Waikato Regional Council Technical Report 2022/44

Investigation of relationships between invertebrates and dissolved nutrient concentrations in New Zealand rivers

www.waikatoregion.govt.nz
ISSN 2230-4363 (Online)

Waikato
REGIONAL COUNCIL
Te Kaunihera ā Rohe o Waikato

Prepared by:
Ton Snelder (LWP Ltd)

For:
Waikato Regional Council
Private Bag 3038
Waikato Mail Centre
HAMILTON 3240

October 2022

Document #: 24671522

Peer reviewed by:
Dr Duncan Gray and Dr Roger Hodson Date Feb 2023

Approved for release by:
Tracey May Date September 2023

Disclaimer

July 20233

Investigation of relationships between invertebrate and dissolved nutrient attributes in New Zealand rivers

July 20233

Prepared By:

Ton Snelder

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: 2022-07

Report Date: April 20233

Quality Assurance Statement

Version	Reviewed By	
Final	Caroline Fraser	

Table of Contents

Figures

Tables

Executive Summary

A requirement of the National Policy Statement for Freshwater Management 2020 (NPS-FM; NZ Government, 2020) is the identification of appropriate nutrient concentration criteria (i.e., concentrations that will achieve ecosystem health objectives) and the implementation of these criteria through setting limits on resource use via regional plans. For rivers, the NPS-FM requires regional councils to determine the appropriate nutrient concentration criteria to "achieve a target attribute state for periphyton, any other nutrient attribute, and any attribute that is affected by nutrients" (NPSFM, Section 3.13). However, guidance provided by MFE for implementation of NPS-FM Section 3.13 highlights that a challenge in defining instream nutrient concentration criteria is accounting for the extent to which environmental factors influence the sensitivity of attributes to nutrient enrichment.

Two NPS-FM attributes that can be expected to be affected by nutrients are the macroinvertebrate community index (MCI) and the quantitative macroinvertebrate community index (QMCI). The values of both indices can be understood as surrogate measures of ecosystem health. The NPS-FM defines bands for both indices from A to D that indicate a scale from excellent to unacceptably poor ecosystem health. Death *et al.* (2018) and Canning *et al.* (2021) derived nutrient criteria that pertain to objectives that are expressed as threshold values for these two indices. The criteria were derived by fitting bivariate linear regression models that express the relationship between the indices and nutrient concentrations observed at river monitoring sites across New Zealand. The linear regressions were fitted to all the available data so did not attempt to account for differences in environmental conditions that exist across the national pool of monitoring sites. Because these models do not account for variation in environmental factors that likely influence the sensitivity of invertebrates to nutrient enrichment, the derived "national criteria" incur a risk of not achieving the desired outcomes. The risk extends in two directions, the national criteria may be under-protective in some systems and over-protective in others.

This study undertook two sets of statistical analyses to investigate the criteria derived by Death *et al.* (2018) and Canning *et al.* (2021). First, simple bivariate linear regression models were used to assess the strength of the relationships underlying the criteria and the precision of the derived criteria. Second, variance partitioning analysis was also used to quantify the strength of relationships between the invertebrate indices and nutrients while considering the extent to which these relations may be overestimated if other environmental variables are not accounted for.

The linear regression models indicated that there is, at best, a weak direct relationship between the invertebrate indices and nutrient concentrations. They also indicate that the relationships vary between different types of rivers. This means that national criteria derived using this approach are likely to be variously under-protective and over-protective. The nutrient concentration criteria defined from these weak relationships were imprecise. This means that a site that is compliant with a nutrient criterion to achieve a given NOF band can reasonably be expected to have actual MCI scores in multiple bands.

Variance partitioning analysis indicated that several environmental factors, including those associated with the segment (elevation, slope, distance to headwaters), and the catchment (climate, topography, geology) are much more strongly associated with variation in MCI and QMCI scores than nutrient concentrations. In addition, the variance partitioning analysis indicated that nutrient concentrations co-vary with environmental factors so that it is unclear whether the direct relationship between invertebrate indices and nutrient concentrations is caused by the nutrients or by other factors.

Together, the two sets of analyses indicate that the relationships from which Death et al.'s (2018) and Canning *et al.'s* (2021) proposed DIN and DRP thresholds are derived are likely confounded by other factors. This does not prove that nutrient concentrations have no effect on invertebrates, but it means that there is high risk that compliance with the nutrient criteria will not bring about the desired changes in ecosystem health. This is consistent with the understanding that environmental and biological factors mediate relationships between nutrients and ecosystem attributes. This study's results indicate that, when environmental factors are not accounted for, direct biology – nutrient relationships are characterised by high variance and low signal to noise ratios (MFE, 2022).

Considerable analytical effort is required to define robust criteria when signals are confounded by multiple environmental factors. For example, Biggs (1996) proposed that periphyton biomass dynamics are a function of the interaction between the supply of resources for periphyton growth (including nutrients, light and temperature) and disturbance that limits biomass accrual (primarily high flows and associated substrate instability, but also grazing). Many studies provide empirical support for this conceptual model of stream periphyton biomass (e.g., Biggs, 2000; Snelder *et al.*, 2019). However, despite the relatively simple conceptual model, and good empirical support, deriving robust nutrient concentration criteria for periphyton has involved the use of much more involved analyses than bivariate linear regression to adequately account for the influence of environmental factors (e.g., Snelder *et al.*, 2022, 2019). Because stream invertebrates are a higher trophic level than periphyton (i.e., many invertebrates feed on periphyton), their relationship to nutrient concentrations is almost certainly more complicated than that of periphyton biomass (e.g., Clapcott and Goodwin, 2014; Collier *et al.*, 2014). It is therefore logical to expect that deriving robust nutrient criteria to achieve ecosystem health outcomes that are defined by invertebrate indices will require at least as much, and probably more, analytical effort as that applied to periphyton.

Defining robust nutrient criteria to achieve ecosystem health outcomes that are defined by invertebrates is sufficiently complex that it is unlikely that significant progress will be made in the short to medium term (i.e., in less than five years). Therefore, in the medium term, it may be more appropriate to assume that ecosystem health objectives, as they pertain to nutrient concentrations, are best achieved by managing primary production (e.g., by managing nutrients to achieve appropriate periphyton outcomes). Data collection and research should continue with the long term aim of improving our confidence in nutrient concentration criteria to achieve ecosystem health outcomes.

1 Introduction

There is concern in New Zealand about impacts of anthropogenic enrichment of nutrients in rivers, streams, lakes, wetlands, estuaries, and aquifers (e.g., MFE & StatsNZ, 2017, 2019; PCE, 2013, 2015). Management of the impacts of nutrient emissions to these aquatic receiving environments is regulated by the Resource Management Act (RMA) and the National Policy Statement for Freshwater Management 2020 (NPS-FM; NZ Government, 2020). A requirement of the NPS-FM is the identification of appropriate nutrient concentration criteria (i.e., concentrations that will achieve ecological health objectives) and the implementation of limits on resource use to comply with these criteria via regional plans.

For rivers, the NPS-FM requires regional councils to determine the appropriate nutrient concentration criteria to “achieve a target attribute state for periphyton, any other nutrient attribute, and any attribute that is affected by nutrients” (NPSFM Section 3.13). Therefore, there have been efforts to provide regional councils with appropriate nutrient criteria, where ‘appropriate’ can be interpreted as criteria that (if complied with) are likely to achieve desired target attribute states. For example, Snelder *et al.* (2019) and Snelder *et al.* (2021) derived nutrient concentration criteria to achieve objectives for maximum periphyton biomass. Death *et al.* (2018) derived national nutrient concentration criteria to achieve ecological health objectives using a ‘multiple lines of evidence’ approach and Canning et al (2021) developed national nutrient criteria for macroinvertebrates. An important difference between criteria for achieving periphyton attribute states and the criteria of Death *et al.* (2018) and Canning *et al.* (2021) is that the former are spatially variable (i.e., they differ between environmentally defined river classes) whereas the latter are applied to all rivers nationally (hereafter, national criteria).

