Drivers of macroinvertebrate communities in Northland streams

Prepared for Northland Regional Council

May 2023

Prepared by:
Elizabeth Graham
Michelle Greenwood

For any information regarding this report please contact:

Elizabeth Graham
Freshwater Ecologist
Freshwater Ecology
+64 7 856 1752

National Institute of Water & Atmospheric Research Ltd
PO Box 11115
Hamilton 3251

Phone +64 7 856 7026

NIWA CLIENT REPORT No: 202384HN
Report date: May 2023
NIWA Project: NRC22202

Quality Assurance Statement		
	Reviewed by:	Rick Stoffels
	Formatting checked by:	Jo Downey
	Approved for release by:	Michael Bruce

Contents

Tables

Figures

Executive summary

The National Policy Statement for Freshwater Management 2020 (NPS-FM) requires regional councils to assess stream water quality and ecosystem health by monitoring a range of freshwater attributes, including macroinvertebrates. The macroinvertebrate attribute is represented by three numeric attribute units: the Quantitative Macroinvertebrate Community Index (QMCI), the Macroinvertebrate Community Index (MCI), and the Average Score Per Metric (ASPM). If metric values are below a national bottom line for the attribute, the NPS-FM requires the regional council to develop an action plan to identify key stressors and attempt to improve macroinvertebrate communities.

Identifying key stressors of macroinvertebrate communities is a significant challenge. Streams are influenced by a wide range of multiple, often correlated, stressors which can have direct and indirect effects on macroinvertebrates. Potential stressors include dissolved nutrients, conductivity, dissolved oxygen, and temperature, substrate (particularly fine sediments), channel morphology and instream habitat, riparian condition, and altered flow regimes. Two stressors of particular concern to Northland Regional Council (NRC) are in-stream nutrient concentrations and drought.

Nutrients are a concern because many of NRC's State of the Environment (SoE) monitoring sites meet both the nutrient toxicity attribute criteria in the NPS-FM and more stringent national nutrient criteria that were recently derived utilising assessments of macroinvertebrate community state (Canning et al. 2021) but still fail to meet macroinvertebrate bottom lines. Therefore, NRC wishes to assess whether national criteria are stringent enough to protect macroinvertebrate communities within its region. Northland has also experienced three drought periods over the past eight years. A previous analysis of drought impacts on macroinvertebrate communities in the region reported correlations between macroinvertebrate metrics and drought in multiple sites but did not investigate potential causal mechanisms such as indirect effects of drought on other environmental drivers (Death et al. 2020).

The objectives of this project were threefold:

1. to investigate the applicability of the national nutrient criteria from Canning et al. (2021) for Northland, and
2. to identify other potential drivers apart from nutrients, including water quality, sediment, algae, and flow, that could be acting as stressors on macroinvertebrate communities, and to compare macroinvertebrate community turnover within and between SoE sites,
3. to identify drought effects on macroinvertebrate community composition and on other potential drivers of community composition.

Objective 1: Nutrient criteria

Are the critical values in Canning et al. (2021) for DIN and DRP sufficient to maintain the NPSFM macroinvertebrate attributes above the national bottom line in Northland rivers?

We repeated the minimisation of mismatch (MoM) analysis from Canning et al. (2021) using Northland SoE data rather than the national dataset to derive Northland-specific nutrient criteria for dissolved inorganic nitrogen (DIN), dissolved reactive phosphorus (DRP), ammoniacal nitrogen (N), nitrate nitrogen (NO_3-N), and Total Kjeldahl's nitrogen (TKN) with NPS-FM 'national bottom lines' for macroinvertebrate attributes (MCI, QMCI, ASPM) as ecological metric targets. The MoM algorithm aims to find a threshold that balances two sets of monitoring records - (a) records with ecological states at or above the target, but with poor nutrient status, and (b) records with ecological states below the target, but with good nutrient status (for this analysis threshold for ecological state is the NPS-FM national bottom line). These two sets of records may be viewed, respectively, as sites that are relatively "over-protected" by the nutrient threshold, given the ecological target, and sites that are relatively "under-protected" by the nutrient threshold, given the ecological target. The intersection point between these two sets of records determines the nutrient threshold at which mismatch is minimised to protect the ecological target state.

To use the MoM algorithm for setting nutrient thresholds, we make certain assumptions about the distributions and relationships of the ecological and nutrient data. The Northland macroinvertebrate and nutrient datasets met some, but not all, of these assumptions. In particular, the distribution of data was uneven, with many sites on the low end of the range in macroinvertebrate metric scores. There were also few significant relationships between macroinvertebrate metrics and nutrients across all sites.

The Northland-derived nutrient criteria for both DIN and DRP were substantially lower than the nationally derived criteria; DIN by an order of magnitude and DRP by approximately half (Table 1-1). However, given that the Northland data did not meet all the MoM assumptions and that GAMM modelling also indicated that nutrients were not key drivers across all sites, it would be useful to undertake further investigation of the role of nutrients in impacting macroinvertebrate communities before large effort or expense was undertaken in reducing nutrient concentrations to below the criteria identified here. Observed relationships with nutrients may be due to another correlated environmental driver.

Table 1-1: Nutrient criteria for Northland. Median and (range) of nutrient criteria developed for Northland compared to national criteria from Canning et al. (2021) for DIN and DRP. The minimisation of mismatch analysis was unable to identify DIN criteria for QMCI or NO_3-N criteria for QMCI or ASPM.

Nutrient	MCI	QMCI	ASPM
Northland			
Amm-N	0.0087 (0.0062 – 0.0111)	0.0062 (0.0041 – 0.0078)	0.0066 (0.0050 – 0.0091)
TKN	0.1704 (0.1457 – 0.1951)	0.1442 (0.1210 – 0.1627)	0.1503 (0.1256 – 0.1719)
NO_3-N	0.1009 (0.036 – 0.1821)		
DIN	0.11 (0.05 – 0.18)		0.06 (0.01 – 0.12)
DRP	0.01 (0.009 – 0.014)	0.01 (0.007 – 0.012)	0.01 (0.007 – 0.013)
National (from Canning et al. 2021)			
DIN	1.07 (0.93 – 1.21)	0.63 (0.45 – 0.77)	1.12 (1.01 – 1.29)
DRP	0.028 (0.025 – 0.03)	0.018 (0.015 – 0.02)	0.028 (0.026 – 0.032)

Are there differences among land use (i.e., predominately pasture, forest, or urban) and geology classes that may affect the applicability of these nutrient criteria?

Both nutrient concentrations and macroinvertebrate metrics varied between streams with different catchment land use. Pastoral and urban streams exceeded the Northland nutrient criteria for all forms of nitrogen, while indigenous forest streams were below the Northland nutrient criteria for all forms of nitrogen. Hard- and soft-bottomed streams were both evenly split between exceeding and meeting nutrient criteria for NO_3, and DIN. More soft-bottomed streams exceeded criteria for ammoniacal N and TKN, perhaps due to poorly drained soils in low-lying floodplains. Sites in all land use and stream type categories exceeded the Northland criteria for DRP. However, DRP concentrations were equally high across land use types, including indigenous forest sites, suggesting the DRP was primarily associated with the volcanic substrates common in Northland, rather than any anthropogenic impact.

Objective 2: Drivers of macroinvertebrate communities and community turnover

What are the predictors of macroinvertebrate community composition in Northland rivers and streams?

We used a full subsets approach to identify other potential drivers of macroinvertebrate community composition in addition to nutrients. Data for a large number of potential drivers or predictors of macroinvertebrate community composition were available (n = 152) across 66 sites over 8 years (although individual predictors varied in the number of sites and time window over which data was available). Generalized additive mixed models (GAMMs) were fit for invertebrate metrics MCI, QMCI, ASPM, and percent EPT taxa and percent EPT abundance, with site as a random effect. A complete set of possible models was created using all combinations of predictor variables. Predictor variables were chosen to represent key stressors on stream ecosystems: nutrients and other water chemistry, flow regimes, habitat and drought (Table 1-2). Spatial attributes (e.g. elevation, slope, catchment area, rainfall) from the River Environment Classification (REC, version 2.5) were also included. The relative importance of predictor variables was assessed by summing the AICc (Akaike Information Criterion) weights for all models containing each variable, while the best model was selected based on the lowest AICc and least number of predictor variables.

Models were fit to the full site by year dataset (time series models) and to selected subsets of the data, including pastoral and indigenous forest streams (urban and exotic forestry sites were excluded due to the low number of sites within each subset (2 and 3, respectively). Models were also fit using the site median values for drivers and metrics (spatial models) to include additional drivers missing too many data points to be included in the time series analysis, such as periphyton percent cover and chlorophyll a, and percent fine sediment cover.

Table 1-2: Potential drivers included in GAMM full subsets analysis.

Driver category	Selected predictors	Model
Nutrients	Ammoniacal N, TKN, DIN, DRP	Time series, spatial
Other water quality	Conductivity, dissolved oxygen, turbidity	Time series, spatial
Periphyton and substrate	Chlorophyll *a*, percent fine sediment cover	Spatial
Habitat	Riparian habitat assessment (RHA) score	Spatial
Flow	Median flow over previous 90 days, base flow index (BFI), days since flow 3 times the long-term median flow (daFRE3)	Time series, spatial
Drought	New Zealand drought index (NZDI)	Time series

Driver category	Selected predictors	Model
Spatial attributes from REC	Elevation, slope, catchment area, variation in rainfall, number of rain days > 10mm, mean air temperature, particle size	Time series, spatial

Apart from nutrients, the key drivers of Northland macroinvertebrate communities identified by the GAMM analyses were river flow metrics, instream habitat condition, climate (such as temperature, rainfall pattern) and topography (Table 1-3). However, this varied between macroinvertebrate metrics as well as catchments across the region.

In particular, macroinvertebrate communities in pastoral streams were associated with different drivers than those in indigenous forest streams. Somewhat surprisingly, there were no notable differences in predictor importance between hard- and soft-bottomed streams, even though differences in MCI scores between the two groups indicate very different macroinvertebrate communities. Some drivers were also more important for one metric than others. In general, MCI, QMCI, and ASPM had stronger associations with drivers than EPT metrics. Overall, environmental drivers explained a greater proportion of variation in macroinvertebrate community metrics between sites than within sites over time.

Table 1-3: Selected predictors from most parsimonious models for MCI, QMCI, and ASPM.

Response	Dataset	Top predictors
	Time series	
MCI	All sites	DIN, conductivity, DO, BFI, NZDI
	Pastoral	BFI, DIN, NZDI, Reporting Year, elevation
	Indigenous Forest	DO, flow, turbidity
QMCI	All sites	Flow, mean air temperature, particle size, turbidity, slope
	Pastoral	BFI, Reporting Year
	Indigenous Forest	Particle size, turbidity
ASPM	All sites	BFI, DO, NZDI, rain days > 10 mm, slope
	Pastoral	NZDI, Reporting Year, rain days > 10 mm
	Indigenous Forest	DO, turbidity
	Spatial	
MCI		DIN, % sand-silt, temperature
QMCI		DRP, temperature
ASPM		Conductivity, daFRE3, DO, Temperature

How does the community composition (temporal species turnover) of MCI scoring taxa change within and among the SoE sites?

Temporal community turnover is the replacement of species over time. Replacement of sensitive taxa such as EPT with less sensitive taxa can be an indication of ecological impairment. Thus, we calculated community turnover to investigate whether sites below the national bottom line for macroinvertebrate metrics had greater displacement of EPT taxa than higher-scoring sites. Turnover was calculated as the percentage of taxa appearing and disappearing each year within a site. Turnover was highly variable within and between sites and did not vary regularly with catchment land use or stream type. The largest contributors to total turnover were common taxa found across many sites, rather than rare species. Taxa disappearances were greater than appearances in the year following a drought in two out of three cases, suggesting that 1) there may be a lagged response to drought and 2) it may take over a year for communities to fully recover. This pattern was observed across all land use categories. However, the 2019-2020 drought was followed by a 1-in-100 year flood event, which may have further impacted community recovery and extended the recovery time. EPT taxa had comparable turnover rates to non-EPT taxa, suggesting that sensitivity to organic pollution was not the main cause of turnover.

Objective 3: drought effects

How do drought conditions impact macroinvertebrate community composition in Northland?

We used linear mixed effects models to investigate direct effects of drought on invertebrate metrics, with site as a random effect. There were significant negative relationships between the drought index and all metrics, although inspection of individual site plots showed that the overall relationship was driven by strong correlations in a small number of sites, with most sites showing no clear relationship.

The influence of drought on macroinvertebrate communities was also tested by including drought as a predictor in the full subsets analysis. Drought had high variable importance scores in models for MCI and ASPM in pastoral streams and across all sites. The drought index also explained a large amount of variation in macroinvertebrate communities in pastoral streams, which also had the strongest associations with other environmental predictors, including temperature, baseflow and water quality, and suggests that already stressed streams are more susceptible to drought effects and/or that drought exacerbates the impact of other environmental stressors.

Do drought conditions impact water quality and environmental variables that may influence macroinvertebrate community composition?

We also used linear mixed effects models, as described above, to investigate effects of drought on other water quality and environmental variables. The drought index had significant relationships with many of the other predictors, though again, overall relationships were primarily driven by strong correlations in a subset of sites. The drought index was negatively related to ammoniacal N, TKN, DIN, DRP, turbidity, temperature, and flow, and positively related to dissolved oxygen, chlorophyll *a*, and daFRE3. The temporary improvements in water quality parameters (nutrients, clarity) were likely associated with reduced surface runoff, or overland flow, in drought years. The positive relationships with chlorophyll *a* and dissolved oxygen indicates increased aquatic plant growth and therefore elevated photosynthesis rates, during drought conditions. Elevated photosynthesis is often associated with oxygen depletion at night-time, suggesting increased stress on aquatic organisms despite high dissolved oxygen levels in daytime when measurements were taken. The mix of positive and negative relationships with other drivers, each of which will in turn influence macroinvertebrate community composition, highlights the complexity of disentangling the causal mechanisms by which drought may impact stream macroinvertebrates.

Summary

The relationship between nutrients and macroinvertebrate communities is difficult to unravel in Northland. The minimisation of mismatch approach resulted in very stringent nutrient criteria; however, these criteria will have been influenced by the uneven distribution of the Northland data, with all sites on the low end of the national range in nutrient concentrations and macroinvertebrate metric scores. Given this influence, we do not recommend use of the newly derived nutrient criteria until the role of nutrients in impacting macroinvertebrate communities in Northland streams has been better quantified. The GAMM analysis showed that nutrients are unlikely to be the main determinant of macroinvertebrate community composition when confounding factors (i.e., other environmental drivers) are taken into account. While nutrients were important in pastoral streams, several other predictors were also selected as important for explaining variation in macroinvertebrate metrics across all streams: temperature, flow, drought index, and dissolved oxygen.

To further understand the influence of nutrients and other drivers on macroinvertebrate communities in Northland, we recommend: 1) investigation into whether the source of organic nitrogen (TKN) is anthropogenic or natural, and whether correlated declines in macroinvertebrate communities are associated with TKN itself, or other drivers which co-vary with TKN (i.e., sediment), 2) continued collection of sediment and periphyton data, as well as continuous temperature and dissolved oxygen data, and re-running the drivers analysis when 5+ years of data is available, 3) careful inspection of the suitability of models generated by the full subsets analysis for predicting macroinvertebrate community composition, 4) incorporating species traits into the taxa turnover analysis to investigate mechanisms of community compositional change in response to environmental stressors, and 5) targeted monthly or bi-monthly macroinvertebrate sampling at a subset of sites immediately following drought and flood events to determine community recovery trajectories.

1 Introduction

1.1 Background

Regional Councils have statutory responsibilities to manage New Zealand's waterways under the Resource Management Act 1991 (RMA) and the National Policy Statement for Freshwater Management 2020 (NPSFM 2020).

The NPS-FM requires regional councils to assess water quality and ecosystem health by monitoring a range of freshwater attributes, including macroinvertebrates. The macroinvertebrate attribute is represented by three attribute units: the Quantitative Macroinvertebrate Community Index (QMCI), the Macroinvertebrate Community Index (MCI) (Stark and Maxted 2007) and Average Score Per Metric (ASPM) (Collier 2008). Each attribute is graded into bands A through D, with band A indicating expected values under nearly pristine conditions and the C/D band cut-off indicating a 'national bottom line' below which values are indicative of degraded ecological state. If sites are below the bottom line, the NPS-FM requires development of an action plan to identify key stressors and attempt to improve macroinvertebrate communities.

Macroinvertebrate communities are influenced by a range of environmental stressors, or drivers. Potential drivers include water quality parameters such as dissolved nutrients, conductivity, dissolved oxygen, temperature, substrate cover (particularly fine sediments), channel morphology, instream habitat, riparian condition, and altered flow regimes. These stressors can have direct as well as indirect effects on macroinvertebrates and most freshwater ecosystems are subject to multiple stressors at any given time. Surrounding topography, geology, and land use may also influence which drivers have the largest effect on a given stream. Stream macroinvertebrate communities also vary naturally between hard-bottomed and soft-bottomed streams. Thus, disentangling the impact of individual stressors on macroinvertebrate communities is likely to be a significant challenge for councils. Northland Regional Council (NRC) has identified two groups of drivers of particular concern for the Northland region: nutrients and drought.

1.1.1 Nutrients

The NPS-FM requires councils to set nutrient criteria to achieve target attribute states for both periphyton and macroinvertebrates. A recent analysis of Northland's State of the Environment (SOE) river monitoring sites found that the majority of SOE sites were within the NPS-FM band C or band D for macroinvertebrate attributes (Death et al. 2020), despite nutrient attributes for the same sites being in mostly above the national bottom line (Rissmann and Pearson, 2020). The majority of Northland sites with poor macroinvertebrate scores also met recently published national nutrient criteria for achieving macroinvertebrate targets for dissolved inorganic nitrogen (DIN) and dissolved reactive phosphorus (DRP)(Canning et al. 2021). The suggested criteria were: ~0.6 mg/L for DIN and ~0.02 mg/L for DRP (Canning et al. 2021).

The first objective of this study was to investigate the applicability of the national nutrient criteria from Canning et al. (2021) for Northland, specifically:

Are the critical values in Canning et al. (2021) for DIN and DRP sufficient to maintain the NPSFM macroinvertebrate attributes above the national bottom line in Northland rivers?

Are there differences among land use (i.e., predominately pasture, forest, or urban) and stream-bed type (i.e., hard-bottomed and soft-bottomed streams) that may affect the applicability of these nutrient criteria?

To test whether the minimisation of mismatch approach was applicable to Northland data (due to the uneven distribution of sites with low nutrients above and below macroinvertebrate metric bottom lines), we derived Northland-specific nutrient criteria for comparison with the national criteria from Canning et al. (2021).

1.1.2 Drivers and community turnover

The second objective of this study was to investigate drivers of macroinvertebrate communities and community turnover in Northland:

What are the predictors of macroinvertebrate community composition in Northland rivers and streams?

The relative importance of potential drivers, including nutrients and drought, was assessed using generalised additive mixed modelling (GAMM) with a full subsets approach. GAMMs were used due to their ability to model non-linear relationships between continuous predictor and response variables, which are common in ecological datasets, and to include random effects to account for spatial or temporal autocorrelation.

NRC was also interested to know how macroinvertebrate communities were changing over time (i.e., temporal species turnover) in their SoE monitoring sites:

How does the community composition (temporal species turnover) of MCI scoring taxa change within and among the SoE sites?

Which taxa were lost or replaced the most?

Turnover was calculated as appearances and disappearances of taxa between years for all sites.

1.1.3 Drought

Northland has experienced two drought periods in recent years (2017, 2020) which may have impacted macroinvertebrate communities via adverse effects on river flows. A preliminary analysis of drought effects on Northland macroinvertebrate communities did not demonstrate any remarkable change in macroinvertebrate communities in response to drought at most of the SOE sites (Death et al. 2020). However, indirect effects of changes in flow under drought conditions via impacts on water quality and other in-stream biophysical variables were not assessed.

The third objective of this study was to investigate drought impacts on macroinvertebrate community composition as well as effects of drought on other potential drivers:

How do drought conditions impact macroinvertebrate community composition in Northland?

Do drought conditions impact water quality and environmental variables that may influence macroinvertebrate community composition?

The relative importance of drought on macroinvertebrate community composition was assessed in the GAMM analyses, along with other potential drivers. Drought effects on other drivers were assessed individually using linear mixed effects models.

1.2 Report roadmap

This report is structured into seven sections, including the introduction (**Section 1**):

Section 2 provides a description of data available at different sites and across varying temporal periods and frequencies for macroinvertebrates, nutrients, drought and other potential drivers. A summary of the collation of the data into a large dataset, separately delivered to NRC, is also provided.

Section 3 describes the state of macroinvertebrate communities across NRC SoE monitoring sites between 2014 and 2021 by comparing macroinvertebrate metric scores (MCI, QMCI, ASPM) calculated using 1) NEMS tolerance values and 2) Northland-specific tolerance values to NPSFM attribute bands.

Section 4 presents the statistical analysis, methods, results, and discussion of the development of nutrient criteria specific to Northland using the minimisation of mismatch (MoM) method.

Section 5 presents statistical analysis, methods, results, and discussion of the analysis of potential drivers, including drought, of macroinvertebrate communities.

Section 6 provides a summary of the community turnover analysis and discussion of observed patterns in total turnover within and between sites and turnover of individual taxa.

Section 7 summarises overall conclusions and recommendations for future work.

Appendix A summarises the results of correlations between median macroinvertebrate metrics and nutrient concentrations across all sites. The modelled versus observed hydrographs for the sites with flow recorders are supplied in Appendix B. Pearson correlation plots and summary statistics between potential drivers of macroinvertebrate communities are in Appendix C. Appendix D contains temporal plots of multiple macroinvertebrate metric scores and the NZDI drought indicator for each site. Temporal plots of potential drivers of macroinvertebrate community composition against the NZDI drought indicator for all sites are provided in Appendix E. Summaries of the turnover of individual macroinvertebrate taxa are in Appendix F. Appendix G contains barplots of the percentage EPT and EPT taxa richness over time within all sites.

2 Data

All data were collected by NRC as part of their State of the Environment (SoE) monitoring between 2014-2021. There were 66 sites monitored, although not all 66 sites were sampled for all parameters each year (Figure 2-1, Table 2-7). Fifty-three of the sites were located in catchments with predominately pastoral land use, eight were located in catchments with predominately indigenous forest, three in catchments with exotic forestry, and two were located in urban catchments. Thirty-nine of the sites are classified as “hard-bottomed’ and 27 as “soft-bottomed”, i.e., with a high proportion of fine sediment on the stream-bed. All but one of the indigenous and exotic forest sites were hard-bottomed, with one soft-bottomed stream in indigenous forest. The two urban streams were also hard-bottomed. The pastoral streams were evenly split, with 27 hard-bottomed and 26 soft-bottomed.

2.1 Macroinvertebrates

The macroinvertebrate data consisted of taxa counts for the 66 sites sampled annually by NRC between December and March from 2014-2021. Samples collected in December were assigned to the following ‘Reporting Year’ in order to avoid splitting data from the same summer into separate calendar years. For example, macroinvertebrate samples taken in December 2019, January 2020, and February 2020 were all designated as Reporting Year 2020. Macroinvertebrate metrics were calculated for each Reporting Year following NEMS (2022) methodology (see section 3.1). Not all sites were sampled in all years (between 33 and 66 sites per year, Table 5-1).

Macroinvertebrate metrics were calculated for each reporting year following the methodology of the National Environmental Monitoring Standards for macroinvertebrates (NEMS 2020) (Table 2-1). Metrics were calculated using both the species tolerance values from the NEMs and Northland-specific tolerance values from Stark (2017). MCI_{HB} tolerance values were used to calculate metrics in sites identified as hard-bottomed by NRC, and MCI_{SB} tolerance values used in soft-bottomed sites.

Table 2-1: Macroinvertebrate metric calculations. Adapted from NEMS (2020).

Metric	Units	Description	Calculation
Macroinvertebrate Community Index (MCI_{HB}, MCI_{SB})		A measure of stream health based on the tolerance of different macroinvertebrate taxa to organic pollution (Stark and Maxted 2007). Each species is assigned a tolerance score from 1 (very tolerant) to 10 (very sensitive). Tolerance values differ for hard-bottomed (HB) and soft-bottomed (SB) streams. Northland-specific tolerance values have also been developed (Stark 2017). MCI is calculated as the sum of tolerance scores for all species in a site.	$MCI = \frac{\sum_{i=1}^{i=S} a_i}{S} \times 20$ Where S = the total number of scoring taxa in a sample and a_i is the tolerance score for the ith taxon.
Quantitative Macroinvertebrate Community Index ($QMCI_{HB}$, $QMCI_{SB}$)		Incorporates abundance of each taxa.	$QMCI = \frac{\sum_{i=1}^{i=S}(n_i \times a_i)}{N}$ Where S = the total number of scoring taxa in a sample, n_i is the abundance of the ith scoring taxon, a_i is the tolerance score for the ith taxa, and N = the total abundance

Metric	Units	Description	Calculation
			for the scoring taxa for the entire sample.
EPT Taxa Richness		EPT (Ephemeroptera – mayflies, Plecoptera – stoneflies, and Trichoptera – caddisflies) are groups known to be sensitive to organic pollution. Caddisflies from the family Hydroptilidae are excluded from EPT metric calculations because they are pollution tolerant.	*Number of EPT taxa*
Percent EPT Taxa Richness	%		$\frac{Number\ of\ EPT\ taxa}{Total\ number\ of\ taxa}$
Percent EPT Abundance	%		$\frac{Number\ of\ EPT\ individuals}{Total\ number\ of\ individuals}$
Average Score Per Metric (ASPM)		A multi-metric index calculated as the mean of three metrics: MCI, EPT taxa richness, and percent EPT abundance (Collier 2008).	Each metric is firstly scaled (normalised) by: $X' = [X - X_{min}]/[X_{max} - X_{min}]$ Where *X'* is the scaled site score, *X* is the raw site score, and X_{min} and X_{max} are: EPT taxa richness (0-29), % EPT Abundance (0-100), MCI (0-200).

2.2 Water quality

The water quality data consisted of monthly samples of 16 parameters from the 66 sites[1] collected by Northland Regional Council between 2014 and 2021 (Table 2-2).

The monthly water quality data for each site was summarised over the twelve months prior to the corresponding macroinvertebrate sampling date for each site to capture the effects of antecedent conditions on macroinvertebrate communities (Table 2-2). A twelve-month period was chosen because most macroinvertebrates spend at least a year as larvae in the aquatic environment, and therefore changes in their community composition due to pollution tolerance levels or habitat preferences will reflect changes in water quality and other environmental stressors over that year (Stark and Maxted 2007). The summarised values were assigned the same Reporting Year as the corresponding macroinvertebrate sample. For example, if the macroinvertebrate sample was collected in January 2020, the summarised water quality data from January 2019-January 2020 was designated as belonging to Reporting Year 2020.

If there was more than 20% missing data (i.e > 2 months) for a site in a given Reporting Year, that Site-Reporting Year combination was excluded from the summary to avoid seasonal bias to annual values (i.e., data only collected in summer, or no winter data collected; Figure 2-2). Censored values reported as less than an analytical detection limit were replaced with half the detection limit, while

[1] The original dataset had 71 sites. However, four sites were slight and/or temporary location shifts due to access issues; in these cases the data was combined under the original site name. One additional site, the Utakura River at Horeke Road, was only sampled in 2014-2015 before the site was permanently moved to Okaka Road in 2016. As no invertebrate samples were collected from the Horeke Road site, it was removed from the dataset.

censored values greater than a detection or reporting limit were replaced with 1.1 times the limit (Helsel 2005, 2012).

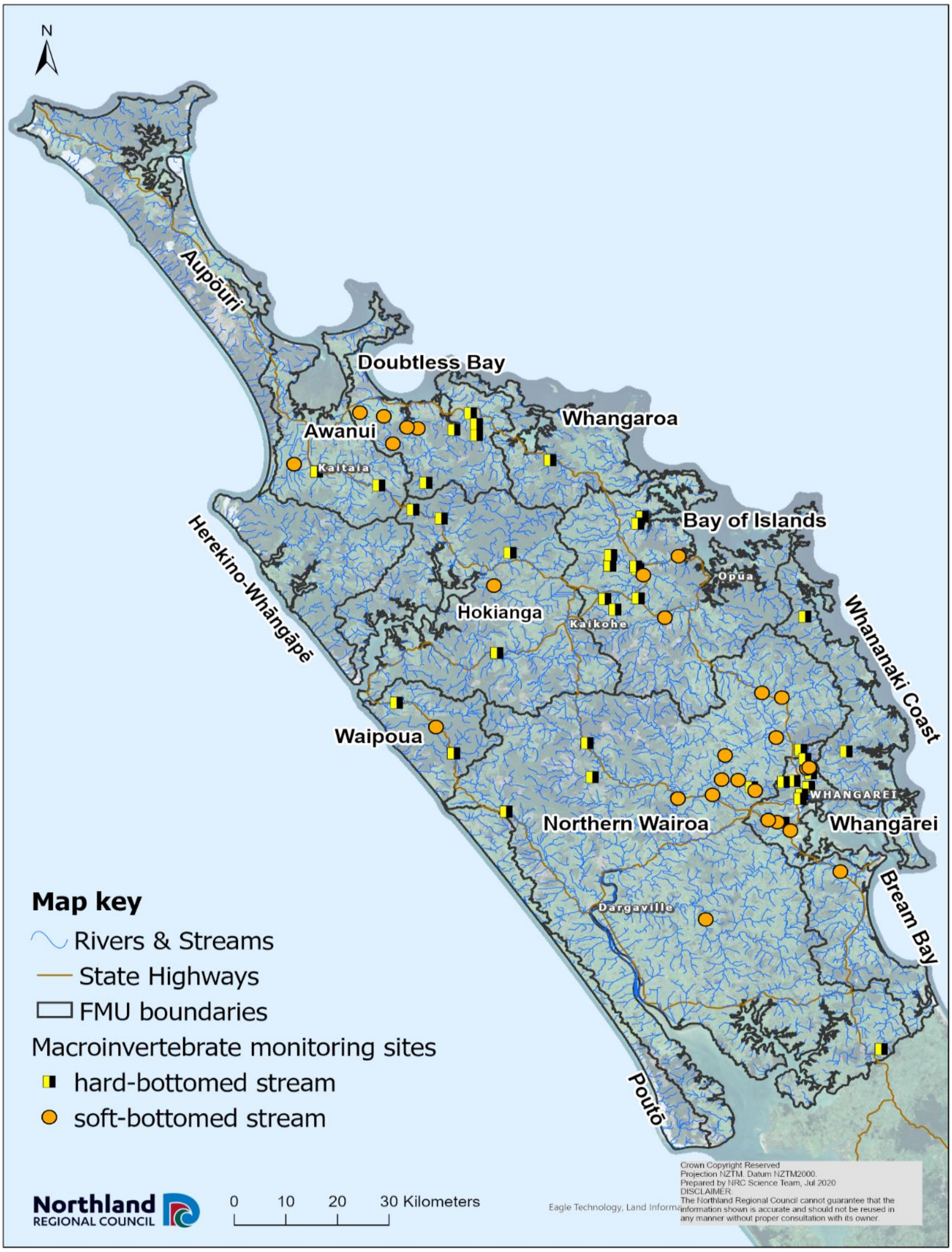

Figure 2-1: Macroinvertebrate sampling sites in Northland.

Table 2-2: Water quality parameters measured monthly by NRC between 2014-2021. Summary statistics were calculated over the year prior to the macroinvertebrate sample being taken.

Variable	Units	Summary statistics
Ammoniacal Nitrogen (Amm-N)	g/m^3	Median Maximum
Black disc	m	Median
Dissolved Inorganic Nitrogen (DIN)	g/m^3	Median 95th percentile
Conductivity	µs/cm	Median
Dissolved Reactive Phosphorus (DRP)	g/m^3	Median 95th percentile
Dissolved Oxygen (DO)	mg/L	Median Minimum
Dissolved Oxygen % saturation (DO %)	%	Median Minimum
Nitrate-Nitrogen (NO_3-N)	g/m^3	Median 95th percentile
Nitrite-Nitrogen (NO_2-N)	g/m^3	Median 95th percentile
Nitrite-Nitrate-Nitrogen (NO_2-NO_3-N)	g/m^3	Median 95th percentile
Temperature (monthly spot measurements)	deg. C	Median 95th percentile
Total Kjeldahl Nitrogen (TKN)	g/m^3	Median 95th percentile
Total Nitrogen (TN)	g/m^3	Median 95th percentile
Total Phosphorus (TP)	g/m^3	Median 95th percentile
Total suspended solids (TSS)	g/m^3	Median
Turbidity	NTU	Median

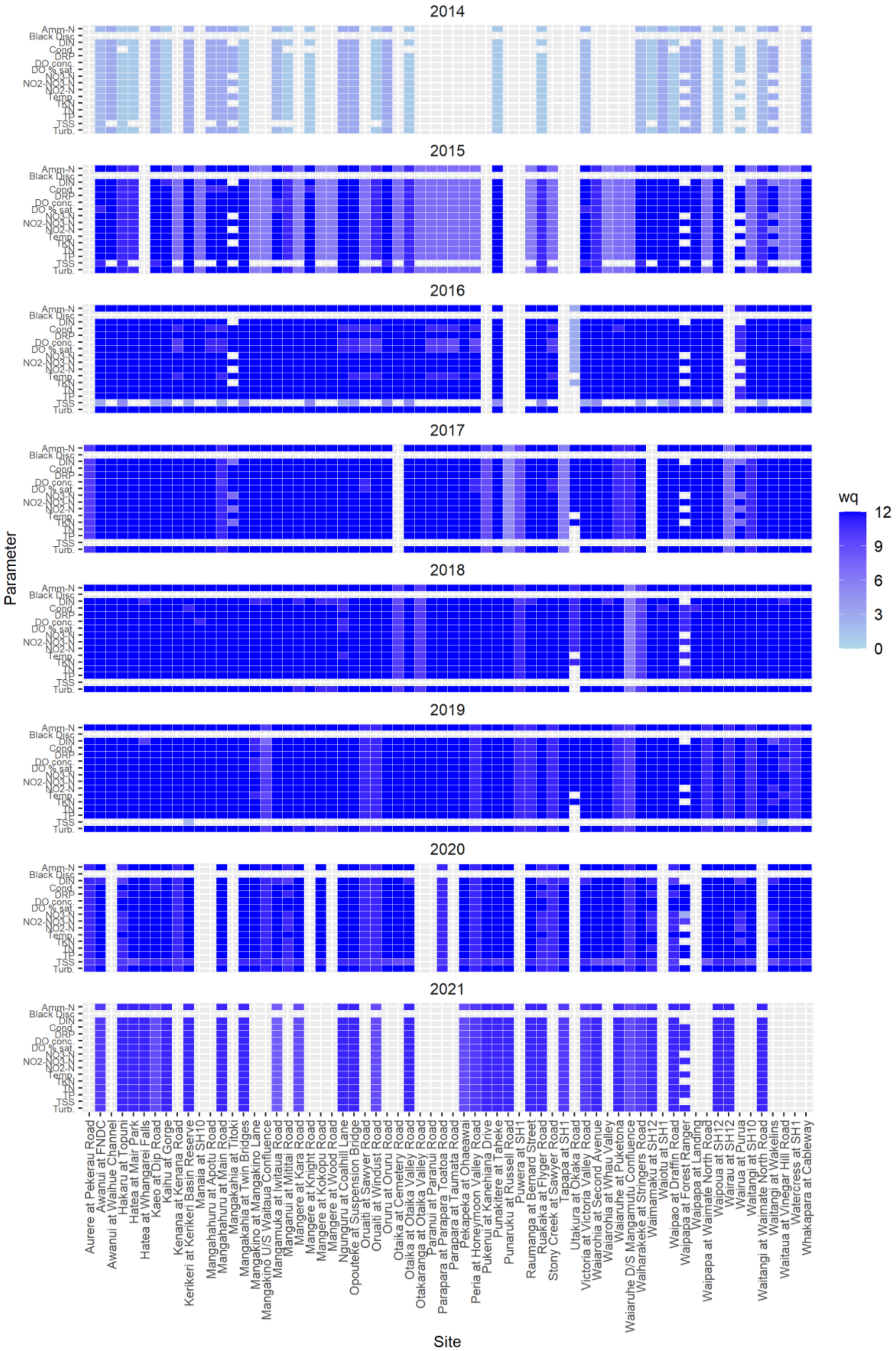

Figure 2-2: Data availability of monthly water quality data at each SoE monitoring site between 2014 and 2021. The darkest blue indicates all 12 monthly samples were collected in a year, paler blues indicate more missing samples. Grey squares show no samples were collected for that parameter during that year at that site.

2.3 Periphyton and substrate

The periphyton data consisted of monthly chlorophyll *a* samples and percentage cover visual assessments from 45 sites collected by Northland Regional Council between 2014 and 2021 (Table 2-3). Three of the sites were not sampled for invertebrates and therefore were removed from the dataset. Not every site was sampled each year (i.e., only 19 sites sampled in 2014, Figure 2-3).

To summarise percentage cover visual assessments of different periphyton types into a single number the Weighted Composite Cover (WCC) of periphyton was calculated following Matheson (2012):

$$WCC = \%\, cover\ by\ filaments + \frac{\%\, cover\ by\ mats}{2}$$

The substrate data consisted of monthly measurements of percent cover by different particle classes (bedrock, boulders, small and large cobbles, small and large gravels, sand, and silt) taken in association with periphyton sampling from 34 sites in 2020 and 31 sites in 2021 (Table 2-3). Due to the lack of temporal data, and because substrate composition is generally less variable than water quality or periphyton, the mean substrate composition from the last two years of sampling was used for all sampling dates.

Summary statistics for various periphyton parameters over the twelve months prior to the macroinvertebrate sampling date were calculated (Table 2-3) following the same missing data exclusion rule as used for the water quality data (see Figure 2-3 for data coverage).

Table 2-3: Periphyton and substate parameters measured monthly by NRC between 2014-2021 and 2020-2021, respectively. Periphyton summary statistics were calculated over the year prior to the macroinvertebrate sample being taken. Mean substrate composition at a site was calculated from the last two years of sampling and used for all time periods due to limited temporal data.

Variable	Units	Summary statistics
Periphyton		
Benthic Chlorophyll *a*	mg/m^2	Median 92^{nd} percentile Maximum
Filaments cover	%	Median Maximum
Films cover	%	Median Maximum
Mats cover	%	Median Maximum
Sludge cover	%	Median Maximum
Total cover	%	Median Maximum

Variable	Units	Summary statistics
Weighted composite cover (WCC)	%	Median 92nd percentile Maximum
Macrophytes % cover	%	Median Maximum
Substrate		
Bedrock cover	%	Mean
Boulder cover	%	Mean
Large cobble cover	%	Mean
Small cobble cover	%	Mean
Gravel cover	%	Mean
Sand cover	%	Mean
Silt cover	%	Mean
Sand + Silt cover	%	Mean
Total Deposited sediment	%	Mean
Embeddedness - Good	%	Mean
Embeddedness - Loose	%	Mean
Embeddedness - Tight	%	Mean
Embeddedness - Moderate	%	Mean

2.4 Habitat

The habitat data consisted of scores from an annual Rapid Habitat Assessment (RHA; Clapcott et al. 2015) conducted each year from 2016 to 2021 in conjunction with the macroinvertebrate sampling. The RHA includes ten physical habitat components: deposited sediment, invertebrate habitat diversity, invertebrate habitat abundance, fish cover diversity, fish cover abundance, hydraulic heterogeneity, bank erosion, bank vegetation, riparian width, and riparian shade. Each component is assigned a score from 1-10 with 1 indicating poor habitat conditions and 10 indicating excellent habitat conditions. The ten scores are then summed for a total RHA score.

