cover is an important predictor of MCI scores at multiple spatial scales (e.g., Death and Collier 2010). However, the proportion of catchment native vegetation was not identified at all in any of our analyses and the only catchment land cover variables that were selected by our models were describing the percentage of horticulture, urban, wetlands or bare area. The reasons for this disparity are unknown, but may be because many of the streams sampled in the Bay of Plenty are draining exotic pine plantation forests. Invertebrate communities in these streams are very similar to those in streams draining native forest (Suren *et al.,* 2017). This finding agrees with observations by Harding *et al.*, (1999). The similarly high ecological health in streams draining catchments dominated by both native and plantation forest may partially explain why percentage of native forest was not selected by any of our models.

5.4 Implications for ongoing monitoring

Out results have shown that micro-scale habitat variables make signifcant contributions to models explaining variation in invertebrate community composition in the Bay of Plenty region. In particular, our models indicate invertebrate communities have associations with substrate (Shuffle_Index, Std_Sed_Depth, Sub_Index), instream hydraulic conditions (Flow_Hetero, perc_Riffle, CV_Vel, Av_Depth, Std_Depth, Std_width, perc_Back), and riparian vegetation (VEG_STR_ADJ, Av_Veg_Over, GRND_ADJ).

The RHA1 score explanatory variable is an index that is made up of many of the micro-scale variables, including deposited sediment, hydraulic heterogeneity, bank vegetation, riparian buffer width, and riparian shade. Our finding that the RHA1 score was most important variable indicates that it is an efficient index of habitat quality. However, the RHA1 score is composed of nine individual habitat variables (Table 2). Some of these variables (e.g., Hydraulic heterogeneity) were also quantified in more detailed ways during the survey, such as measuring width at 5 transects, or water depth and velocity at 15 locations to derive the average and standard deviation of these factors. The RHA1 score was not directly correlated to these individual measurements of the component variables, but it was identified as the most important predictor variable suggesting that this combined assessment of habitat quality performs better than individual assessments of the component variables. This suggests that composite variables such as the RHA1 score may be a more cost-effective and robust way of assessing habitat quality than measuring individual components of habitat quality.

These results have implications for managing ecosystem health. As part of implementing the NPS-FM, a number of attribute states have been defined that council must use to set numerical objectives. For rivers, there are only three compulsory attributes that are relevant to ecosystem health: nitrogen concentration (as nitrate and ammonia) for toxicity, and periphyton biomass. The inherent assumption of these attributes is that increasing nitrogen in streams can have both direct effects (toxic) and indirect effects, via periphyton proliferation. Our results highlight that additional variables that characterise physical habitat

are also strongly associated with ecosystem health, and in particular microscale variables. It is therefore appropriate to consider that physical habitat considerations are also important to achieving ecosystem health objectives, and as such, need to be monitored and managed. Our results also suggest that Regional Councils need to monitor habitat characterisitics to better understand observed changes in ecological health. In particular, important micro-scale variables include composite assessments of habitat quality using Rapid Habitat Assessment scores, as well as direct measurements of fine sediment deposition and the amount of overhanging vegetation. These latter findings have important implications for BoPRC's ongoing riparian protection work, highlighting the importance of riparian shade and sediment deposition on the ecological health of streams. It may be possible to develop criteria for some micro-scale habitat variables to support ecological health objectives. This would result in more focussed habitat management and surveys in the future.

Acknowledgements

Each year, Bay of Plenty Regional Council (BOPRC) employs summer students to collect invertebrate samples, and measure habitat variables throughout the region. This report, and the data it contains, would not have been possible without their help and perseverance over the years. We also thank all the landowners throughout the region who provided access to the monitoring sites during the summer monitoring period.

References

Akaike, H., 1973. Information Theory and an Extension of the Maximum Likelihood Principle. B. N. Petrov and F. Csaki (Editors). Springer Verlag, ed. Akademiai Kiado: Budapest., pp. 267–281.

Allan, J.D. 2007. Landscapes and riverscapes: the influence of land use on stream ecosystems. *Annual Review of Ecology and Systematics 35*: 257-284.

Angermeier, P.L. and M.R. Winston, 1999. Characterizing Fish Community Diversity across Virginia Landscapes: Prerequisite for Conservation. Ecological Applications 9:335–349.

Biggs, B., 1985. The Use of Periphyton in the Monitoring of Water Quality. Water & Soil Miscellaneous Publication-National Water and Soil Conservation Authority.

Biggs, B.J.F., 2000. Eutrophication of Streams and Rivers: Dissolved Nutrient-Chlorophyll Relationships. Journal of the North American Benthological Society. 19:17–31.