The national criteria of Death *et al.* (2018) were considered by the Freshwater Science and Technical Advisory Group (STAG) during development of the NPSFM (2020). STAG was convened to, among other things, consider whether criteria for dissolved inorganic nitrogen (DIN) and dissolved reactive phosphorus (DRP) in rivers should be mandated at the national level as NPS-FM attributes. The approach to defining criteria taken by Death *et al.* (2018) involved fitting linear regression models to all the available data without attempting to account for differences in environmental conditions that exist across the national pool of monitoring sites. There are three potential problems with this. First, variation in invertebrate indices can be expected due to variation in environmental conditions such as local habitat, flow regimes and light and temperature (e.g., Clapcott *et al.*, 2017). If natural variation in the indices is not accounted for, it could confound the assessed relationships with DIN and DRP. Second, the response of invertebrate indices to DIN and DRP can be expected to be mediated by environmental factors such as flow regimes and light and temperature. Therefore, the relationships can be expected to vary due to variation in environmental conditions. Third, the relationships are correlative and if variables other than DIN and DRP are among the causative agents but are not accounted for, any action to manage these nutrients will not produce the desired change in ecological health. Although Canning *et al.* (2021) did not use regression to define their criteria, the approach was also applied to the national pool of monitoring sites without attempting to account for variation in environmental conditions and therefore suffers from the same potential problems.

The approach taken by Death *et al.* (2018) was reviewed by the Ministry for the Environment (MFE) and the potential problems set out above concerning the lack of accounting for environmental variation across the sites were investigated (MFE, 2019). That review concluded that because environmental variation across the sites was not accounted for, the evidence supporting the proposed criteria was weak.

The aim of this report is to update the review undertaken by MFE using a more up to date dataset. In this study, ecological health is represented by the two invertebrate indices that are attributes in the NPS-FM: the macroinvertebrate community index (MCI) and the quantitative macroinvertebrate community index (QMCI). The present study has investigated the relationship between these two indices and concentrations of the dissolved forms of nitrogen and phosphorus. The findings are relevant to Regional Councils who are responsible for implementation of NPS-FM Section 3.13 and provide information that is relevant to the robust definition of instream concentration thresholds (ICT; MFE, 2022) to achieve NPSFM attribute states.

2 Data

2.1 Invertebrate and nutrient concentration data

Observations of two invertebrate indices: macro-invertebrate community index (MCI) and quantitative macro-invertebrate community index (QMCI), at long term monitoring sites across New Zealand were obtained from the Land Air Water Aotearoa (LAWA) database (**https://www.lawa.org.nz/**). These sites represent river monitoring carried out by regional and district councils and monitoring carried out by the National Institute of Water and Atmosphere (NIWA) associated with the National River Water Quality Network (NRWQN).

For each site for the period 2016 to 2020, the median values of MCI and QMCI, the number of observations and the number of unique years in which observations were made were extracted from this dataset. It is noted that the NPS-FM included Macroinvertebrate Average Score Per Metric (ASPM) in 2020 but these data were not available for all sites over the time period used by this study. The data was filtered to retain combinations of sites and indices with at least three annual observations (i.e., observations of MCI or QMCI in unique years within the period 2016 to 2020). Here after these data are referred to as the 'invertebrate indices dataset'.

Water quality data pertaining to long term monitoring sites across New Zealand were obtained from analyses prepared for the Ministry for the Environment and Statistics New Zealand by Whitehead *et al.* (2021). Observations of several water quality variables had been made at each site in this data set at either monthly or quarterly sampling intervals for the period from 2016 to 2020. The median values of observations of nitrate nitrite nitrogen (NNN) and dissolved reactive phosphorus (DRP) were extracted for all sites that passed the filtering rules used by Whitehead *et al.* (2021). NNN was used in this study because it is available at more sites than dissolved inorganic nitrogen (DIN; DIN = NNN + NH4-N). NNN and DIN are very similar because ammoniacal nitrogen is a small proportion of DIN in most rivers. The filtering rules define the acceptable proportion of missing observations (i.e., data gaps) and how these are distributed across sample intervals so that site grades are assessed from comparable data (e.g., Larned *et al.*, 2018). The filtering rules restricted site × variable combinations to those with measurements for at least 90% of the sampling intervals in the period (at least 56 of 60 months or 18 of 20 quarters).

The invertebrate indices dataset was joined to the water quality dataset based on the unique site identification code used by LAWA. This resulted in 496 and 403 sites with at least three annual observations of MCI and QMCI, respectively and site median values of NNN and DRP (Figure 1).

The sites were associated with a national digital river network that comprises 560,000 segments (defined by upstream and downstream confluences) with a mean length of ~700m (Snelder and Biggs, 2002). Every site was checked manually using metadata describing the site to ensure the association with the network was accurate (see Whitehead *et al.*, 2021 for details). The digital network is contained within a Geographic Information System (GIS) and each segment is associated with many independent variables including characteristics of the segment and its catchment and membership of river classification systems.

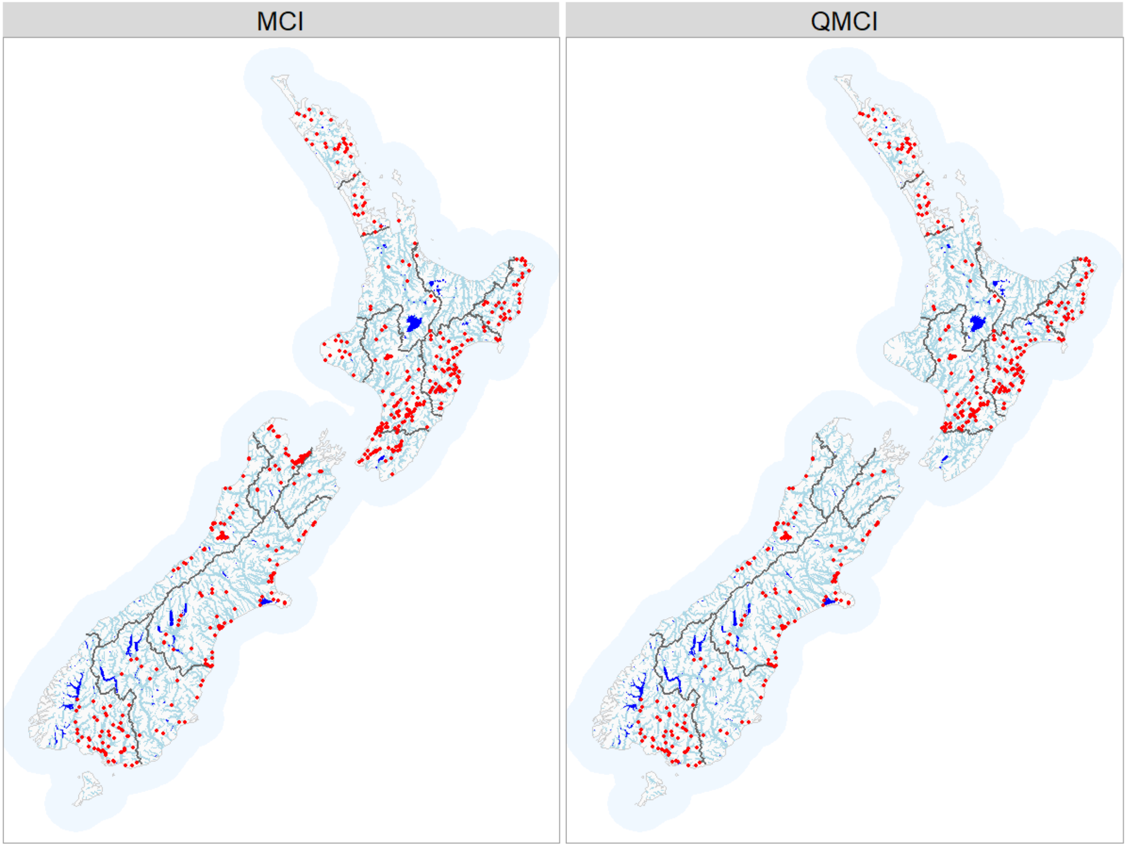

Figure 1. Location of the river water quality monitoring sites included in the study.

2.2 Classification of monitoring sites

The sites were classified using the national river classification of rivers provided by the River Environment Classification (REC; Snelder and Biggs, 2002). The REC is a classification that was developed to provide resource managers with a multi-level hierarchical classification of rivers, and a biophysical basis for catchment management (Pyle *et al.*, 2001). The first two levels of the REC are referred to collectively as Source-of-flow classes and define river classes based on differences in catchment climate and topography. REC Source-of-flow classes have been shown to broadly discriminate water quality, invertebrate communities, hydrology and river morphology (Snelder *et al.*, 2004, 2005).