Figure 2-3: Data availability of monthly periphyton and substrate measurements at each SoE monitoring site between 2014 and 2021. Darkest grey and green indicate all 12 monthly samples were collected at each site within a year. Lighter green and grey indicate fewer monthly samples were collected (see legend).

2.5 Flow

Continuous flow data (at 5-minute intervals) was available for 31 sites and summarised as daily mean flow. Daily mean flows for the remaining 35 sites were estimated using NIWA's national hydrology model, TopNet. Observed and modelled flows from the 31 sites with monitoring data corresponded well, although TopNet often underestimated the highest peaks observed in the measured flow records (Appendix B).

The daily flow data were summarised as minimum, mean, median, and maximum flow over the one year, 3 months, 1 month, and 1 week prior to the macroinvertebrate sampling date (Table 2-4). Additional antecedent flow metrics were calculated for each invertebrate sampling date including days since last flow of 3 and 10 times the long-term median flow, long-term (over entire reporting period, 2014-2021) mean and median flow, annual low flow (ALF), yearly base flow index (BFI), and long-term base flow index (Table 2-4).

Table 2-4: Calculated flow measurements and antecedent flow metrics for each Northland SoE site between 2014 and 2021. Measured flows data were used when available, otherwise estimated flows from the TopNet national hydrology model were used.

Variable		Units	Summary statistics
Daily flow			
Daily flow		m^3/s	Minimum
1 week			Mean
1 month	Prior to macroinvertebrate sampling		Median
3 months			Maximum
1 year			
Long-term (2014-2021) flow		m^3/s	Mean
			Median
Antecedent flows			
Days since last flow 3 x long-term median flow (daFRE3)		days	-
Days since last flow 10 x long-term median flow (daFRE10)		days	-
Annual low flow (ALF)		m^3/s	Median
Yearly base flow index (BFI)			-
Long-term (2014-2021) base flow index (BFI)			-

2.6 Drought

Two drought indices were available: The New Zealand Drought Index (NZDI) and the Standardised Discharge Index (SDI). The NZDI is a regional index developed by NIWA which combines four climatological drought indicators: the Standardised Precipitation Index, the Soil Moisture Deficit, the Soil Moisture Deficit Anomaly, and the Potential Evapotranspiration Deficit (Mol et al. 2017).

The NZDI is calculated daily using data from approximately 100 climate stations around New Zealand which is the interpolated to produce national NZDI values at an approximately 500 m grid resolution (Mol et al. 2017). The Northland macroinvertebrate sampling sites fall within three different NZDI districts: Far North (includes 35 of NRC's macroinvertebrate sampling sites), Kaipara (3 sites), and Whangarei (29 sites). The annual maximum, median, and mean NZDI was calculated for each Site and Reporting Year combination between 2014 and 2021. Higher NZDI values indicate more severe drought (Table 2-5).

The Standardised Discharge Index is a local hydrological drought indicator developed by NRC based on mean monthly flow data from river flow stations (Pham et al. 2022). The SDI was provided by NRC for a three-month window during which droughts are most common (December-January) each year for 24 river sites with flow monitoring stations. Lower values of SDI indicate more severe drought (Table 2-5).

Table 2-5: Drought categories for the New Zealand Drought Index (NZDI) and Standardised Discharge Index (SDI). Note higher NZDI and lower SDI values indicate more severe drought.

Index	Value	Category
NZDI	0.75	Dry
	1.00	Very dry
	1.25	Extremely dry
	1.50	Drought
	1.75	Severe drought
SDI	SDI ≥ 0.0	Near normal
	-1.0 ≤ SDI < 0.0	Mild drought
	-1.5 ≤ SDI < 1.0	Moderate drought
	-2.0 ≤ SDI < 1.5	Severe drought
	SDI < -2.0	Extreme drought

2.7 Spatial attributes

For each site, catchment geography and topography, climate and geology data were extracted from the New Zealand River Environment Classification (REC2.5) spatial layer (Table 2-6). Land cover information was derived from the national Land Cover Database 5 (LCDB5) spatial layer (Table 2-6).

Table 2-6: Catchment and local attributes from the River Environment Classification spatial layer and the Land Cover database.

Spatial Layer	Attribute class	Attribute description	Abbreviation	Units
REC	Geography and topography	Mean elevation above sea level of catchment	elev	m
		Mean slope of the catchment	slope	degrees
		Catchment area	catarea	m2

Spatial Layer	Attribute class	Attribute description	Abbreviation	Units
	Climate	Coefficient of variation of annual catchment rainfall	rnvar	mm
		Mean number of catchment rain days greater than 10 mm/month	rd10	days/month
		Mean local air temperature	mat	degrees C
	Geology	Mean local particle size	psize	ordinal
LCDB	Land cover	Dominant (by proportion) land cover in catchment - Pastoral - Indigenous Forest - Exotic forest - Urban	-	-

2.8 Collation

The seven datasets (Section 2.1 to 2.7) were aligned by Site Name and Reporting Year and combined into a single spreadsheet (provided separately to NRC). The full collated dataset contained eight invertebrate metrics, 152 environmental parameters, and two categorical variables (dominant land use in catchment and stream type) for 66 sites (Table 2-7). Note that the spatial and temporal coverage of data varied between parameters leading to gaps in full collated dataset (Table 2-6). This dataset was further summarised and filtered as needed for specific analyses (described in the Methods of Sections 3 to 6).

Table 2-7: Number of monitoring sites and frequency and duration of sampling for each provided dataset.

Dataset	Number of sites	Frequency	Duration
Monitoring			
Invertebrates	66	Annually (once between Dec-Feb)	2014-2021
Water quality: nutrients, conductivity, turbidity, DO, temperature, sediment	66	Monthly (varies by site)	2014-2021 (varies by site)
Periphyton: chlorophyll *a*, WCC	45	Monthly	2014-2021 (19 sites) 2015-2021 (45 sites)
Sediment % cover	34	Monthly	2020-2021

Dataset	Number of sites	Frequency	Duration
In-stream habitat (RHA)	66	Annually	2016-2021
Flow – measured	31	Every 5-mins	2014-2021
Flow – modelled	35	Daily	2014-2021
Drought			
NZDI	3 regions	Daily	2014-2021
SDI	24	Annually (3-month window Dec-Feb)	2014-2021
Spatial information			
REC	Spatial GIS layer	-	-
Land cover	Spatial GIS layer	-	-

3 NPS-FM attribute state

The objective of this section was to examine the state of macroinvertebrates in Northland in accordance with the NPS-FM 2020. Nutrient toxicity and periphyton attributes were also examined to assess their potential as drivers of macroinvertebrate community composition in the Northland region.

3.1 Methods

For macroinvertebrate attributes (MCI, QMCI, and ASPM), the NPS-FM stipulates that state should be calculated as the median score of the previous five years. When less than five years of data were available (e.g., for 2014-2017), the median of available years was used, and the number of years noted. For example, 2014 values were used for the state in 2014 and the median of 2014, 2015, 2016 and 2017 for the state in 2017. MCI and QMCI metrics were calculated using both: 1) tolerance values from the NEMS and 2) Northland-specific tolerance values provided in Stark (2017). MCI_{HB} - tolerance values were used to calculate metrics for sites identified as hard-bottomed by NRC and MCI_{SB} tolerance values for soft-bottomed sites. Metrics were compared to the NPS-FM attribute bands for MCI, QMCI and ASPM. Attribute values are assigned to four bands, with Band A indicating minimal organic pollution, high ecological integrity or near pristine conditions while the boundary between Bands C and D is the national bottom line, with values in Band D indicative of severe organic pollution or nutrient enrichment and severe loss of ecological integrity.

Nutrient toxicity attributes (Amm-N, NO_3-N) were calculated as the annual median of monthly data. Attribute values were compared to the NPS-FM attribute bands and national bottom lines, with Band A indicating little stress or observed effect on freshwater species, Band B indicating impacts on sensitive species, Band C indicating increased impacts and reduced survival of sensitive species, and Band D indicating acute impacts (risk of death) for sensitive species. DRP was also calculated as the annual median of monthly data, with Band A indicating no adverse effects and DRP similar to reference conditions, Band B indicating slight DRP elevation above reference condition and slight impacts, including loss of sensitive species, Band C indicating moderate DRP elevation and loss of sensitive species, and Band D indicating substantial DRP elevation and changes in biotic communities.

Periphyton state was calculated as the 92nd percentile of monthly data, with Band A indicating negligible nutrient enrichment, Band B indicating low nutrient enrichment and occasional algal blooms, Band C indicating moderate nutrient enrichment and periodic nuisance blooms, and Band D indicating high nutrient enrichment and regular or extended-duration nuisance blooms.

3.2 Results

3.2.1 Macroinvertebrate state

In 2021, the most recent year of analysis, 34 sites were sampled for macroinvertebrates and only 2-3 (depending on metric) of these sites were in the A band of the NPSFM, indicating minimal degradation, for any of the three macroinvertebrate attributes calculated using NEMS tolerance values (Figure 3-1, Table 3-1). The majority of sites were in the D band, below the national bottom line and indicative of degraded communities, for MCI and QMCI (16 and 21, respectively), although there were approximately equal number of sites in the B, C, and D bands for ASPM (10, 12, and 9, respectively). The proportion of sites in the A and B bands was higher in 2021 than in the previous seven years for all three macroinvertebrate attributes (Figure 3-2).

Using Northland-specific tolerance values to calculate macroinvertebrate metrics, rather than the national tolerance values provided in the NEMS, resulted in higher metric scores for many sites. The greatest improvement was in MCI scores, with a higher proportion of sites moving into the A band and fewer sites in the D band (Figure 3-1, Table 3-1).

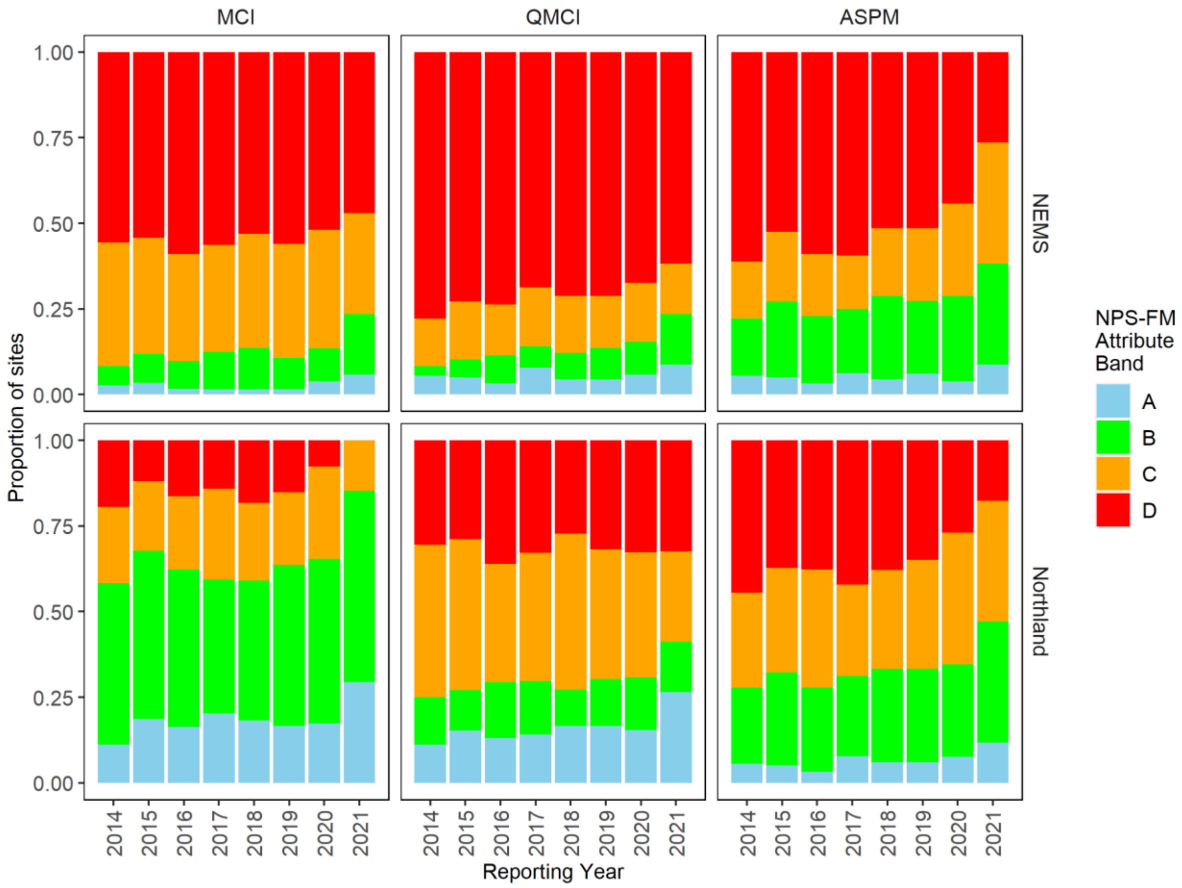

Figure 3-1: Proportion of NRC SoE monitoring sites with macroinvertebrate attribute scores in NPS-FM bands A, B, C, and D each year 2014-2021. Metrics for 2018-2021 were calculated based on 5-year median values, metric scores for 2014-2017 were calculated as the median of the years of available data.

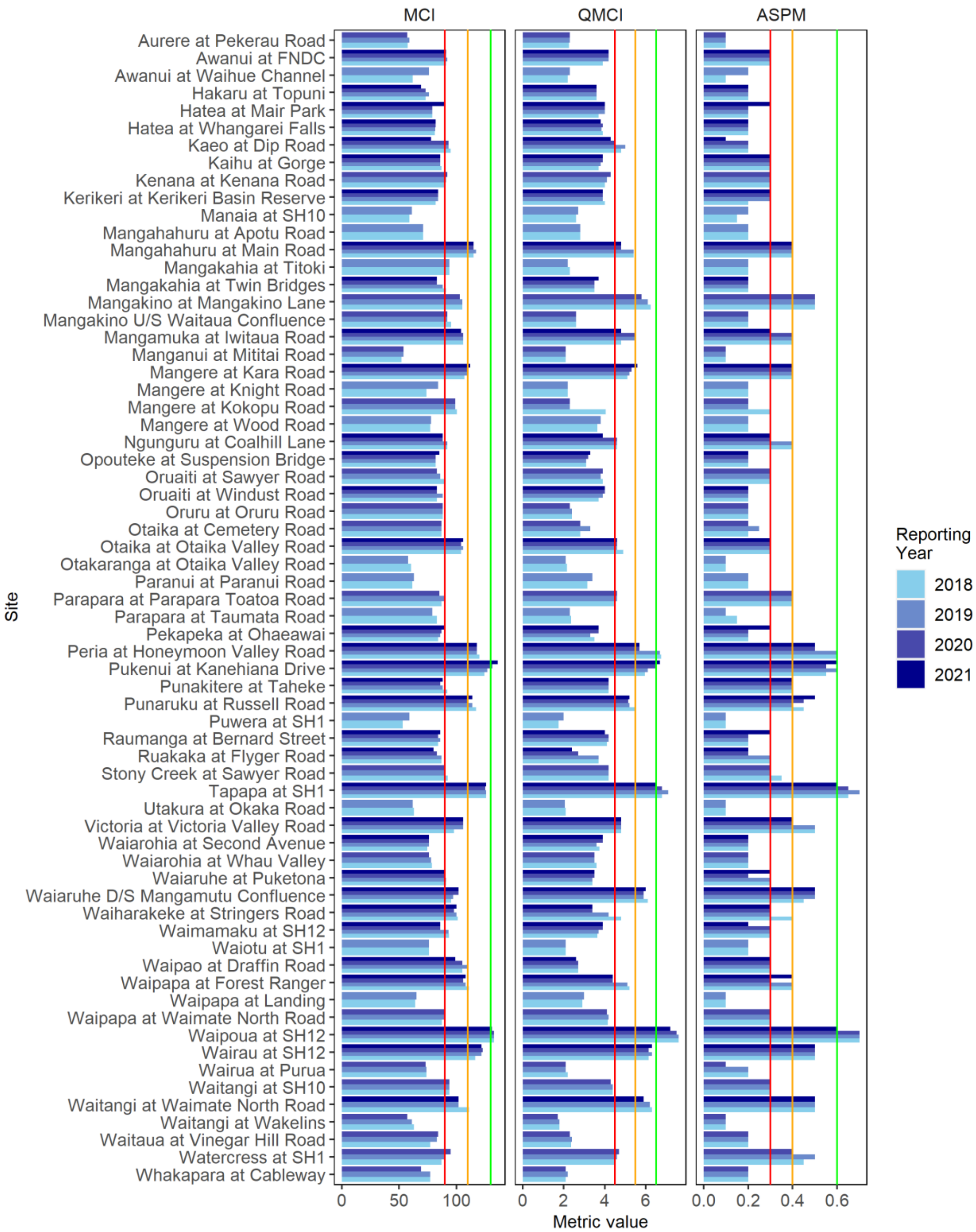

Figure 3-2: Macroinvertebrate attribute states of each SoE monitoring site between 2018 and 2021. State in each year was calculated as median of the previous inclusive five years using the species tolerance values from NEMS (2020). Vertical lines indicate NPSFM attribute bands: green - A/B band boundary, orange - B/C band boundary, red – C/D band boundary = national bottom line.

Table 3-1: Number (percent) of sites in each NPSFM attribute band each year for attribute site calculated using either tolerance values from the NEMS or Northland-specific tolerance values. Note that 'current state' in the NPSFM for each year is the median score over the past 5 years; therefore, only scores for 2018-2021 were able to be calculated over data from the previous 5 years and are true 'current state.' Scores for 2014-2017 were calculated as the median of the years of available data.

Reporting Year	# sites (years)	Tolerance Values	MCI				QMCI				ASPM			
			A	B	C	D	A	B	C	D	A	B	C	D
2014	36 (1)	NEMS	1 (3)	2 (6)	13 (36)	20 (56)	2 (6)	1 (3)	5 (14)	28 (78)	2 (6)	6 (17)	6 (17)	22 (61)
		Northland	4 (11)	17 (47)	8 (22)	7 (19)	4 (11)	5 (14)	16 (44)	11 (31)	2 (6)	8 (22)	10 (28)	16 (44)
2015	59 (2)	NEMS	2 (3)	5 (8)	20 (34)	32 (54)	3 (5)	3 (5)	10 (17)	43 (73)	3 (5)	13 (22)	12 (20)	31 (53)
		Northland	11 (19)	29 (49)	12 (20)	7 (12)	9 (15)	7 (12)	26 (44)	17 (29)	3 (5)	16 (27)	18 (31)	22 (37)
2016	61 (3)	NEMS	1 (2)	5 (8)	19 (31)	36 (59)	2 (3)	5 (8)	9 (15)	45 (74)	2 (3)	12 (20)	11 (18)	36 (59)
		Northland	10 (16)	28 (46)	13 (21)	10 (16)	8 (13)	10 (16)	21 (34)	22 (36)	2 (3)	15 (25)	21 (34)	23 (38)
2017	64 (4)	NEMS	1 (2)	7 (11)	20 (31)	36 (56)	5 (8)	4 (6)	11 (17)	44 (69)	4 (6)	12 (19)	10 (16)	38 (59)
		Northland	13 (20)	25 (39)	17 (27)	9 (14)	9 (14)	10 (16)	24 (38)	21 (33)	5 (8)	15 (23)	17 (27)	27 (42)
2018	66 (5)	NEMS	1 (2)	8 (12)	22 (33)	35 (53)	3 (5)	5 (8)	11 (17)	47 (71)	3 (5)	16 (24)	13 (20)	34 (52)
		Northland	12 (18)	27 (41)	15 (23)	12 (18)	11 (17)	7 (11)	30 (45)	18 (27)	4 (6)	18 (27)	19 (29)	25 (38)
2019	66 (5)	NEMS	1 (2)	6 (9)	22 (33)	37 (56)	3 (5)	6 (9)	10 (15)	47 (71)	4 (6)	14 (21)	14 (21)	34 (52)
		Northland	11 (17)	31 (47)	14 (21)	10 (15)	11 (17)	9 (14)	25 (38)	21 (32)	4 (6)	18 (27)	21 (32)	23 (35)
2020	52 (5)	NEMS	2 (4)	5 (10)	18 (35)	27 (52)	3 (6)	5 (10)	9 (17)	35 (67)	2 (4)	13 (25)	14 (27)	23 (44)
		Northland	9 (17)	25 (48)	14 (27)	4 (8)	8 (15)	8 (15)	19 (37)	17 (33)	4 (8)	14 (27)	20 (38)	14 (27)
2021	34 (5)	NEMS	2 (6)	6 (18)	10 (29)	16 (47)	3 (9)	5 (15)	5 (15)	21 (62)	3 (9)	10 (29)	12 (35)	9 (26)
		Northland	10 (29)	19 (56)	5 (15)	0 (0)	9 (26)	5 (15)	9 (26)	11 (32)	4 (12)	12 (35)	12 (35)	6 (18)

The majority of Northland SoE sites had median metric scores (over the eight years, 2014-2021) below the NPSFM national bottom lines for each metric (Figure 3-3). When metric scores were calculated using the Northland-specific tolerance values, more sites were above the national bottom lines for MCI and QMCI, and to a lesser extent, ASPM (Figure 3-3). Whereas in a national dataset of measured macroinvertebrate metric scores compiled by Canning et al. (2021) for the five years 2012-2016, the majority of sites were above the national bottom lines for all three metrics (Figure 3-3). However, metric scores in the national dataset also spanned a much a larger range.

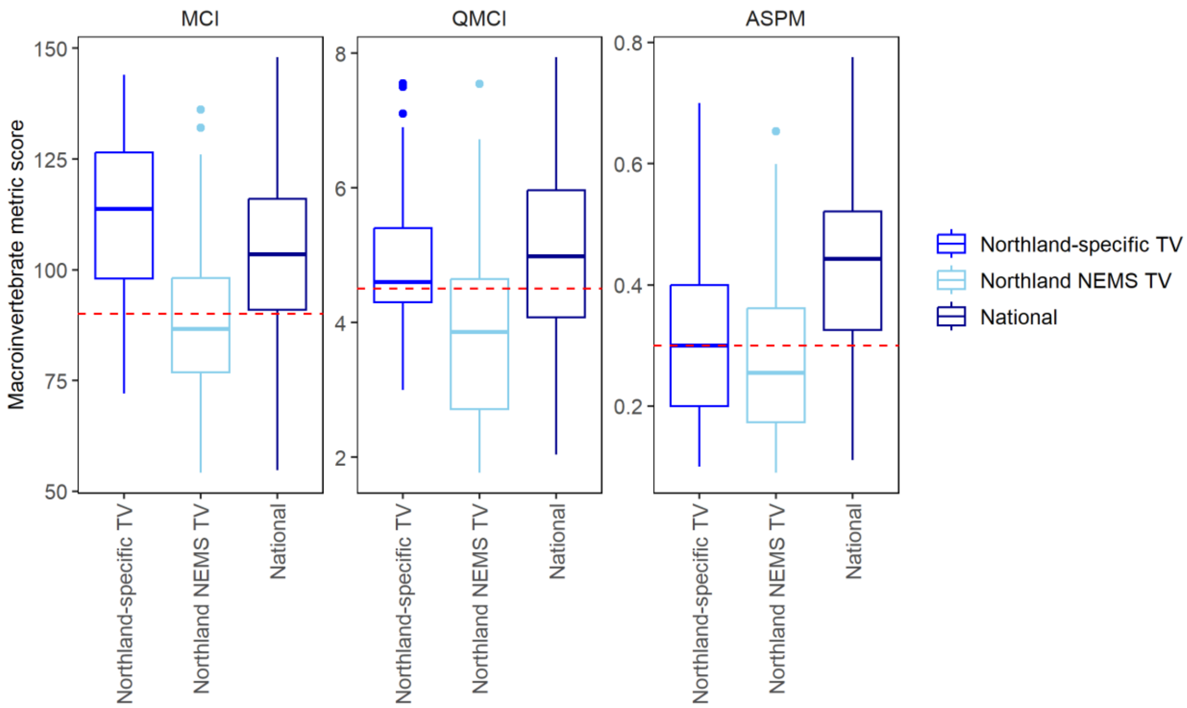

Figure 3-3: Comparison of the distribution of macroinvertebrate metric scores across Northland sites between 2014-2021 compared to the distribution of national metric scores between 2012-2016 from Canning et al. (2021). Macroinvertebrate metric scores for Northland were calculated using either tolerance values from the NEMS (light blue) or Northland-specific tolerance values (blue). Macroinvertebrate metric scores for the national dataset were calculated using NEMS tolerance values. The red dashed line is the NPS-FM national bottom line for each macroinvertebrate metric.

3.2.2 Nutrient and periphyton state

Nutrient toxicity attributes (ammonia and nitrate) were primarily in the A or B band in all years, indicating minimal impact on freshwater species (Figure 3-4). DRP was more variable between sites, with the majority of sites falling in the C or D band in most years (Figure 3-4), indicating moderate impact on freshwater communities and conditions favouring eutrophication. Correspondingly, many (but not all) sites with higher DRP also had higher chlorophyll *a* (Figure 3-4). Three sites were below the national bottom line for chlorophyll *a* (Awanui at FNDC, Hakaru at Topuni, and Waiharakeke at Stringers Road), indicating regular nuisance blooms and possible nutrient enrichment. However, the majority of sites were in the A or B band for chlorophyll *a* in most years, indicating negligible nutrient enrichment and rare algal blooms.

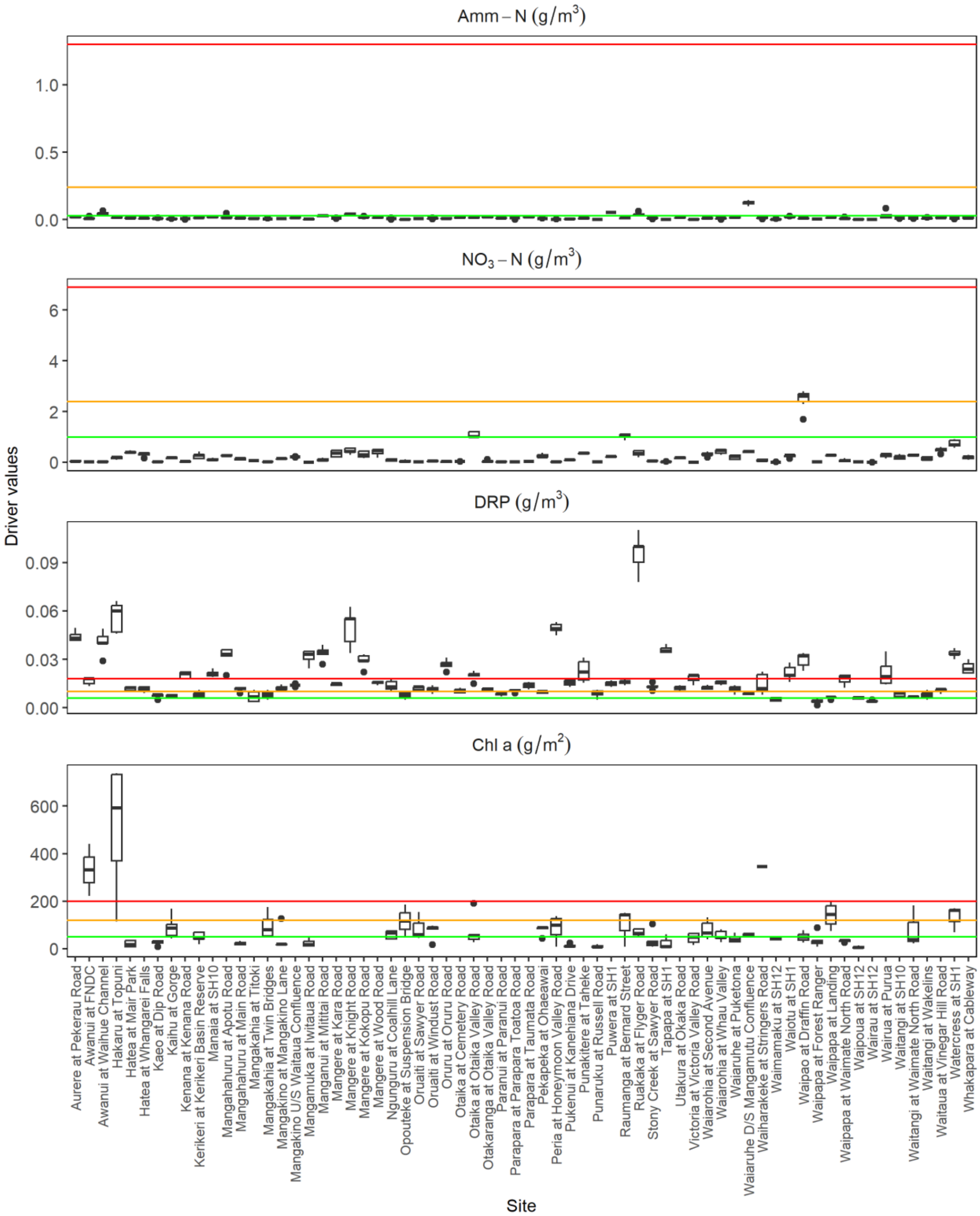

Figure 3-4: Annual median values of NPS-FM attributes in Northland SoE sites. Coloured horizontal lines indicate NPSFM attribute bands: green - A/B band boundary, orange - B/C band boundary, red – C/D band boundary. The C/D boundary is the national bottom line for periphyton (Chl *a*) while the B/D boundary is the national bottom line for ammonia (Amm-N) and nitrate (NO_3-N).

4 Nutrient criteria

The objective of this section was to investigate the applicability of national nutrient criteria developed by Canning et al. (2021) for Northland, specifically answering two questions:

Are the critical values in Canning et al. (2021) for DIN and DRP applicable in Northland rivers?

Are there differences among land use (i.e., predominately pasture, forest, or urban) and stream-bed type (hard-bottomed and soft-bottomed) that may affect the applicability of these nutrient criteria?

4.1 Methods

Datasets of median nutrient concentrations and median macroinvertebrate metric scores per site were developed and used for a minimisation of mismatch analysis following Canning et al. (2021). Nutrient concentrations were summarised as the median of all monthly values over the entire eight-year period for each site, excluding any values from within a year with more than 20% of months missing data. Median concentrations were used to represent typical environmental conditions within a site and for consistency with the analysis in Canning et al. (2021). Invertebrate metric scores were summarised as the median of all annual values over the entire eight-year period for each site. Invertebrate metrics included were QMCI, MCI and ASPM.

4.1.1 Minimisation of mismatch

Nutrient criteria were developed for dissolved inorganic nitrogen (DIN), dissolved reactive phosphorus (DRP), ammoniacal nitrogen (Amm-N) and total Kjeldahl nitrogen (TKN) following the 'minimisation of mismatch' (MoM) method utilised in Canning et al. (2021), which was in turn adapted from the European Union's 'Best practice for establishing nutrient concentrations to support good ecological status' guidelines (Phillips et al. 2018, 2019). Given an ecological target and a set of data (ordered pairs of nutrient-ecological response), the MoM algorithm aims to find a nutrient threshold that balances the numbers of sites in two sets: (a) sites with ecological states at or above the target, but with nutrient concentrations above the threshold; and (b) sites with ecological states below the target, but with nutrient concentrations below the threshold (Figure 4-1). These two sets may be viewed, respectively, as sites that are relatively "over-protected" by the nutrient threshold, given the ecological target, and sites that are relatively "under-protected" by the nutrient threshold, given the ecological target (MfE 2022). Following Canning et al. (2021), ecological metric targets were set as the NPS-FM 'national bottom lines' for macroinvertebrate attributes (e.g., 90 for MCI, 4.5 for QMCI, 0.3 for ASPM).

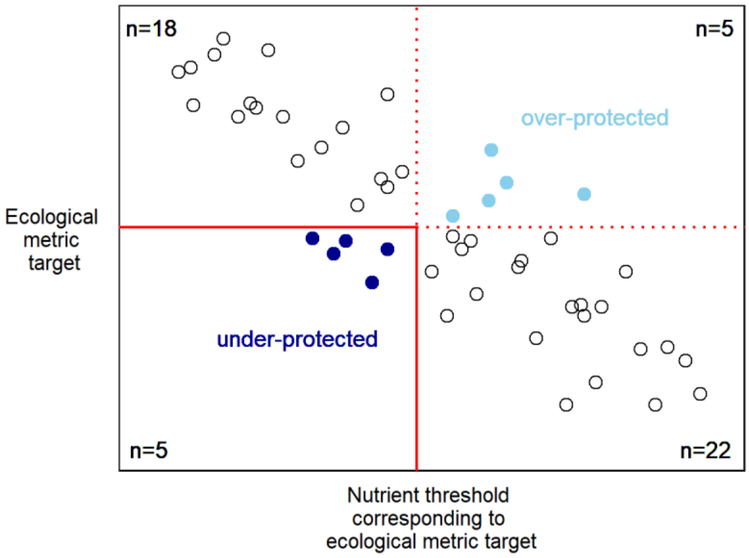

Figure 4-1: Illustration of minimisation of mismatch (MoM) approach. Adapted from MfE 2022.

To calculate the mismatch point, ecological metrics are given a binary classification (good, poor) based on whether the metric falls above or below a designated threshold (in this analysis, NPS-FM national bottom line). Nutrient concentrations are split into a series of bins equating to different potential nutrient boundary values. The percentage of records with either the same or different ecological and nutrient classifications is then calculated for each bin. One loess curve is fit to records which have good ecological status but poor nutrient status, and a second loess curve is fit to records which have poor ecological status but good nutrient status. The intersection of the two curves is the concentration at which mismatch is minimised. Uncertainty in the estimated nutrient criteria is assessed via boot-strapping; the analysis was repeated 500 times with a random sub-sample of 75% of the total data to obtain the mean, median, quantiles, and range of the estimate.

The minimisation of mismatch analysis was run using the site medians from all sites in R using a bespoke R script adapted from that available online from Phillips et al. (2018). Boxplots and bar charts were used to compare how often streams of different land use (pasture, plantation forest, and urban) or substrate types (hard-bottomed or soft-bottomed) exceeded either national or Northland nutrient criteria.

4.2 Results

4.2.1 Appropriateness of the method with NRC data

When we apply the MoM algorithm to setting nutrient thresholds, we assume that:

1. the ecological and nutrient data are normally or uniformly distributed and span the range of values represented within the spatial region of concern (Northland, in this case),

2. there is a significant negative or curvilinear relationship between the stressor and the ecological response,

3. the residuals of the above relationship are centred on zero, and

4. the stressor of concern is not strongly correlated with other environmental drivers.

The NRC macroinvertebrate and nutrient datasets met some, but not all, of these assumptions. Firstly, while all three macroinvertebrate metrics contained values that spanned all four NPS-FM attribute bands, proportionally fewer sites were above the national bottom line, or ecological metric target in the mis-match analysis, than in the national data set used by Canning et al. (2021). Long-term median values (2014 to 2021) indicated that 39%, 27% and 30% of NRC sites were above the ecological metric targets for MCI, QMCI and APSM, respectively. However, in the Canning et al. (2021) data set these values were considerably higher; 78%, 42%, and 69% of sites were above the ecological metric target for MCI, QMCI, and ASPM, respectively. Concentrations of DIN in Northland streams were lower than the national medians reported in Canning et al. (2021) - 0.15 mg/L in Northland vs 0.24 mg/L nationally. DRP, on the other hand, was slightly higher in Northland (median 0.013 mg/L vs 0.0095 mg/L). The range of both DIN and DRP concentrations measured in Northland were smaller than in the national dataset (Table 4-1).

Secondly, while macroinvertebrate metrics were negatively related to nutrient concentrations across sites (Figure 4-2) with similar slopes to those in Canning et al. (2021) from the national dataset, only the relationships with TKN (which was not included in the national analysis) were statistically significant (Table A-1).

Additionally, ranges of nutrient concentrations in Northland sites which were either above or below the NPS-FM national bottom line for each invertebrate metric showed considerable overlap, except for TKN (and to a lesser extent Amm-N), where concentrations were higher in sites below the national bottom line (Figure 4-3). According to Phillips et al. (2018), to calculate ecologically meaningful nutrient criteria, the ranges of nutrient values in each ecological class should not substantially overlap. Failing to meet these assumptions means that the derived criteria should be interpreted with caution, and that the minimisation of mismatch analysis may not be the most suitable method for determining nutrient thresholds for Northland.

Table 4-1: Summary statistics for national and Northland nutrient and macroinvertebrate data. The minimum, median, mean, maximum, 25th and 75th percentiles from Northland SoE monitoring sites compared to measured[2] national data used in Canning et al. (2021).

Metric/ Nutrient	# sites	Minimum	25^{th} percentile	Median	Mean	75^{th} percentile	Maximum
Northland							
MCI	66	54.22	77.1	86.6	88.11	97.97	136.19
QMCI	66	1.8	2.7	3.9	3.8	4.6	7.5
ASPM	66	0.09	0.17	0.25	0.28	0.36	0.65
DIN (mg/L)		0.007	0.044	0.151	0.249	0.333	2.434
DRP (mg/L)		0.0045	0.0095	0.0133	0.0186	0.0213	0.0933
Amm-N (mg/L)		0.003	0.007	0.011	0.014	0.015	0.128
NO_3-N (mg/L)		0.004	0.036	0.140	0.236	0.320	2.425
TKN (mg/L)		0.050	0.150	0.187	0.219	0.273	0.560
National							

[2] Canning et al. (2021) also summarised modelled national nutrient data; this comparison is only for the measured data.