Biggs, B.J.F., M.J. Duncan, A.M. Suren, and J.R. Holomuzki, 2001. The Importance of Bed Stability to Benthic Ecosystems. Gravel-Bed Rivers. In: Mosley, M.P. (Editor). New Zealand Hydrological Societ, Christchurch, New Zealand, pp. 423–449.

Booker, D.J. and T.H. Snelder, 2012. Comparing Methods for Estimating Flow Duration Curves at Ungauged Sites. Journal of Hydrology.

Borcard, D., P. Legendre, and P. Drapeau, 1992. Partialling out the Spatial Component of Ecological Variation. Ecology 73:1045–1055.

ter Braak, C.J.F., 1986. Canonical Correspondence Analysis: A New Eigenvector Technique for Multivariate Direct Gradient Analysis. Ecology 67:1167–1179.

Clapcott JE 2013. Rapid habitat assessment workshop. Prepared for Hawkes Bay Regional Council. Cawthron Report No. 2445. 7 p.Clapcott, J.E., 2015. National Rapid Habitat Assessment Protocol Development for Streams and Rivers. Cawthron Institute report, Cawthron Institute, Nelson, New Zealand.

Clapcott, J.E., E.O. Goodwin, T.H. Snelder, K.J. Collier, M.W. Neale, and S. Greenfield, 2017. Finding Reference: A Comparison of Modelling Approaches for Predicting Macroinvertebrate Community Index Benchmarks. New Zealand Journal of Marine and Freshwater Research 51:44–59.

Crawley, M.J., 2002. Statistical Computing: An Introduction to Data Analysis Using S-Plus. John Wiley & Sons Inc, Chichester, United Kingdom.

Death, R.G., 1995. Spatial Patterns in Benthic Invertebrate Community Structure: Products of Habitat Stability or Are They Habitat Specific? Freshwater Biology 33:455–467.

Death, R.G., A. Canning, R. Magierowski, and J. Tonkin, 2018. Why Aren't We Managing Water Quality to Protect Ecological Health? Occasional Report 31. Palmerston North, NZ., p. .

Death RG, Collier KJ.2010. Measuring stream macroinvertebrate responses to gradients of veg-etation cover: when is enough enough? Freshwater Biology. 55:1447–1464.

Ferreira, W.R., R. Ligeiro, D.R. Macedo, R.M. Hughes, P.R. Kaufmann, L.G. Oliveira, and M. Callisto, 2014. Importance of Environmental Factors for the Richness and Distribution of Benthic Macroinvertebrates in Tropical Headwater Streams. Freshwater Science 33:860–871.

Fortin, M.J. and M.R.T. Dale, 2005. Spatial Analysis – A Guide for Ecologists. Cambridge University Press, Cambridge.

Frissell, C.A., W.L. Liss, C.E. Warren, and M.C. Hurley, 1986. A Hierarchical Framework for Stream Habitat Classification, Viewing Streams in a Watershed Context. Environmental Management 10:199-214.

Greenwood, M.J. and D.J. Booker, 2015. The Influence of Antecedent Floods on Aquatic Invertebrate Diversity, Abundance and Community Composition. Ecohydrology 8:188–203.

Grömping, U., 2006. Relative Importance for Linear Regression in R: The Package Relaimpo. Journal of Statistical Software 17:1–27.

Guégan, J.F., S. Lek, and T. Oberdorff, 1998. Energy Availability and Habitat Heterogeneity Predict Global Riverine Fish Diversity. Nature 391:382–384.

Hall, M.J., G.P. Closs, and R.H. Riley, 2001. Relationships between Land Use and Stream Invertebrate Community Structure in a South Island, New Zealand, Coastal Stream Catchment. New Zealand Journal of Marine and Freshwater Research 35:591–603.

Harding, J.S., 2009. Stream Habitat Assessment Protocols for Wadeable Rivers and Streams in New Zealand. University of Canterbury, School of Biological Sciences.

Harding, J.S. and M.J. Winterbourn, 1995. Effects of Contrasting Land Use on Physio-Chemical Conditions and Benthic Assemblages of Streams in a Canterbury (South Island, New Zealand) River System. New Zealand Journal of Marine and Freshwater Research 29:479–492.

Harding, J.S.; Young, R.G.; Hayes, J.W.; Shearer, K.A.; Stark, J.D. 1999. Changes in agricultural intensity and river health along a river continuum. *Freshwater Biology 42*: 345-357.

Horwitz, R.J., 1978. Temporal Variability Patterns and the Distribution Patterns of Stream Fishes. Ecological Monographs 48,:307–321.