All monitoring sites were allocated to REC Source-of-flow classes based on their location on the digital river network. Some Source-of-flow classes had poor representation within the monitoring network (i.e., < 10). Some REC classes with poor representation were aggregated into the class that was closest in environmental terms (Table 1). Hereafter the aggregated REC Source-of-flow classes, as defined by Table 1, are referred to as 'REC classes'.

Table 1. REC classes used in the analysis.

REC class	Aggregated REC Source-of-flow classes	Description – catchment dominated by:
WW/L	WW/L, WW/H, WW/Lk, WX/L	Warm wet lowlands, including extremely wet lowland and lake-fed
WD/L	WD/L, WD/Lk	Warm dry lowlands, including lake-fed
CW/H	CW/H	Cool wet hills
CX/H	CX/H, CX/M	Cool extremely wet hills and mountains
CW/L	CW/L, CW/Lk	Cool wet lowland and lake fed
CD/H	CD/H, CD/M	Cool dry hills and mountains
CW/M	CW/M	Cool wet mountains
CD/L	CD/L, CD/Lk	Cool dry lowlands and lakes
CX/L	CX/L, CW/Lk	Cool extremely wet lowlands and lake fed

2.3 Additional environmental explanatory variables for sites

Additional environmental descriptors for each monitoring site were obtained from the freshwater environments of New Zealand (FENZ) database. Each segment is associated with many descriptors that were derived by intersecting the network with other spatial layers as described by Wild *et al.* (2005).

Two groups of environmental explanatory variables in the FENZ database were chosen to represent the character of the segment and the upstream catchment of each monitoring site. These additional explanatory variables were derived from spatial layers including a terrain model, mapped climate data, and geological maps. The variables were chosen based on previous analyses that have demonstrated their association with ecological characteristics (e.g., Leathwick *et al.*, 2011). The additional explanatory variables included catchment climate (usAvTWarm and usRainDays10), catchment topography (usAveSlope and usLake) and the character of the catchment surface geology (usHard). Three variables were chosen to represent the characteristics of the segment including its position in the network as indicated by the distance to the headwater (HeadwaterDistance), its elevation (segAveElev), and slope (segSlope).

Table 2. Additional environmental explanatory variables. See Wild et al., 2005 for details.

Type of variable	Variable name	Description (units)
Catchment variables	usAvTWarm	Mean January air temperature (^oC x 10)
	usRainDays10	Catchment rain days (greater than 10mm/month) (days/month)
	usAveSlope	Average slope of catchment calculated from 30m DEM grid (m/m)
	usLake	Lake index (dimensionless)
	usHard	Catchment average of hardness (induration) of surface geology (ordinal)
	usPhos	Catchment average of phosphorous in surface geology (ordinal)
Segment variables	segAveElev	Average segment elevation (m. asl)
	HeadwaterDistance	Distance to the headwater (m)
	segSlope	Average segment slope (m/m)

3 Methods

3.1 Assessment of relationships between invertebrate indices and nutrients at the national scale

Linear regression was used to quantify the relationship of NNN and DRP with the invertebrate indices (i.e., MCI and QMCI) for the whole dataset (national scale). The values of both nutrients were log transformed to normalise the regression model residuals and achieve a better model fit. Transformation of the explanatory variables (NNN and DRP) was consistent with Death *et al.* (2018). It is noted that Death *et al.* (2018) fitted models with both log transformed and untransformed response variables (i.e., MCI and QMCI). In this study, transformation of MCI and QMCI made little difference to the model or its performance and was not therefore performed.

The performance of the models was evaluated using the coefficient of determination (R^2 value) and the overall significance was determined by the model *p*-value. The R^2 values indicate the goodness-of-fit of the linear regression models. R^2 values are interpreted as how strongly the invertebrate index is associated with the nutrient concentration. Low R^2 values (e.g., <20%) indicate that a wide range of the invertebrate index can occur at a given nutrient concentration. Low *p*-values (e.g., $p < 0.05$) are interpreted as strong evidence there is a relationship (i.e., a correlation) between nutrient and the invertebrate index in the population (i.e., that the signal obtained from the sample would be unlikely if there were not a relationship). It is important to consider the significance (i.e., the *p*-value) alongside the R^2 value because when sample size is large, a significant result can be obtained even if the R^2 value is small.

Criteria corresponding to three NPS-FM attribute states for MCI and QMCI were derived as nutrient concentrations that are consistent with the lower thresholds for bands A, B and C defined by the NPS-FM (Table 3). The criteria were derived by inverting the models to obtain the nutrient concentrations that were associated with values of the index corresponding to the lower thresholds for bands A, B and C. This is consistent with the method used by Death *et al.* (2018).

The precision of predictions obtained from the fitted models were quantified by the 90% prediction interval. Wide prediction intervals indicate that values of an index, predicted by the model for a given nutrient concentration, cover a wide range of values and that nutrient concentration is therefore a poor predictor of the index. Poor predictions means that there is high risk that the desired outcomes (i.e., the MCI and QMCI attribute states) will not be achieved, despite complying with the nutrient criteria. To demonstrate the precision of the models, the fitted linear regression models were used to estimate the 90% prediction interval for the two invertebrate indices for concentrations of NNN and DRP that were consistent with the derived criteria.

Table 3. Lower thresholds for NPS-FM attribute state bands defined by macroinvertebrate indices.

Macroinvertebrate index	Lower thresholds for NPS-FM attribute state bands		
	A	B	C
MCI	130	110	90
QMCI	6.5	5.5	4.5

3.2 Assessment of relationship between invertebrate indices and nutrients within river classes

Linear regression models were used to quantify the relationship of NNN and DRP with the invertebrate indices (i.e., MCI and QMCI) for data pertaining to classes defined by the REC. This analysis had two objectives. First, to investigate whether different environmental conditions result in different relationships between invertebrate indices and nutrients. Second to assess whether a restricted set of environmental conditions, as defined by REC classes, increases the precision of predictions made by the models. Only REC classes that had at least eight representative sites were analysed. Model performance, 90% prediction intervals, and the best estimates of criteria corresponding to the three NPS-FM attribute states for MCI and QMCI (Table 3) were obtained as for the national scale models.

3.3 Variance partitioning analysis

Nutrient concentrations generally increase in the downstream direction within catchments in response to the increasing influence in upstream land use (Larned *et al.*, 2018). However, many other variables that influence invertebrates and invertebrate indices also vary longitudinally (e.g., local habitat, flow regime and temperature). Many of these variables are correlated because they share strong hierarchical relationships and because they tend to vary monotonically as a function of position in the river network (e.g., Poff, 1997; Vannote *et al.*, 1980). Therefore, several habitat variables co-vary with nutrient concentrations measured at monitoring sites. Correlation between these environmental variables may lead to overestimating the strength of relationships between nutrients and invertebrate indices if covariance is not accounted for. To avoid overestimating the strength of a relationship, it is good practice to account for co-variance between the explanatory variable of interest and other potential explanatory variables (Borcard *et al.*, 1992; Legendre and Legendre, 1998).

Variance partitioning analysis (Borcard *et al.*, 1992; Legendre and Legendre, 1998) was used to quantify the strength of relationships between the invertebrate indices and nutrients while considering the extent to which these relations may be overestimated if other environmental variables are not accounted for. The analysis included a minimal set of additional environmental variables, other than nutrients, that may explain variation in the invertebrate indices. The additional environmental variables included those describing characteristics of the upstream catchment and characteristics of the river segment shown in Table 2.

The analysis used a procedure that is based on multiple linear regression to partition the total explained variation in the invertebrate indices into 15 components that included the individual, shared and unique contributions of the three sets of variables representing the factors: nutrient concentrations, catchment variables and segment variables (Borcard *et al.*, 1992). The significance of all components was tested using permutation tests. The significance of the unique fractions was tested by running the other set of variables 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. The method of Peres-Neto *et al.* (2006) was used to adjust the estimate of variation explained (i.e., R^2) by each set of variables to make valid comparisons between sets of variables of differing size. All analyses and variance partitioning were performed in R using the ‘vegan’ package (R Core Team, 2021).