Metric/ Nutrient	# sites	Minimum	25th percentile	Median	Mean	75th percentile	Maximum
MCI	450	54.8	91	103.5	103.2	116	148
QMCI	294	2	4.1	5	5.1	6	7.9
ASPM	389	0.11	0.33	0.44	0.42	0.52	0.78
DIN (mg/L)		0.001	0.0515	0.241	0.5673	0.67	10.5788
DRP (mg/L)		0.0003	0.005	0.0095	0.0161	0.016	0.25

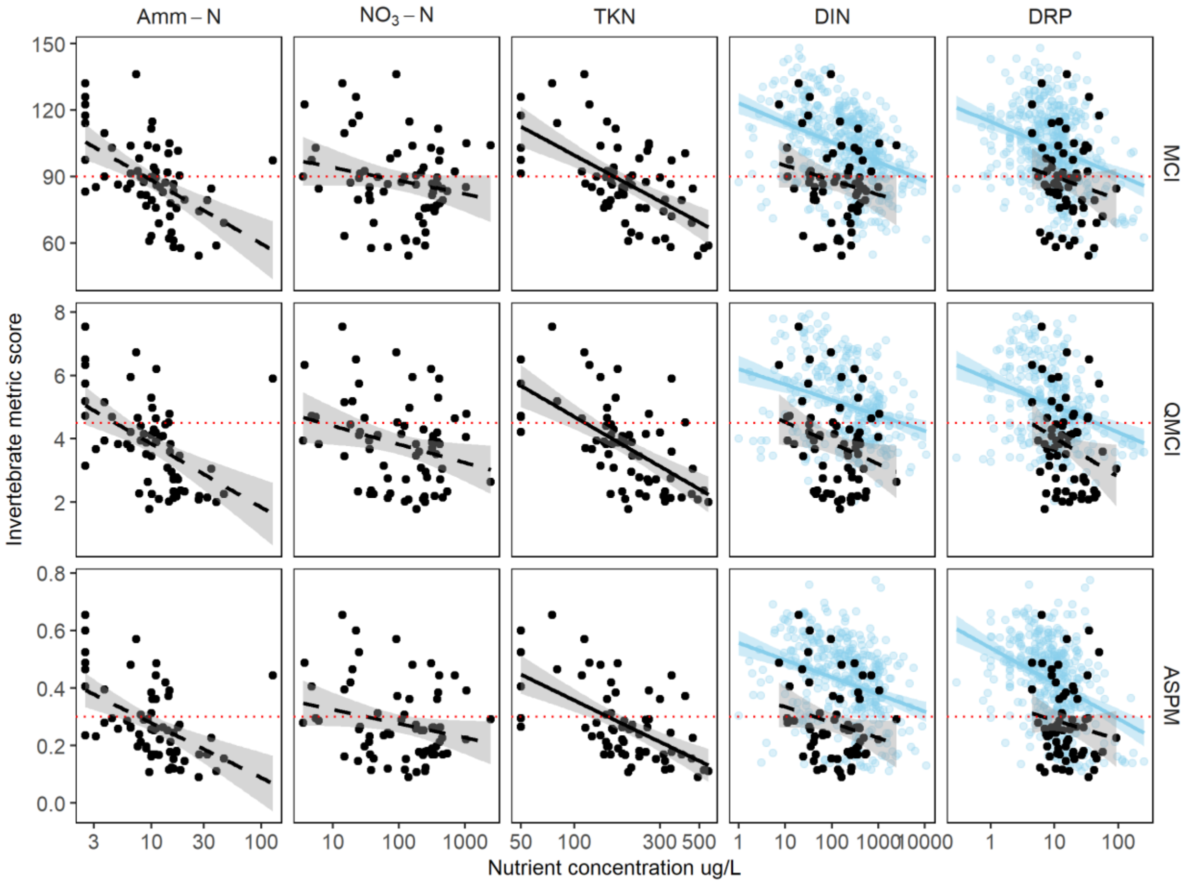

Figure 4-2: Linear regressions between 8-year median macroinvertebrate metrics and nutrients across Northland SoE sites. Solid lines indicate a significant relationship, dashed lines indicate non-significant relationships. The national data used in Canning et al. (2021) for DIN and DRP is shown in blue. Dotted red lines indicate the national bottom line for each macroinvertebrate metric.

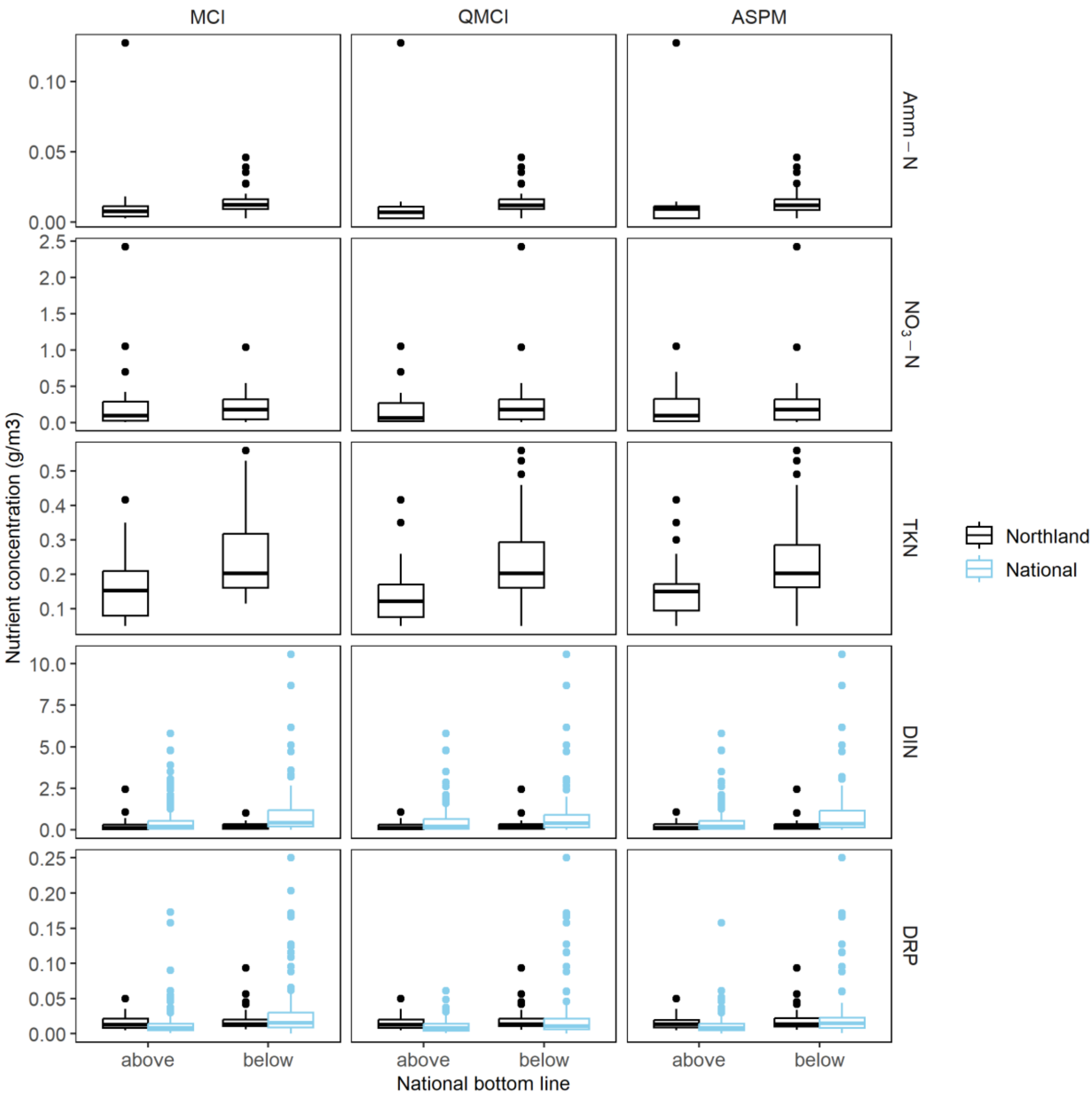

Figure 4-3: Median nutrient concentrations over the eight-year sampling period in Northland sites above or below NPSFM bottom lines for macroinvertebrate attributes. The national nutrient concentration data used in Canning et al. (2021) for DIN and DRP is shown in blue for comparison.

4.2.1 Results of mis-match analysis

The DIN criteria for achieving national bottom lines for MCI and ASPM derived from the Northland dataset were 0.12 (0.05-0.18) mg/L and 0.06 (0.01-0.12) mg/L, respectively (Table 4-2). These criteria values are an order of magnitude lower than the corresponding national criteria estimated by Canning et al. (2021), which were 1.07 mg/L for MCI and 0.63 mg/L for ASPM. Unfortunately, we were unable to identify the mismatch point for QMCI due to high uncertainty in estimates likely arising from the small range of metric scores in the dataset. The DRP criteria for MCI, QMCI, and ASPM were 0.012 mg/L (0.0099-0.015), 0.009 mg/L (0.007-0.012), and 0.01 mg/L (0.007-0.013), respectively (Table 4-2). These criteria values were approximately half the value of the nationally derived criteria (0.03 mg/L for MCI, 0.02 mg/L for QMCI, and 0.03 mg/L for ASPM) (Table 4-2).

The ammoniacal N (Amm-N) criteria to achieve national bottom lines in Northland was 0.009 (0.006-0.011) mg/L for MCI, 0.006 (0.004-0.008) mg/L for QMCI, and 0.007 (0.005-0.009) mg/L for ASPM (Table 4-2). The criteria for NO_3-N to achieve the MCI national bottom line was 0.1 (0.04 – 0.18) mg/L, while the TKN criteria for MCI, QMCI, and ASPM were 0.18 (0.15-0.20) mg/L, 0.14 (0.12-0.16) mg/L, and 0.15 (0.13-0.17) mg/L, respectively (Table 4-2). It was not possible to identify the minimisation of mismatch point for either QMCI or ASPM and NO_3-N. The percentage of sites which remained mis-classified at the new criteria was between 10-20% (Figure 4-4).

Table 4-2: Nutrient criteria derived for Northland using the minimisation of mismatch approach. Summary statistics indicate the range of uncertainty around each nutrient criteria estimate (median value). The national criteria presented in Canning et al. (2021) are included for comparison.

Metric	Nutrient	Minimum	1st Qu.	Median	Mean	3rd Qu.	Maximum
Northland							
MCI	Amm-N	0.0062	0.0083	0.0087	0.0087	0.0091	0.0111
QMCI	Amm-N	0.0041	0.0058	0.0062	0.0061	0.0066	0.0078
ASPM	Amm-N	0.0050	0.0062	0.0066	0.0067	0.0070	0.0091
MCI	NO_3-N	0.036	0.0847	0.1009	0.1035	0.1171	0.1821
MCI	TKN	0.1457	0.1642	0.1704	0.1701	0.1750	0.1951
QMCI	TKN	0.1210	0.1380	0.1442	0.1433	0.1488	0.1627
ASPM	TKN	0.1256	0.1457	0.1503	0.1502	0.155	0.1719
MCI	DIN	0.0480	0.1028	0.1130	0.1150	0.1292	0.1779
ASPM	DIN	0.0074	0.0561	0.0642	0.0679	0.0805	0.1211
MCI	DRP	0.0099	0.0116	0.0121	0.0122	0.013	0.015
QMCI	DRP	0.0070	0.0090	0.0096	0.0096	0.0102	0.0119
ASPM	DRP	0.0073	0.0096	0.0102	0.0103	0.0107	0.0130
National							
MCI	DIN	0.93	1.04	1.07	1.07	1.1	1.21
QMCI	DIN	0.45	0.57	0.63	0.62	0.67	0.77
ASPM	DIN	1.01	1.09	1.12	1.13	1.16	1.29
MCI	DRP	0.025	0.027	0.028	0.028	0.028	0.03
QMCI	DRP	0.015	0.017	0.018	0.018	0.019	0.02
ASPM	DRP	0.026	0.028	0.028	0.028	0.029	0.032

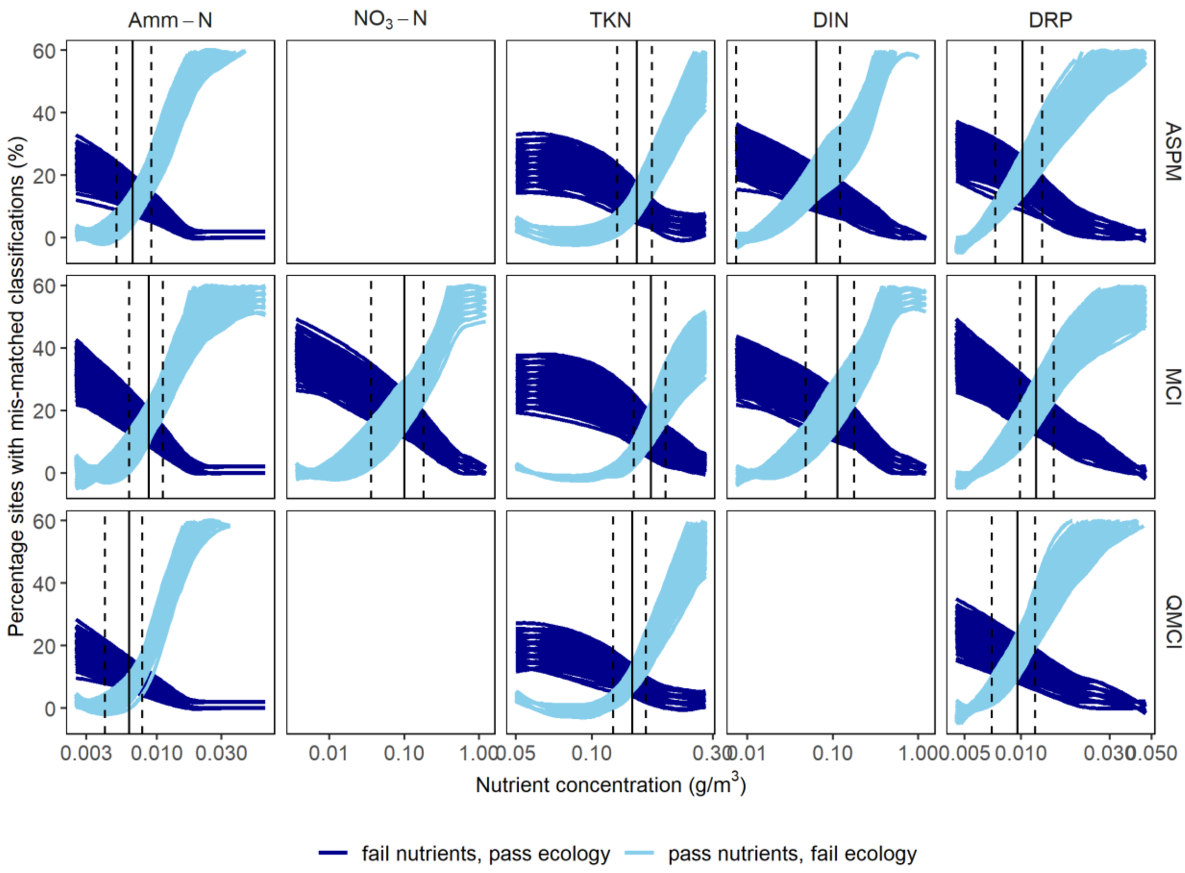

Figure 4-4: Minimisation of mismatch analysis to derive nutrient criteria for Northland. Dark blue lines indicate the proportion of sites which did not meet nutrient targets but passed macroinvertebrate targets (national bottom lines), light blue lines indicate the proportion of sites which met nutrient targets but failed macroinvertebrate targets. The solid vertical black line indicates the median estimated criteria value, while the dashed vertical lines indicate the maximum and minimum ranges of the criteria estimate. Blank panels indicate metric-nutrient combinations for which it was not possible to derive nutrient criteria.

4.2.2 Influence of catchment land use and stream type

Northland streams surrounded by predominately indigenous forest (8 sites) had higher macroinvertebrate metric scores than sites in catchments with exotic forestry (3 sites), which in turn had higher metric scores than streams in pastoral catchments (53 sites, Figure 4-5). Urban streams (2 sites) had the lowest metric scores (Figure 4-5). The reverse was true for DIN concentrations, which were lowest in indigenous forest sites and highest in urban sites (Figure 4-5). Median DRP concentrations were similar across streams in catchments with different land uses, although there was a large range in the indigenous forest sites (Figure 4-5).

All eight indigenous forest sites were below the national nutrient criteria value for maintaining macroinvertebrate attributes for DIN, and seven were below the Northland-derived nutrient criteria (Figure 4-6, Figure 4-7). As mentioned, some of the indigenous forest sites had high DRP concentrations and exceeded both the national and Northland-specific nutrient criteria (Figure 4-6, Figure 4-7). Streams in catchments with exotic forestry also had relatively high metric scores overall, though some sites were below the national bottom lines for all three metrics. Forestry streams had high TKN concentrations, with median concentrations around the Northland criteria value.

The Amm-N and NO_3 concentrations in forestry sites were also around the criteria value for those nutrients, while all three forestry sites had median DIN concentrations above the Northland criteria value. Streams in catchments dominated by pastoral land use had the lowest QMCI metric scores and the highest TKN, while urban streams had the lowest MCI and ASPM scores, as well as the highest nitrate and DIN. Nutrient concentrations in pastoral streams were largely below the national DIN and DRP criteria but above the Northland criteria. Given that there were only two urban streams, one must have been above the national DIN criteria and the other below. Both urban streams (Raumanga at Bernard Street and Waiarohia at Second Avenue) were above the Northland DIN and DRP criteria.

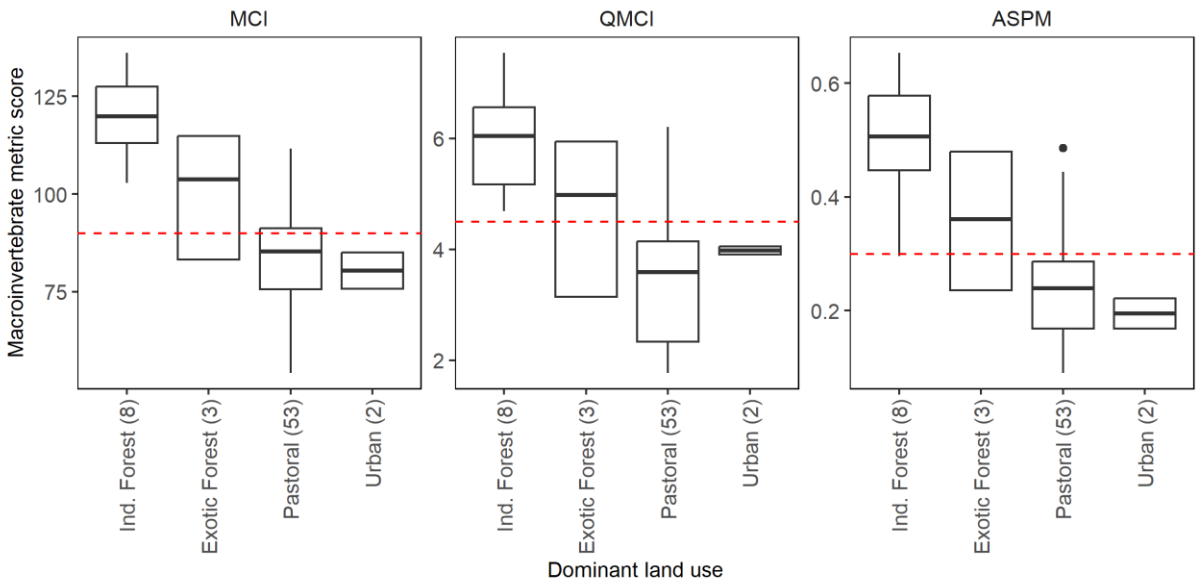

Figure 4-5: Macroinvertebrate metric scores (median over eight years) in Northland sites by catchment land use category. The dashed red line indicates the national bottom line for each metric. The number of sites in each category is given in brackets.

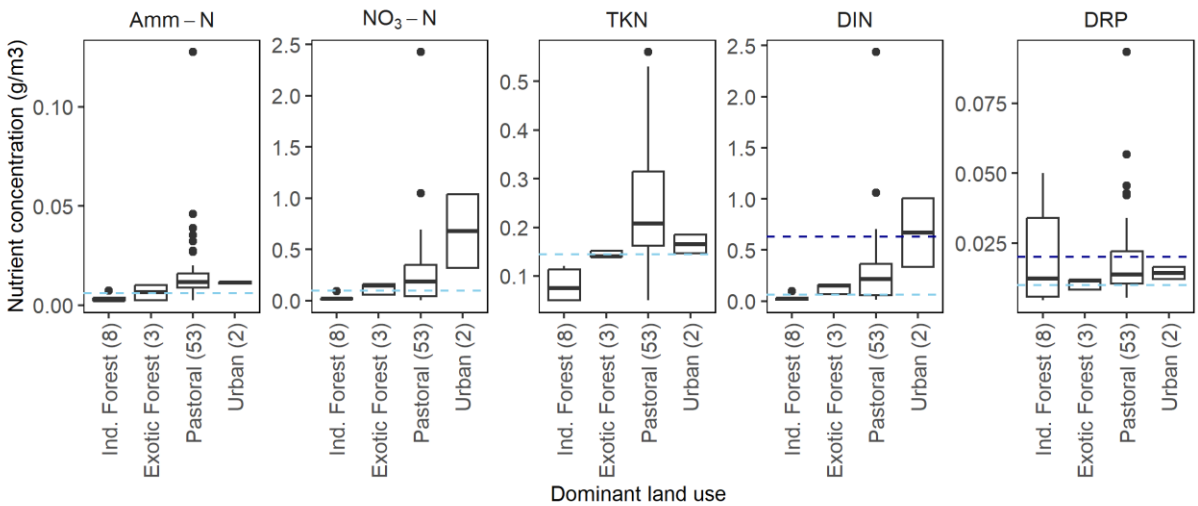

Figure 4-6: Nutrient concentrations (median over eight years) in Northland sites by catchment land use category. The dashed light blue line indicates the Northland-specific nutrient criteria derived in this project; the dashed dark blue line indicates the national criteria derived in Canning et al. (2021). The number of sites in each category is given in brackets.

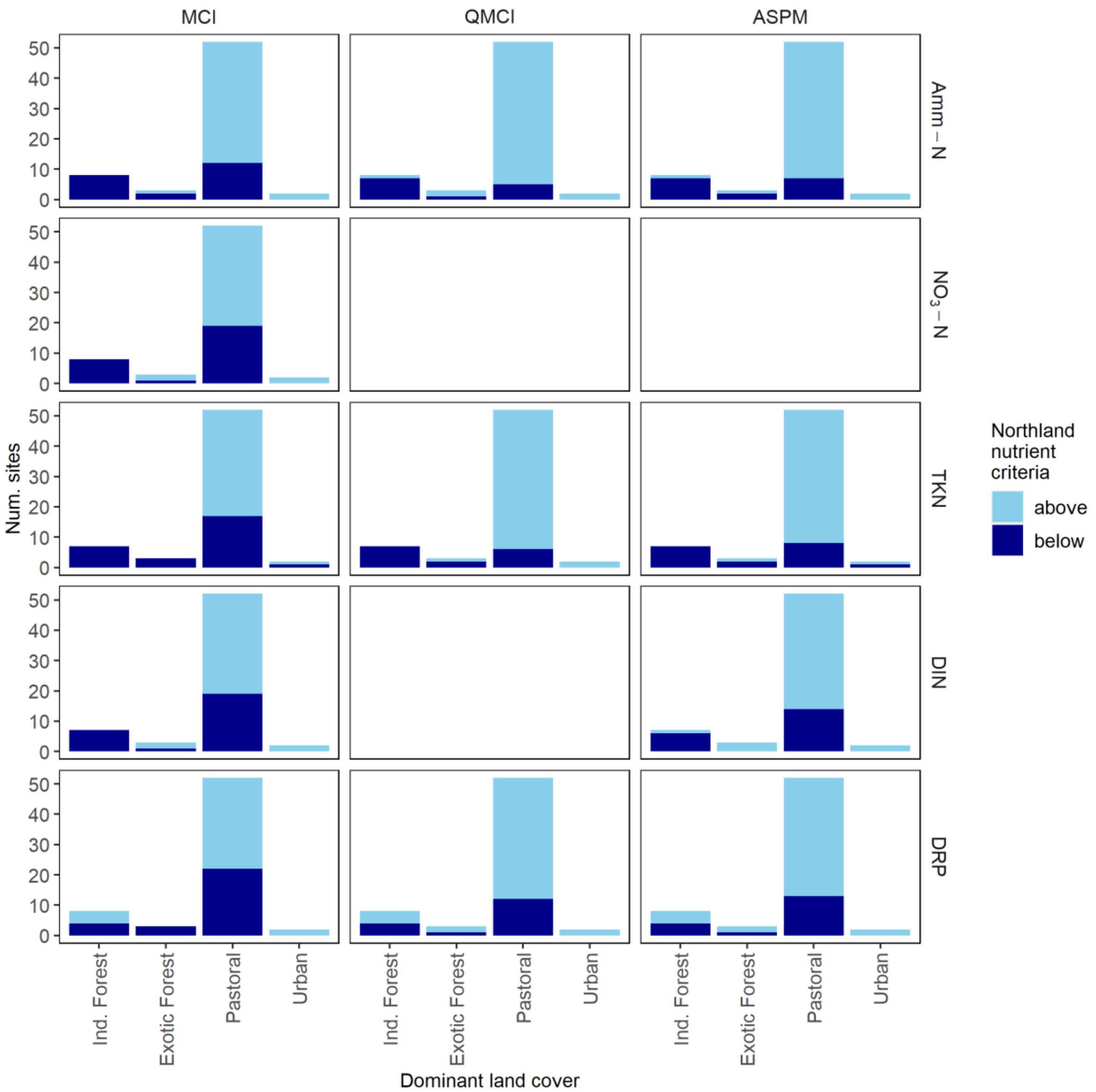

Figure 4-7: Number of sites with median nutrient concentrations over the eight-year sampling period above and below the Northland-derived nutrient criteria within each land use category. . Blank panels indicate metric-nutrient combinations for which it was not possible to derive nutrient criteria.

Macroinvertebrate metric scores were higher in hard-bottomed streams than soft-bottomed streams; about half the hard-bottomed streams scored above the NPSFM national bottom line for each metric, but the majority of soft-bottomed streams were below (Figure 4-8). Northland hard-bottomed and soft-bottomed streams had similar concentrations of DIN, NO_3, and to a lesser extent DRP (Figure 4-9, Figure 4-10). Both stream types were below the national DIN criteria; hard-bottomed streams were also below the national DRP criteria, but some soft-bottomed sites were above (Figure 4-9, Figure 4-10). Both stream types were largely above the Northland-specific DIN and DRP criteria. Slightly over half of the sites in both stream types were above the Northland NO_3-N criteria.

Soft-bottomed streams had higher concentrations of TKN and ammoniacal N than hard-bottomed streams, and most soft-bottomed sites were above the Northland-specific criteria for these nutrients.

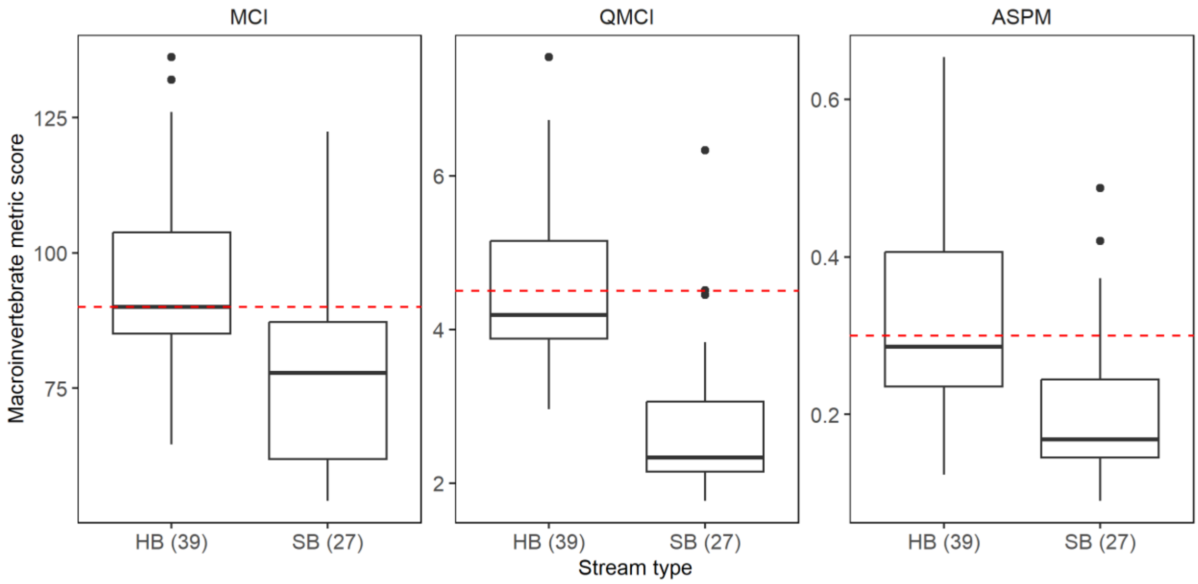

Figure 4-8: Macroinvertebrate metric scores (median over eight years) in Northland sites by stream type. HB is hard-bottomed streams, SB is soft-bottomed streams. The number of sites in each category is given in brackets. The dashed red line indicates the national bottom line for each metric.

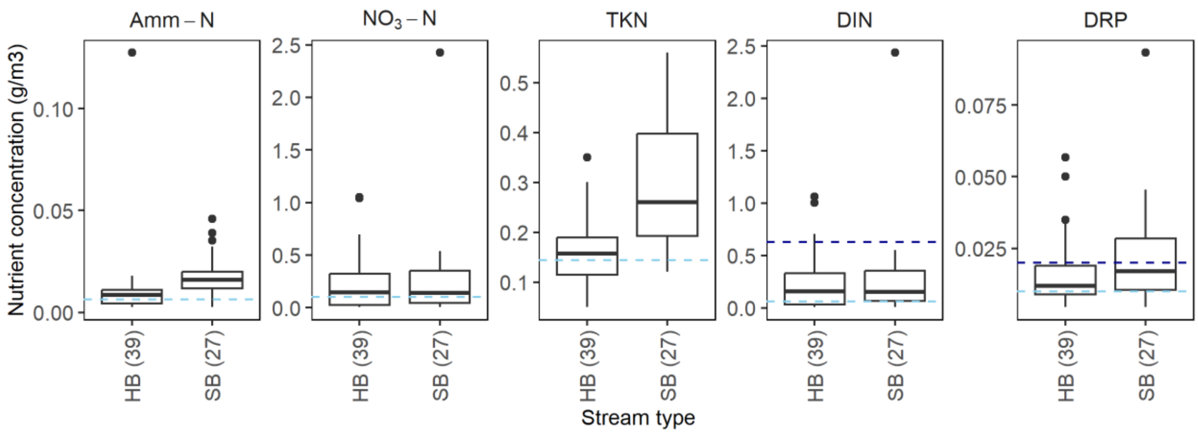

Figure 4-9: Nutrient concentrations (median over eight years) in Northland sites of different stream type. HB is hard-bottomed streams, SB is soft-bottomed streams. The number of sites in each category is given in brackets. The dashed light blue line indicates the Northland-specific nutrient criteria derived in this project; the dashed dark blue line indicates the national criteria derived in Canning et al. (2021).

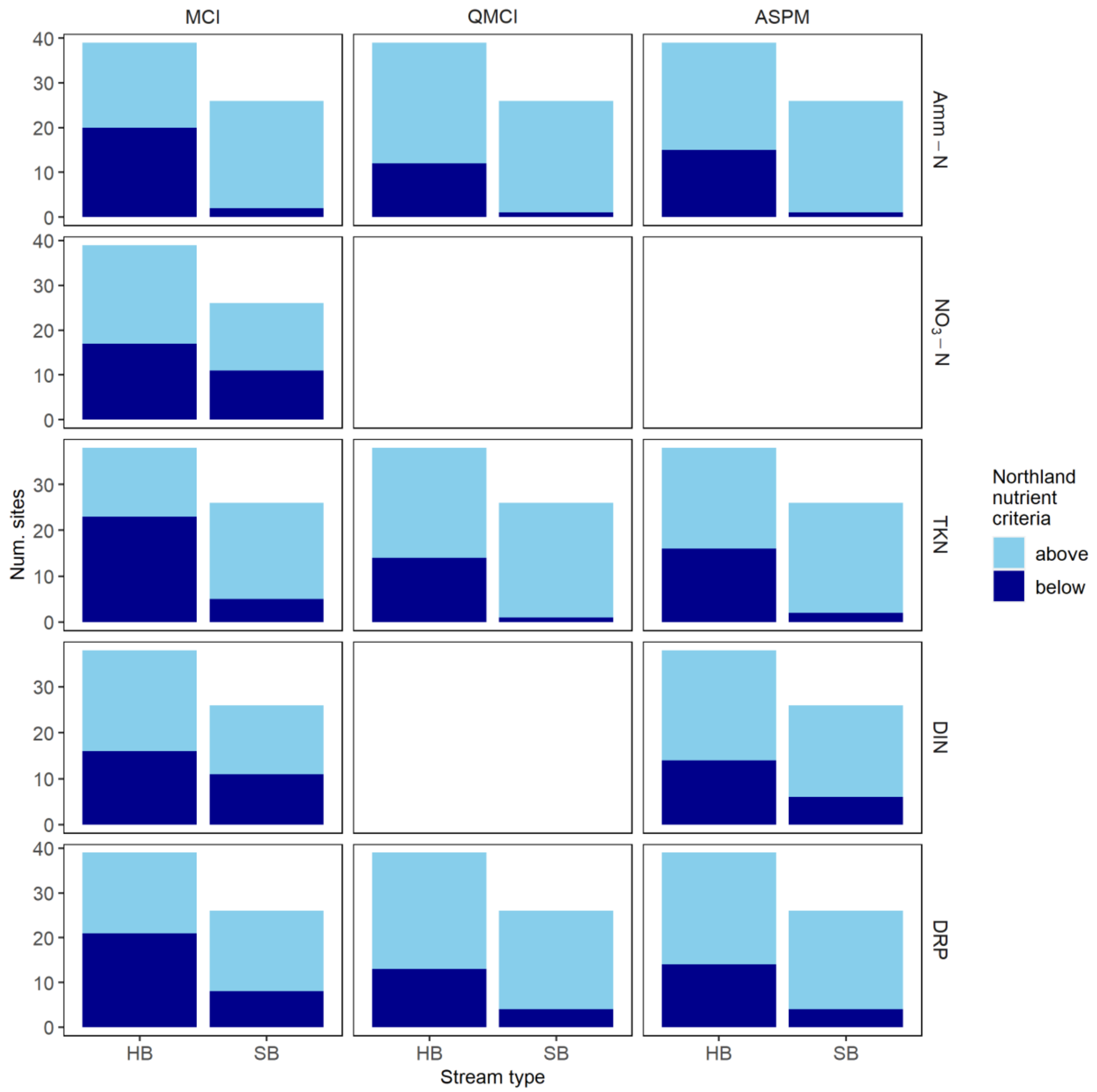

Figure 4-10: Number of sites with median nutrient concentrations over the eight-year sampling period above and below the Northland-derived nutrient criteria by stream type. HB is hard-bottomed streams, SB is soft-bottomed streams. Blank panels indicate metric-nutrient combinations for which it was not possible to derive nutrient criteria.

4.3 Discussion

The NRC dataset meets some, but not all, of the assumptions required for the analysis of mismatch approach:

1. Macroinvertebrate metric scores span all four NPSFM attribute bands from minimal to severe impact of organic pollution or loss of integrity, but the distribution is uneven, with an over-representation of points with both low scores and relatively lower nutrient values. Over-representation of any combination of nutrient and ecological metric targets (i.e., pass nutrients, fail ecology or pass ecology, fail nutrients) will influence the MoM results.

2. Only TKN shows significant linear relationships with the macroinvertebrate metrics. The other nutrients do not.
3. Nutrient concentrations in sites that either exceed or are below the ecological targets (macroinvertebrate metric national bottom lines) show considerable overlap in DIN and DRP. The Canning et al. (2021) dataset also showed some, although to a lesser degree, overlap in nutrient concentrations between sites that either meet or did not meet the ecological target. In reality, multiple stressors impacting macroinvertebrate communities will likely make this assumption difficult to meet using field data.

The fact that the NRC data set did not fully meet the assumptions of the analysis means that care should be taken in interpreting the output of the analysis.

Are the critical values in Canning et al. (2021) for DIN and DRP sufficient to maintain the NPSFM macroinvertebrate attributes above the national bottom line in Northland rivers?

The Northland nutrient criteria for both DIN and DRP were substantially lower (i.e., more stringent) than the nationally criteria; DIN by an order of magnitude and DRP by approximately half. However, the low criteria are likely an artifact of the over-representation of sites with low nutrients and low metric scores, rather than an ecologically meaningful threshold, as there is no strong evidence that these nutrients are a key stressor of macroinvertebrate communities across the region.

Sites with high TKN concentrations commonly had more degraded macroinvertebrate communities, indicating a potential role of TKN as a stressor. TKN and ammoniacal N make approximately two-thirds of the organic nitrogen load in Northland pastoral streams (Rissmann et al. 2020). However, pastoral streams are often impacted by multiple stressors in addition to nutrients which commonly affect macroinvertebrate communities. Many of the pastoral streams in Northland are soft-bottomed and located in poorly-drained low-lying floodplains (Rissmann et al. 2018). Correlations between potential stressors of macroinvertebrate communities can make it hard to disentangle the role of nutrients.

Causation is also difficult to identify when multiple correlated potential stressors are present. For example, TKN frequently co-varies with sediment (Vidon et al. 2008), which is also known to negatively impact macroinvertebrate community composition (Burdon et al. 2013). Moreover, while anthropogenic sources of organic nitrogen, such as stock effluent or fertiliser (Vidon et al. 2008), may be causing higher TKN concentrations in pastoral and/or soft-bottomed streams, natural sources of organic carbon, such as upstream wetlands or macrophyte beds, may be likely in some locations. TKN has also been found to be associated with peat and lacustrine deposits, as well as lignite and mudstone geology, in Northland (Rissmann and Pearson 2020). Additionally, the low-lying floodplains which cover 56% of Northland are poorly drained and high in organic matter and therefore organic N. Further investigation into whether the source of organic nitrogen is anthropogenic or natural would be valuable.

DRP concentrations in Northland are also strongly associated with the underlying geology, particularly volcanic substrates (Rissmann et al. 2019, Rissmann and Pearson 2020, LAWA 2022), rather than anthropogenic impacts. Median DRP concentrations were similar across land use categories, including indigenous forest sites, and several sites with high DRP also had high macroinvertebrate metric scores, indicating that DRP was not affecting macroinvertebrate communities.

Freshwater macroinvertebrate communities are generally influenced by in-stream nutrients via two main pathways. Firstly, very high nutrient concentrations can cause direct toxic effects on macroinvertebrates. However, 95% of the sites in NRC had median nitrate concentrations within band A of the nitrate toxicity attribute under the NPSFM, and 92% of sites had median ammoniacal N concentrations within band A of the ammonia toxicity attribute, indicating minimal likely toxic effects at almost all sites. Secondly, high nutrient concentrations, combined with a lack of other limiting factors such as adequate light, warm temperature and infrequent scouring floods, can lead to excessive periphyton growth such as blooms of filamentous algae or thick mats. Such changes in the biomass or growth form of periphyton can reduce habitat suitability and alter food availability for macroinvertebrates, resulting in shifts in community composition (Tonkin et al. 2014), generally leading to a decline in biodiversity and macroinvertebrate metric scores. Periphyton data was not available for all sites at all dates, however 45% of sites with periphyton measurements had median mg/m^2 of chlorophyll *a* within band A and 39% within band B of the NPS-FM, also indicating likely limited impacts of periphyton on macroinvertebrate community communities at the majority of sites. A small number of sites, however, have either nutrient concentrations at levels that may begin to impact macroinvertebrate communities directly (e.g., nitrate at Waipao at Draffin Road) or have high periphyton biomass (e.g., Awanui at FNDC, Hakaru at Topuni, and Waiharakeke at Stringers Road). Further investigation and perhaps management of nutrient concentrations at these sites could be beneficial.

Our results suggest that the minimisation of mismatch approach may not be the most suitable method for setting nutrient criteria in Northland, and that further investigation of the role of nutrients in impacting macroinvertebrate communities is needed before large effort or expense is undertaken in reducing nutrient concentrations to below the criteria identified here. There may be other more important stressors to concentrate on first.

Are there differences among land use (i.e., predominately pasture, forest, or urban) and geology classes that may affect the applicability of these nutrient criteria?

Most pastoral and urban streams exceeded the Northland nutrient criteria for all forms of nitrogen, while hard- and soft-bottomed streams were evenly split between exceeding and meeting nutrient criteria for NO_3 and DIN. More soft-bottomed streams exceeded criteria for ammoniacal N and TKN. Indigenous forest sites were below nitrogen criteria levels. Reducing nitrogen concentrations in soft-bottomed pastoral streams to meet the criteria levels could enable improvements in macroinvertebrate communities in these sites, if nutrients are a key driver of community composition. However, as discussed above, it will be first be necessary to confirm whether the source of organic nitrogen is natural or anthropogenic, and investigate the influence of potential covariates, particularly sediment. Except for a few indigenous forest sites, sites in all land use categories exceeded the Northland criteria for DRP. However, given that median DRP was similar across land use categories, it is likely that DRP concentrations are primarily associated with the volcanic substrates common in Northland, rather than any anthropogenic impact, and cannot be reduced by management.