Jowett, I.G., 1993. A Method for Objectively Identifying Pool, Run, and Riffle Habitats from Physical Measurements. New Zealand Journal of Marine and Freshwater Research 27:241–248.

Jowett, I. and J. Richardson, 1995. Habitat Preferences of Common, Riverine New Zealand Native Fishes and Implications for Flow Management. New Zealand Journal of Marine and Freshwater Research 29:13–23.

Kutner, M.H., C. Nachtsheim, and J. Neter, 2004. Applied Linear Regression Models. McGraw-Hill/Irwin.

Lamouroux, N., J.-M. Olivier, H. Persat, M. Pouilly, Y. Souchon, and B. Statzner, 1999. Predicting Community Characteristics from Habitat Conditions: Fluvial Fish and Hydraulics. Freshwater Biology 42:275–299.

Leathwick, J., T. Snelder, W. Chadderton, J. Elith, K. Julian, and S. Ferrier, 2011. Use of Generalised Dissimilarity Modelling to Improve the Biological Discrimination of River and Stream Classifications. Freshwater Biology 56:21–38.

Legendre, P. and L. Legendre, 1998. Numerical Ecology. Elsevier, Amsterdam, The Netherlands.

Legendre, P. and M. Troussellier, 1988. Aquatic Heterotrophic Bacteria: Modeling in the Presence of Spatial Autocorrelation. Limnology and Oceanography 33:1055–1067.

Lenat, D.R. and J.K. Crawford, 1994. Effects of Land Use on Water Quality and Aquatic Biota of Three North Carolina Piedmont Streams. Hydrobiologia 294:185–199.

Leps, M., J.D. Tonkin, V. Dahm, P. Haase, and A. Sundermann, 2015. Disentangling Environmental Drivers of Benthic Invertebrate Assemblages: The Role of Spatial Scale and Riverscape Heterogeneity in a Multiple Stressor Environment. Science of the Total Environment 536:546–556.

Marzin, A., P.F. Verdonschot, and D. Pont, 2013. The Relative Influence of Catchment, Riparian Corridor, and Reach-Scale Anthropogenic Pressures on Fish and Macroinvertebrate Assemblages in French Rivers. Hydrobiologia 704:375–388.

Ministry for the Environment, 2017. National Policy Statement for Freshwater Management 2014 (Amended 2017).

Minshall, G.W., 1984. Aquatic Insect-Substratum Relationships. The Ecology of Aquatic Insects. Praeger, New York, pp. 358–400.

Mykrä, H., J. Heino, and T. Muotka, 2007. Scale-Related Patterns in the Spatial and Environmental Components of Stream Macroinvertebrate Assemblage Variation. Global Ecology and Biogeography 16:149–159.

Parkyn, S.M., 2004. Review of Riparian Buffer Zone Effectiveness. MAF technical paper, Ministry of agriculture and forestry, Wellington, New Zealand.

Paul, M.J. and J.L. Meyer, 2001. Streams in the Urban Landscape. Annual Review of Ecology and Systematics 32:333–365.

Peres-Neto, P.R., P. Legendre, S. Dray, and D. Borcard, 2006. Variation Partitioning of Species Data Matrices: Estimation and Comparison of Fractions. Ecology 87:2614–2625.

Piggott, J.J.; Lange, K.; Townsend, C.R.; Matthaei, C.D. 2012. Multiple Stressors in Agricultural Streams: A Mesocosm Study of Interactions among Raised Water Temperature, Sediment Addition and Nutrient Enrichment. *PLoS ONE 7*: e49873.

Poff, N.L., 1997. Landscape Filters and Species Traits: Towards Mechanistic Understanding and Prediction in Stream Ecology. Journal of the North American Benthological Society 16:391–409.

Quinn, J.M.; Cooper, A.B. 1997. Land-water interactions at Whatawhata, New Zealand: introduction and synthesis. *New Zealand Journal of Marine and Freshwater Research 31*: 569-577.

Quinn, J.M., A.B. Cooper, R.J. Davis-Colley, J.C. Rutherford, and R.B. Williamson, 1997. Land Use Effects on Habitat, Water Quality, Periphyton, and Benthic Invertebrates in Waikato, New Zealand, Hill-Country Streams. New Zealand Journal of Marine and Freshwater Research 31:579–597.

R Development Core Team, 2004. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org.

Resh, V.H. and D.M. Rosenberg, 1993. Freshwater Biomonitoring and Benthic Macroinvertebrates. Chapman & Hall, New York.