4 Results

4.1 Relationships of NNN and DRP with invertebrate indices at the national scale

Relationships at the national scale between the $\log_{10}$ transformed nutrients and the MCI index indices were significant (p <0.05) but weak (R^2 < 15%; Figure 2). Nutrient criteria were not defined where the threshold values were outside the range of the fitting data (i.e., where the regression lines shown in Figure 2 do not intersect the thresholds indicated by the red horizontal lines). The 90% prediction interval associated with the defined NNN and DRP criteria covered a wide range in MCI values (Table 4). For example, the C-band criteria for MCI (i.e., to achieve an MCI score of 90) ranged from 61 to 119 and 62 to 118 for NNN and DRP, respectively. This indicates that a site that is compliant with a nutrient criterion to achieve an MCI C-band can reasonably be expected to have actual MCI scores in the B, C or D bands (Table 3).

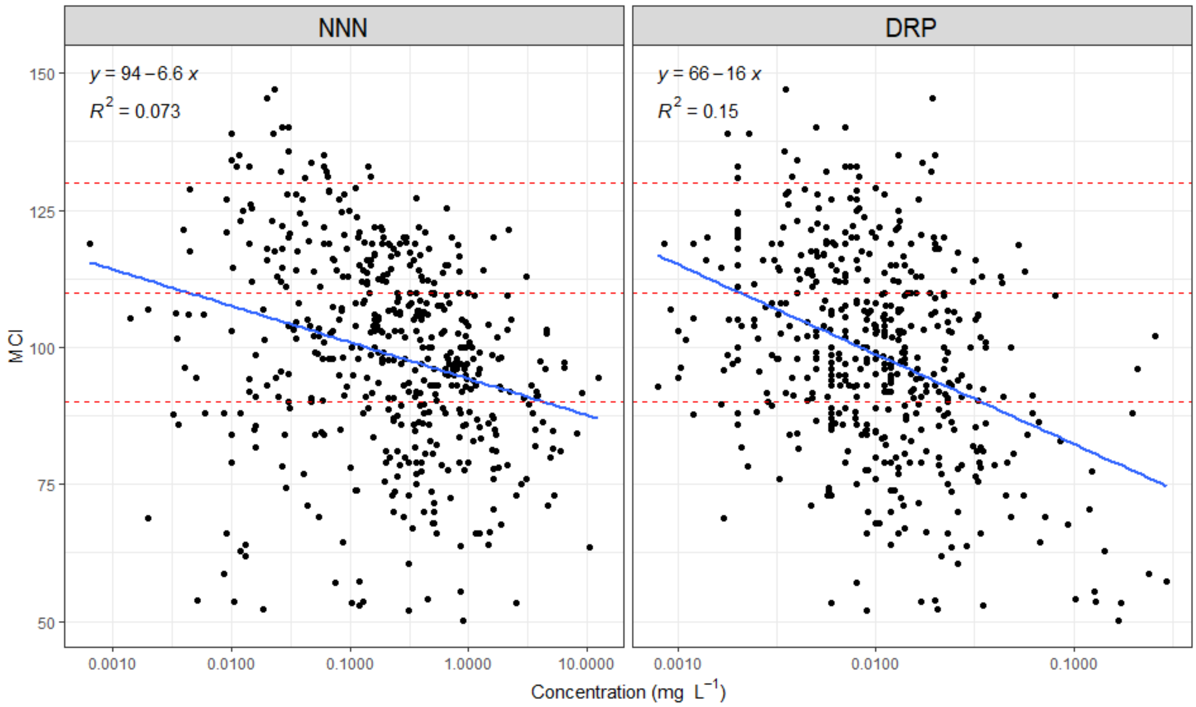

Figure 2. Relationships of site median values of NNN and DRP with site median MCI at national scale. The blue lines represent linear regression models of the invertebrate indices against the nutrient concentration. Note that the x-axis (nutrient concentration) is log (base 10) scale. The red lines indicate MCI scores of 90, 100 and 130, which are the thresholds for the NPS-FM A, B and C bands for the MCI attribute.

Relationships at the national scale between the $\log_{10}$ transformed nutrients and the QMCI index indices were significant ($p < 0.05$) but weak ($R^2 < 20\%$; Figure 3). Consequently, the 90% prediction interval associated with the NNN and DRP criteria covered a wide range in QMCI values (Table 5). For example, the C-band criteria for QMCI (i.e., to achieve a QMCI score of 4.5) ranged from 2.2 to 6.8 and 2.3 to 6.7 for NNN and DRP, respectively. This indicates that a site that is compliant with a nutrient criterion to achieve the QMCI C-band can reasonably be expected to have actual QMCI scores in the A, B, C or D bands (Table 3).

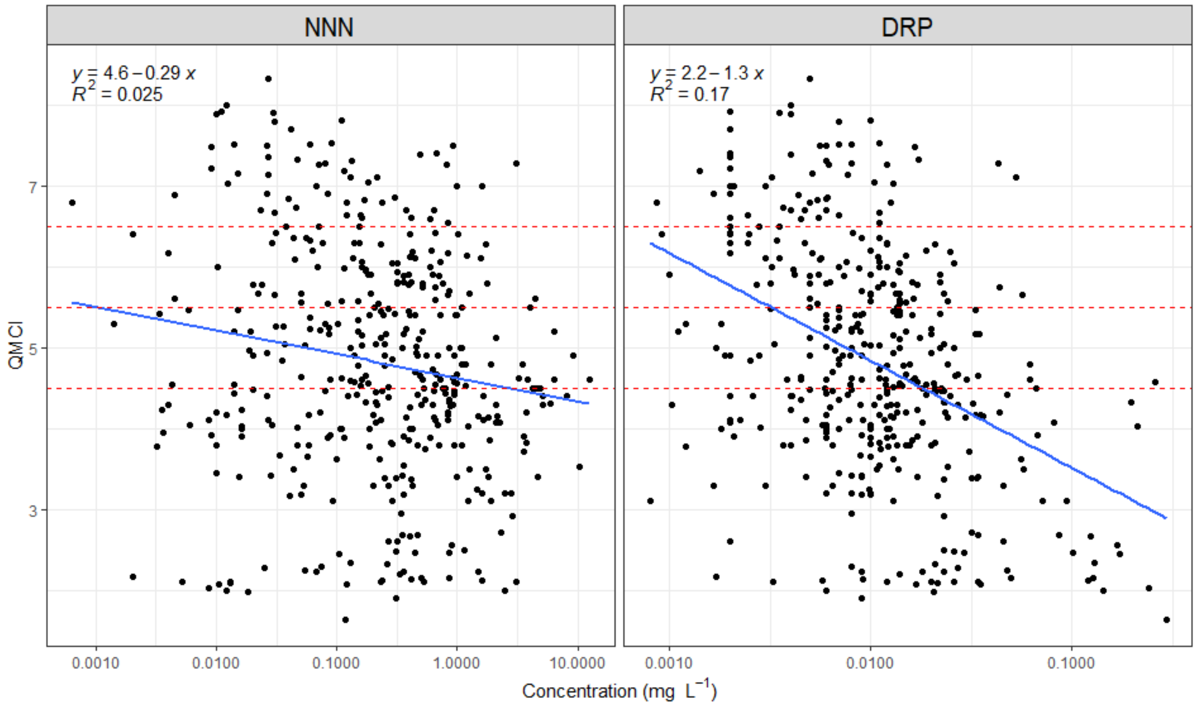

Figure 3. Relationships of site median values of NNN and DRP with site median QMCI at national scale. The blue lines represent linear regression models of the invertebrate indices against the nutrient concentration. Note that the x-axis (nutrient concentration) is log (base 10) scale. The red lines indicate nominated QMCI scores of 4.5, 5.5 and 6.5, which are the thresholds for the NPS-FM A, B and C bands for the QMCI attribute.

Table 4. Results of linear regression modelling of MCI as a function of the nutrient concentration at the national scale. Significant p-values are shown as bold text. The nutrient criteria are derived from the models for the lower thresholds for MCI A, B and C bands (Table 3). The 90% prediction intervals pertain to the derived A, B and C band nutrient criteria. NA values indicate the nutrient criteria could not be evaluated because it was outside the range of the fitting data.

Nutrient	N	Intercept	Coefficient	R^2 (%)	*p*-value	Criteria			90% Prediction interval		
						A	B	C	A	B	C
NNN	496	94.3	-6.64	7	**<0.001**	NA	0.004	4.399	NA	80 - 140	61 - 119
DRP	496	66	-16.36	15	**<0.001**	NA	0.002	0.034	NA	82 - 138	62 - 118

Table 5. Results of linear regression modelling of QMCI as a function of the nutrient concentration at the national scale. See caption of Table 4 for details.