5 Drought and other drivers

The objective of this Section was to investigate drivers of macroinvertebrate communities in Northland, specifically:

1. *How do drought conditions impact macroinvertebrate community composition in Northland?*

2. *Do drought conditions impact water quality and environmental variables that may influence macroinvertebrate community composition?*

3. *What are the key drivers of macroinvertebrate communities of Northland streams?*

4. *Do drivers vary between streams with different stream type (hard- or soft-bottomed) or surrounding land use (pastoral or indigenous forest)?*

The influence of drought directly on macroinvertebrate communities and indirectly through changes on potential drivers of macroinvertebrate communities were assessed in several ways. Firstly, national and regional drought metrics were used to identify drought years. Secondly, values of macroinvertebrate metrics and potential driver were visualised in drought and non-drought years. Thirdly, time series datasets and spatial datasets were analysed using GAMM models to investigate relationships between potential drivers of macroinvertebrate communities, including drought metrics, and macroinvertebrate metrics, and to assess the relative importance of each driver. Details of data processing and analyses are provided below.

5.1 Methods

5.1.1 Data processing and driver selection

A large dataset with 152 potential predictors of macroinvertebrate community composition across 66 sites was available (see Section 2 for details and collation information). Individual predictors varied in the number of sites and time window over which data was available. In particular, periphyton and substrate information was limited spatially and temporally. This led to the selection of two datasets for analyses:

1. A time series data set that incorporated data from as many years at as many sites as possible. Periphyton and substrate were not included in this data set as both lacked spatial and temporal coverage.

2. A spatial dataset where single values of predictors were summarised as overall medians for each site. Substrate and periphyton information were included in this dataset, and drought (NZDI) was not (as it represented temporal rather than spatial effects).

Each dataset contained too many potential drivers to include in modelling analyses (Table 5-1). We therefore attempted to select the most relevant potential drivers using a combination of methods:

1. Expert opinion gathered from discussion with NRC staff and their local knowledge, as well as knowledge of likely mechanisms of stressor impact on macroinvertebrates.

2. Examination of variation in NPS-FM attributes across Northland.

3. Principal components analysis (PCA) to identify groupings of related variables.

4. Identification of correlated variables using pairwise scatterplots and Pearson's correlation coefficients.
5. An automatic selection method (full subset selection, FSS) of statistical models to identify most important predictors.

The PCA was run using both the time series dataset and the spatial data set with the rda function from the 'vegan' package (Oksanen et al. 2022) in R. Based on the groupings from the PCA, subsets of potential drivers were filtered for correlated variables using pairwise scatterplots and Pearson's correlation coefficients (Appendix C). Many variables were inherently correlated, such as proportions of periphyton or sediment cover and flow values summarised over windows of varying duration (i.e., one month prior to sampling, 3 months prior to sampling). Consequently, a single variable was chosen to represent the influence of that driver group (Table 5-1). Likewise, the different summary statistics of each driver were typically highly correlated (i.e., median DIN correlated with 95th percentile DIN). Therefore, annual medians were used for all drivers except chlorophyll *a*, where the annual 92nd percentile was used. The reduced set of drivers chosen for the model selection analysis are listed in (Table 5-1).

Table 5-1: Selected drivers for GAMM models. Drivers were selected based on data availability (number of years and sites sampled) and to minimize inclusion of correlated variables. s indicates drivers included in the spatial model, t indicates drivers included in the time series models.

Driver/Response category	Indicators	Units	Frequency	years with data	Sites/year	Correlations	Final selection
Response							
Macroinvertebrate community metrics	QMCI, MCI, ASPM, %EPT taxa, %EPT abun.		Annually	8	34-66	All	All
Potential drivers >50 sites							
Nutrients	Ammoniacal N	g/m^3	Monthly	7	28 - 66	TKN, TN	Amm-N (s, t)
	DIN	g/m^3		7	27 - 63	TN	DIN (s, t)
	Nitrate-N	g/m^3		7	27 - 66	Nitrate-Nitrite-N, TN	
	Nitrite-N	g/m^3		7	27 - 65	Nitrate-Nitrite-N	
	Nitrate-Nitrite-N	g/m^3		7	28 - 66	Nitrate-N, TN	
	TKN	g/m^3		7	27 - 65	Amm-N, TN	TKN (s, t)
	TN	g/m^3		7	28 - 65	DIN	
	DRP	g/m^3		7	28 - 66	TP	DRP (s, t)
	TP	g/m^3		7	28 - 65	DRP	

Driver/Response category	Indicators	Units	Frequency	years with data	Sites/year	Correlations	Final selection
Other water chemistry	Conductivity	µs/cm	Monthly	7	28 - 65		Conductivity (s, t)
	Turbidity - median	NTU		7	28 - 65	Black disc, TSS	Turbidity (s, t)
	Black disc	M		7	24 - 60	Turbidity	
	TSS	g/m^3		3	23 - 41	Turbidity, VSS	
	VSS	g/m^3		2	27 - 41	TSS	
	DO	mg/L, % sat		7	28 - 65		DO (s, t)
	Temperature	deg. C		7	28 - 64		Temperature (s, t)
Drought metrics	NZDI	-	Daily	8	34 - 66		NZDI (t)
	SDI	-	Annually	7	18 – 21		
Flow metrics (measured and modelled data)	Long term flow	m^3/s	Annually	8	66	All flow, ALF	
	Days since last flow (daFRE) 3 and 10 x long term median flow	days	Daily	8	66		daFRE3 (s, t)
	Annual low flow (ALF)	m^3/s	Annually	8	66	All flow, long-term flow	
	Base flow index (BFI)	-	Daily	8	66		BFI (s, t)
	Flow - min/mean/max/median over 7, 30, 90, 365 days prior	m^3/s	Daily	8	66	All flow, ALF	Flow 90 days prior (s, t)
In-stream habitat	RHA	-	Annually	6	33 - 65		RHA score (s)

Driver/Response category	Indicators	Units	Frequency	years with data	Sites/year	Correlations	Final selection
Catchment characteristics	Us slope	degrees	-	8	66		Us slope (s, t)
	Us elev	m		8	66		Us elev (s, t)
	Us rnvar	mm		8	66		Us rnvar (s, t)
	Us rd10	days/month		8	66		Us rd10 (s, t)
	Us_catarea	m^2		8	66	Flow	
	Loc_psize	-		8	66		Loc_psize (s, t)
	Loc_mat	degrees C		8	66		Loc_mat (s, t)
Potential drivers <50 sites							
Periphyton and macrophytes	Chlorophyll *a*	mg/m^2	Monthly	6	12 - 33	All % cover, WCC	Chlorophyll *a* (s)
	WCC	%		6	6 - 28	All % cover, Chl *a*	
	Cover of different categories	%		6	6 - 28	Chl *a*, WCC	
	Macrophytes	%		2	12-23		
Deposited fine sediment and substrate	% cover substrate categories	%	Monthly	2	31 - 34	Embeddedness, % cover	% sand-silt (s)
	Embeddedness categories	%		2	31 - 34	Embeddedness, % cover	
	Total deposited sediment cover	%		1	16	% cover	

5.1.2 Generalized additive models

Generalized additive mixed models (GAMMs) with a full subsets approach (Fisher et al. 2018) was used to identify potential drivers of macroinvertebrate community composition using both the time series and the spatial dataset. Drought metrics were included in all analyses using the time series dataset. GAMMs were chosen because they can model nonlinear relationships, which are common in ecological data, using smooth functions, or splines. GAMMs can also include random effects to account for the non-independence of time series data collected from the same site and accommodate a variety of distributions (i.e., normal, Poisson, or binomial). Models were fit for invertebrate metrics MCI, QMCI, ASPM, and percent EPT taxa and abundance. The MCI and QMCI models were fit using a Gaussian error distribution, the percent EPT taxa and percent EPT abundance models were fit using a binomial distribution, and the ASPM models were fit using a beta error distribution (because values were bounded between 0 and 1). Predictor variables were log or square-root transformed as necessary to improve distribution. Site was included as a random effect in models using the time series dataset. Reporting year was included as a fixed effect in the full models to account for temporal differences in macroinvertebrate metrics. It was not included as a random term due to relatively small number of years (eight) and because we were interested in assessing the influence of sampling year on the macroinvertebrate metrics.

A complete set of possible models was created using all combinations of predictor variables. Predictors were fit as cubic regression splines with a maximum of five knots. Correlated predictors were excluded from being in the same model but retained in the model set (Fisher et al. 2018). Models were compared using the Akaike Information criterion (AIC) for small sample sizes (AICc). The relative importance of predictor variables was assessed by summing the AICc weights for all models containing each variable, while the most parsimonious model was selected based on the lowest AICc and least number of predictor variables (Burnham and Anderson 2002). All GAMM and full subset analyses were done using functions from the 'mcgv' (Wood 2011, Wood et al. 2016), and 'FSSgam' (Fisher et al. 2018) packages in R. For time-series data sets the variation in metric scores explained by temporal variation in predictor values once between site differences had been accounted for was assessed by subtracting the null model (random site effect only) R^2 from each model R^2.

The FSSgam analysis was run using the full spatial and timeseries datasets, as well as using subsets of hard-bottomed streams only, soft-bottomed streams only, streams with primarily pastoral land use in the catchment, and streams with primarily indigenous forest in the catchment. Due to the small numbers of urban streams (2) and streams with primarily exotic forestry (3) in the catchment it was not possible to test models for these subsets.

Drought

The effects of drought on macroinvertebrate communities were assessed by including the New Zealand drought index (NZDI) and flow parameters as predictors in the generalized additive models using the time series dataset. The effects of drought on other environmental drivers were also assessed by fitting GAMMs for each individual driver and mean NZDI, with Site as a random effect.

5.2 Results

5.2.1 Principal components analysis

The PCA identified two general groups in the principal components analysis: flow-related parameters along axis one and water quality and nutrients along axis 2 (Figure 5-1). Therefore, we tested for correlations between individual drivers within each group (Appendix C) and chose a representative subset of drivers from each group to include in the modelling analyses. Spatial attributes were largely associated with axis 2 as well, except for catchment area (i.e., us_catarea) which clustered with the flow parameters and variation in annual catchment rainfall (i.e., us_rnvar), which was on axis one in the PCA of site medians. Sites clustered together by land use, particularly the pastoral sites, and to a lesser degree by stream type with land use (Figure 5-1).

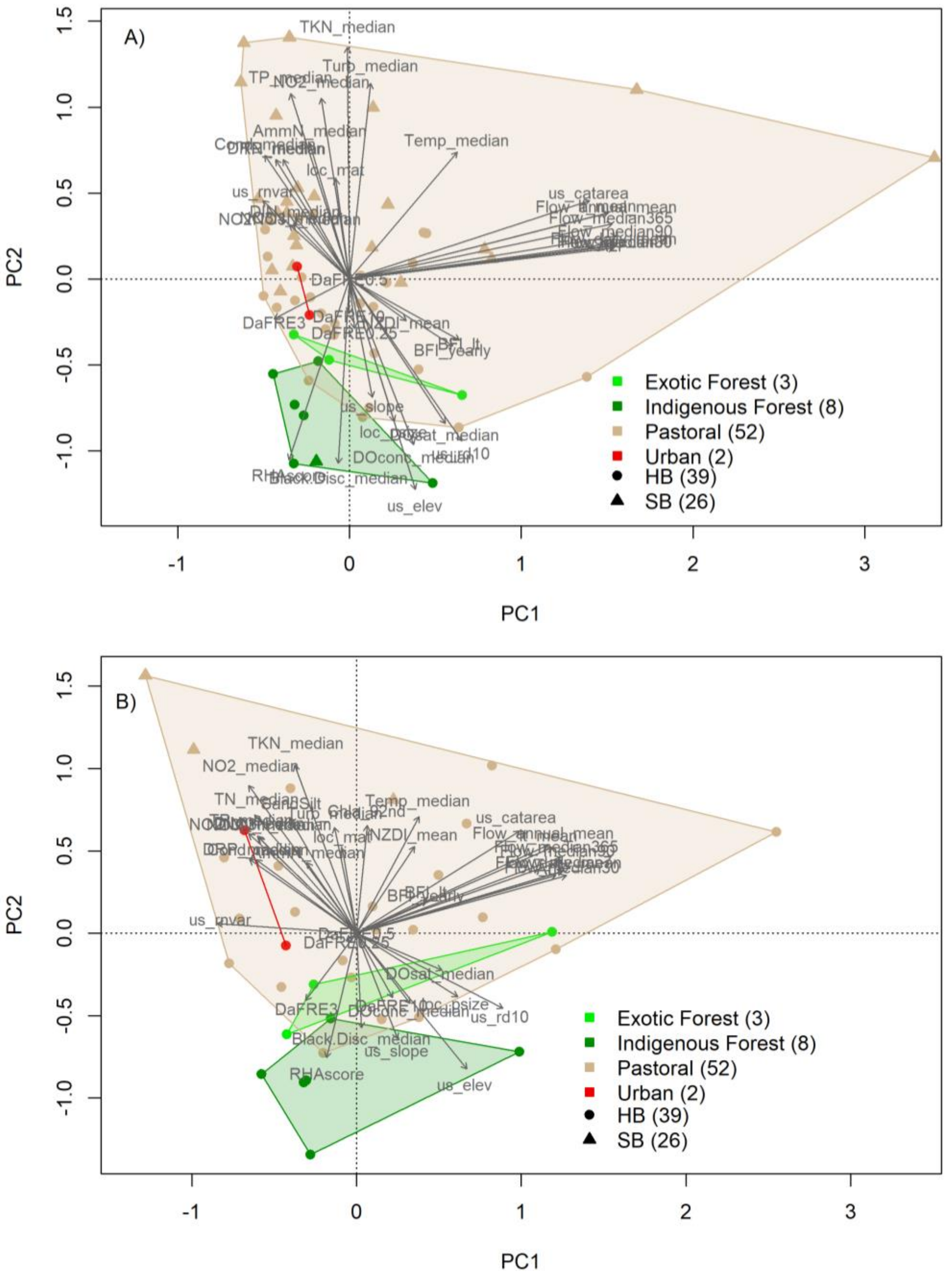

Figure 5-1: Principal components analyses with potential environment drivers across all sites. A) Full time-series dataset, B) Spatial dataset (site median values for each driver). The colour of the points and polygons indicates the land use type: light green – exotic forest, dark green – indigenous forest, light brown – pastoral, red – urban. The shape of the points indicates the stream type: circle – hard-bottomed, triangle – soft-bottomed.

5.2.2 Drought

There were three periods of climatic drought (NZDI > 1.5) across all three Northland regions during the time period covered by the macroinvertebrate dataset: summers 2014, 2017, and 2020 (Figure 5-2, Table 5-2). There was also a brief drought period in summer 2021 in the Far North District, and two additional drought periods in the Kaipara district – late summer 2014 and early summer 2018. Drought periods identified by the NZDI also corresponded to low river flows (Table 5-3).

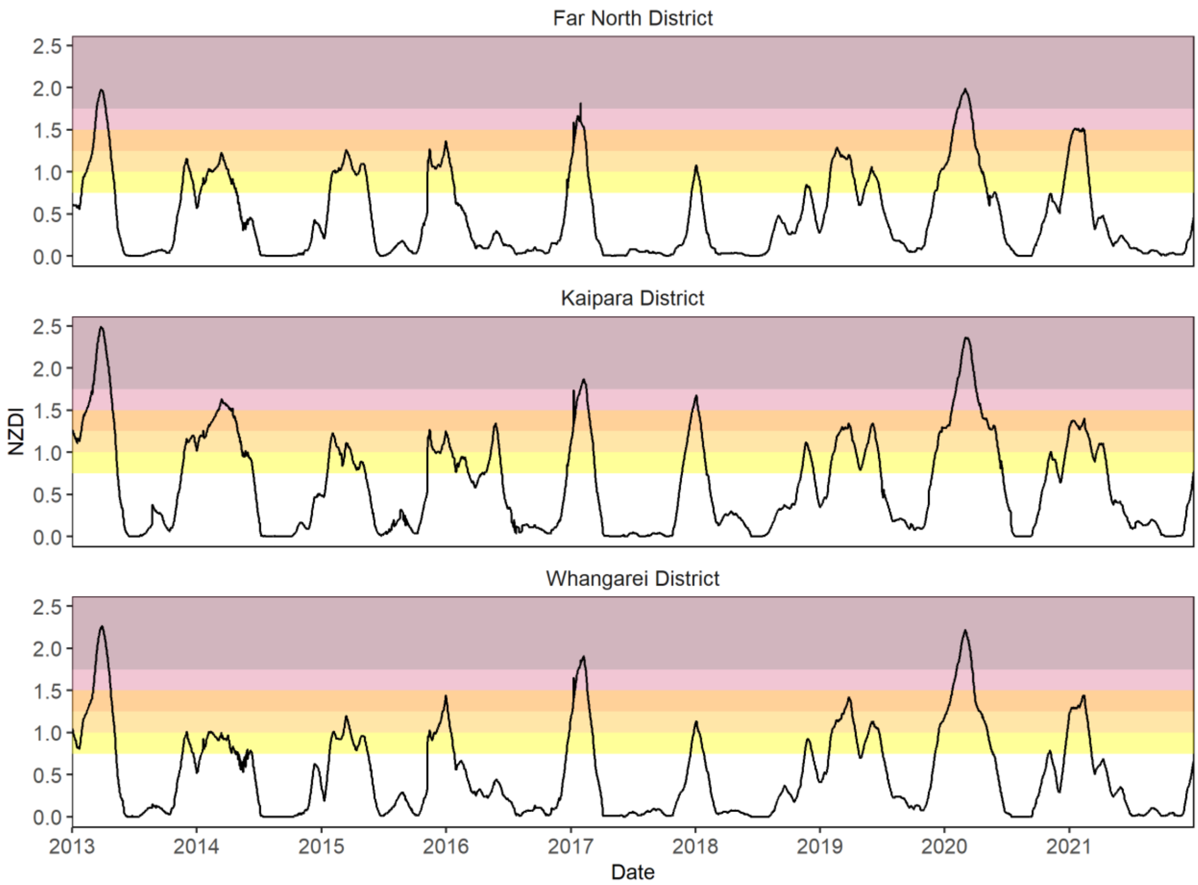

Figure 5-2: Daily New Zealand Drought Index (NZDI) values for the 3 Northland regions between 2013 and 2022. Coloured bars indicate drought categories: yellow - dry, light orange – very dry, dark orange – extremely dry, pink - drought, purple – severe drought.

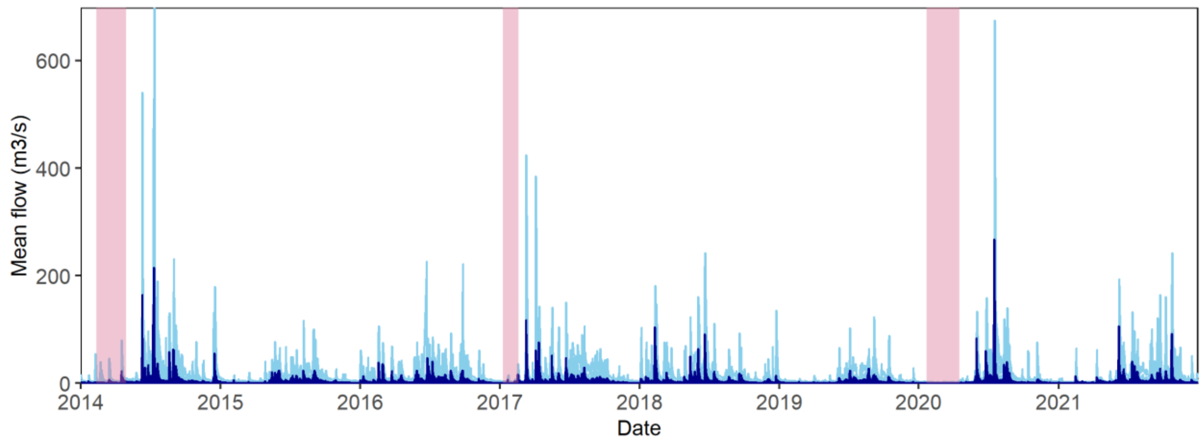

Figure 5-3: Hydrograph of mean daily flows for all Northland sites between 2014 and 2021. Light blue lines are sites with measured flow data, the dark blue lines represent the flow data estimated from the TopNet hydrologic model. Pink shaded areas indicate periods of drought conditions identified by the NZDI (NZDI > 1.5).

Table 5-2: Periods of drought across the 3 Northland regions over the time period corresponding to the macroinvertebrate dataset (2014-2021). Drought periods were classified as days when NZDI > 1.5.

District	Drought periods – NZDI	Reporting Year
Far North	8/3/2013 – 16/4/2013	2014
	15/1/2017 – 9/2/2017	2017
	26/1/2020 – 25/3/2020	2020
	12/1/2021 – 13/2/2021	2021
Kaipara	11/2/2013 – 29/4/2013	2014
	3/3/2014 – 15/4/2014	2015
	9/1/2017 – 19/2/2017	2017
	23/12/2017 – 9/1/2018	2018
	22/1/2020 – 16/4/2020	2020
Whangarei	4/3/2013 – 23/4/2013	2014
	9/1/2017 – 17/2/2017	2017
	23/1/2020 – 27/3/2020	2020

The SDI, which is based on measured stream flow data, was loosely correlated with the NZDI (Pearson's r = -0.28). The SDI also identified the summers 2017 and 2020 as drought periods at all monitored sites (Figure 5-4). The SDI indicated that the 2020 drought was more severe than the 2017 drought at most sites, although drought severity varied between sites. The majority of sites also experienced mild to moderate drought in 2019. The summer 2014 drought was mild at most sites, with the exception of Manganui at Mititai Road. As the SDI was only available for 22 of the 66 sites, the NZDI was chosen to represent drought in all future analyses.

Figure 5-4: Standardised Discharge Index (SDI) values for 22 Northland sites between 2014-2020. The SDI was calculated for a three-month window from December-February. Coloured bars indicate the drought classification: green – near normal, yellow – mild drought, orange – moderate drought, pink – severe drought, red – extreme drought.

ASPM scores and percent EPT taxa scores were lower across all sites sampled during the identified drought years. There were no obvious differences in median MCI, QMCI, and percent EPT metric scores across during those years, although the range of observed scores was smaller than in many other years (Figure 5-5). Scores for all metrics were higher in 2021, which was also identified as drought year for the Far North district sites.

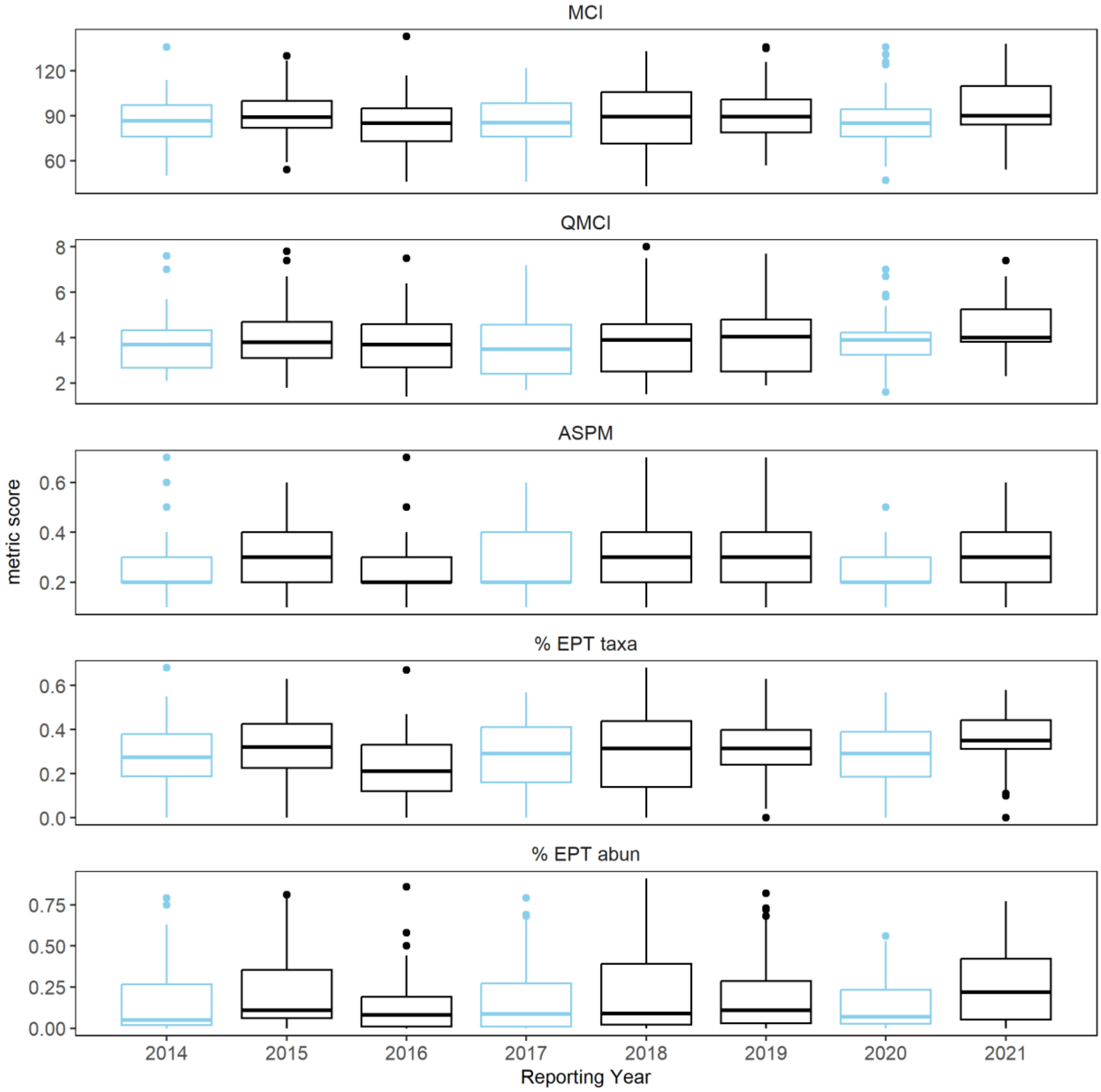

Figure 5-5: Distributions of macroinvertebrate metric scores across all Northland sites per year. Years with drought periods (NZDI > 1.5) indicated in blue.

Individual mixed effects models fit for each macroinvertebrate metric with mean NZDI as a predictor (fixed effect) and Site as a random effect all showed a significant negative effect of drought on metric scores (Table 5-3). Plots of metrics vs mean NZDI for each site show indicate that some sites appear to have strong relationships while others do not (Appendix D).

Table 5-3: Significance of mean NZDI in linear mixed effects models fit for each macroinvertebrate metric. Site was included as a random effect in each model.

Metric	Coefficient of mean NZDI	SE	DF	t	p-value
MCI	-11.54	2.3600	368.93	-4.89	< 0.001
QMCI	-0.38	0.1761	369.30	-2.15	< 0.05
ASPM	-0.09	0.0187	369.65	-4.83	< 0.001
% EPT taxa	-0.12	0.0222	370.11	-5.26	< 0.001
% EPT abun	-0.14	0.0353	372.84	-3.84	< 0.001

The majority of selected potential environmental drivers did not vary distinctly between drought and non-drought years (Figure 5-6). TKN was slightly lower in 2017 and 2020, turbidity and temperature were also lower in 2020. Days since last flow greater than 3 x median was greater in drought years.

However, individual mixed effects models with drivers as the response, mean NZDI as the predictor, and site as a random effect indicated significant relationships between the drought index and all drivers except conductivity and BFI (Table 5-4). The plots of relationships between drivers and NZDI in individual sites indicate the relationships are strong in some sites but indeterminate in others (Appendix E). The relationships with dissolved oxygen, chlorophyll *a*, and DaFRE3 were positive, all others were negative.

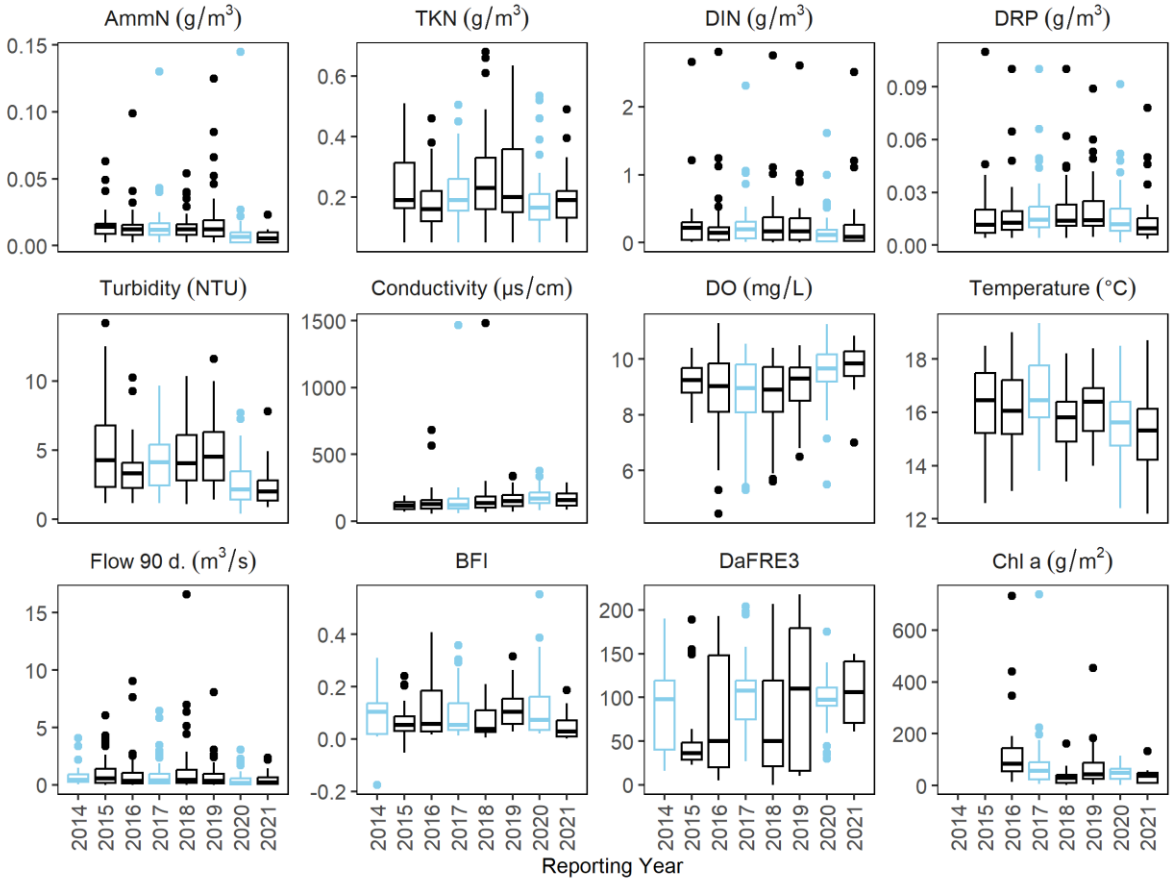

Figure 5-6: Distributions of selected drivers across all Northland sites each year. Years with drought periods indicated in blue.

Table 5-4: Significance of mean NZDI in linear mixed effects models fit for each potential driver. Site was included as a random effect in each model.

Driver	Coefficient of mean NZDI	SE	DF	t	p-value
Amm-N	-0.01	0.0017	304.96	-5.53	< 0.001
TKN	-0.10	0.0139	292.80	-7.06	< 0.001
DIN	-0.15	0.0238	288.17	-6.32	< 0.001
DRP	-0.01	0.0009	304.80	-9.10	< 0.001
Cond.	34.60	26.1000	319.44	1.33	0.19
Turb	-4.13	0.3972	309.45	-10.40	< 0.001
DO	0.60	0.1540	304.26	3.88	< 0.001
Temp.	-0.81	0.2279	307.36	-3.56	< 0.001
Chl a	1.02	0.3000	118.50	3.40	< 0.001
Flow	-1.08	0.1920	382.50	-5.64	< 0.001
DaFRE3	41.08	13.3810	414.15	3.07	< 0.01
BFI	-0.02	0.0131	385.45	-1.21	0.23

5.2.3 GAMM analysis

The top spatial models for MCI, QMCI, and ASPM explained between 70-80% of the variation in metric scores across sites (Table 5-5). DIN, DRP, conductivity, dissolved oxygen, temperature, percent fine sediment cover (sand-silt), and daFRE3 were the most important (i.e., explained the most variation) variables for MCI, QMCI, and ASPM (Figure 5-7). Variable importance scores for percent EPT taxa and percent EPT abundance were low and relatively evenly distributed, indicating uncertainty in which of the potential drivers was associated with these metrics across sites.

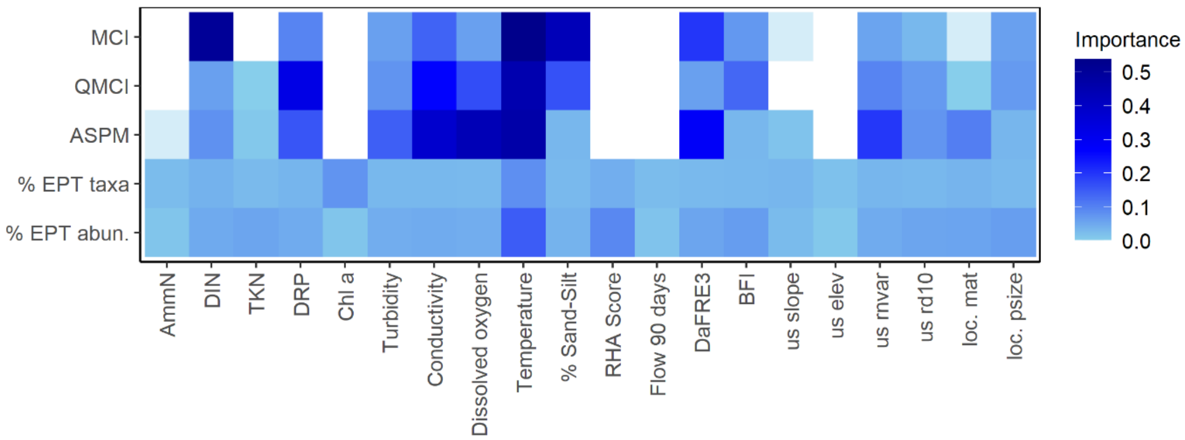

Figure 5-7: Relative importance of drivers in explaining the variation in macroinvertebrate metric scores, based on full-subsets analysis using site medians.

Table 5-5: Top fitting (ΔAICc < 2) generalised models for each macroinvertebrate metric using full subsets analysis on spatial dataset (site medians).

Response	# models < 2AICc	Selected drivers	AICc	BIC	wi.AICc	wi.BIC	r2	edf
MCI	2	DIN % Sand-Silt Temperature	260.74	267.63	0.21	0.17	0.82	9.29
		DIN DaFRE3 % Sand-Silt Temperature	262.46	268.85	0.09	0.09	0.82	10.04
QMCI	5	DRP Temperature	82.01	89.03	0.09	0.10	0.67	5.82
		Conductivity Temperature	82.27	89.30	0.08	0.09	0.67	5.85
		DRP DO Temperature	82.32	89.63	0.08	0.07	0.69	6.82
		Conductivity % Sand-Silt Temperature	82.68	90.04	0.07	0.06	0.70	7.52
		BFI DRP Temperature	83.15	90.23	0.05	0.05	0.72	8.90
ASPM	5	Conductivity DaFRE3 DO Temperature	-76.11	-68.75	0.14	0.09	0.73	6.82
		DRP DO Temperature us rnvar	-75.24	-69.12	0.09	0.11	0.79	9.62
		Conductivity DO Temperature	-74.93	-67.59	0.08	0.05	0.73	6.82
		Conductivity DO Temperature us rnvar	-74.53	-67.34	0.06	0.04	0.74	7.67
		DIN DaFRE3 loc. mat Turbidity	-74.13	-68.03	0.05	0.06	0.79	9.74
% EPT taxa	4	Temperature	35.40	38.21	0.03	0.05	0.46	2.00
		Chl a	35.46	38.26	0.02	0.05	0.34	2.00
		RHA Score	36.90	39.70	0.01	0.02	0.17	2.00
		null	37.19	38.65	0.01	0.04	0.00	1.00
% EPT abun	4	Temperature	28.87	31.68	0.04	0.10	0.46	2.00
		RHA Score	29.98	32.78	0.03	0.06	0.33	2.00

Response	# models < 2AICc	Selected drivers	AICc	BIC	wi.AICc	wi.BIC	r2	edf
		loc. psize Temperature	30.24	34.24	0.02	0.03	0.49	3.00
		Temperature us rd10	30.87	34.87	0.02	0.02	0.49	3.00

In the time series models, the unique variation explained by environmental drivers over time was determined by subtracting the variation associated with site differences, represented by the null model with a random site term only. Environmental drivers explained 10-30% of the variation in MCI, QMCI, and ASPM within sites over time, depending on the dataset (Table 5-6), while between site differences explained 60-70% of the variation, similar to the spatial models.

The most important variables in the time series analysis using all sites were conductivity, dissolved oxygen, baseflow, and drought (Figure 5-8). DIN was also important in explaining variation in MCI. Turbidity and median flow over the previous 90 days were important variables for QMCI, as well as spatial attributes local mean air temperature and particle size. Important variables in ASPM models also included catchment slope and number of rain days >10 mm.

DIN, ammoniacal N, temperature, flow, BFI, and the drought index were the most important variables in pastoral streams (Figure 5-8). Reporting Year was important in pastoral streams as well, indicating metric scores varied more over time in those sites. Surprisingly, turbidity was important in forest streams, along with dissolved oxygen, median flow, days since a flushing flow, and local particle size. Variable importance scores were low and fairly evenly distributed in the hard-bottomed and soft-bottomed analyses, which indicates high model uncertainty, as do the large numbers of candidate models with comparable fit (ΔAICc < 2) and low model weights (i.e., $\leq$ 0.02; Table 5-6).

Few of our selected drivers were important in explaining variation in percent EPT taxa, which also had low variable importance and model weights in all subset analyses. In fact, the most parsimonious model for percent EPT taxa in soft-bottomed, indigenous forest, and pastoral streams was the null model with only the random site term. Percent EPT abundance, on the other hand, was associated with temperature and drought in pasture streams (and across all sites as well; Figure 5-8).