Richards, C., L.B. Johnson, and G.E. Host, 1996. Landscape-Scale Influences on Stream Habitats and Biota. Canadian Journal of Fisheries and Aquatic Sciences 53:295–311.

Sandin, L. and R.K. Johnson, 2004. Local, Landscape and Regional Factors Structuring Benthic Macroinvertebrate Assemblages in Swedish Streams. Landscape Ecology 19:501–514.

Scarsbrook, M.R., I.G. Boothroyd, and J.M. Quinn, 2000. New Zealand's National River Water Quality Network: Long-Term Trends in Macroinvertebrate Communities. New Zealand Journal of Marine and Freshwater Research. 34:289–302.

Snelder, T.H. and B.J.F. Biggs, 2002. Multi-Scale River Environment Classification for Water Resources Management. Journal of the American Water Resources Association 38:1225–1240.

Snelder, T., U. Shankar, K. Dey, and H. Hurren, 2006. Development of Variables for Freshwater Environments of New Zealand (FWENZ): Lakes. Christchurch.

Stark, J.D., 1985. A Macroinvertebrate Community Index of Water Quality for Stony Streams. Water & Soil Miscellaneous Publications.

Stark, J.D., I.K.G. Boothroyd, J.S. Harding, J.R. Maxted, and M.R. Scarsbrook, 2001. Protocols for Sampling Macroinvertebrates in Wadeable Streams. Cawthron Institute. http://www.pams.canterbury.ac.nz/ferg/documents/ProtocolsManual.pdf. Accessed 11 Mar 2017.

Stark, J.D. and J.R. Maxted, 2007. A Biotic Index for New Zealand's Soft-Bottomed Streams. New Zealand Journal of Marine and Freshwater Research 41:43–61.

Statzner, B. and B. Higler, 1986. Stream Hydraulics as a Major Determinant of Benthic Invertebrate Zonation Patterns. Freshwater Biology 16·:127-139.

Suren, A.M., B.J.F. Biggs, M.J. Duncan, L. Bergey, and P. Lambert, 2003. Benthic Community Dynamics during Summer Low-flows in Two Rivers of Contrasting Enrichment 2. Invertebrates. New Zealand Journal of Marine and Freshwater Research 37:71–83.

Suren, A.M., Carter, R (2018). Ecological and water quality conditions of drains and land drainage canals in the Rangitaiki and Kaituna plains. Environmental Publication 2018/05. ISSN: 1175-9372 (Print) ISSN: 1179-9471 (Online). 116p.

Suren, A.M. and S. Elliot, 2004. Impacts of Urbanisation on Streams. J. S. Harding, P. Mosley, C. Pearson, and B. Sorell (Editors). Freshwaters of New Zealand. New Zealand Hydrological and Limnological Societies, Christchurch, New Zealand, pp. 35.1-35.17.

Suren, A.M. and I.G. Jowett, 2006. Effects of Floods versus Low Flows on Invertebrates in a New Zealand Gravel-Bed River. Freshwater Biology 51:2207–2227.

Suren, A.M. and T. Riis, 2010. The Effects of Plant Growth on Stream Invertebrate Communities during Low Flow: A Conceptual Model. Journal of the North American Benthological Society 29:711–724.

Suren, A.M., D. Van Nistelrooy, and V. Fergusson, 2017. State and Trends in River Health (1992 – 2014) in the Bay of Plenty: Results from 22 Years of the NERMN Stream Biomonitoring Program. Bay of Plenty Regional Council, Whakatane, New Zealand.

Vannote, R.L., G.W. Minshall, K.W. Cummins, J.R. Sedell, and C.E. Cushing, 1980. The River Continuum Concept. Canadian Journal of Fisheries and Aquatic Sciences 37:130–137.

Walsh, C.J.; Roy, A.H.; Feminella, J.W.; Cottingham, P.D.; Groffman, P.M.; Morgan, R.P.I. 2005. The urban stream's syndrome: current knowledge and the search for a cure. *Journal of the North American Benthological Society 24*: 706-723.

Wagenhoff A, Liess A, Pastor A, Clapcott JE, Goodwin EO, Young RG 2017. Thresholds in ecosystem structural and functional responses to agricultural stressors can inform limit setting in streams. Freshwater Science 36: 178-194

Wolman, M.G., 1954. A Method of Sampling Coarse River-Bed Material. EOS, Transactions American Geophysical Union 35:951–956.

Zar, J.H., 1999. Biostatistical Analysis. Prentice-Hall, New Jersey.

Appendix A Importance and directions of relationships between the explanatory variables and the individual biological indices