Nutrient	N	Intercept	Coefficient	R^2 (%)	*p*-value	Criteria			90% Prediction interval		
						A	B	C	A	B	C
NNN	403	4.6	-0.29	2	**<0.001**	NA	0.001	2.705	NA	3.1 - 7.9	2.2 - 6.8
DRP	403	2.2	-1.33	17	**<0.001**	NA	0.003	0.018	NA	3.3 - 7.7	2.3 - 6.7

4.2 Relationship of NNN and DRP with invertebrate indices within REC classes

Relationships between the $\log_{10}$ transformed nutrients and the MCI index within REC classes were sometimes significant ($p < 0.05$) but generally weak ($R^2 \leq 44\%$ across all models and <20% for 14 of the 18 models; Figure 4). Nutrient criteria were not defined where the threshold index value was outside the range of the fitting data (i.e., where the regression lines shown in Figure 4 do not intersect the MCI thresholds indicated by the red horizontal lines). Where criteria could be derived for REC classes (Table 6), they varied between classes and were appreciably different to the criteria derived from the national scale model (Table 4).

The 90% prediction intervals associated with the derived criteria covered a wide range in MCI values (Table 6). This indicates that a site that is compliant with a nutrient criterion to achieve a given NOF band can reasonably be expected to have actual MCI scores in multiple bands. For one class (WD/L), the relationship with NNN was statistically significant but with positive coefficients, indicating that larger MCI indices were associated with higher NNN concentrations.

Relationships between the $\log_{10}$ transformed nutrients and the QMCI index within REC classes indices were sometimes significant ($p < 0.05$) but generally weak ($R^2 < 45\%$ across all models and ≤ 20% for 16 of the 18 models; Figure 5). Nutrient criteria were not defined where the threshold QMCI value was outside the range of the fitting data (i.e., where the regression lines shown in Figure 5 do not intersect the thresholds indicated by the red horizontal lines). Where criteria could be derived for REC classes (Table 7), they varied between classes and were appreciably different to the criteria derived from the national scale model (Table 5). The 90% prediction interval associated with the NNN and DRP criteria covered a wide range in QMCI values (Table 7).

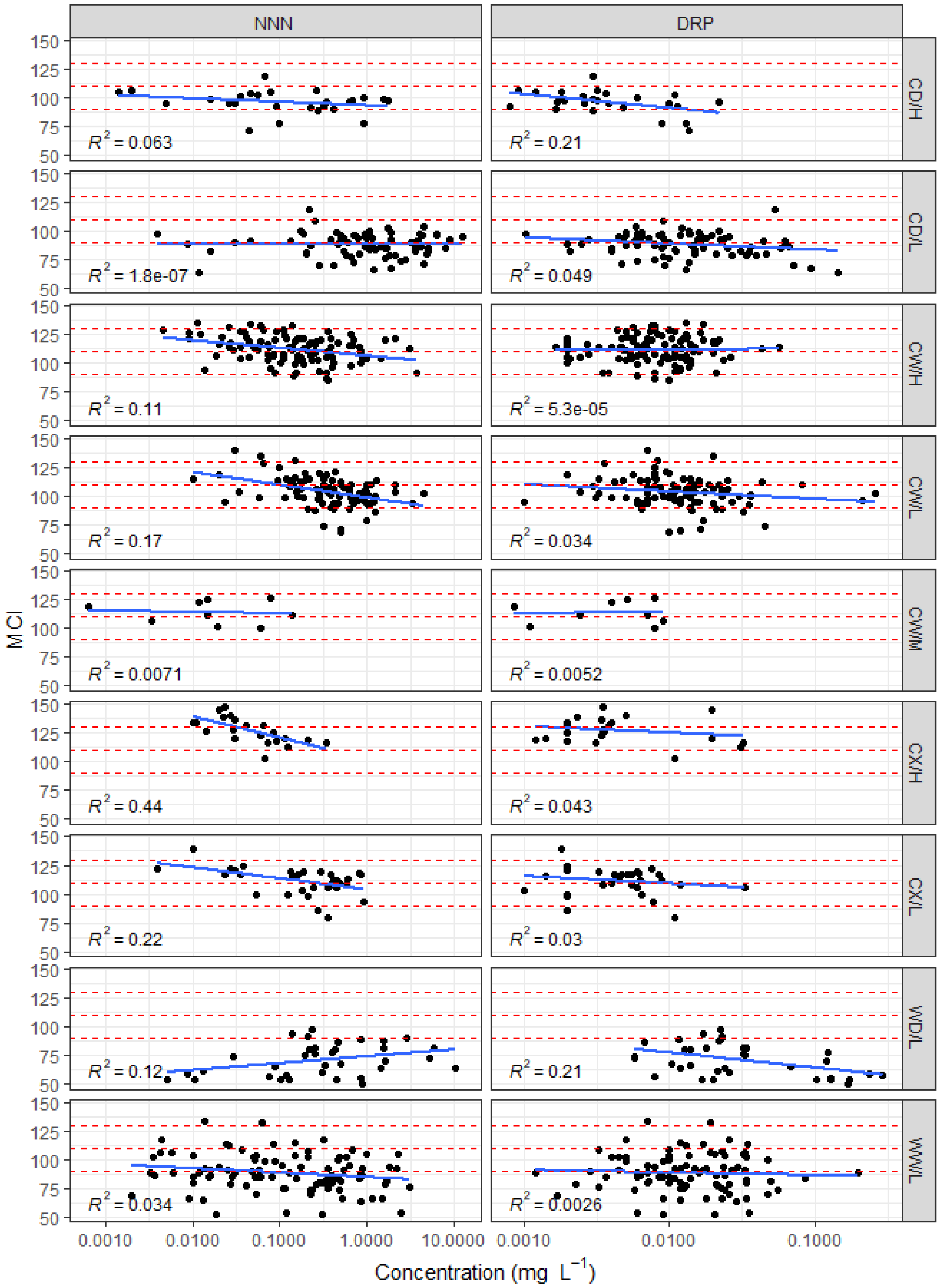

Figure 4. Relationships of site median values of NNN and DRP with site median MCI within REC classes. Note that class WX/H has been excluded because it has fewer than eight sites. See caption of Figure 2 for details.

Table 6. Results of linear regression modelling of MCI as a function of the nutrient concentration within REC classes. See caption of Table 4 for details. Note that class WX/H was excluded because it had fewer than eight sites.

Nutrient	**REC Class**	**N**	**Intercept**	**Coefficient**	**R^2 (%)**	***p*-value**	**Criteria**			**90% Prediction interval**		
							A	**B**	**C**	**A**	**B**	**C**
	CD/H	26	93.2	-3	6	0.22	NA	NA	NA	NA	NA	NA
	CD/L	78	87.8	-0.01	0	1	NA	NA	NA	NA	NA	NA
	CW/H	103	106.4	-6.65	11	**<0.001**	NA	0.287	NA	NA	92 - 128	NA
	CW/L	94	98.4	-11.02	17	**<0.001**	NA	0.09	NA	NA	90 - 130	NA
NNN	CW/M	9	111.9	-1.18	1	0.83	NA	NA	NA	NA	NA	NA
	CX/H	21	102.1	-18.4	44	**<0.001**	0.03	NA	NA	114 - 146	NA	NA
	CX/L	33	103.9	-9.51	22	0.01	NA	0.226	NA	NA	92 - 128	NA
	WD/L	35	74.2	6.01	12	**0.04**	NA	NA	NA	NA	NA	NA
	WW/L	94	84.4	-3.8	3	0.07	NA	NA	0.034	NA	NA	63 - 117
	CD/H	26	67.2	-11.9	21	**0.02**	NA	NA	0.012	NA	NA	74 - 106
	CD/L	78	77.8	-5.3	5	**0.05**	NA	NA	0.005	NA	NA	73 - 107
	CW/H	103	112.5	0.27	0	0.94	NA	NA	NA	NA	NA	NA
	CW/L	94	90.9	-6.34	3	0.08	NA	NA	NA	NA	NA	NA
DRP	CW/M	9	118.6	1.88	1	0.85	NA	NA	NA	NA	NA	NA
	CX/H	21	113.7	-5.5	4	0.37	NA	NA	NA	NA	NA	NA
	CX/L	33	96	-6.47	3	0.34	NA	0.007	NA	NA	90 - 130	NA
	WD/L	35	51.2	-13.07	21	**0.01**	NA	NA	NA	NA	NA	NA
	WW/L	94	83.9	-2.22	0	0.62	NA	NA	0.002	NA	NA	62 - 118