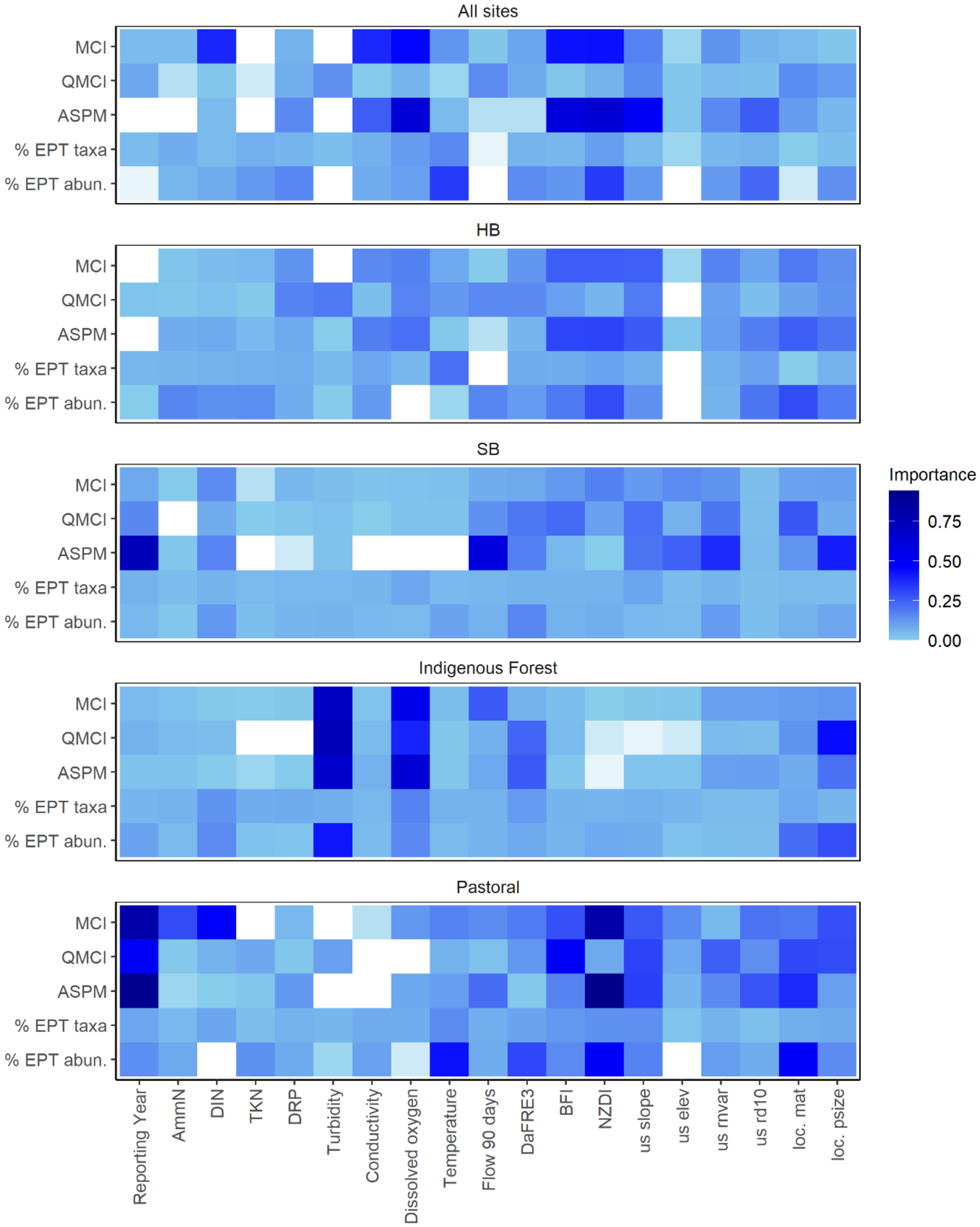

Figure 5-8: Relative importance of drivers in explaining the variation in macroinvertebrate metric scores, based on full-subsets analysis. Each panel indicates the subset of data used for the analysis.

Table 5-6: Best generalised model for each macroinvertebrate metric using full subsets analysis with different datasets. Only the top model is shown for brevity; however, the number of models <2AICc indicates the number of alternative top models with equally good fit for each response x dataset combination. The unique R2 is the variance explained by the model above what is explained by the null model alone (in this case, the random site term).

Response	Dataset	# models < 2AICc	Selected drivers	AICc	BIC	wi.AICc	wi.BIC	r2	unique r2	edf
MCI	All sites (n=66)	4	BFI DIN Conductivity DO NZDI	2 427.39	2 659.29	0.09	0.00	0.81	0.04	70.41
	HB (n=39)	7	BFI DRP loc. mat NZDI us rnvar us slope	1 547.49	1 657.38	0.02	0.23	0.80	0.03	35.91
	SB (n=27)	46	NZDI us elev us rnvar	872.93	924.45	0.01	0.05	0.74	0.02	23.05
	Pastoral (n=53)	17	BFI DIN NZDI Reporting Year us elev	1 948.00	2 115.76	0.03	0.01	0.73	0.07	53.85
	Indigenous Forest (n=8)	1	DO Flow 90 days Turbidity	237.23	237.81	0.20	0.47	0.85	0.23	11.88
QMCI	All sites	11	Flow 90 days loc. mat loc. psize Turbidity us slope	742.32	958.66	0.02	0.00	0.77	0.01	64.47
	HB	33	DaFRE3 Flow 90 days loc. mat Turbidity us slope	425.75	546.17	0.01	0.01	0.74	0.03	40.27
	SB	6	BFI DaFRE3 loc. mat us rnvar	287.31	333.81	0.03	0.03	0.60	0.02	19.81
	Pastoral	35	BFI Reporting Year	551.64	720.41	0.01	0.00	0.72	0.02	54.28
	Indigenous Forest	2	loc. psize Turbidity	81.92	87.67	0.22	0.09	0.74	0.35	8.77
ASPM	All sites	5	BFI DO NZDI	-769.43	-557.21	0.11	0.00	0.74	0.04	61.67

Response	Dataset	# models < 2AICc	Selected drivers	AICc	BIC	wi.AICc	wi.BIC	r2	unique r2	edf
			us rd10 us slope							
	HB	21	BFI loc. mat NZDI us slope	-485.33	-375.16	0.02	0.11	0.71	0.03	35.15
	SB	6	Flow 90 days loc. psize Reporting Year	-284.06	-228.52	0.09	0.13	0.72	0.07	25.33
	Pastoral	11	NZDI Reporting Year us rd10	-658.07	-498.15	0.05	0.09	0.66	0.06	49.92
	Indigenous Forest	9	DO Turbidity	-59.82	-55.22	0.10	0.07	0.74	0.25	8.75
% EPT taxa	All sites	20	Amm-N Temperature	253.98	265.32	0.02	0.06	0.22	0.22	4.00
	HB	9	Temperature	195.95	202.68	0.02	0.32	0.27	0.27	3.00
	SB	9	null	57.66	60.34	0.02	0.23	0.00	0.00	2.00
	Pastoral	8	null	177.60	181.17	0.01	0.37	0.00	0.00	2.00
	Indigenous Forest	11	null	47.55	48.91	0.03	0.09	0.00	0.00	2.00
% EPT abundance	All sites	4	NZDI Temperature us rd10	218.41	233.51	0.06	0.19	0.27	0.22	5.00
	HB	35	loc. mat NZDI us rd10	157.19	182.21	0.02	0.00	0.39	0.39	8.55
	SB	18	DIN DaFRE3	57.09	67.17	0.01	0.04	0.37	0.23	4.83
	Pastoral	8	loc. mat NZDI Temperature	148.48	166.95	0.02	0.01	0.18	0.18	6.22
	Indigenous Forest	1	loc. psize Turbidity	39.14	42.80	0.13	0.15	0.50	0.45	4.00

5.3 Discussion

What are the predictors of macroinvertebrate community composition in Northland rivers and streams?

Identifying drivers of stream macroinvertebrate communities is always challenging, due to the nature of stream ecosystems and the metrics themselves. Streams are subject to a wide range of multiple, correlated stressors, and macroinvertebrate metrics are affected similarly by many of them. While the full subsets modelling approach enabled us to explore the relative importance of a wide range of potential predictors (more than could be tested in a single stepwise model selection approach), like any automated model selection technique, it also has drawbacks. Firstly, it only identifies correlations, not causality. Secondly, there is always the possibility that we failed to include an additional predictor which is in fact the main driver of observed responses. Thirdly, it can easily generate many more models than can reasonably be rigorously sense-checked. For example, we obtained 337 'top' models with ΔAICc <2 across our 25 metric and dataset combinations. Therefore, we will restrict our discussion to variable importance across all models rather than detailed investigation of individual models.

The key drivers identified by the full subsets analysis for pastoral streams were DIN, ammoniacal N, temperature, drought index, and several flow metrics – median flow over the previous 90 days, base flow index (BFI), and days since last flow greater than 3 times the long-term median flow (daFRE3). Spatial attributes including upstream catchment area, rainfall variation, air temperature, and particle size, were also moderately important for at least one metric. The drivers identified for pastoral streams were consistent with known relationships between pastoral land cover/land use and stream water quality and habitat; namely positive correlations with nutrient (N and P) concentrations (Larned et al. 2019) and fine sediment cover (Niyogi et al. 2007). Nevertheless, it should be noted that while variables were identified as important across all models, selected drivers only explained around 3-4% of the unique variation in metrics in the most parsimonious model for each metric.

Key drivers in indigenous forest sites were turbidity, dissolved oxygen, flow, and local particle size. The high relative importance of turbidity was surprising; while turbidity is often high in planted forests due to soil erosion associated with harvest activities, unstable stream banks, or unsealed access roads (Quinn et al. 1997, Quinn and Stroud 2002, Boothroyd et al. 2004), native forests typically have low turbidity and high visual clarity (Quinn et al. 1997, Quinn and Stroud 2002). However, a recent analysis of landscape stability and susceptibility to mass wasting (i.e., erosion due to gravity) across the Northland region showed that the ancient basement rocks which underlie the Brynderwyn Hills and Omahuta Forest areas are inherently unstable and prone to high rates of mass wasting, erosion, and sediment yields despite being covered by indigenous vegetation (Rissman et al. 2019). Selected drivers explained a larger proportion of the variation (20-30%) in invertebrate metrics across indigenous forest streams. While this could suggest that macroinvertebrate communities in indigenous forest streams are highly sensitive to environmental drivers, it could also be an artefact due to the small size of the dataset (only eight indigenous forest streams) and little variation in metric scores (e.g., see Figure 4-7).

It was more difficult to determine key drivers for soft-bottomed or hard-bottomed streams, likely due to co-occurring differences in catchment land use (i.e., approximately equal numbers of hard- and soft-bottomed pastoral streams, all but one forest stream also being hard-bottomed). The most important variables in the analysis using all sites were largely a combination of those identified for pastoral and forest streams.

There was higher uncertainty in models for percent EPT taxa and percent EPT abundance than for MCI, QMCI, or ASPM, suggesting that these metrics are less useful for assessing impacts of environmental drivers (or at least the subset of environmental drivers selected for this analysis) on macroinvertebrate community composition in the Northland dataset. We suspect this is likely due to high variation in percent EPT taxa and abundance within streams over time (Figure G-1). This is supported by the lower proportion of variation explained by the site-only null model for these metrics. An alternate, or additional, explanation is that species replacement maintained overall proportions of EPT whilst still impacting MCI and QMCI scores. Several EPT taxa were high contributors to overall turnover (see section 5.3 for further discussion), which also lends support to this conjecture.

It is interesting that TKN did not explain a large amount of variation in any of the models (although it did explain relatively more in the pastoral dataset than in any of the other datasets), as TKN makes up two-thirds of the total N in the Northland water quality dataset (due to underlying geology and poorly drained floodplain soils) and the nutrient criteria analysis indicated TKN was strongly correlated with macroinvertebrate metric scores. This suggests that macroinvertebrates may not be responding to TKN itself, but instead to another driver correlated with TKN that was not included in our analyses. Sediment is a logical possibility; both sediment and TKN are associated with pastoral land use, particularly livestock (Vidon et al. 2008, Rissmann et al. 2020).

Although sediment, along with chlorophyll *a*, was not sampled frequently enough to be included in the time series analysis, it was included in the full subsets analysis using median values for all sites. That analysis showed that percent fine sediment cover was an important driver of macroinvertebrate composition between sites, in addition to DIN, DRP, conductivity, dissolved oxygen, temperature, and daFRE3. Other than DRP, these drivers were also all important in the time series analysis, indicating that sediment would likely also have been important if included. Furthermore, Death et al. (2020) also found that deposited sediment was strongly correlated with differences in macroinvertebrate communities between pastoral and indigenous forest streams in Northland.

The fact that DRP was only important for explaining variation between sites, but not within sites over time, suggests that DRP concentrations have remained fairly stable across the region over the last eight years, consistent with the main source of phosphorus being geological. Surprisingly, chlorophyll *a* was not important in explaining variation in median metric scores across sites, even though it is a main food source for many grazing macroinvertebrates.

How do drought conditions impact macroinvertebrate community composition in Northland?

The New Zealand drought index, was identified as an important variable in the FSSgam analysis for macroinvertebrate metrics across all sites, as well as in the hard-bottomed and pastoral stream subsets. The importance of drought in explaining variation in macroinvertebrate communities in pastoral streams, which also had the strongest associations with other environmental predictors, aligns with global research showing that already stressed streams are more susceptible to drought effects and/or that drought exacerbates the impact of other environmental stressors (Mosley 2015).

The drought index also had significant negative linear relationships with all five macroinvertebrate metrics, although the majority of the variation explained was due to the random site term. Examination of relationships between metrics and drought within each site showed that only a handful of sites had strong negative relationships between macroinvertebrate metrics and the drought index, while the majority were indeterminate. This is consistent with a previous analysis by

Death et al. (2020), who examined correlations between MCI and SQMCI and drought indices in Northland using an older dataset, and found a mix of positive, negative, and no correlations across sites.

Do drought conditions impact water quality and environmental variables that may influence macroinvertebrate community composition?

The drought index also had significant negative relationships with many of the selected drivers, indicating that there may also be indirect effects of drought on macroinvertebrate communities.

Some of the observed effects of drought on other drivers could in fact be beneficial for macroinvertebrates, such as decreased nutrient concentrations. This could be the result of less run-off during drought periods reducing the transport of contaminants from land into streams. However, this may only be a short-term effect, as the first rainfall after a drought will carry a large load of built-up contaminants (Lisboa et al. 2020). Habitat diversity could also increase in some hard-bottomed streams as more riffles are exposed under low flows, although drying will also reduce habitat availability, particularly along bank margins.

The drought index was positively related to periphyton, likely in conjunction with the positive relationship between drought and DaFRE3, or days since a flushing flow (and therefore days of periphyton accrual). Increased periphyton can have positive or negative impacts on macroinvertebrate communities, depending on the composition of the community and whether it is resource limited. Increases from small to moderate amounts of periphyton can support a larger macroinvertebrate population, and perhaps even enhance diversity. However, studies have shown that highly eutrophic streams with excessive in-stream plant and algal growth can become dominated by generalist consumers, which then outcompete other more sensitive taxa, reducing biodiversity (Graham et al. 2015, Barrett et al. 2021).

Interestingly, temperature and dissolved oxygen had the opposite relationships with the drought index than what is usually observed. There was an overall positive relationship between the drought index and dissolved oxygen and a negative relationship between temperature and the drought index (although the relationships were not consistent across all sites). Typically, dissolved oxygen declines with drought, as lower flows result in warmer water temperatures and reduced turbulence, both of which also reduce oxygen concentrations. The contradictory relationships could be an artefact of when the streams were sampled, as all oxygen and temperature data were from spot measurements rather than continuous records. For example, the increased periphyton growth also associated with drought may result in increased oxygen concentrations due to photosynthesis during the day, but greater diurnal fluctuations and low oxygen conditions overnight. The negative temperature relationship could also be due to thermal stratification and limited mixing of cooler bottom water in wide, slow-flowing streams, which includes many of Northland's soft-bottomed streams, during drought conditions.

The small set of relationships described above illustrates the complexity of possible combinations and interactions of indirect effects that drought may have on other environmental variables which will in turn influence macroinvertebrate community composition.

6 Community turnover

6.1 Methods

Community turnover was assessed as the percentage of taxa appearing and disappearing over time. Three components of taxa turnover were calculated using the R package 'codyn' (Hallet et al. 2016):

1. total turnover, or the proportion of species which differ between time points,
2. appearances, or the proportion of new species not present at the previous time point, and
3. disappearances, or the proportion of species no longer present compared to the previous time point.

Individual taxa turnover was determined by summing the number of times a taxa went from 'present' to 'absent' or vice versa within a site between years.

6.2 Results

Turnover was fairly similar across years (Figure 6-1). Interestingly, there were more disappearances than appearances across all sites in all land use categories in 2018 and 2021, the years following a drought year. However, a similar pattern was not observed in 2015 after the 2014 drought. The disappearances may have also been related to the large flood events which occurred soon after the end of each drought period.

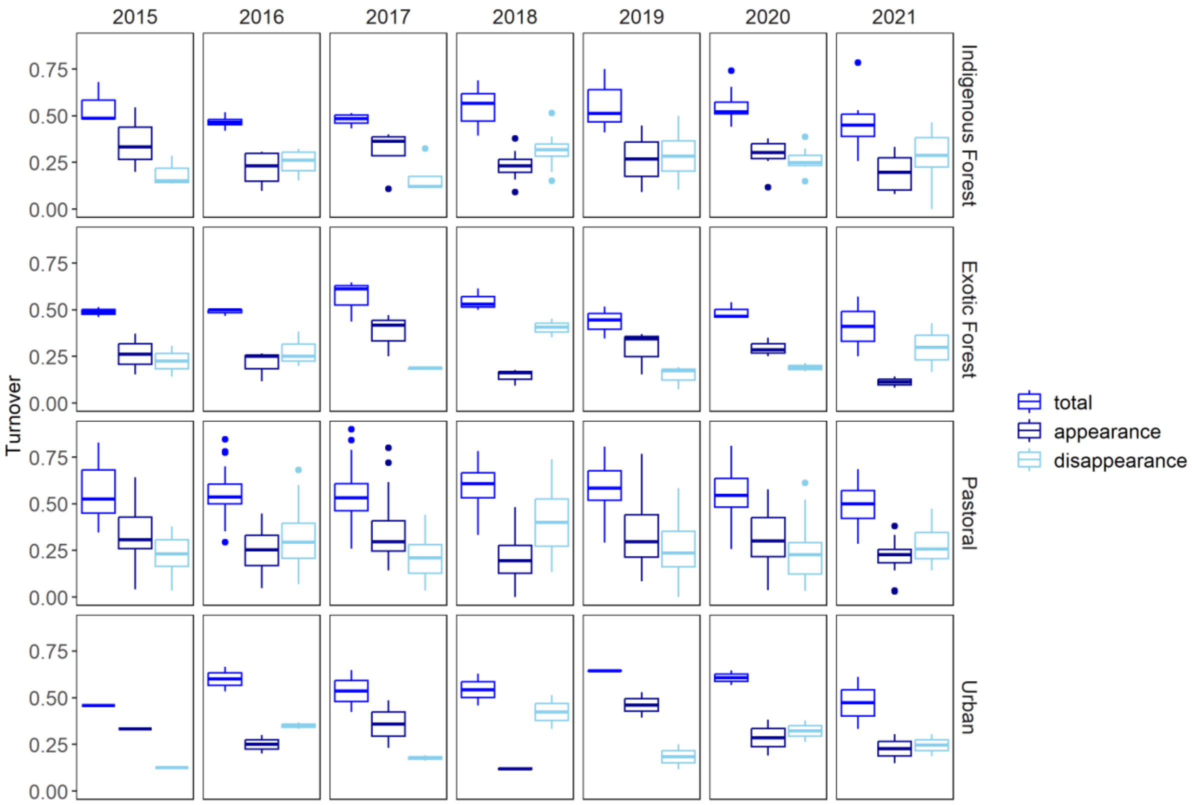

Figure 6-1: Community turnover across Northland sites in different land use categories each year. Total turnover is the sum of appearances and disappearances of individual taxa.

Within individual sites, however, turnover was more variable between years (Figure 6-3). Some sites had consistently lower turnover than others, such as Awanui at FNDC and Punakitere at Taheke, while others were routinely high (i.e., Kaeo at Dip Road). A few sites had noticeable changes in turnover over time, such as Hakaru at Topuni or Kenana at Kenana Road. The ratio of appearances to disappearances also varied considerably between years in some sites, notably Aurere at Pekerau Road and Whakapara at Cableway. Those two sites, along with several others (e.g., Mangere at Kokopu Road, Orauiti at Sawyer Road, Waitangi at Wakelins) had much higher proportions of disappearances in 2016 and 2018 than other years. Mean total turnover per site was higher in soft-bottomed streams than hard-bottomed streams, but did not differ significantly between streams with different catchment land use (Figure 6-2, Figure 6-3).

Taxa present in low abundance across many sites often had the highest turnover (Figure 6-4, Table F-1). *Austrosimulium* (sandflies) had the greatest number of total appearances and disappearances across all sites. Whereas *Potamopyrgus* (New Zealand mud snail), which was present and abundant in all sites, was highly persistent, with, on average, only one appearance and disappearance per site over the eight years.

Several EPT (Ephemeroptera – mayflies, Plecoptera – stoneflies, Trichoptera – caddisflies; taxa known to be sensitive to organic pollution) were also persistent across many sites, including *Pycnocentrodes* and *Aoteapsyche* caddisflies. Rare species which were found only in a handful of sites did not persist between years, and were found only once or twice during the eight years. Overall, EPT taxa had similar turnover rates to non-EPT taxa (Figure 6-4).

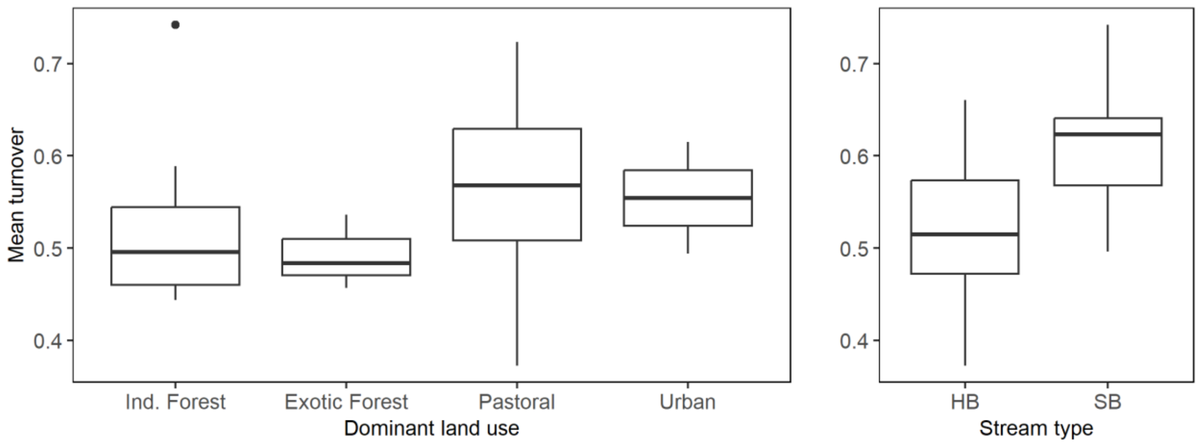

Figure 6-2: Mean turnover per site in streams with different catchment land use or substrate type. HB = hard-bottomed, SB = soft-bottomed.

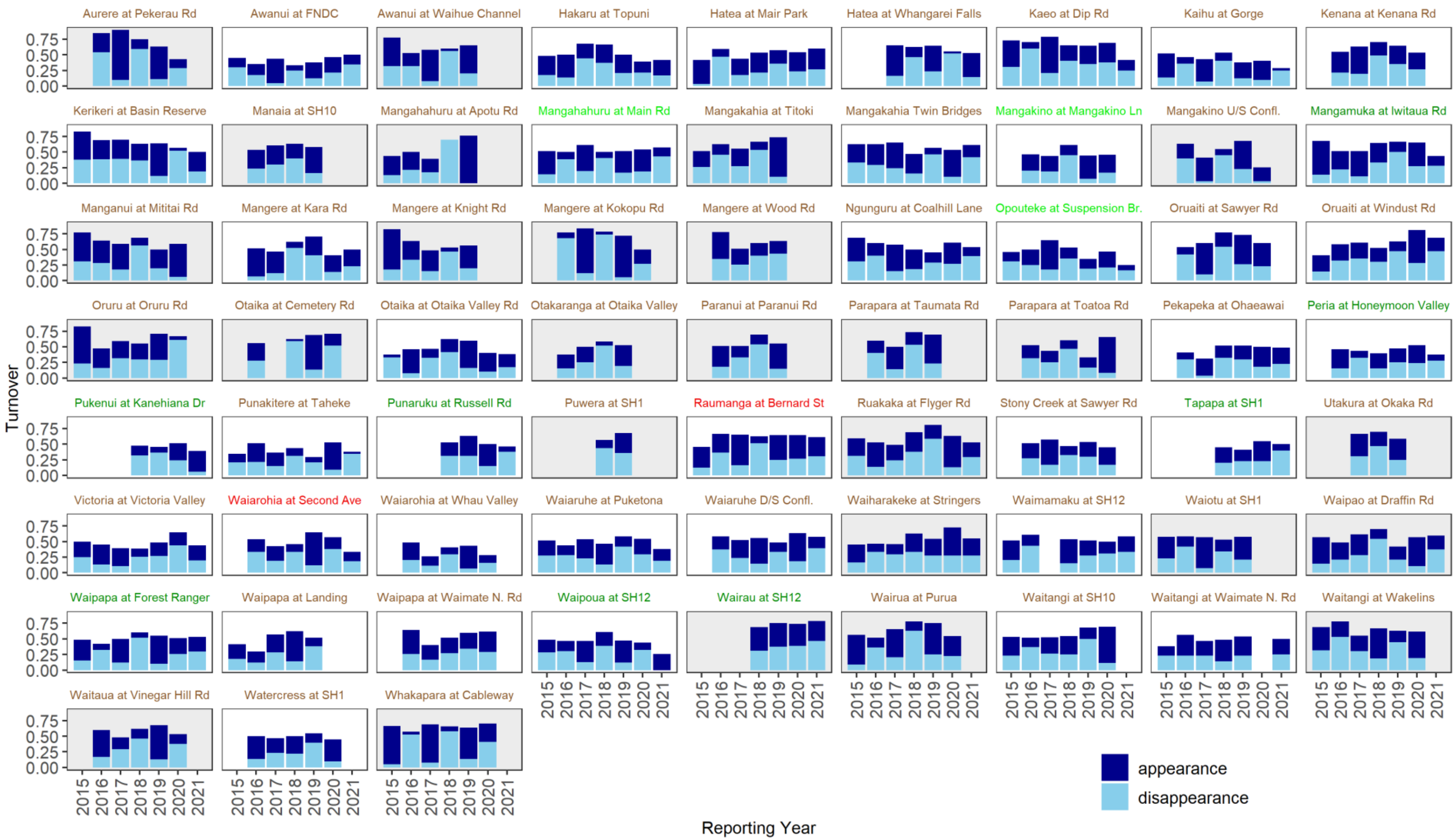

Figure 6-3: Community turnover (total number of appearances and disappearances of taxa) in each site over time. The background of each panel indicates the stream type: grey – soft-bottomed, white – hard-bottomed. The colour of the site names indicates the dominant land use in the catchment: dark green – indigenous forest, light green – exotic forest, brown – pastoral, red – urban.

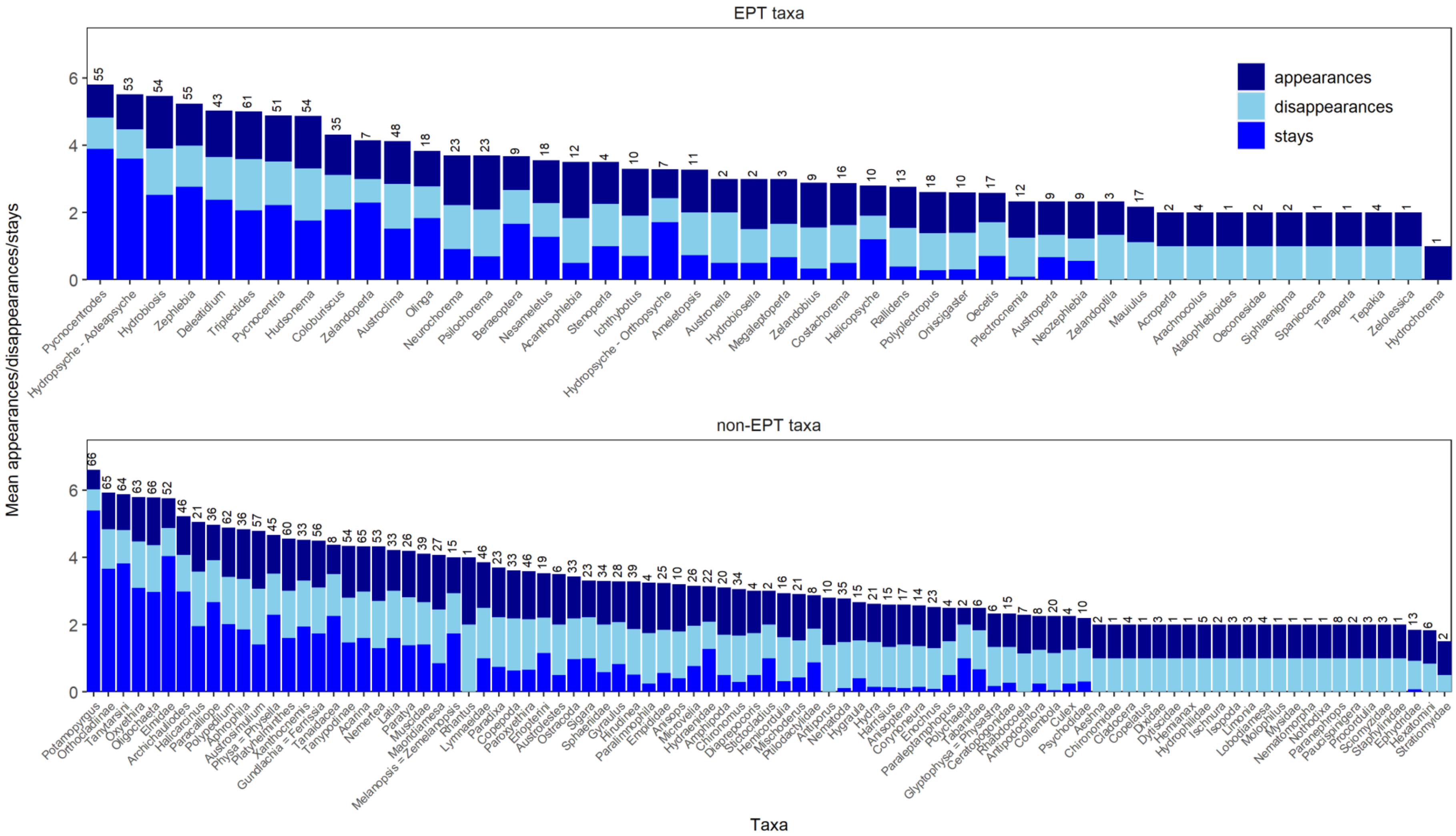

Figure 6-4: Taxa turnover (mean appearances, disappearances, and stays) across all sites 2014-2021. The numbers above each bar are the number of sites (out of 66) in which the taxa was found. The top plot shows EPT taxa and the bottom non-EPT taxa.

6.3 Discussion

How does the community composition (temporal species turnover) of MCI scoring taxa change within and among the SoE sites?

The proportion of disappearances across all sites was greater than appearances the year after a drought for two of three droughts (2017 and 2020 but not 2014), suggesting that there may be a lag between a drought occurrence and detectable changes in invertebrate communities, and that it may take over a year for communities to recover. However, many of the drought events in Northland during our analysis period were followed by high flow events, which complicates determining the recovery of the macroinvertebrate community. The subsequent floods likely deposited large amounts of fine sediment on the streambeds and delayed reestablishment of slower-colonising taxa.

Turnover varied considerably within and between sites and was higher on average in soft-bottomed sites than hard-bottomed sites. However, given that the drivers analysis found few strong associations between macroinvertebrate community metrics and potential drivers in soft-bottomed sites, turnover in these streams was either not correlated with our measured drivers or changes in community composition did not affect metric scores, i.e., species which disappeared were replaced by others with similar MCI tolerance values.

Correspondingly, high-scoring EPT taxa, did not have higher turnover rates than non-EPT taxa. This could indicate that turnover is associated with species traits other than sensitivity to organic pollution. Traits such as body shape and size, mobility, feeding mode and dietary preference, egg-laying behaviour, generation time, and dispersal capability have all been linked to environmental stressors, including fine sediment, eutrophication, acid mine drainage, floods, and droughts (Dolodec et al. 2005, Aspin et al. 2018, Barrett et al. 2022). For example, small caddisflies and specialist predators both disappeared from stream mesocosms under drying conditions, while small grazers and aerial dispersing Diptera became more common (Aspin et al. 2018). Adding a traits component to future turnover analyses could provide additional insight into the mechanisms by which environmental drivers, including drought, are impacting macroinvertebrate communities in Northland.

7 Conclusions and recommendations

The key findings from each component of this project are as follows:

Nutrient criteria

The nutrient criteria analysis identified nutrient concentrations that were much lower than national criteria for protecting macroinvertebrates. This was due to greater proportion of Northland sites with low nutrient concentrations (above the national bottom line) but with macroinvertebrate community metrics well below the national bottom line. This may indicate that nutrients are not the main drivers of macroinvertebrate community patterns in Northland streams.

Macroinvertebrate communities are influenced by in-stream nutrient concentrations either via direct toxic/physiological effects or through nutrient mediated changes in periphyton communities. Nutrient concentrations in Northland waterways rarely exceeded toxicity guidelines and periphyton biomass was not identified as a strong driver of macroinvertebrate community composition in the drivers analysis. However, more periphyton data, particularly in years with stable river flows, would assist in elucidating the mechanisms by which nutrients may be influencing macroinvertebrate communities. In comparison to the Canning et al. (2021) national dataset, the Northland data had a higher proportion of sites with macroinvertebrate metrics below the national bottom line, and of sites with lower nutrient concentrations. The low macroinvertebrate metric values may be caused by a stressor other than nutrient concentrations, therefore introducing a bias in the mis-match analysis leading to lower nutrient criteria.

Drivers

The importance (i.e., amount of variation in macroinvertebrate community composition explained) of environmental predictors over time varied between pastoral streams and indigenous forest streams, but not between hard- and soft-bottomed streams. MCI, QMCI, and ASPM had stronger associations with drivers than EPT metrics. Overall, the amount of variation explained by environmental drivers in the time series models was low compared to that explained by site differences alone. The spatial models identified additional drivers which explained variation in metrics between sites but not over time. The main predictors of differences between sites included nutrients, dissolved oxygen, temperature, fine sediment cover, and flow metrices. The main predictors of differences over time within sites included conductivity, dissolved oxygen, baseflow, and the drought index, as well as nutrients in pastoral streams only.

The drought index was a key driver of macroinvertebrate communities both in the full subsets analysis and in individual regressions. It also had significant independent linear relationships with many of the other environmental drivers.

Community turnover

Turnover was highly variable within and between sites. The largest contributors to total turnover were taxa found across many sites, rather than rare species. Taxa disappearances increased the year following a drought in two out of three cases, suggesting that impacts are longer-term and it may take over a year for communities to fully recover. However, recovery times may have also been affected by high flow events following most drought periods.

Recommendations

We do not recommend use of the newly derived nutrient criteria until the role of nutrients in driving macroinvertebrate community composition in Northland streams has been better quantified or a more robust MoM approach has been developed to handle unevenly distributed data.

A combination of approaches will be required to determine the relationships between nutrients, other potential stressors, and macroinvertebrate communities:

1. Further investigation into whether the source of organic nitrogen (TKN) is anthropogenic or natural, and whether correlated declines in macroinvertebrate communities are associated with TKN itself, or other drivers which co-vary with TKN (i.e., sediment).

2. Continued collection of monthly periphyton and sediment data. We recommend that continued collection of periphyton and sediment data be a priority for NRC, so that both can be included in future driver analyses. Sediment has been shown to adversely impact macroinvertebrate communities (Matthaei et al. 2010, Burdon et al. 2013) and many areas of Northland are erosion prone (Rissmann et al. 2019). Periphyton is one of the main mechanisms by which nutrients impact macroinvertebrate communities, via changes in resource and habitat availability (Tonkin et al. 2014). The drivers analysis should be repeated once 5+ years of periphyton sediment data are available.

3. Continued collection of continuous dissolved oxygen and temperature data at all sites (at the time of this report, only six sites had continuous data available, which was not enough for inclusion in the GAMM analysis). NPS-FM attributes based on continuous data (i.e., dissolved oxygen minima) should be included in the next drivers analysis.

4. Testing and validation of GAMMs. The top models from the full subsets approach should be individually assessed. First, the shape of the splines in each model should be checked for over-fitting (e.g., complex 'wiggliness' is unlikely to be ecologically meaningful). Second, the predictive ability of the models should be tested via cross-validation. It would also be useful to test the models on a dataset from outside Northland.

5. Further investigate patterns in community and taxa turnover. Species traits could be used to compare taxa with high and low contributions to turnover. Turnover could also be used as a response variable in full subsets models to identify any environmental drivers associated with turnover.

6. More frequent (i.e., monthly or bi-monthly) sampling at a subset of sites immediately following drought and flood events to investigate macroinvertebrate community recovery trajectories. The timing of annual SoE monitoring is not ideal for separating the impacts of these two types of stressors, as sites are likely to be sampled at the beginning of the summer before the peak drought occurs, followed by floods during autumn and winter before the next annual sample is collected.

8 Acknowledgements

Many thanks to the Northland Regional Council team – Manas Chakraborty, Preston Ferreira, and Brandon Ruehle – for many productive discussions, questions, and advice throughout the course of the project. Thank you to Mike Pingram (WRC) for sharing the taxa turnover code. This report was greatly improved by reviews by Rick Stoffels, Manas Chakraborty, and Brandon Ruehle.

9 Glossary of abbreviations and terms

AIC_c	Akaike Information Criterion for small sample sizes
Amm-N	Ammoniacal nitrogen
ASPM	Average Score Per Metric
BFI	Base Flow Index
daFREX	Days since last flow exceeding X times the long-term median flow
DIN	Dissolved Inorganic Nitrogen
DRP	Dissolved Reactive Phosphorus
EPT	Ephemeroptera (mayflies), Plecoptera (stoneflies), Trichoptera (caddisflies); taxa sensitive to organic pollution
FSSgam	Full subsets GAM analysis
GAMM	Generalised Additive Mixed Model
MCI	Macroinvertebrate Community Index
MoM	Minimisation of mismatch
NO_3-N	Nitrate-nitrogen
NPS-FM	National Policy Statement for Freshwater Management
NRC	Northland Regional Council
NZDI	New Zealand Drought Index
QMCI	Quantitative Macroinvertebrate Community Index
RHA	Riparian Habitat Assessment
SDI	Standardised Discharge Index
SoE	State of the Environment
TKN	Total Kjeldahl's Nitrogen

10 References

ANZECC, ARMCANZ (2000) Australian and New Zealand guidelines for fresh and marine water quality.

Aspin, T.W.H., Matthews, T.J., Khamis, K., Milner, A.M., Wang, Z., O'Callaghan, M.J., Ledger, M.E. (2018) Drought intensification drives turnover of structure and function in stream invertebrate communities. *Ecography*, 41(12): 1992-2004.

Barrett, I.C., McIntosh, A.R., Febria, C.M., Graham, S.E., Burdon, F.J., Pomeranz, J.P.F., Warburton, H.J. (2022) Integrative analysis of stressor gradients reveals multiple discrete trait-defined axes underlie community assembly. *Ecosphere*, 13(7): e4164.

Barrett, I.C., McIntosh, A.R., Febria, C.M., Warburton, H.J. (2021) Negative resistance and resilience: biotic mechanisms underpin delayed biological recovery in stream restoration. *Proceedings of the Royal Society B: Biological Sciences*, 288(1947): 20210354.

Boothroyd, I.K.G., Quinn, J.M., Langer, E.R., Costley, K.J., Steward, G. (2004) Riparian buffers mitigate effects of pine plantation logging on New Zealand streams: 1. Riparian vegetation structure, stream geomorphology and periphyton. *Forest Ecology and Management*, 194(1): 199-213.

Burdon, F.J., McIntosh, A.R., Harding, J.S. (2013) Habitat loss drives threshold response of benthic invertebrate communities to deposited sediment in agricultural streams. *Ecological Applications*, 23(5): 1036-1047.