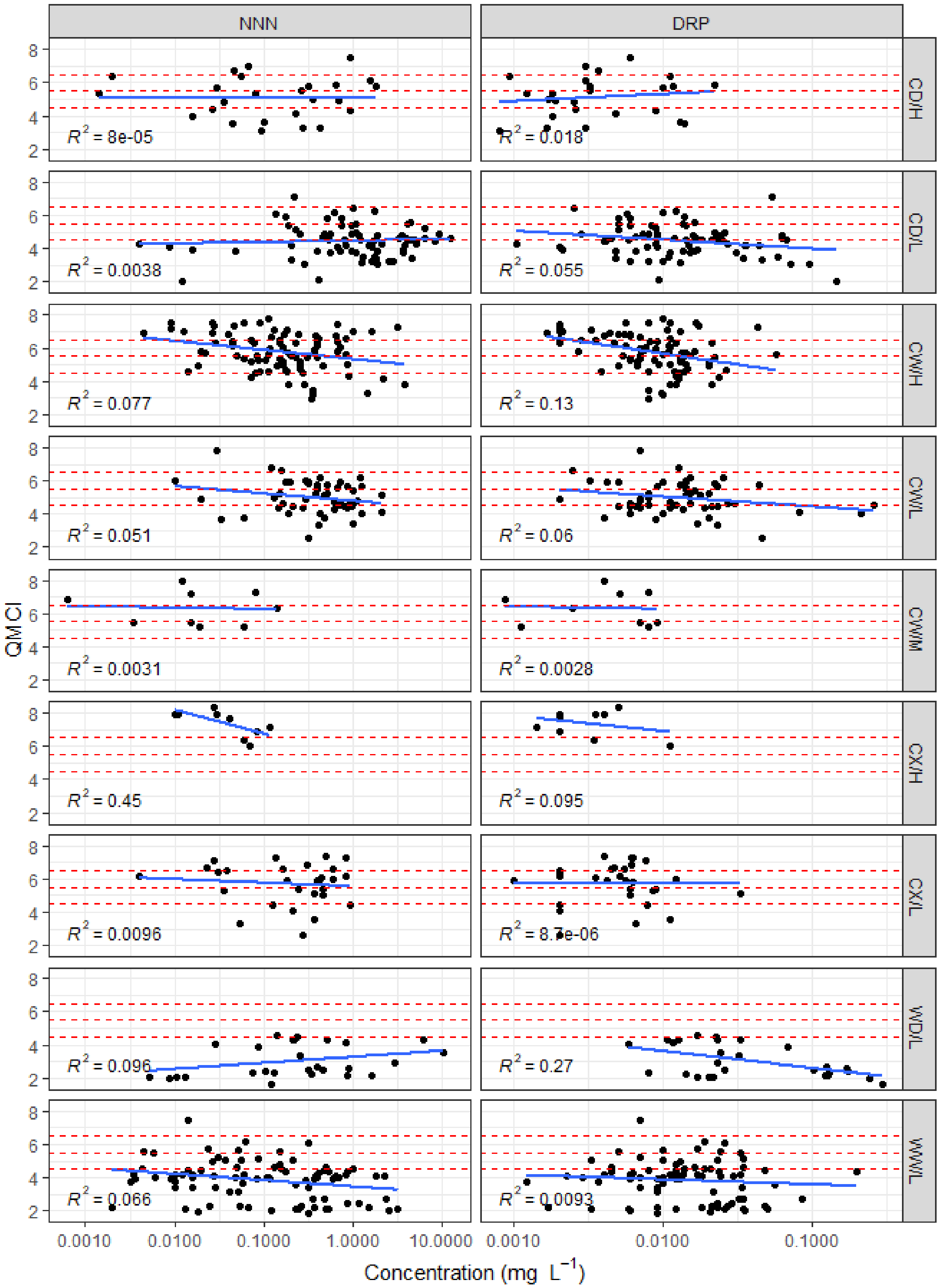

Figure 5. Relationships of site median values of NNN and DRP with site median QMCI within REC classes. Note that class WX/H has been excluded because it had only three sites. See caption of Figure 2 for details.

Table 7. Results of linear modelling of QMCI as a function of nutrient concentration within REC classes. Note that class WX/H was excluded because it had only three sites. See caption of Table 4 for details.

Nutrient	REC Class	N	Intercept	Coefficient	R^2 (%)	*p*-value	Criteria			90% Prediction interval		
							A	B	C	A	B	C
NNN	CD/H	25	5.1	0.01	0	0.97	NA	NA	NA	NA	NA	NA
	CD/L	74	4.5	0.09	0	0.6	NA	NA	2.945	NA	NA	2.9 - 6.1
	CW/H	86	5.3	-0.55	8	**0.01**	0.007	0.467	NA	4.6 - 8.4	3.6 - 7.4	NA
	CW/L	56	4.8	-0.44	5	0.09	NA	0.024	NA	NA	3.8 - 7.2	NA
	CW/M	9	6.2	-0.08	0	0.89	NA	NA	NA	NA	NA	NA
	CX/H	9	5.3	-1.43	45	**0.05**	NA	NA	NA	NA	NA	NA
	CX/L	29	5.5	-0.21	1	0.61	NA	NA	NA	NA	NA	NA
	WD/L	27	3.3	0.37	10	0.12	NA	NA	NA	NA	NA	NA
	WW/L	85	3.5	-0.38	7	**0.02**	NA	NA	NA	NA	NA	NA
DRP	CD/H	25	6.1	0.43	2	0.52	NA	NA	NA	NA	NA	NA
	CD/L	74	3.4	-0.54	5	**0.04**	NA	NA	0.01	NA	NA	2.9 - 6.1
	CW/H	86	3	-1.32	13	**<0.001**	0.002	0.013	NA	4.7 - 8.3	3.7 - 7.3	NA
	CW/L	56	3.8	-0.59	6	0.07	NA	NA	0.078	NA	NA	2.9 - 6.1
	CW/M	9	6	-0.14	0	0.89	NA	NA	NA	NA	NA	NA
	CX/H	9	5.1	-0.91	9	0.42	NA	NA	NA	NA	NA	NA
	CX/L	29	5.7	0.01	0	0.99	NA	NA	NA	NA	NA	NA
	WD/L	27	1.6	-1.02	27	**0.01**	NA	NA	NA	NA	NA	NA
	WW/L	85	3.3	-0.29	1	0.38	NA	NA	NA	NA	NA	NA

4.3 Partitioning of explained variance for MCI

The results of the variance partitioning analysis performed using the national MCI data are shown graphically on Figure 6. The outer box represents the total variation in the site MCI scores. The Venn-diagram within the box represents the total explained variation (R^2 = 55%) in MCI scores (residual unexplained variation = 45%). The variation explained by each group of variables is represented by the sum of the values lying within each of the three circles that represent the nutrient, catchment and segment variable groups.

The unique contribution of each of the groups of variables is shown by the parts of the circles that do not overlap with the other circles (labelled Nutrients, Segment and Catchment). The explained variation that is shared is shown by the intersection areas of the three circles (labelled Nutrients & Segment, Nutrients & Catchment, Segment & Catchment, and Nutrients & Segment & Catchment).