Burnham, K.P., Anderson, D.R. (2002) *Model selection and multimodel inference: a practical information-theoretic approach*. Springer, New York, NY.

Canning, A.D., Joy, M.K., Death, R.G. (2021) Nutrient criteria to achieve New Zealand's riverine macroinvertebrate targets. *PeerJ*, 9: e11556.

Clapcott, J. (2015) National rapid habitat assessment protocol development for streams and rivers. *Cawthron Report, Prepared for Northland Regional Council*: 2649.

Collier, K.J. (2008) Average score per metric: An alternative metric aggregation method for assessing wadeable stream health. *New Zealand Journal of Marine and Freshwater Research*, 42(4): 367-378.

Death, R.G., Nicholson, C., Chakraborty, M., Pohe, S. (2020) Ecological health of stream invertebrate communities in Northland: can state of the environment data identify drought impacts? *Northland Regional Council Report*: 93.

Dolédec, S., Phillips, N., Scarsbrook, M., Riley, R.H., Townsend, C.R. (2006) Comparison of structural and functional approaches to determining landuse effects on grassland stream invertebrate communities. *Journal of the North American Benthological Society*, 25(1): 44-60.

Fisher, R., Wilson, S.K., Sin, T.M., Lee, A.C., Langlois, T.J. (2018) A simple function for full-subsets multiple regression in ecology with R. *Ecology and Evolution*, 8(12): 6104-6113.

Graham, S.E., O'Brien, J.M., Burrell, T.K., McIntosh, A.R. (2015) Aquatic macrophytes alter productivity-richness relationships in eutrophic stream food webs. *Ecosphere*, 6(6): art89.

Hallett, L.M., Jones, S.K., MacDonald, A.A.M., Jones, M.B., Flynn, D.F.B., Ripplinger, J., Slaughter, P., Gries, C., Collins, S.L. (2016) codyn: An r package of community dynamics metrics. *Methods in Ecology and Evolution*, 7(10): 1146-1151.

Helsel, D. (2012) *Statistics for Censored Environmental Data Using Minitab and R*. Wiley.

Helsel, D.R. (2005) *Nondetects and data analysis: statistics for censored environmental data*. Wiley-Interscience.

Larned, S., Moores, J., Gadd, J., Baillie, B., Schallenberg, M. (2019) Evidence for the effects of land use on freshwater ecosystems in New Zealand. *New Zealand Journal of Marine and Freshwater Research*, 54: 1-41.

LAWA (2022) Groundwater quality. https://www.lawa.org.nz/explore-data/groundwater-quality/

Lisboa, M.S., Schneider, R.L., Sullivan, P.J., Walter, M.T. (2020) Drought and post-drought rain effect on stream phosphorus and other nutrient losses in the Northeastern USA. *Journal of Hydrology: Regional Studies*, 28: 100672.

Matheson, F., Quinn, J., Hickey, C. (2012) Review of the New Zealand instream plant and nutrient guidelines and development of an extended decision making framework: Phases 1 and 2 final report. *Prepared for the Ministry of Science and Innovation EnviroLink Fund*.

Matthaei, C.D., Piggott, J.J., Townsend, C.R. (2010) Multiple stressors in agricultural streams: interactions among sediment addition, nutrient enrichment and water abstraction. *Journal of Applied Ecology*, 47(3): 639-649.

Menezes, S., Baird, D.J., Soares, A.M.V.M. (2010) Beyond taxonomy: a review of macroinvertebrate trait-based community descriptors as tools for freshwater biomonitoring. *Journal of Applied Ecology*, 47(4): 711-719.

Ministry for the Environment (2020) National Policy Statement for Freshwater Management 2020.

Ministry for the Environment (2022) Setting instream nutrient concentration thresholds for nutrient-affected attributes in rivers: guidance on implementing Clause 3.13 of the NPS-FM. Wellington.

Mol, A., Tait, A., Macara, G. (2017) An automated drought monitoring system for New Zealand. *Weather and Climate*, 37(1): 23-36.

Mosley, L.M. (2015) Drought impacts on the water quality of freshwater systems; review and integration. *Earth-Science Reviews*, 140: 203-214.

NEMS (2022) National Environmental Monitoring Standards - Macroinvertebrates: Collection and Processing of Macroinvertebrate Samples from Rivers and Streams.

Niyogi, D.K., Koren, M., Arbuckle, C.J., Townsend, C.R. (2007) Longitudinal changes in biota along four New Zealand streams: Declines and improvements in stream health related to land use. *New Zealand Journal of Marine and Freshwater Research*, 41(1): 63-75.

Oksanen, J., Simpson, G.L., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., Solymos, P., Stevens, M.H.H., Szoecs, E., Wagner, H., Barbour, M., Bedward, M., Bolker, B., Borcard, D., Carvalho, G., Chirico, M., Caceres, M.D., Durand, S., Evangelista, H.B.A., FitzJohn, R., Friendly, M., Furneaux, B., Hannigan, G., Hill, M.O., Lahti, L., McGlinn, D., Ouellette, M.-H., Cunha, E.R., Smith, T., Stier, A., Braak, C.J.F.T., Weedon, J. (2022) vegan: community ecology package. https://CRAN.R-project.org/package=vegan

Pham, H., Kitto, S., Ruehle, B., Delport, C., Perquin, J.-C., Naidu, R. (2022) Northland drought 2018-2020 assessment. Northland Regional Council, Whangārei, New Zealand 0110. Report No: TR2022/SWQty/01. *Northland Regional Council Report*: 78.

Phillips, G., Kelly, M., Teixeira, H., Salas, F., Free, G., Leujak, W., Solheim, A., Varbiro, G., Poikane, S. (2018) Best practice for establishing nutrient concentrations to support good ecological status. *Technical Report EUR 29329 EN*: 142.

Phillips, G., Teixeira, H., Poikane, S., Salas Herrero, F., Kelly, M.G. (2019) Establishing nutrient thresholds in the face of uncertainty and multiple stressors: A comparison of approaches using simulated datasets. *Science of The Total Environment*, 684: 425-433.

Tonkin, J., Death, R., Barquín, P. (2014) Periphyton control on stream invertebrate diversity: Is periphyton architecture more important than biomass? *Marine and Freshwater Research*, 65: 818–829.

Quinn, J., Stroud, M. (2002) Water Quality and Sediment and Nutrient Export From New Zealand Hill-Land Catchments of Contrasting Land Use. *New Zealand Journal of Marine and Freshwater Research*, 36: 409-429.

Quinn, J.M., Cooper, A.B., Davies-Colley, R.J., Rutherford, J.C., Williamson, R.B. (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(5): 579-597.

Rissmann, C., Pearson, L. (2020) Physiographic controls over water quality state for the Northland region. *Land and Water Science Report 2020/05*: 149.

Rissmann, C., Pearson, L., Lindsay, J., Couldrey, M. (2018) Sediment process-attribute layer for Northland. *Land and Water Science Report 2018/35*: 73.

Rissmann, C., Pearson, L., Lindsay, J., Couldrey, M., Lovett, A. (2018) Application of physiographic science to the Northland region: preliminary hydrological and redox process attribute layers. *Land and Water Science Report 2018/11*: 88.

Stark, J.D. (2017) Macroinvertebrate biotic indices for soft-bottomed streams in the Northland Region. *Stark Environmental Report, Prepared for Northland Regional Council*: 47.

Stark, J.D., Maxted, J.R. (2007) A user guide for the macroinvertebrate community index. *Cawthron Report, Prepared for the Ministry for the Environment*, 1166: 58.

Vidon, P., Campbell, M.A., Gray, M. (2008) Unrestricted cattle access to streams and water quality in till landscape of the Midwest. *Agricultural Water Management*, 95(3): 322-330.

Wood, S.N. (2011) Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society (B)* 73(1): 3-36.

Wood, S.N., Pya, N., Saefken, B. (2016) Smoothing parameter and model selection for general smooth models (with discussion). *Journal of the American Statistical Association*, 111: 1548-1575.

Appendix A Macroinvertebrate metric – nutrient regressions

Table A-1: Regressions between metrics and nutrients across Northland SoE sites.

Metric	Nutrient	Coef.	R^2	F-stat	DF	p-value
MCI	Amm-N	-0.24	0.05	3.22	1, 64	0.08
	NO_3-N	0.00	0.002	0.13	1, 64	0.72
	TKN	-0.09	0.32	29.97	1, 63	<0.001
	DIN	0.00	0.003	0.16	1, 63	0.69
	DRP	-0.13	0.01	0.78	1, 64	0.38
QMCI	Amm-N	-0.01	0.02	1.17	1, 64	0.28
	NO_3-N	0.00	0.02	1.08	1, 64	0.30
	TKN	-0.01	0.32	29.96	1, 63	<0.001
	DIN	0.00	0.02	1.00	1, 63	0.32
	DRP	-0.02	0.04	2.44	1, 64	0.12
ASPM	Amm-N	0.00	0.02	1.07	1, 64	0.31
	NO_3-N	0.00	0.001	0.07	1, 64	0.80
	TKN	0.00	0.27	22.94	1, 63	<0.001
	DIN	0.00	0.001	0.05	1, 63	0.83
	DRP	0.00	0.01	0.65	1, 64	0.42

Appendix B Measured vs modelled flow relationships

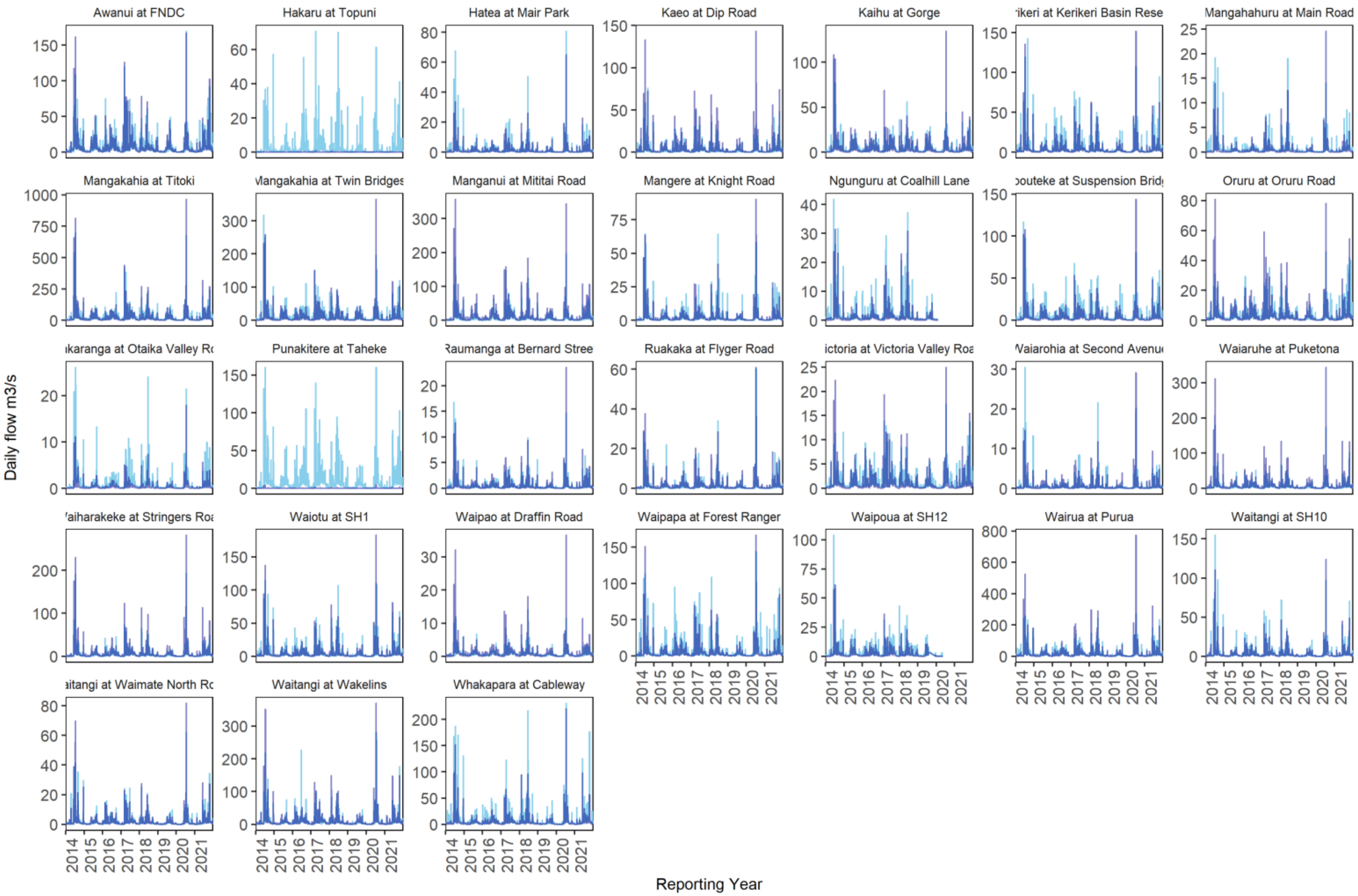

Figure B-1: Hydrographs for Northland SoE from measured flows vs TopNet estimate flows in Northland SoE sites . Measured flows are in light blue, TopNet modelled flows are in dark blue.

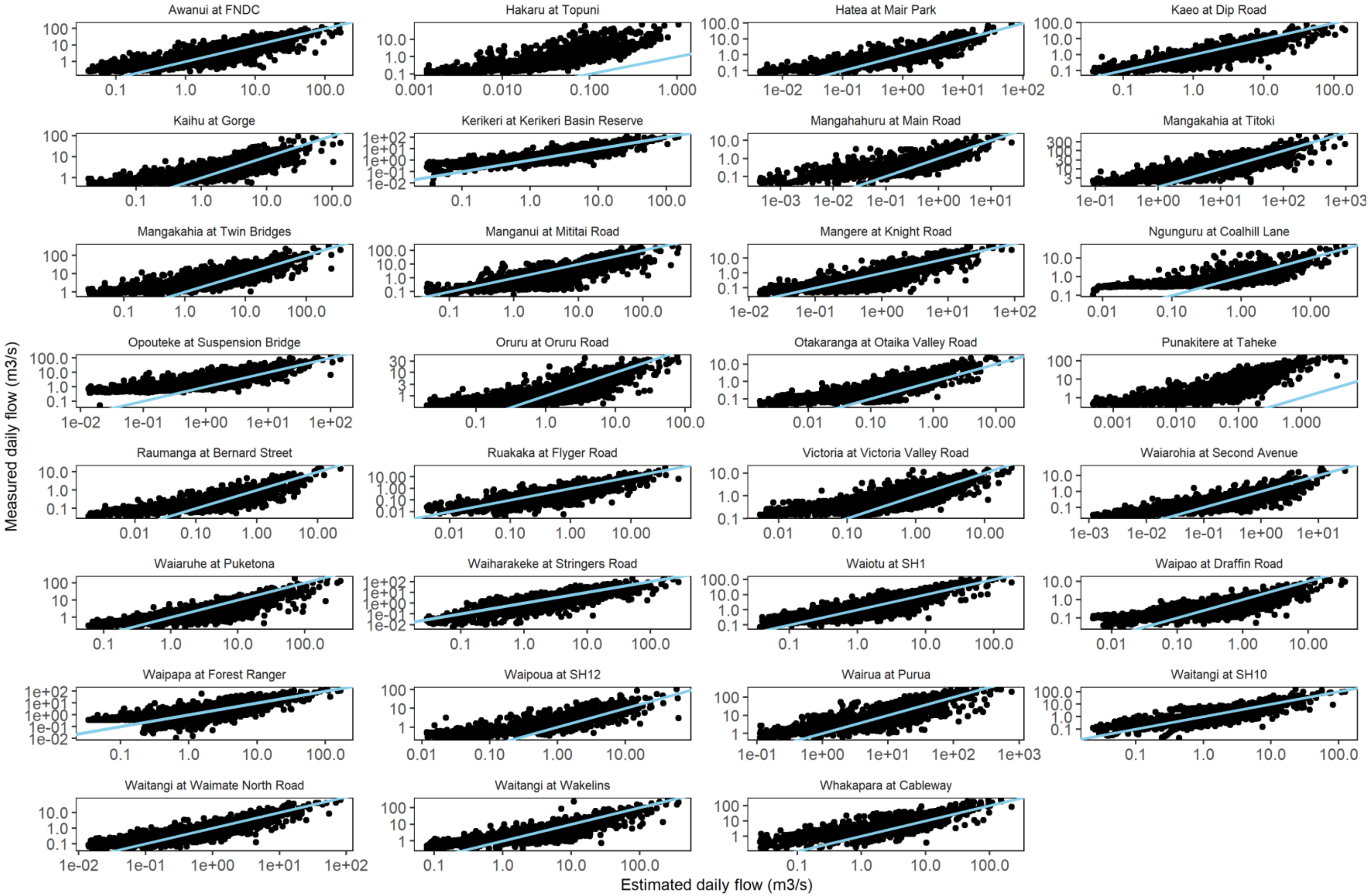

Figure B-2: Measured vs TopNet modelled flows for Northland Soe sites. The 1:1 line is shown in light blue.

Appendix C Correlations between environmental drivers

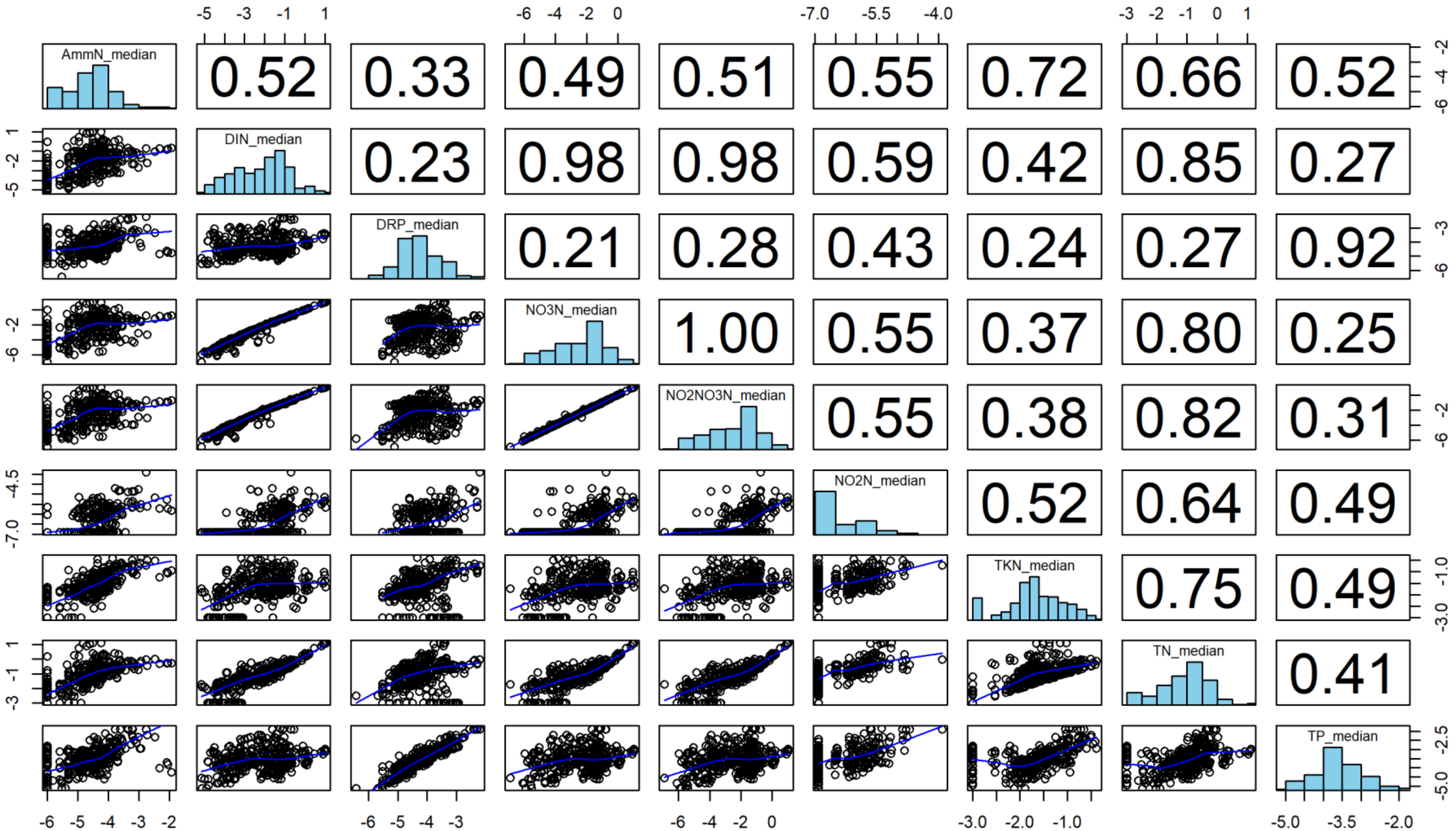

Figure C-1: Pearson's correlation coefficients and pairwise relationships between nutrient measurements.

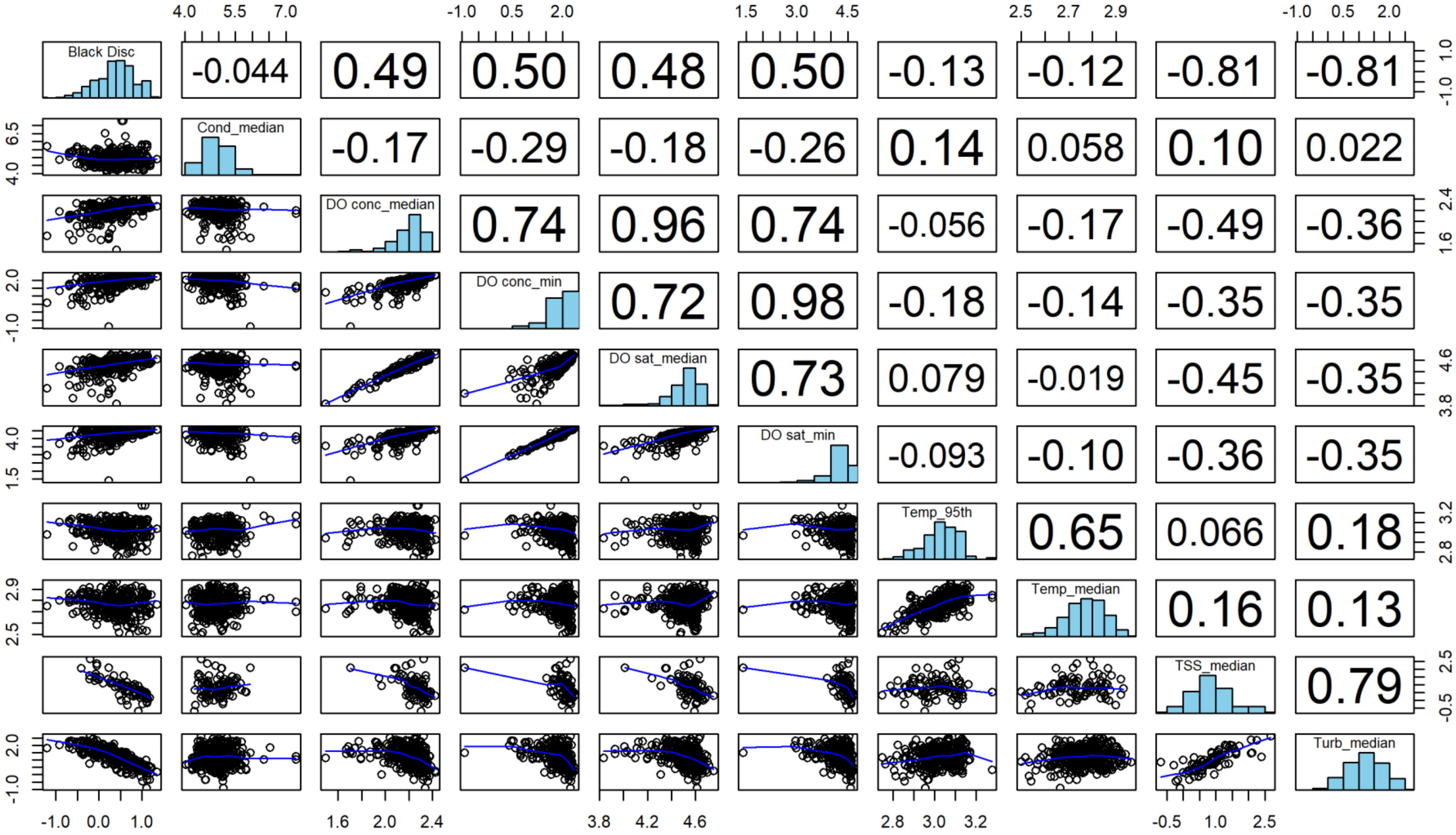

Figure C-2: Pearson's correlation coefficients and pairwise relationships between water quality parameters.

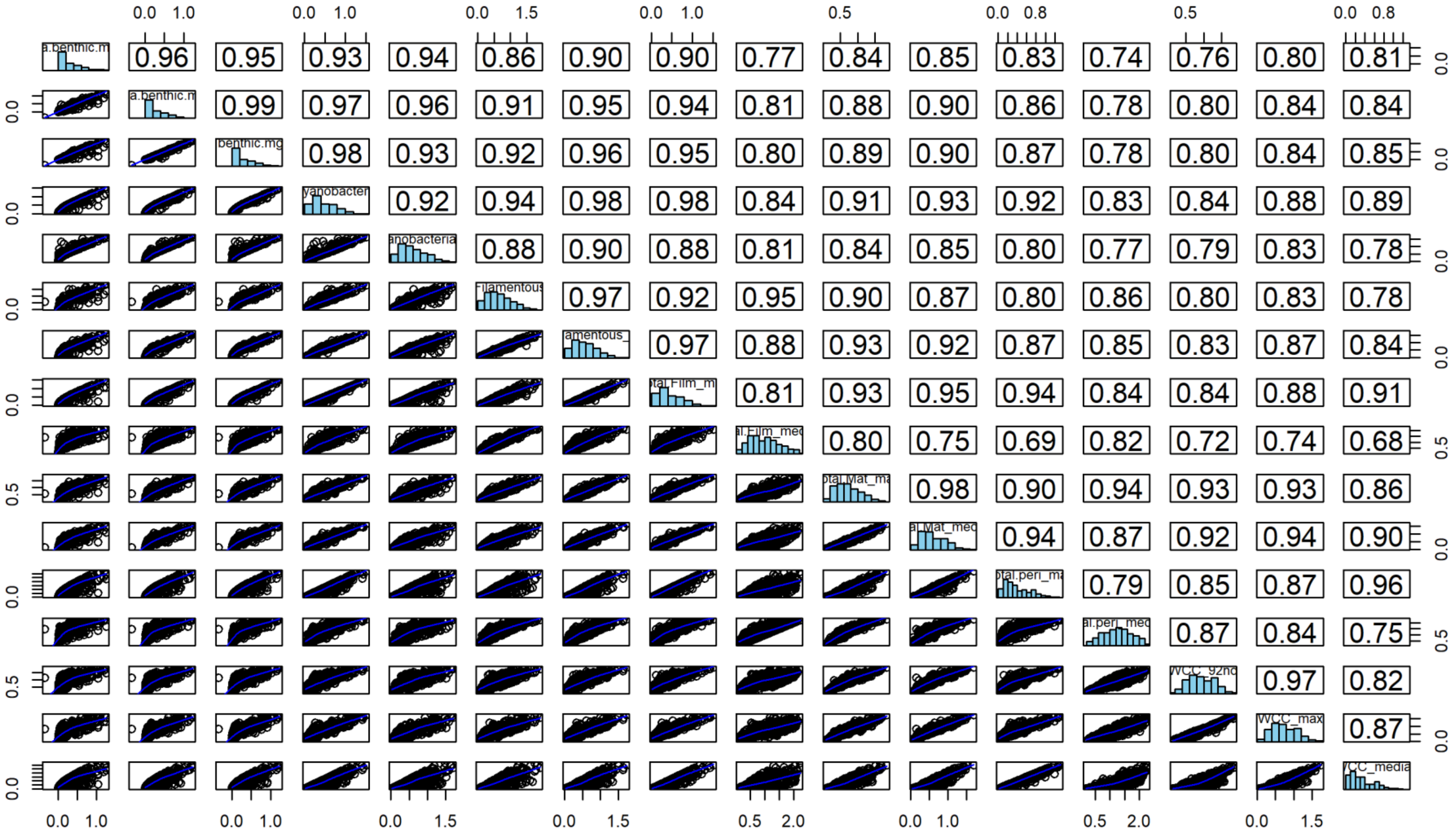

Figure C-3: Pearson's correlation coefficients and pairwise relationships between periphyton measurements (Chl *a* and % cover).

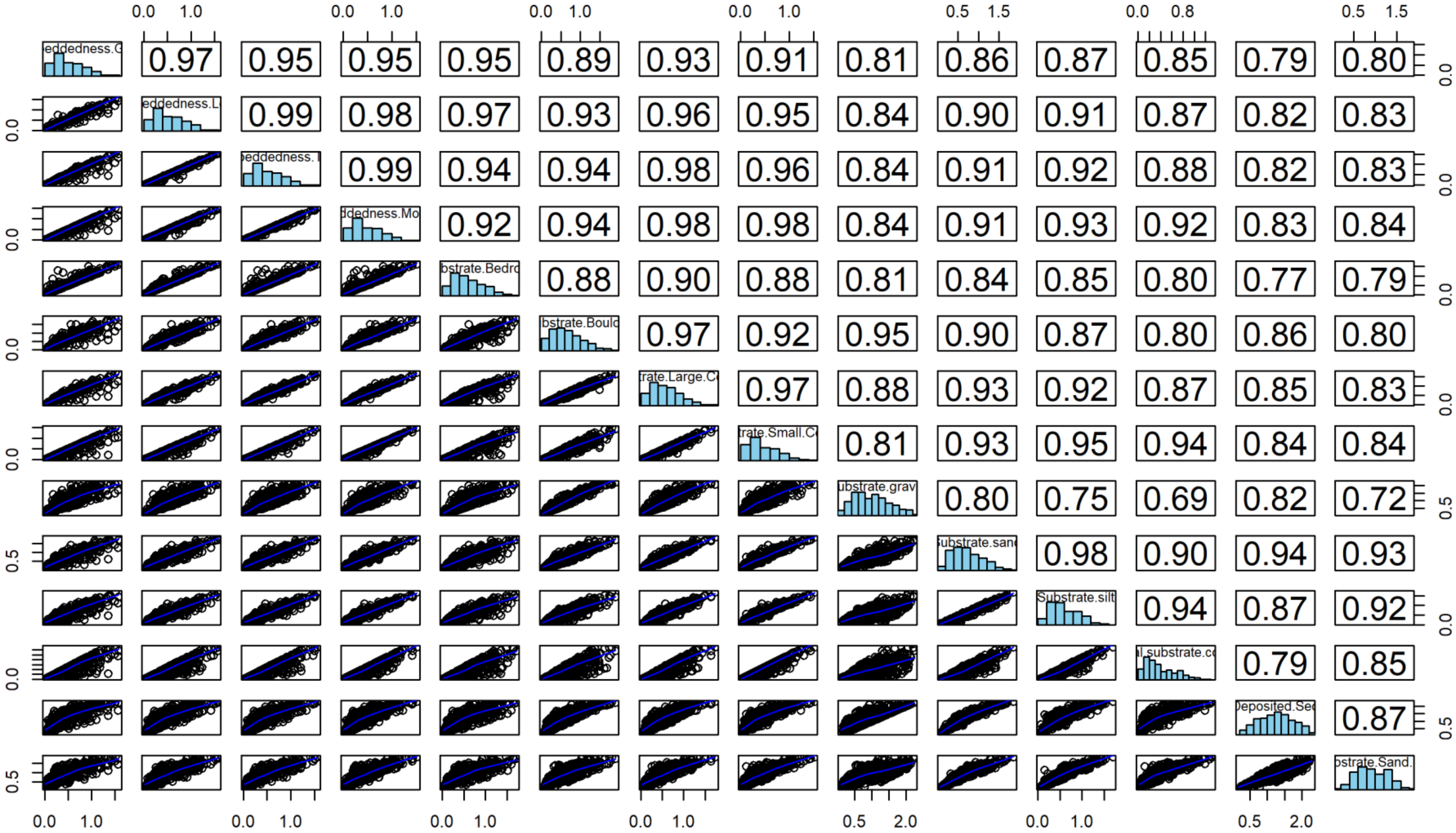

Figure C-4: Pearson's correlation coefficients and pairwise relationships between substrate measurements.

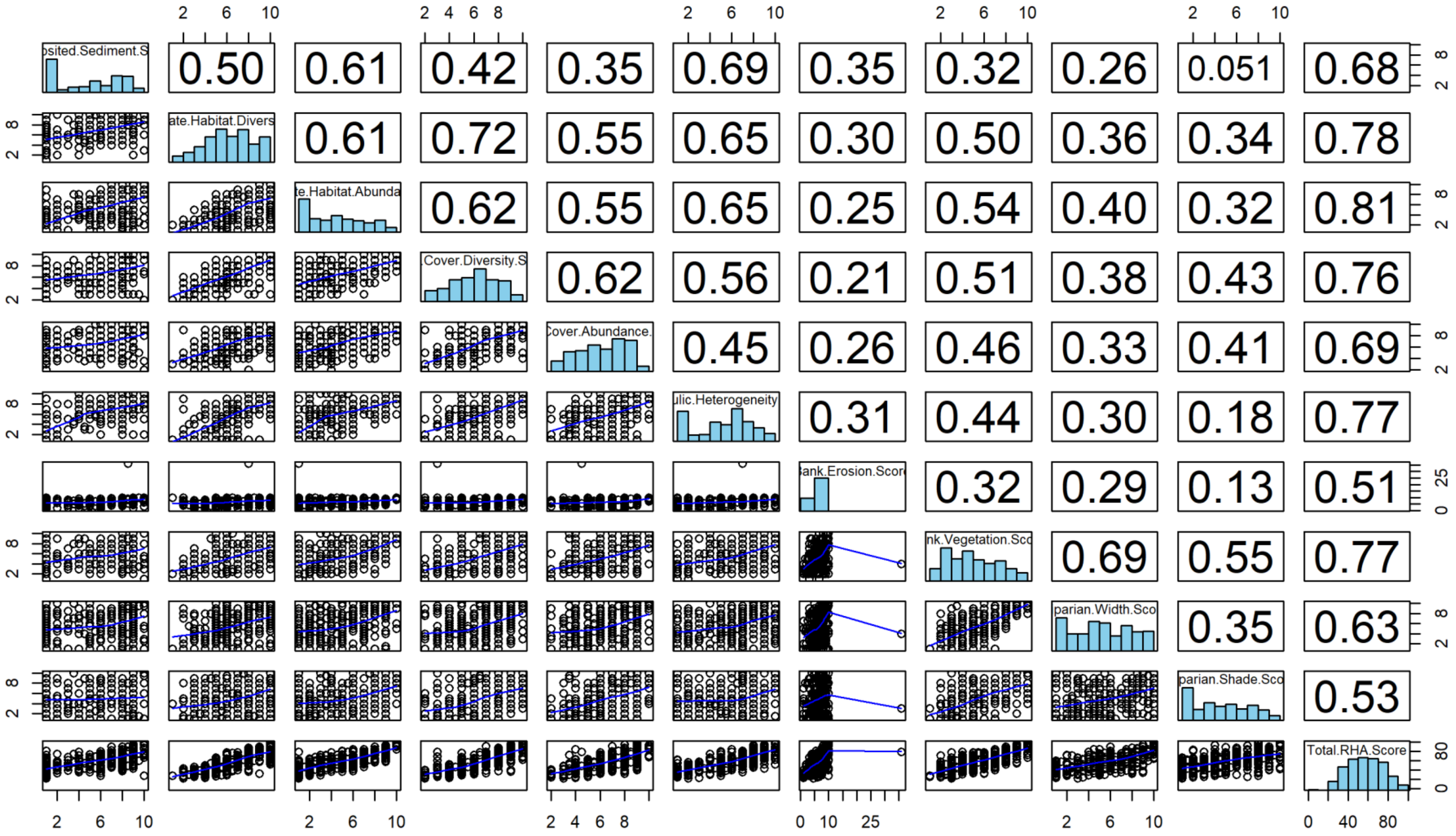

Figure C-5: **Pearson's correlation coefficients and pairwise relationships between habitat assessment scores.**

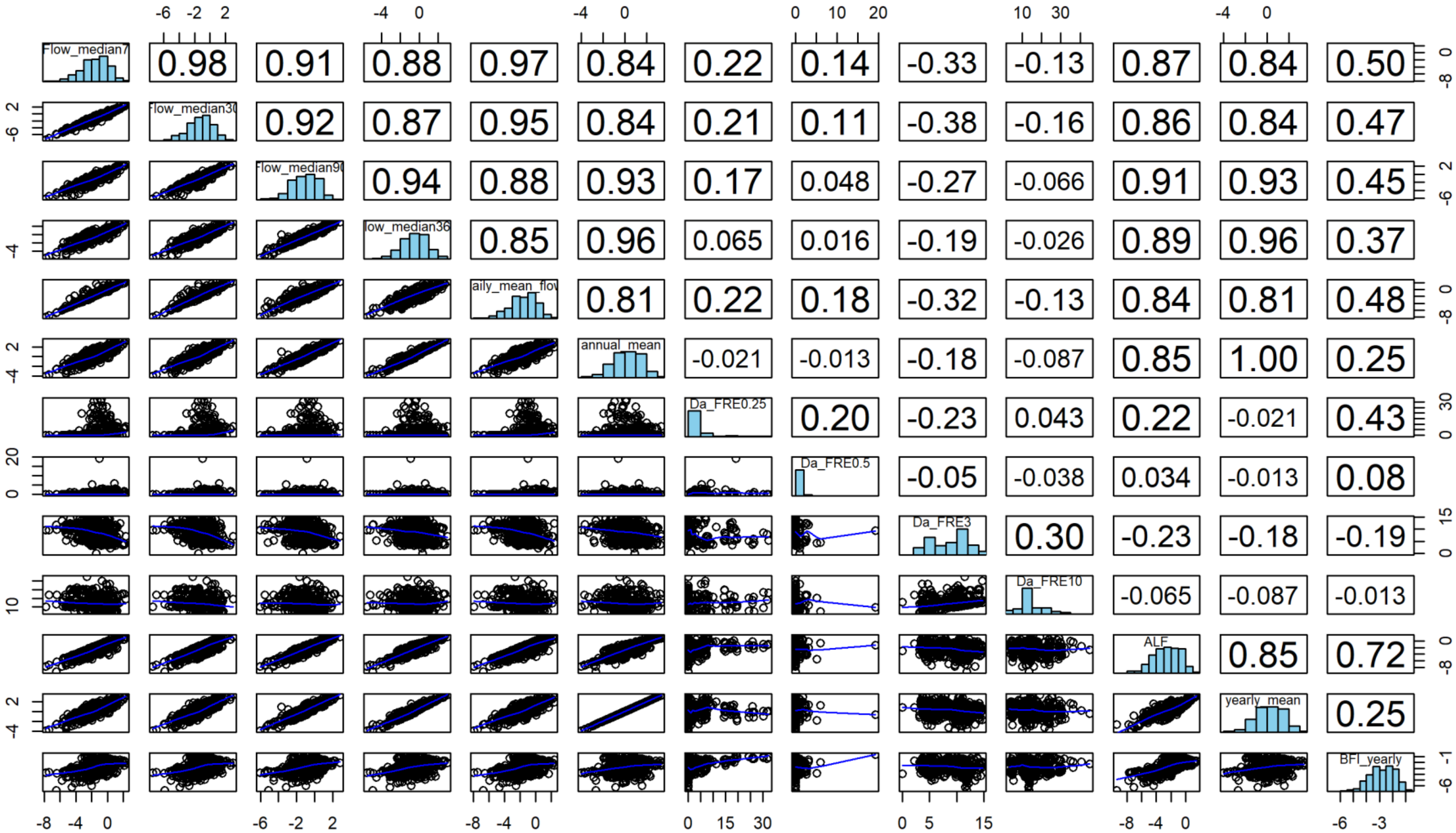

Figure C-6: Pearson's correlation coefficients and pairwise relationships between flow measurements .

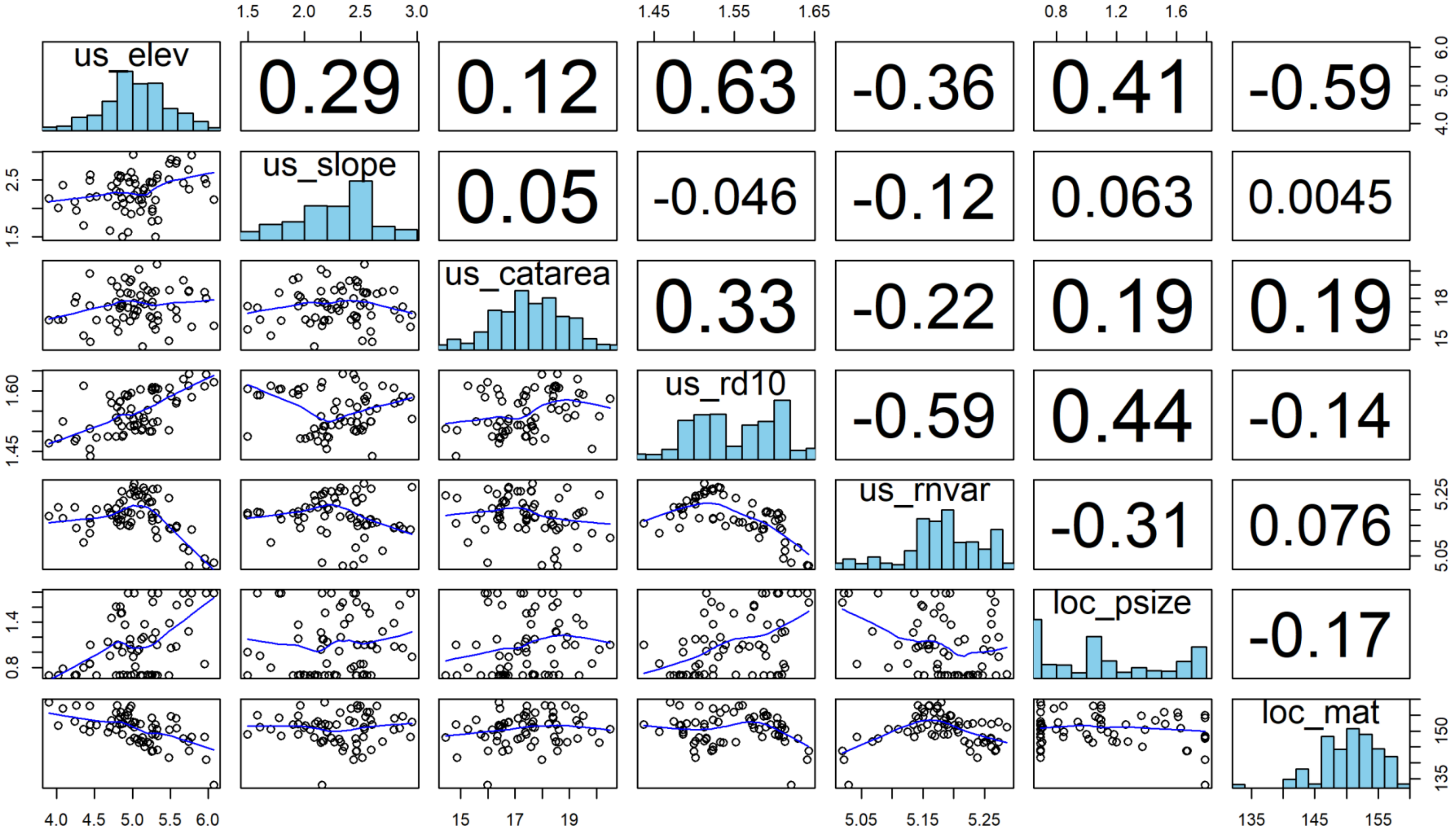

Figure C-7: Pearson's correlation coefficients and pairwise relationships between REC spatial attributes .

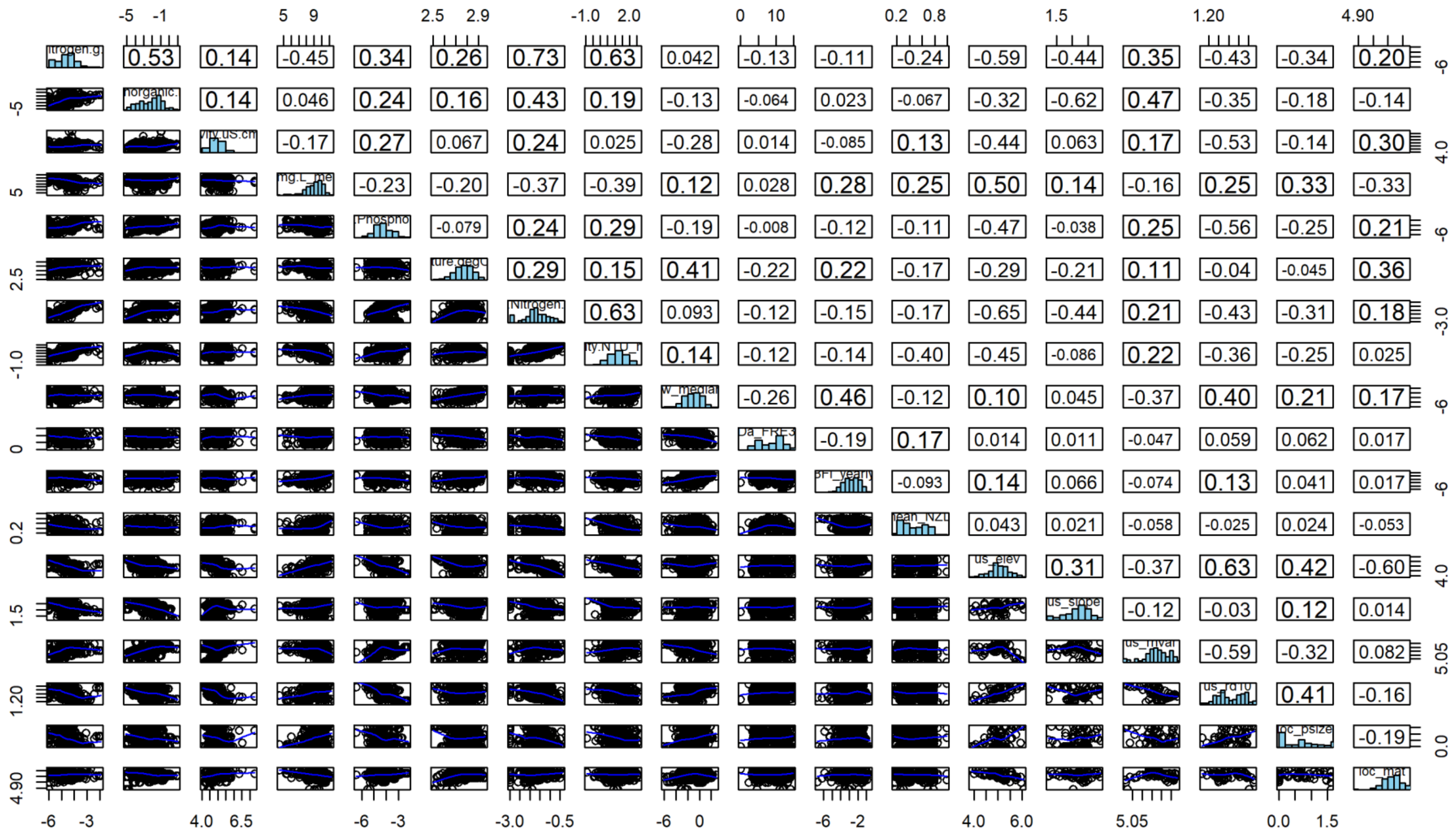

Figure C-8: Pearson's correlation coefficients and pairwise relationships between drivers selected for full subsets analysis.

Appendix D Metrics vs mean NZDI

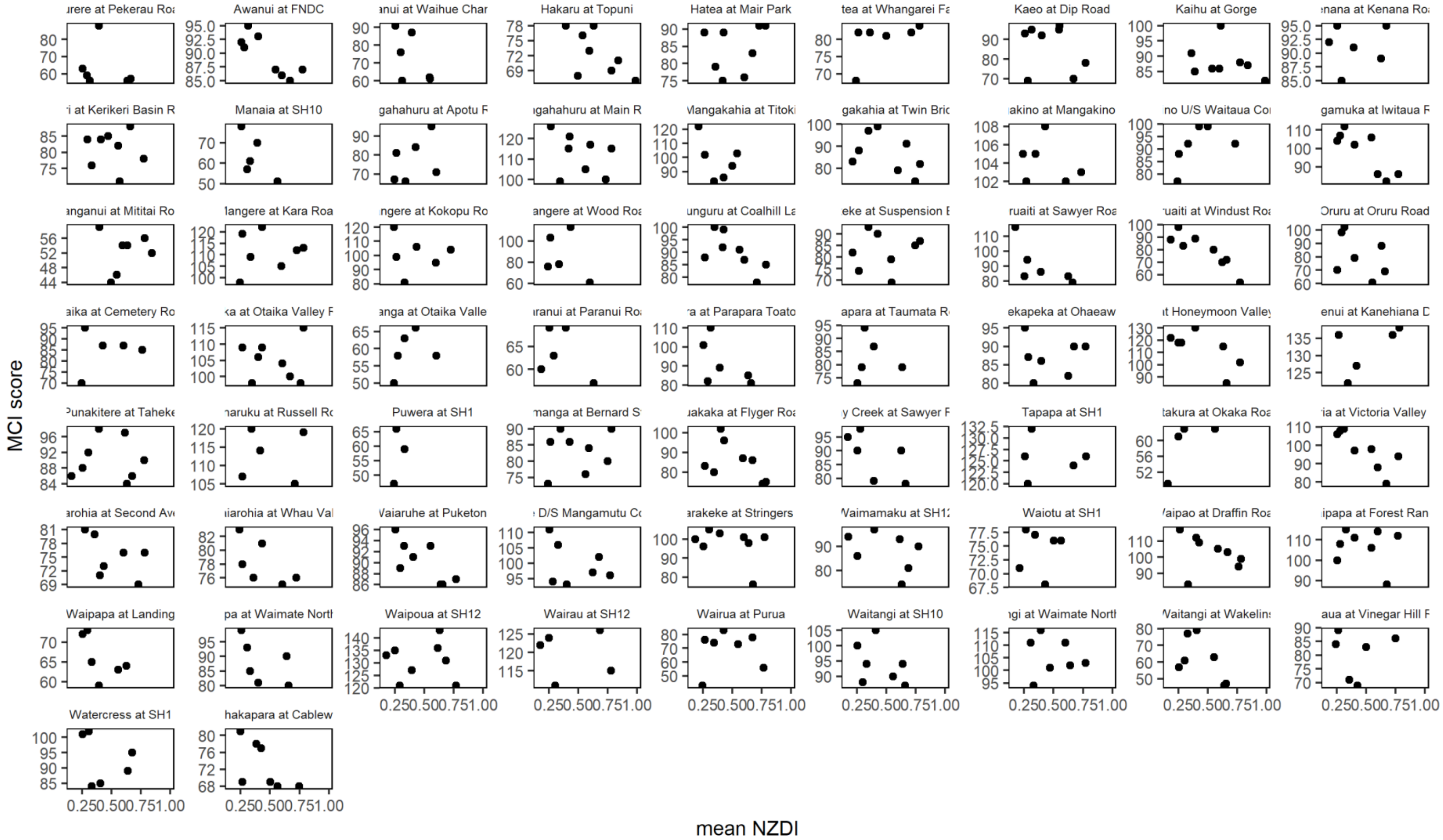

Figure D-1: MCI scores vs mean NZDI for all Northland SoE monitoring sites.

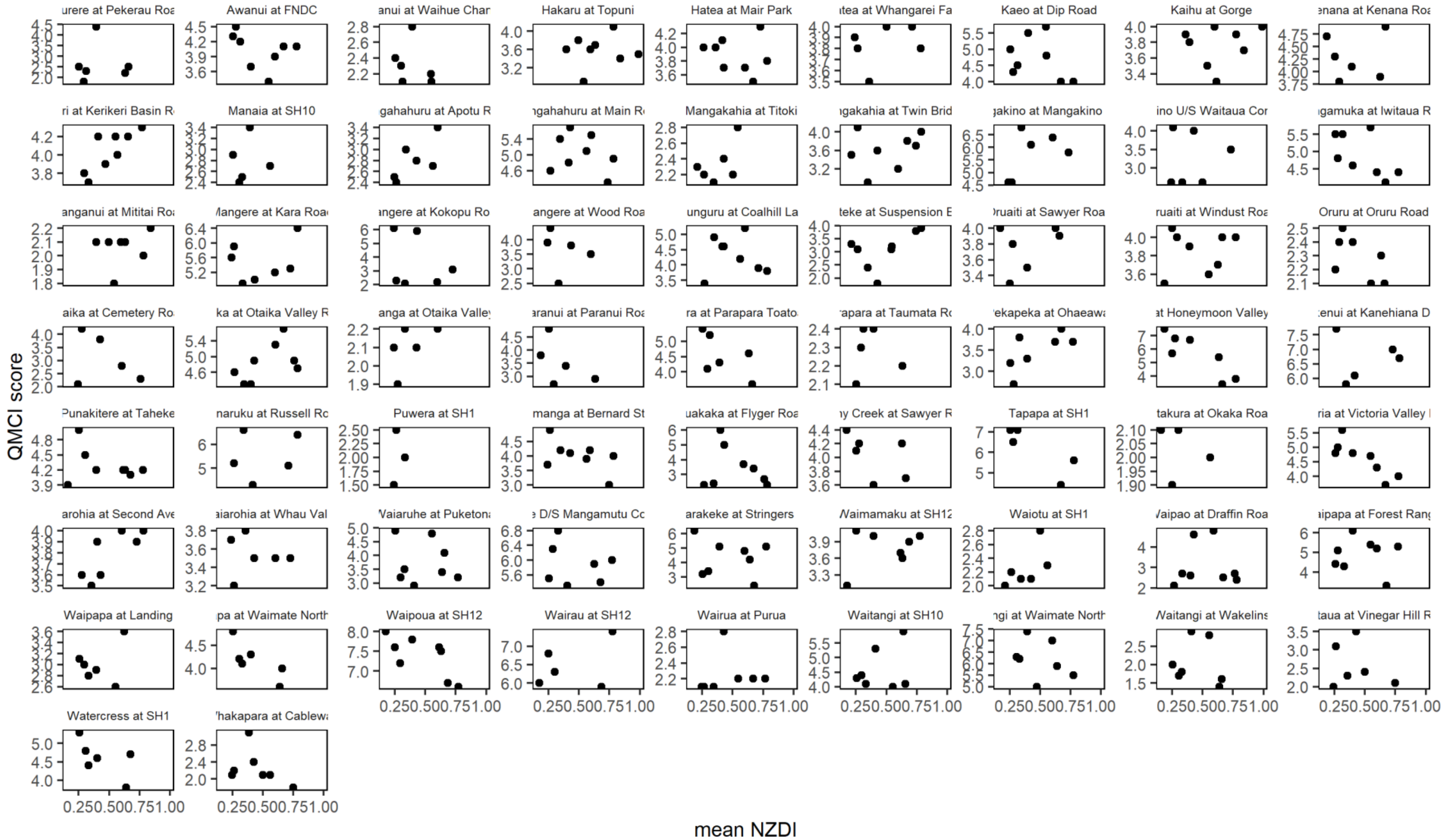

Figure D-2: QMCI scores vs mean NZDI for all Northland SoE monitoring sites.

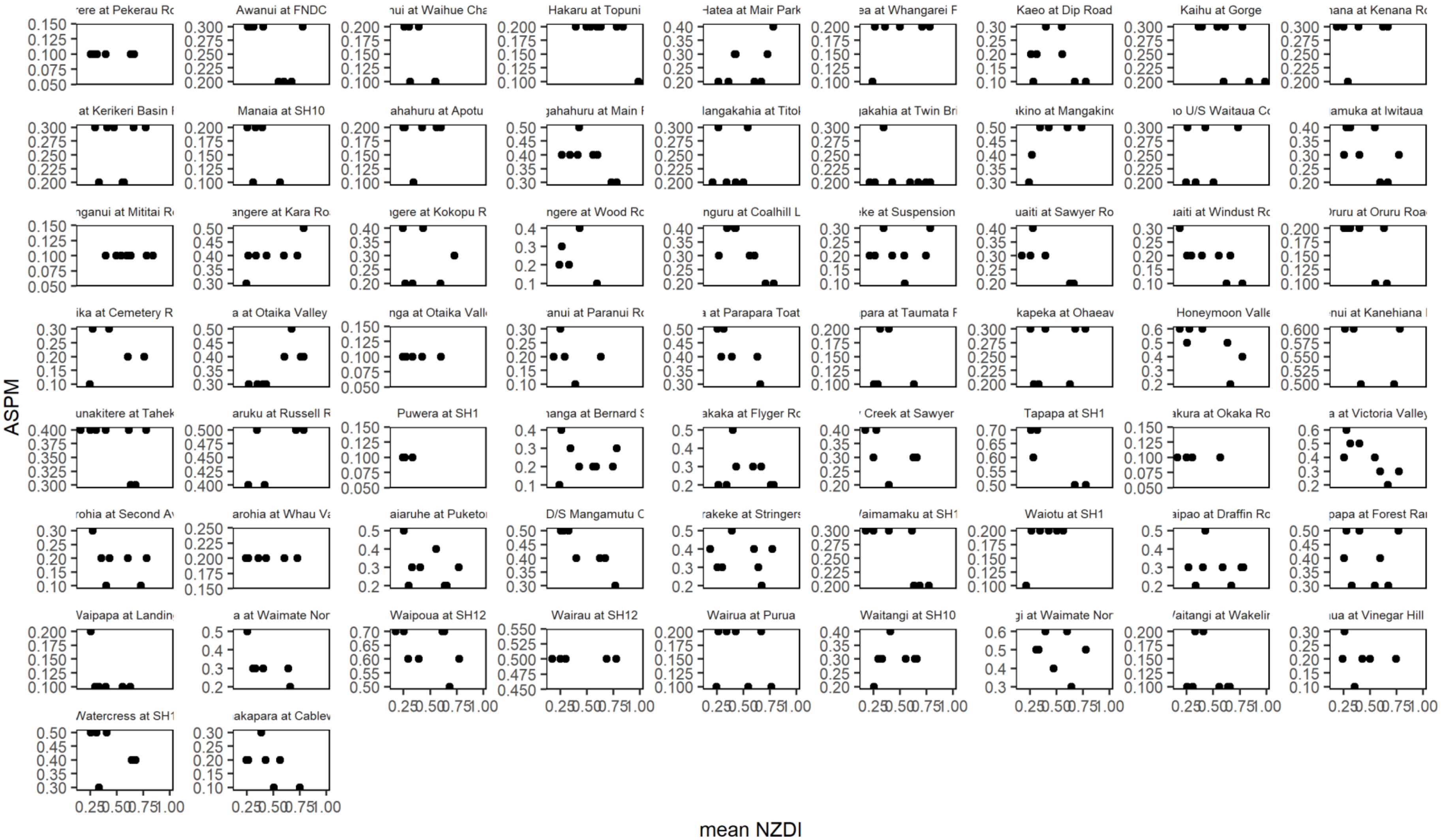

Figure D-3: ASPM vs mean NZDI for all Northland SoE monitoring sites.

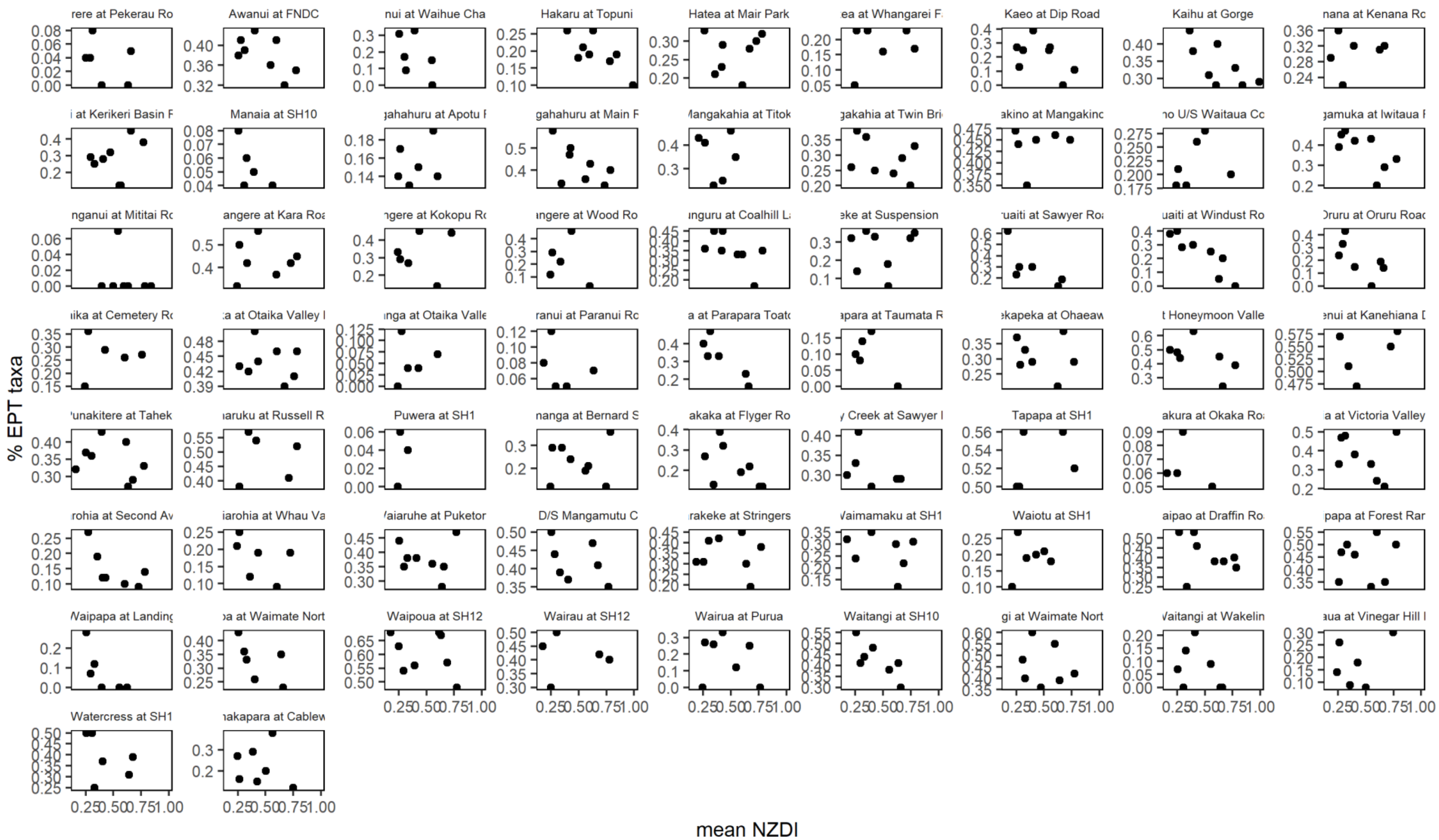

Figure D-4: % EPT taxa vs mean NZDI for all Northland SoE monitoring sites.

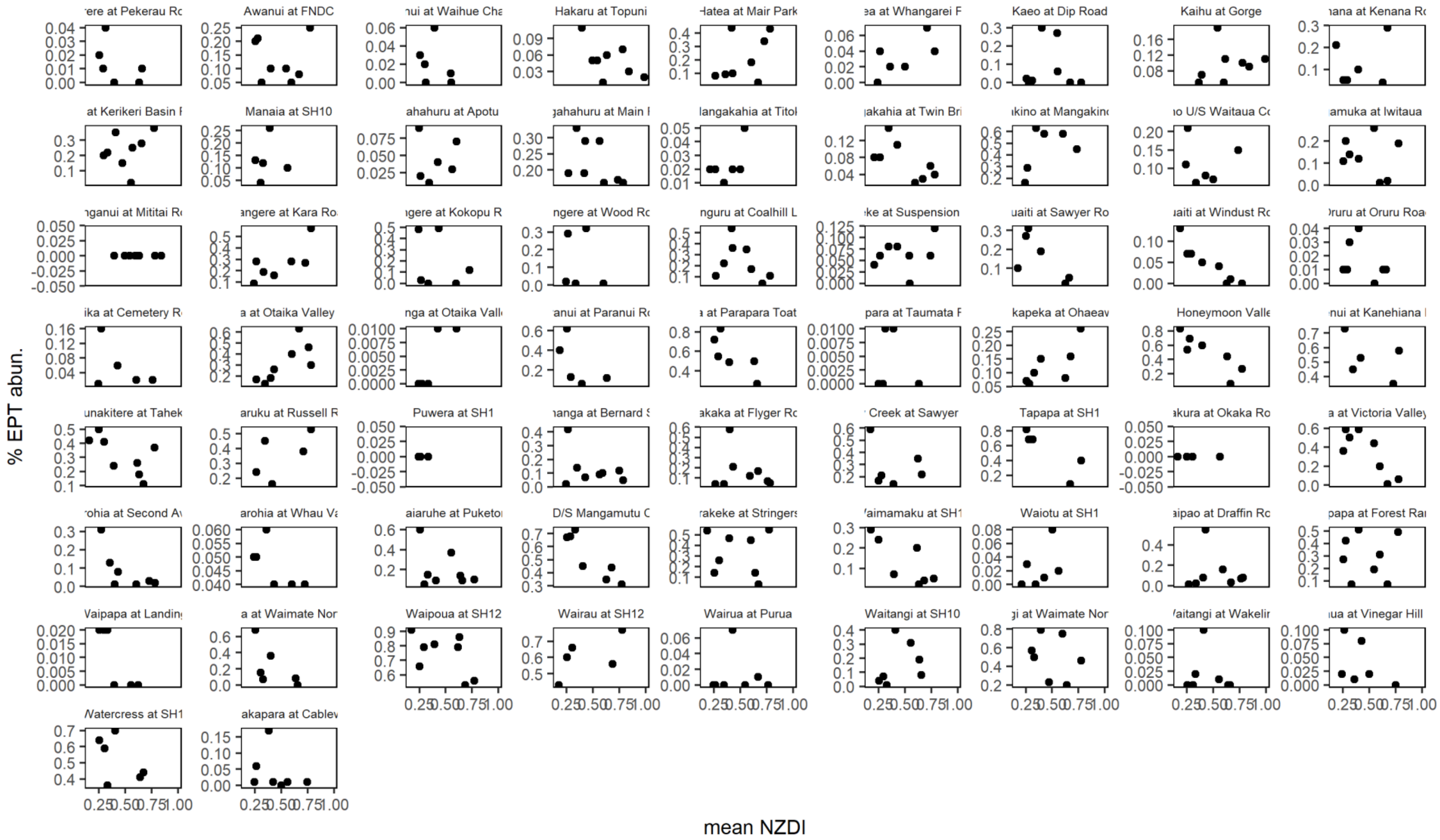

Figure D-5: % EPT abundance vs mean NZDI for all Northland SoE monitoring sites.

Appendix E Drivers vs mean NZDI

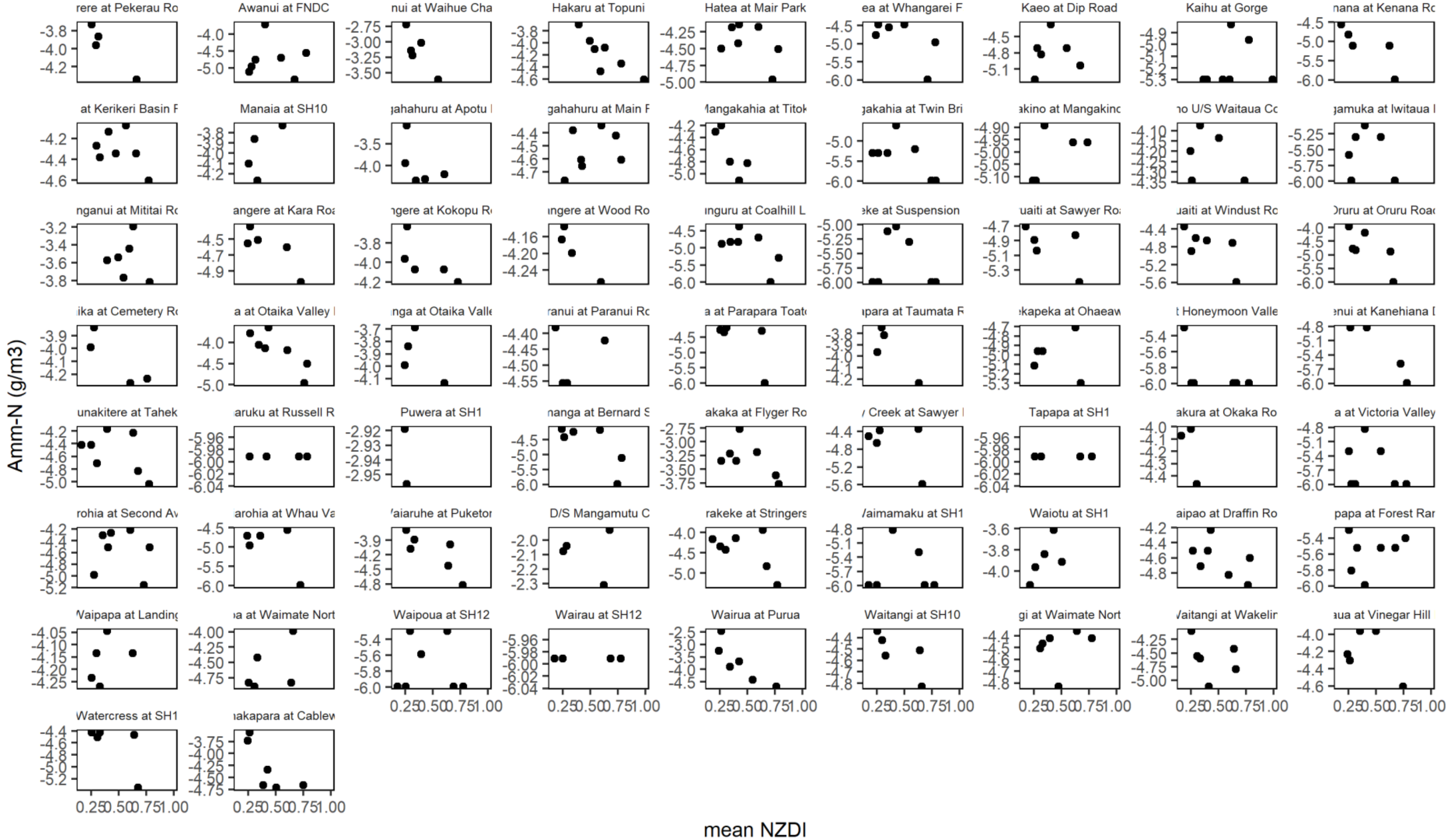

Figure E-1: Ammoniacal nitrogen vs mean NZDI for all Northland SoE monitoring sites.

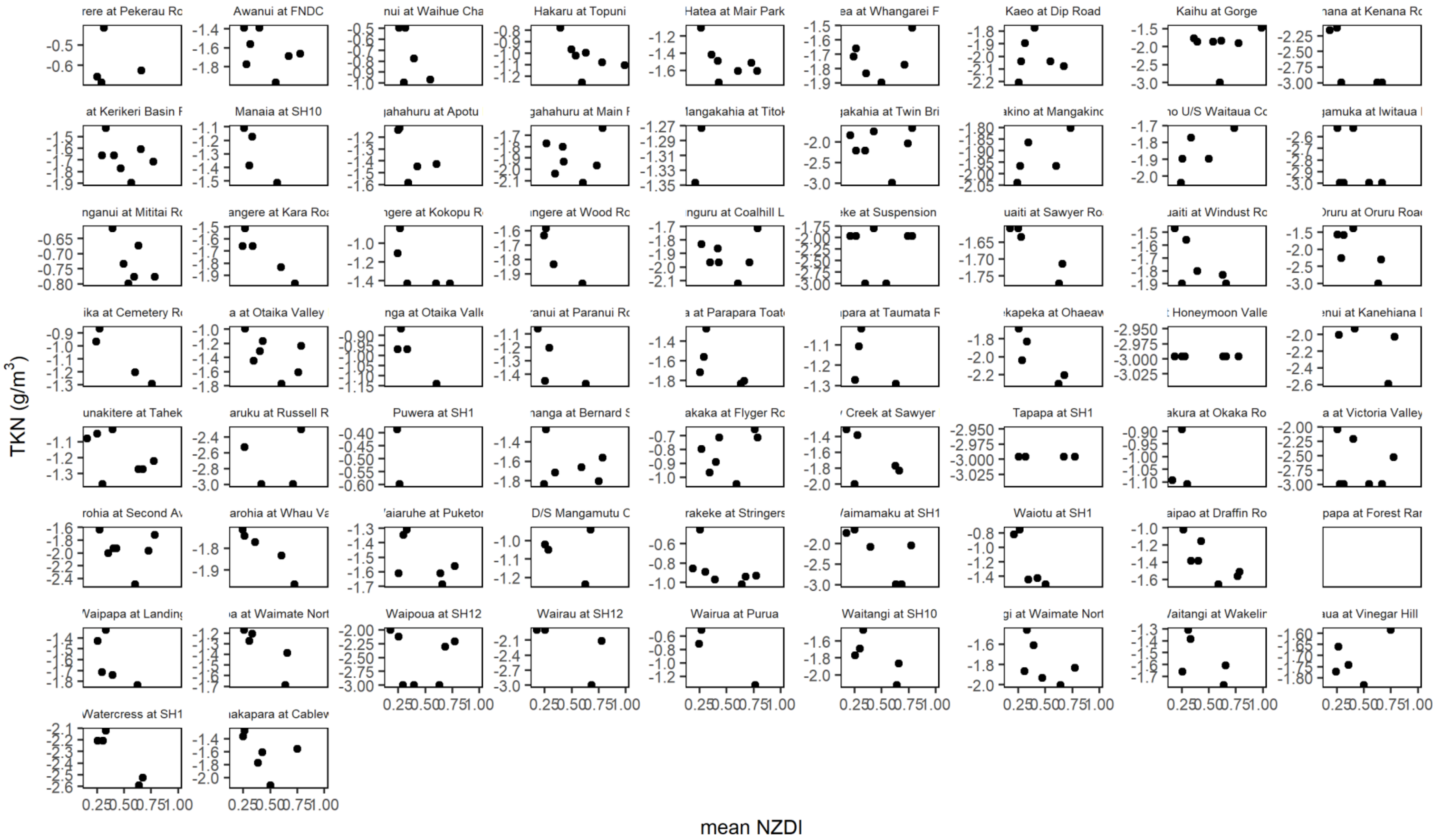

Figure E-2: Total Kjeldahl nitrogen vs mean NZDI for all Northland SoE monitoring sites.

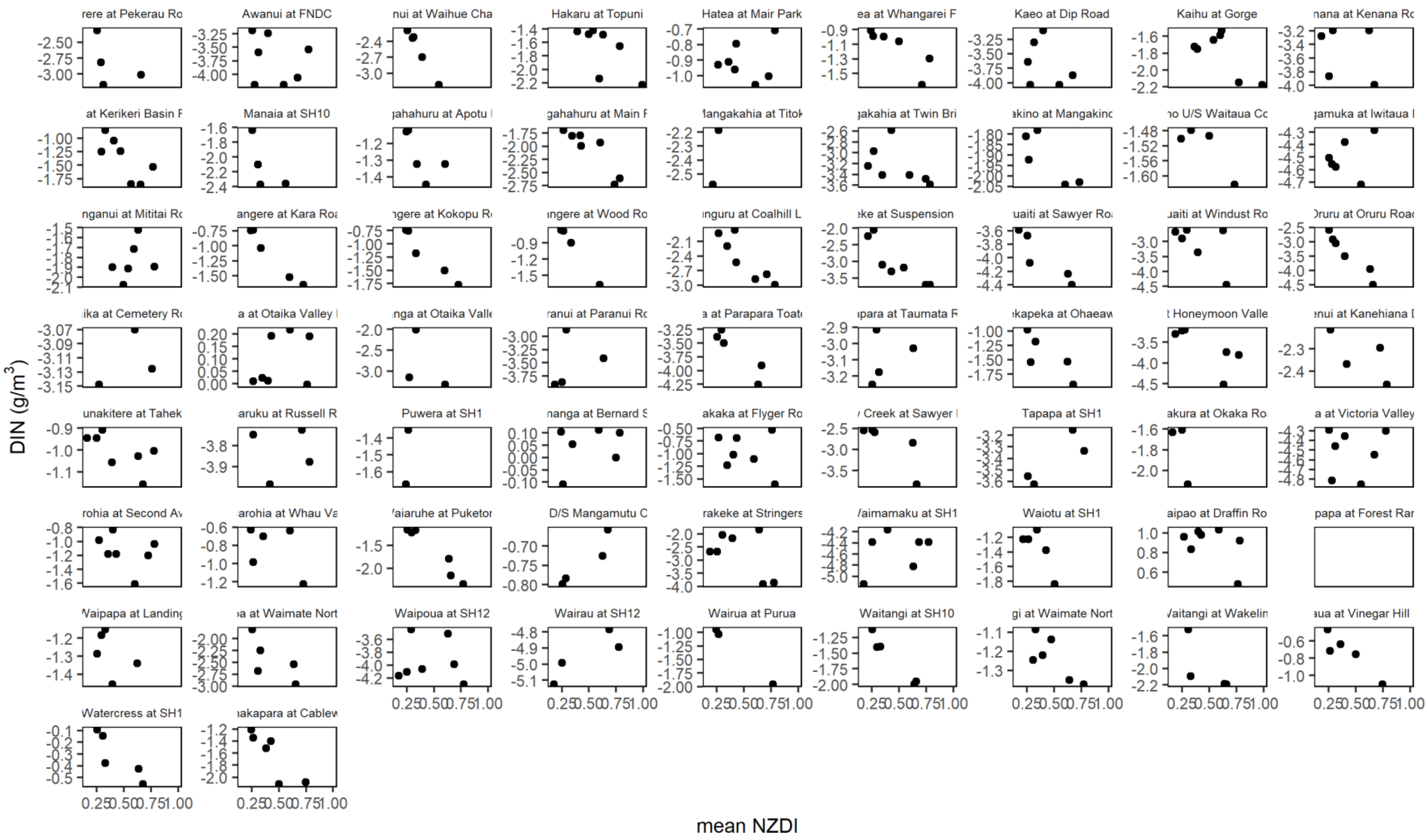

Figure E-3: Dissolved inorganic nitrogen vs mean NZDI for all Northland SoE monitoring sites.