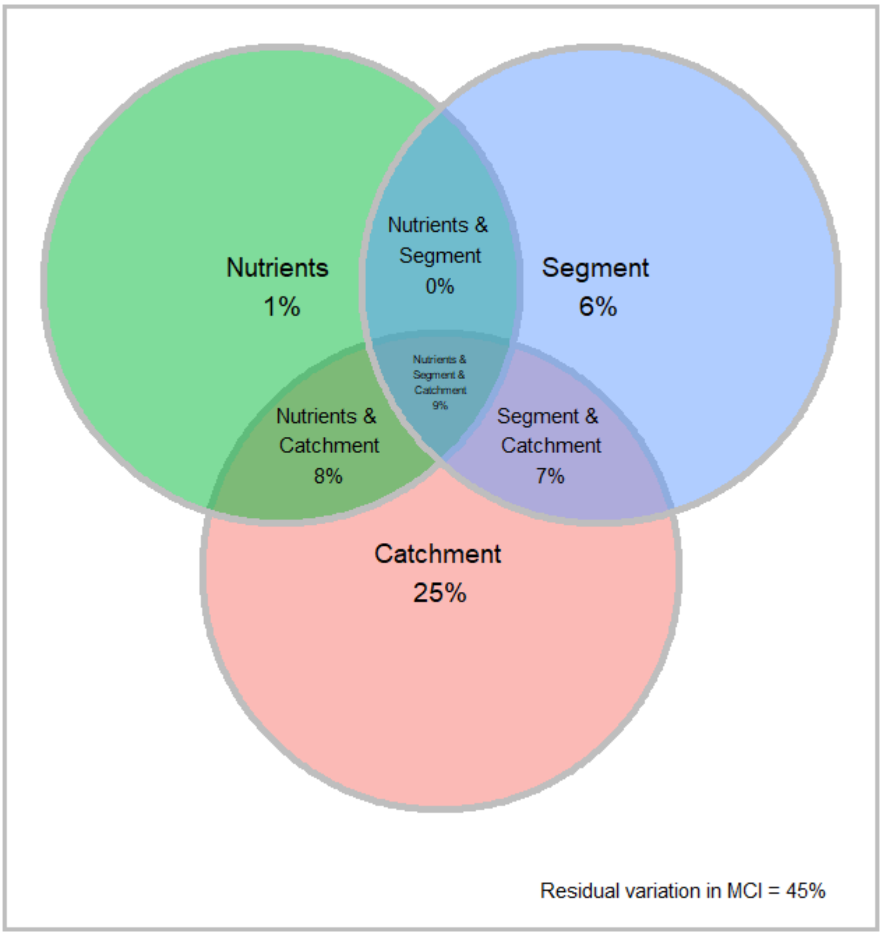

Figure 6. Schematic diagram of all components of variation in MCI indices provided by the variance partitioning.

Nutrients (i.e., the combination of NNN and DRP) individually explained 18% of the variation (i.e., Nutrients, Nutrients & Segment, Nutrients & Catchment and Nutrients & Segment & Catchment in Figure 6). The segment and catchment variables individually explained 20% and 49% of the variation respectively. Nutrients uniquely explained 1% of the variation in MCI scores (i.e., Nutrients in Figure 6). This means that, after accounting for variation explained by the segment and catchment variables, nutrients explained a further 1% of the variation. Segment and catchment uniquely explained 6% and 25% of the variation respectively.

Permutation tests indicated that all components of variation that were testable were statistically significant ($p < 0.001$).

4.4 Partitioning of explained variance for QMCI

The results of the variance partitioning analysis performed using the national QMCI data are shown graphically on Figure 7. The outer box represents the total variation in the site QMCI scores. The Venn-diagram within the box represents the total explained variation (R^2 = 44%) in QMCI scores (residual unexplained variation = 56%). Figure 7 has the same interpretation as Figure 6.

Nutrients (i.e., the combination of NNN and DRP) individually explained 17% of the variation (i.e., Nutrients, Nutrients & Segment, Nutrients & Catchment and Nutrients & Segment & Catchment in Figure 7. The segment and catchment variables individually explained 13% and 41% of the variation respectively. Nutrients uniquely explained 2% of the variation in QMCI scores (i.e., Nutrients in Figure 7). This means that, after accounting for variation explained by the segment and catchment variables, nutrients explained a further 2% of the variation. Segment and catchment uniquely explained 2% and 21% of the variation respectively. Permutation tests indicated that all components of variation that were testable were statistically significant ($p < 0.001$).

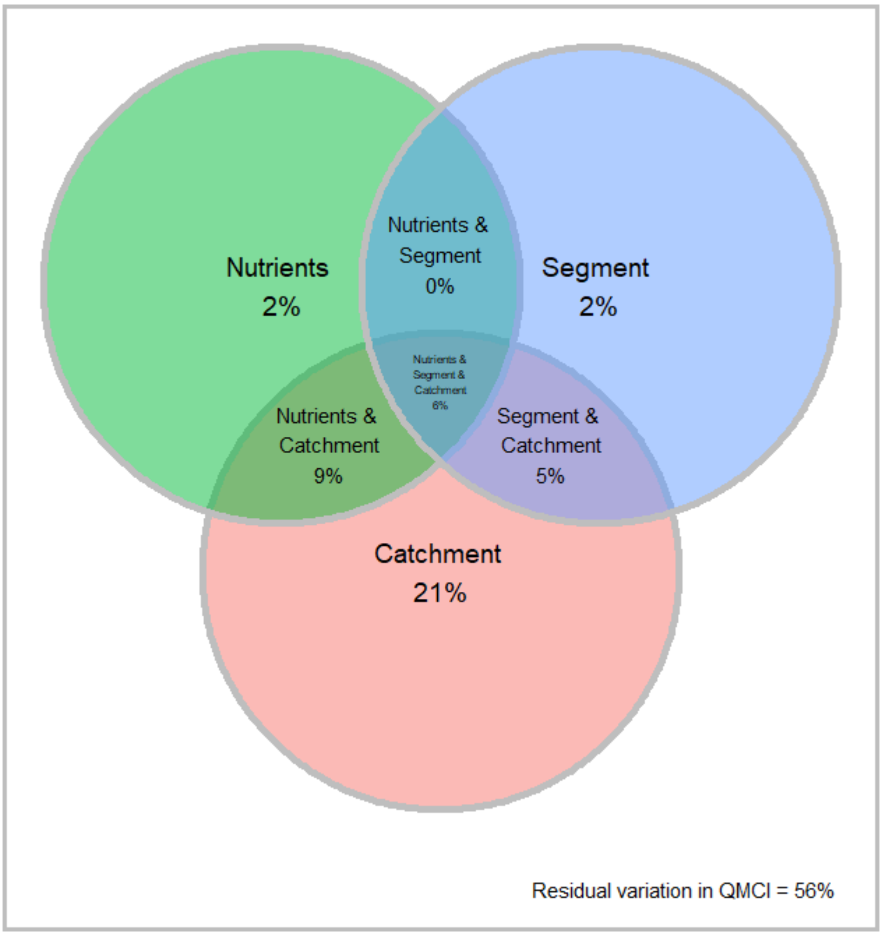

Figure 7. Schematic diagram of all components of variation in QMCI indices provided by the variance partitioning.

5 Discussion

Linear regression models relating invertebrate communities (MCI and QMCI) to nutrient concentrations (DIN and DRP), without accounting for differences in environmental conditions that exist across the national pool of monitoring sites, had low R^2 and low precision. Linear regression models relating MCI and QMCI to concentrations of DIN and DRP within REC classes were not statistically significant for some classes. In addition, when relationships were significant, the criteria derived for different REC classes differed substantially between classes. These results indicate that there is, at best, a weak direct relationship between the invertebrate indices and nutrient concentrations. The varying relationships between different types of rivers means that there is substantial risk that NNN and DRP criteria derived from analyses that do not account for variation in environmental conditions will be inappropriate in some systems. In other words, national nutrient concentration criteria incur a risk of not achieving the desired outcomes. The risk extends in two directions, the national criteria may be under-protective in some systems and over-protective in others.

Variance partitioning showed that associations between nutrients and MCI and QMCI are very weak when the effect of other variables is accounted for. In addition, the variance partitioning analysis indicated that nutrient concentrations co-vary with environmental factors so that it is unclear whether the direct relationship between MCI and QMCI and nutrient concentrations is caused by the nutrients or other factors. However, the analyses indicate that variation in MCI and QMCI is reasonably well explained by models that include environmental variables as well as nutrient concentrations. This finding is consistent with the generally accepted conceptualisation of ecosystems as the outcomes of biological organisms interacting with a wide range of attributes of the physical environment (e.g., Clapcott *et al.*, 2018).

The variance partitioning results were consistent with the linear regression modelling results indicating that for a given nutrient concentration, MCI and QMCI vary widely. The relevant explanation is that there are multiple drivers of variation in invertebrate communities. Therefore, for any given nutrient concentration, there exists a wide range of potential MCI and QMCI states, which are determined by the environmental context. In other words, environmental variation (i.e., catchment and segment factors) is a stronger cause of the observed variation in MCI and QMCI than nutrient concentration.

Together, the results of both sets of analyses undertaken by this study indicate that the relationships from which Death et al. (2018) and Canning *et al.* (2021) derived nutrient criteria are confounded by other factors. This does not prove that nutrient concentrations have no effect on invertebrate communities, however it reduces confidence that achieving the criteria will bring about the desired changes in MCI and QMCI and ecosystem health.

In the earlier study carried out for MFE, there was general agreement between the variance partitioning analyses performed using the national dataset (i.e., regional council and NRWQN sites) and just the NRWQN dataset. That result is evidence that the analyses are not compromised or confounded by use of data collected by multiple agencies.

Canning *et al.* (2021) derived national criteria and suggested that it was not necessary to develop “eco-regionalised nutrient criteria” (i.e., criteria that vary spatially) because New Zealand rivers share a common core assemblage of macroinvertebrates and also because a New Zealand macroinvertebrate eco-region is considerably smaller than those in other countries. These statements are not supported by the evidence provided by this study. The general expectation that the biological community response to a stressor is mediated by the environment is independent of whether the community comprises a common core assemblage of macroinvertebrates or not. In addition, a comparison of sizes of New Zealand ecoregions to elsewhere in the world is not a robust basis for deciding on the need for ecoregions for two reasons. First, the characteristic size of ecoregions is a cartographic choice that reflects the objectives and spatial resolution of the regionalisation. Second, given the objective of a regionalisation, the characteristic size of the individual regions is determined by the rates of change in space of the defining environmental characteristics (environmental gradients). New Zealand has high rates of change in space of many environmental characteristics that drive ecoregional variation (for example steep gradients in climatic, topographic and geological variation). This means that, ecoregions with the same objectives will be characteristically smaller in New Zealand compared to (for example) continental areas where environmental gradients are less steep and consequently, ecoregions are large.

It is broadly understood that environmental and biological factors mediate relationships between nutrients and ecosystem attributes. For example, Biggs (1996) proposed that periphyton biomass dynamics are a function of the interaction between the supply of resources for periphyton growth (including nutrients, light and temperature) and disturbance that limits biomass accrual (primarily high flows and associated substrate instability but also grazing). Many studies provide empirical support for this conceptual model of stream periphyton biomass (e.g., Biggs, 2000; Snelder *et al.*, 2019). However, despite the relatively simple conceptual model, and good empirical support, deriving robust nutrient concentration criteria for periphyton has involved the use of much more involved analyses than bivariate linear regression to adequately account for the influence of environmental factors (e.g., Snelder *et al.*, 2022, 2019).

This study's results indicate that, when environmental factors are not accounted for, direct biology – nutrient relationships are characterised by high variance and low signal to noise ratios (MFE, 2022). As shown by the efforts required to define nutrient criteria to achieve periphyton objectives, considerable analytical effort is required to define robust criteria when signals are confounded by multiple environmental factors. Because stream invertebrates are a higher trophic level than periphyton (i.e., many invertebrates feed on periphyton), their relationship to nutrient concentrations is almost certainly more complicated than that of periphyton biomass. Conceptual models of the determinants of invertebrate communities in stream and rivers are invariably complex (e.g., Clapcott and Goodwin, 2014; Collier *et al.*, 2014). It is therefore logical to expect that deriving robust nutrient criteria to achieve ecosystem health outcomes that are defined by invertebrate indices will require at least as much, and probably more, analytical effort as that applied to periphyton.

Defining robust nutrient criteria to achieve ecosystem health outcomes that are defined by invertebrates is sufficiently complex that it is unlikely that significant progress will be made in the short to medium term (i.e., in less than five years). Therefore, in the medium term, it may be more appropriate to assume that ecosystem health objectives, as they pertain to nutrient concentrations, are best achieved by managing primary production (e.g., by managing nutrients to achieve appropriate periphyton outcomes). Data collection and research should continue with the long term aim of improving our confidence in nutrient concentration criteria to achieve ecosystem health outcomes.

6 Acknowledgements

Thanks are due to Mike Scarsbrook who provided review comments on an early draft of this report and to Amy Whitehead and Caroline Fraser for provision of the water quality data. I also thank Roger Hodson and Duncan Gray for insightful reviews that improved this report. My thanks are also due to the many regional council staff whose hard work provided the long-term water quality and invertebrate monitoring data.

7 References

Biggs, B.J.F., 1996. Hydraulic Habitat of Plants in Streams. Regulated Rivers: Research and Management 12:131–144.

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

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

Canning, A.D., M.K. Joy, and R.G. Death, 2021. Nutrient Criteria to Achieve New Zealand's Riverine Macroinvertebrate Targets. PeerJ 9:e11556.

Clapcott, J. and E. Goodwin, 2014. Relationships between Macroinvertebrate Community Index and Environmental Drivers. Cawthrone Report, Cawthrone 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.

Clapcott, J., R. Young, J. Sinner, M. Wilcox, R. Storey, J. Quinn, C. Daughney, and A. Canning, 2018. Freshwater Biophysical Ecosystem Health Framework. Technical Report prepared for Ministry for the Environment of New Zealand.

Collier, K.J., J. Clapcott, and M.W. Neale, 2014. A Macroinvertebrate Attribute to Assess Ecosystem Health for New Zealand Waterways for the National Objectives Framework – Issues and Options. Report, Environmental Research Institute. University of Waikato, Hamilton, New Zealand.

Death, R.G., A. Canning, R. Magierowski, and J. Tonkin, 2018. Why Aren't We Managing Water Quality to Protect Ecological Health. Farm Environmental Planning—Science, Policy and Practice. Massey University, New Zealand: Occasional Report 31.

Larned, S., A. Whitehead, C.E. Fraser, T. Snelder, and J. Yang, 2018. Water Quality State and Trends in New Zealand Rivers. Analyses of National-Scale Data Ending in 2017. Prepared for Ministry for the Environment, NIWA.

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.

MFE, 2019. Science and Technical Advisory Group. Meeting Agenda. 16 April 2019. 16 April – priority paper compilation, Ministry for the Environment., Wellington, New Zealand.

MFE, 2022. Setting Instream Nutrient Concentration Thresholds for Nutrient-Affected Attributes in Rivers: Guidance on Implementing Clause 3.13 of the NPS-FM. Guidance Report, Wellington, New Zealand.

MFE & StatsNZ, 2017. Our Fresh Water 2017. Ministry for the Environment & Statistics NZ, Wellington, New Zealand.

MFE & StatsNZ, 2019. Environment Aotearoa 2019. Environmental Reporting Series, Ministry for Environment and Statistics New Zealand, Wellington, New Zealand.

Ministry for Environment, 2020. National Policy Statement for Freshwater Management 2020.

NZ Government, 2020. New Zealand's Greenhouse Gas Inventory: 1990–2018. Ministry for the Environment, Wellington, New Zealand.

PCE, 2013. Water Quality in New Zealand: Land Use and Nutrient Pollution. Parliamentary Commissioner for the Environment, Wellington, New Zealand.

PCE, 2015. Update Report. Water Quality in New Zealand: Land Use and Nutrient Pollution. Parliamentary Commissioner for the Environment, Wellington, New Zealand.

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.

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.

Pyle, E., R. Ward, G. McBride, and B. Huser, 2001. Establishing Watershed Management in Law: New Zealand's Experience. Journal of the American Water Resources Association 37:783–792.

R Core Team, 2021. R: A Language and Environment for Statistical Computing. https://www.R-project.org/.

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.H., F. Cattanéo, A.M. Suren, and B.J.F. Biggs, 2004. Is the River Environment Classification an Improved Landscape-Scale Classification of Rivers? Journal of the North American Benthological Society. 23:580–598.

Snelder, T., C. Kilroy, and D. Booker, 2022. Derivation of Nutrient Criteria for Periphyton Biomass Objectives. Using Regional Council Monitoring Data. LWP Client Report, LWP and NIWA Ltd, Christchurch, New Zealand.

Snelder, T.H., C. Moore, and C. Kilroy, 2019. Nutrient Concentration Targets to Achieve Periphyton Biomass Objectives Incorporating Uncertainties. JAWRA Journal of the American Water Resources Association 55:1443–1463.

Snelder, T.H., R. Woods, and B.J.F. Biggs, 2005. Improved Eco-Hydrological Classification of Rivers. River Research and Applications 21:609–628.

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.

Whitehead, A., C.E. Fraser, T.H. Snelder, K. Walter, S. Woodward, and C. Zammit, 2021. Water Quality State and Trends in New Zealand Rivers. Analyses of National Data Ending in 2020. NIWA, Christchurch.

Wild, M., T. Snelder, J. Leathwick, U. Shankar, and H. Hurren, 2005. Environmental Variables for the Freshwater Environments of New Zealand River Classification. Christchurch.

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