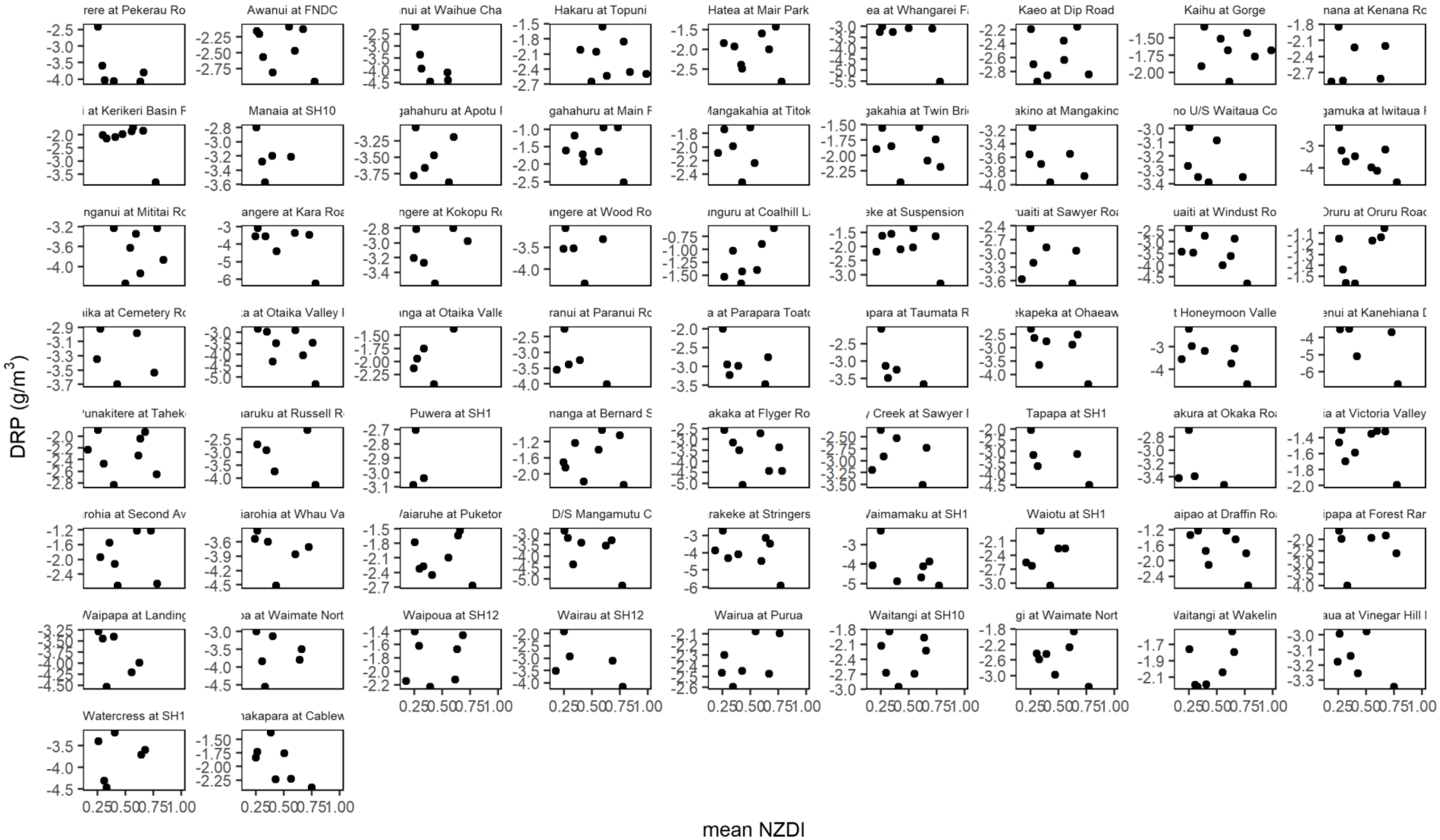

Figure E-4: Dissolved reactive phosphorus vs mean NZDI for all Northland SoE monitoring sites.

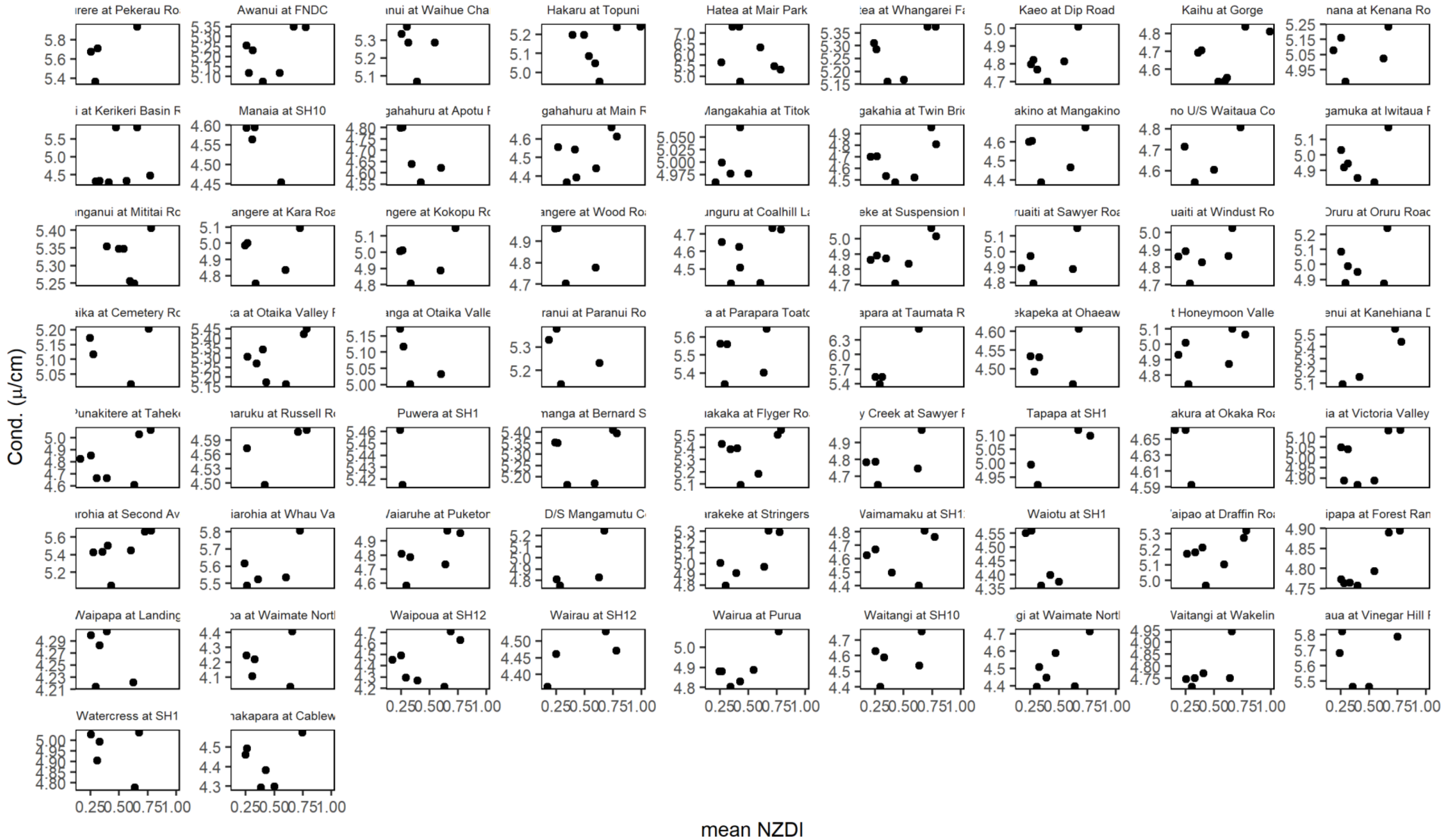

Figure E-5: Conductivity vs mean NZDI for all Northland SoE monitoring sites.

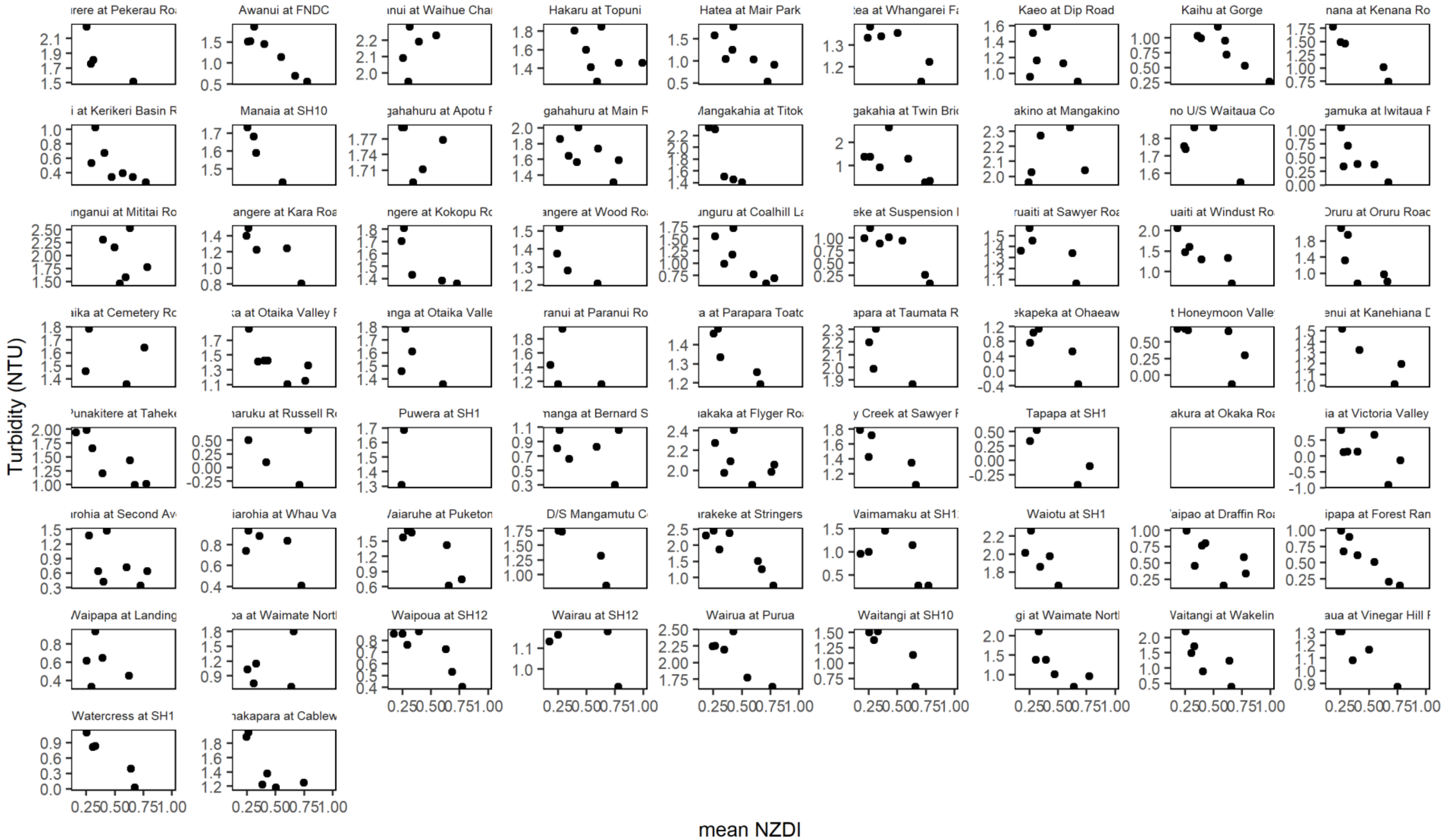

Figure E-6: Turbidity vs mean NZDI for all Northland SoE monitoring sites.

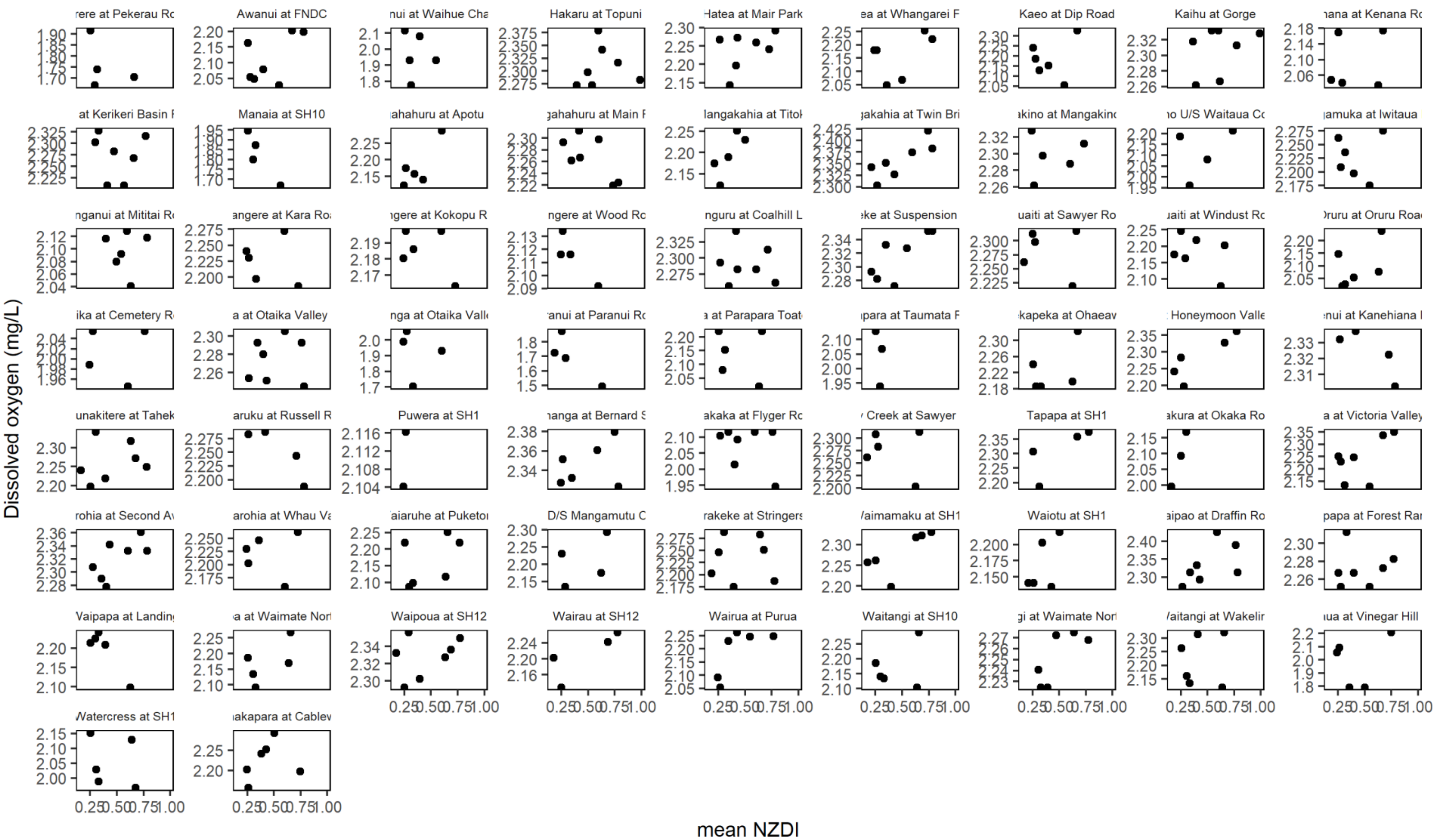

Figure E-7: Dissolved oxygen vs mean NZDI for all Northland SoE monitoring sites.

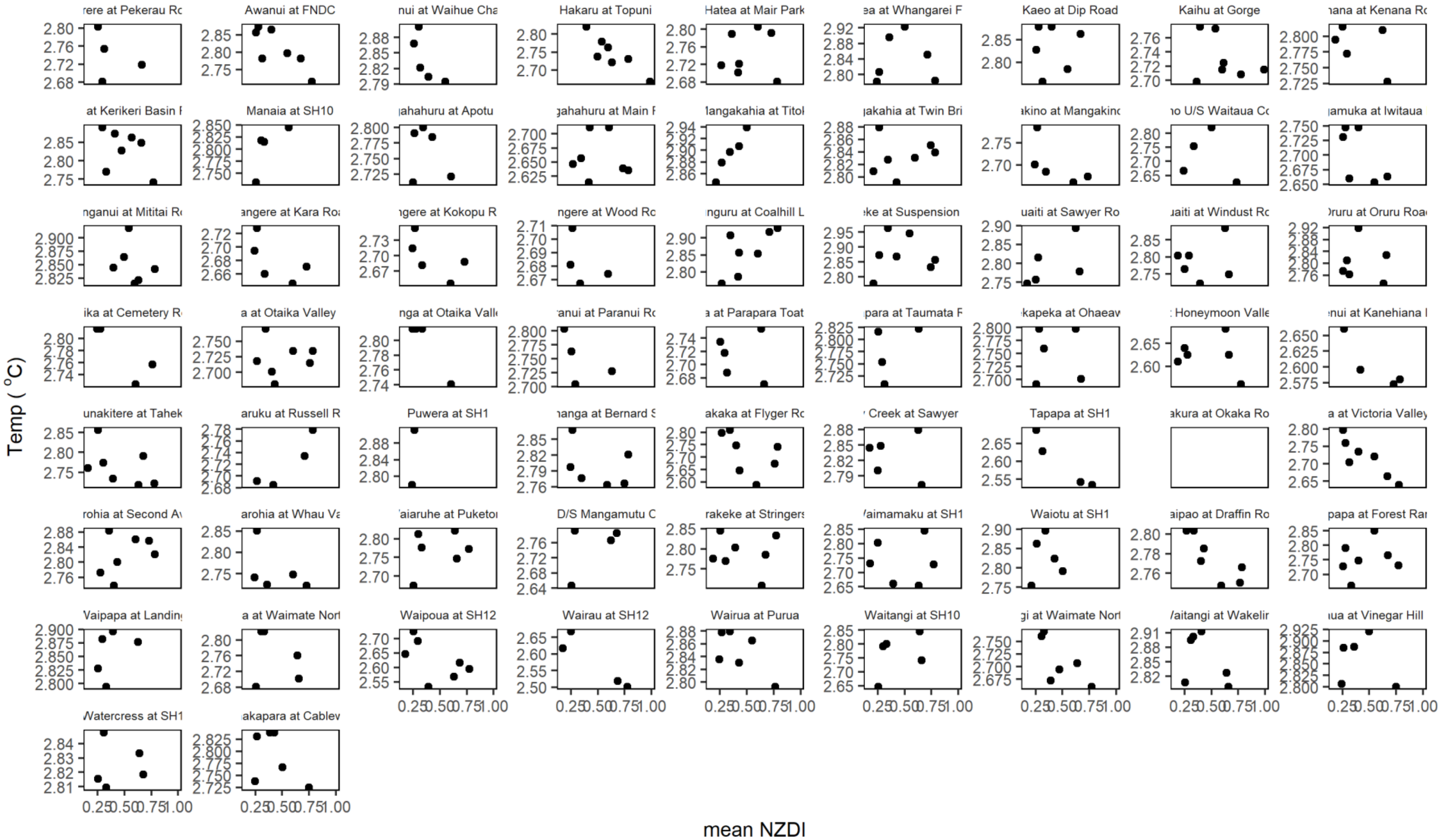

Figure E-8: Temperature vs mean NZDI for all Northland SoE monitoring sites.

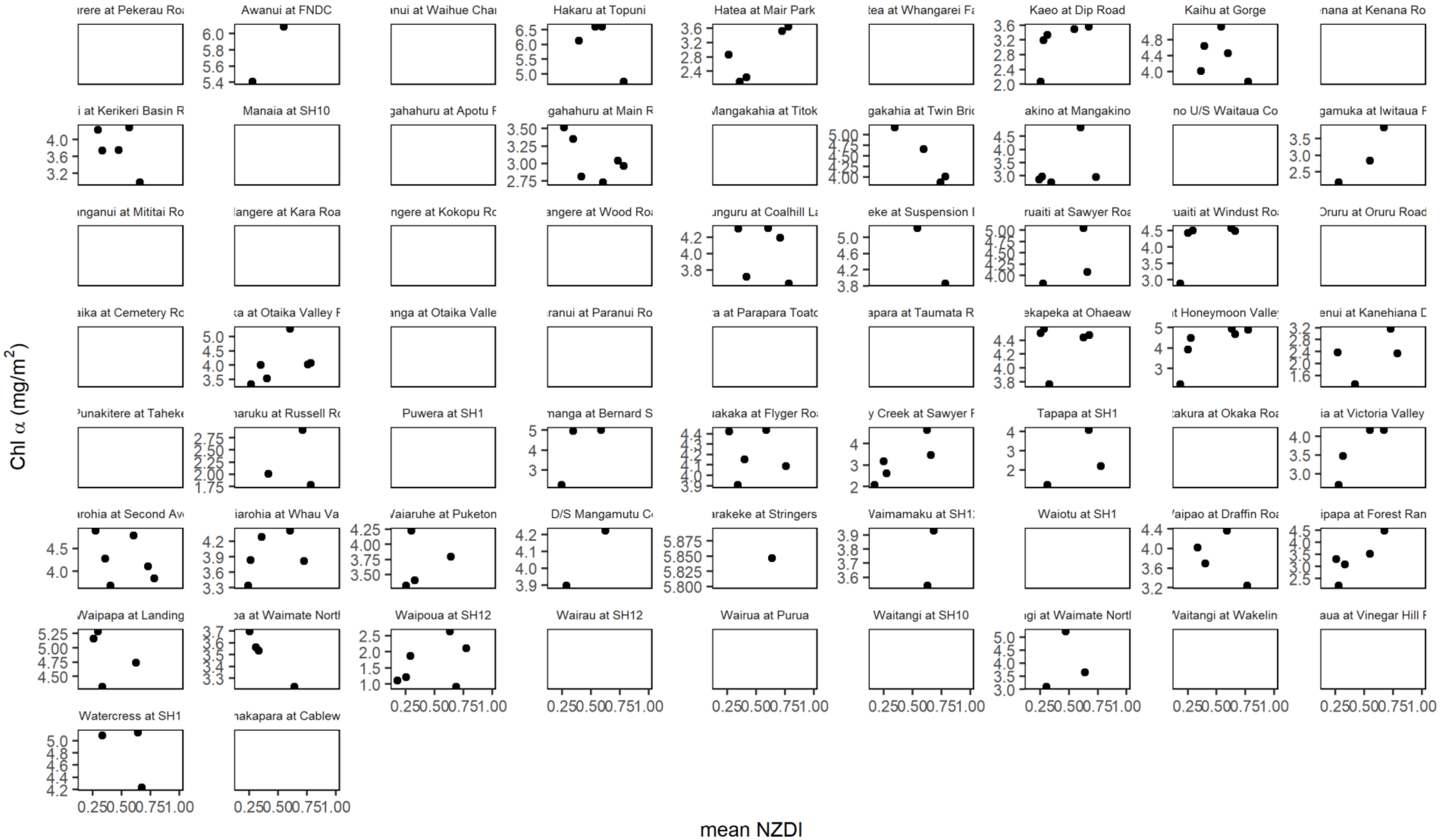

Figure E-9: **Chlorophyll *a* vs mean NZDI for all Northland SoE monitoring sites.**

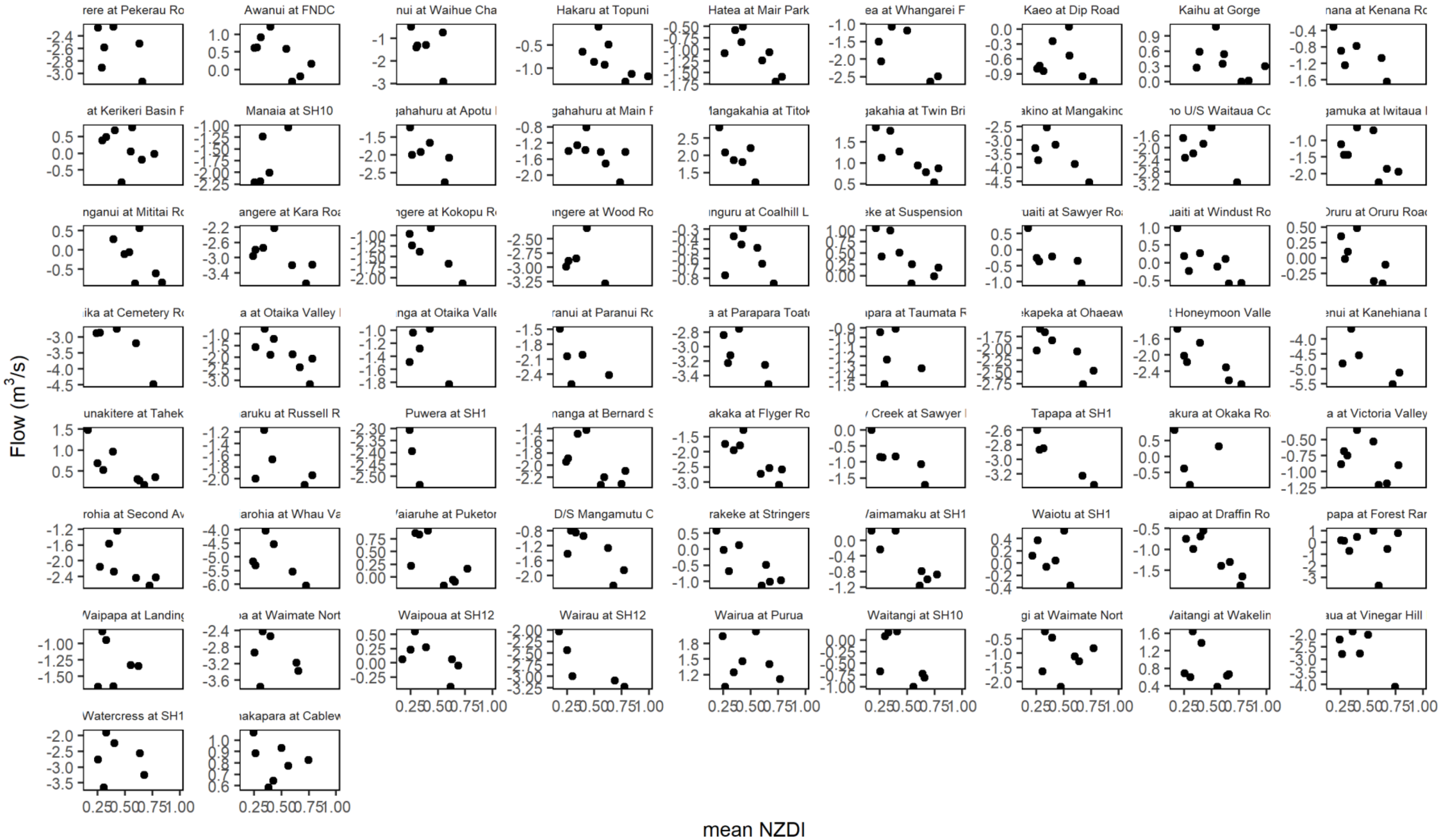

Figure E-10: Flow vs mean NZDI for all Northland SoE monitoring sites.

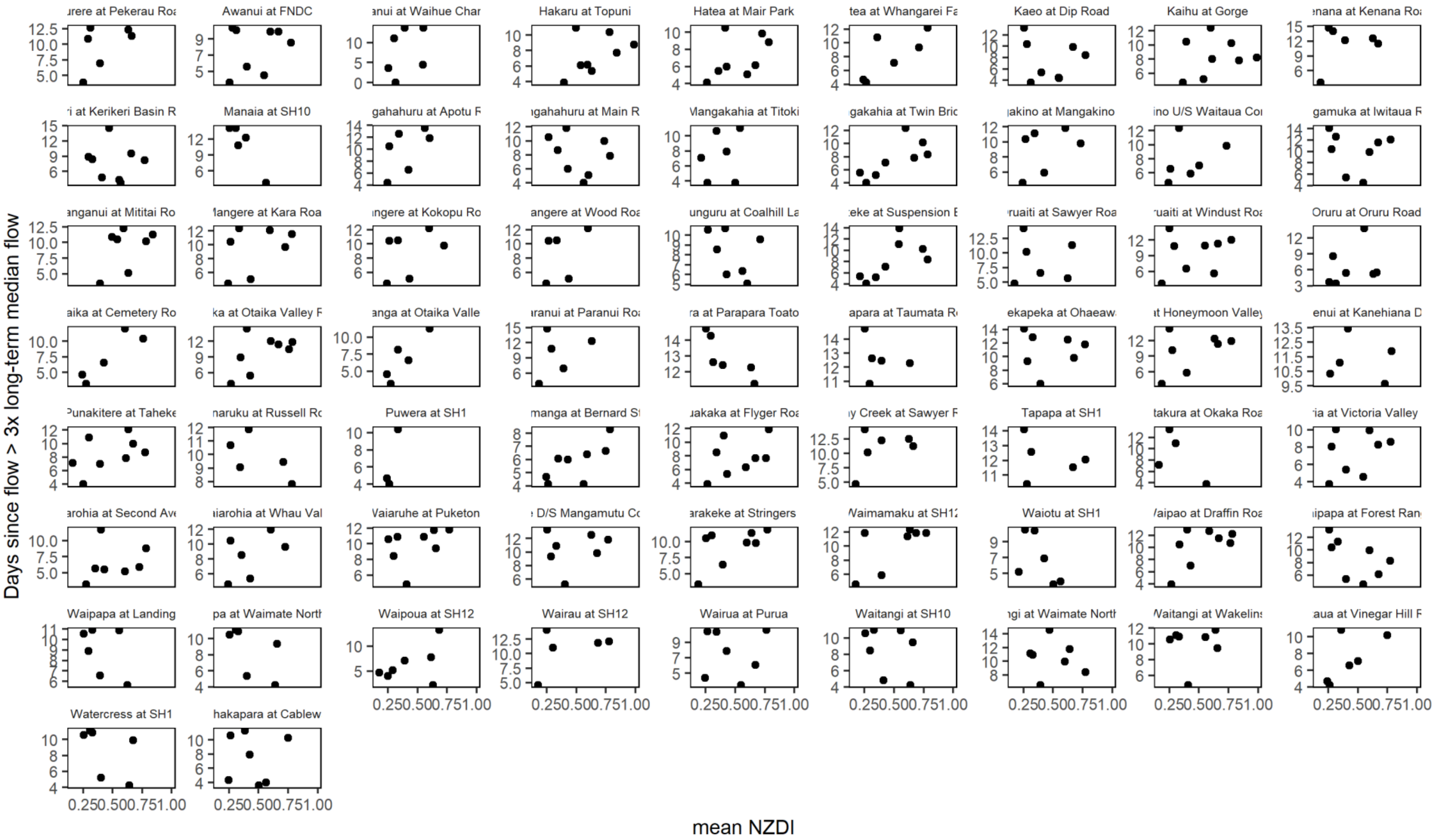

Figure E-11: Days since last flow greater than 3 times the long-term median flow (DaFRE3) vs mean NZDI for all Northland SoE monitoring sites.

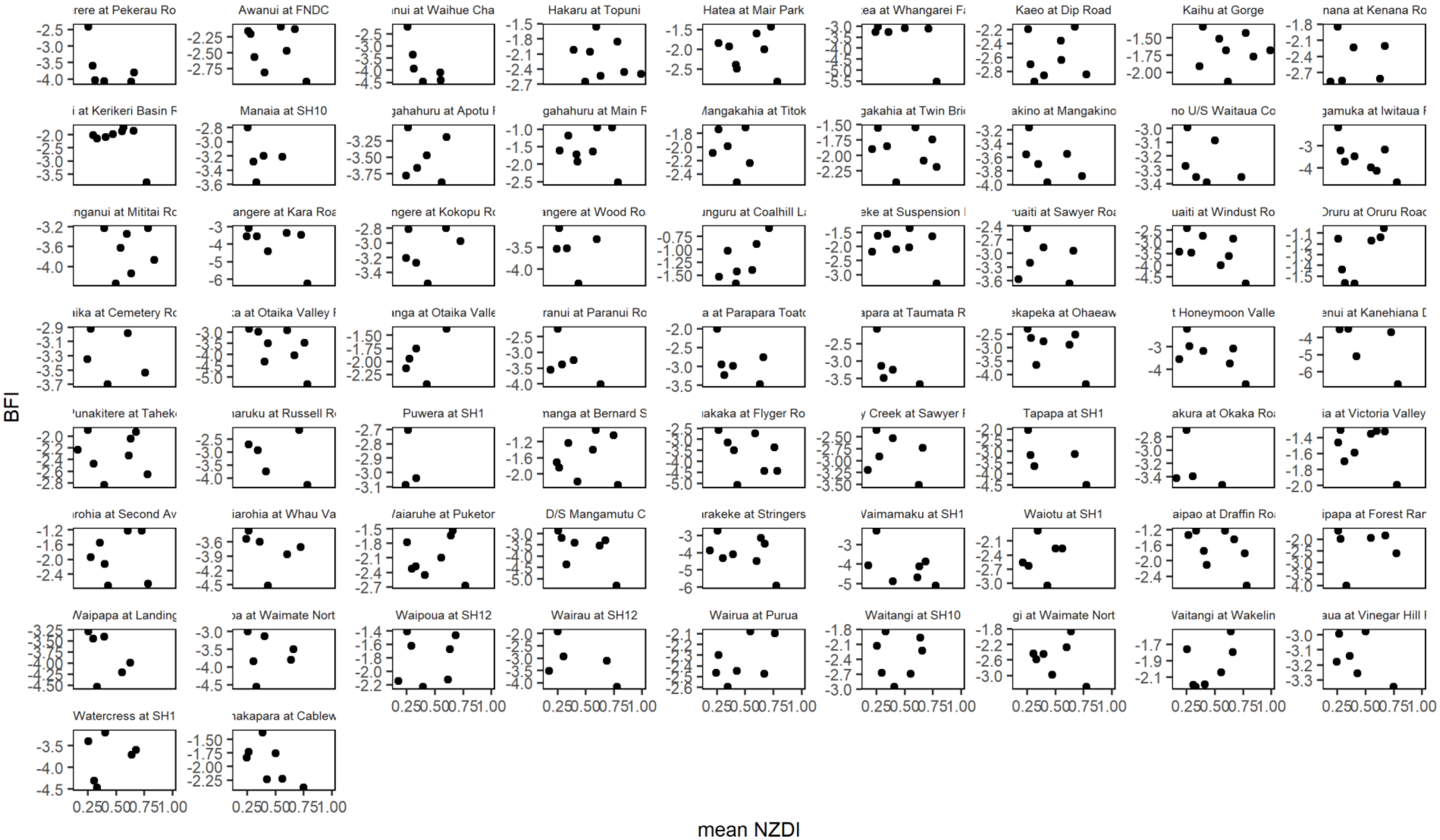

Figure E-12: Base flow index vs mean NZDI for all Northland SoE monitoring sites

Appendix F Taxa turnover

Table F-1: Total turnover by individual taxa. 'Stays' indicates the taxa remained present in the site. Mean abundance is per site.

Taxa	EPT	Appearances	Disappearances	Stays	Total turnovers	# Sites	Mean abundance
Austrosimulium		98	95	80	193	57	18
Oligochaeta		93	92	196	185	66	82
Triplectides	T	86	93	126	179	61	40
Polypedilum		91	87	125	178	62	32
Acarina		87	90	104	177	65	16
Platyhelminthes		93	84	96	177	60	34
Oxyethira	(T)	83	87	195	170	63	45
Hudsonema	T	84	84	95	168	54	16
Nemertea		86	74	69	160	53	14
Hydrobiosis	T	84	75	136	159	54	11
Gundlachia = Ferrissia		78	77	97	155	56	32
Tanypodinae		83	72	79	155	54	10
Orthocladiinae		71	76	238	147	65	70
Pycnocentria	T	70	66	113	136	51	37
Zephlebia	E	69	67	152	136	55	38
Paroxyethira	(T)	66	69	30	135	46	22
Tanytarsini		68	64	244	132	64	116
Lymnaeidae		62	69	46	131	46	15
Austroclima	E	61	64	73	125	48	21
Deleatidium	E	59	55	102	114	43	59
Hirudinea		55	53	20	108	39	7
Aphrophila		53	54	67	107	36	9
Physa = Physella		52	55	103	107	45	42
Muscidae		56	49	55	105	39	9
Pycnocentrodes	T	54	51	214	105	55	81
Archichauliodes		53	50	137	103	46	13
Hydropsyche - Aoteapsyche	T	55	46	191	101	53	52
Copepoda		47	51	21	98	33	59
Chironomus		47	47	10	94	34	88
Nematoda		45	48	4	93	35	9
Sphaeriidae		44	48	20	92	34	27
Elmidae		46	43	210	89	52	84
Maoridiamesa		44	43	23	87	27	17
Latia		40	46	53	86	33	16

Taxa	EPT	Appearances	Disappearances	Stays	Total turnovers	# Sites	Mean abundance
Xanthocnemis		40	45	64	85	33	63
Paracalliope		38	45	96	83	36	156
Ostracoda		41	40	32	81	33	64
Potamopyrgus		39	41	356	80	66	1981
Coloburiscus	E	42	36	73	78	35	31
Paratya		36	37	36	73	26	23
Gyraulus		34	35	23	69	28	71
Psilochorema	T	37	32	16	69	23	4
Paradixa		34	34	17	68	23	17
Empididae		35	32	14	67	25	7
Halicarcinus		31	34	41	65	21	7
Neurochorema	T	34	30	21	64	23	4
Microvelia		31	31	20	62	26	11
Enochrus		28	28	2	56	23	9
Sigara		25	28	23	53	23	11
Amphipoda		28	24	10	52	20	66
Hydra		24	28	3	52	21	18
Mischoderus		29	23	9	52	21	4
Eriopterini		25	20	22	45	19	3
Collembola		22	22	1	44	20	9
Anisoptera		20	22	2	42	17	13
Hemicordulia		21	21	5	42	16	4
Polyplectropus	T	22	20	5	42	18	11
Hydraenidae		23	18	28	41	22	8
Nesameletus	E	23	18	23	41	18	25
Costachorema	T	20	18	8	38	16	5
Harrisius		19	18	2	37	15	7
Mauiulus	E	18	19	0	37	17	26
Acanthophlebia	E	20	16	6	36	12	5
Olinga	T	19	17	33	36	18	24
Corynoneura		17	17	2	34	14	35
Hygraula		17	17	6	34	15	4
Melanopsis = Zemelanopsis		16	18	26	34	15	24
Oecetis	T	15	17	12	32	17	11
Ceratopogonidae		15	16	4	31	15	4
Rallidens	E	16	15	5	31	13	4
Ameletopsis	E	14	14	8	28	11	5
Anisops		14	14	4	28	10	10

Taxa	EPT	Appearances	Disappearances	Stays	Total turnovers	# Sites	Mean abundance
Antiporus		14	14	0	28	10	3
Plectrocnemia	T	13	14	1	27	12	11
Ichthybotus	E	14	12	7	26	10	4
Ephydridae		12	11	1	23	13	5
Oniscigaster	E	12	11	3	23	10	2
Zelandobius	P	12	11	3	23	9	3
Psychodidae		9	10	3	19	10	9
Austrolestes		9	9	3	18	6	3
Beraeoptera	T	9	9	15	18	9	37
Tanaidacea		7	10	18	17	8	75
Antipodochlora		8	8	2	16	8	2
Helicopsyche	T	9	7	12	16	10	22
Neozephlebia	E	10	6	5	16	9	23
Paranephrops		8	8	0	16	8	46
Ptilodactylidae		8	8	7	16	8	4
Rhabdocoela		8	8	0	16	7	20
Austroperla	P	9	6	6	15	9	2
Glyptophysa = Physastra		6	7	1	13	6	5
Zelandoperla	P	8	5	16	13	7	12
Paralimnophila		6	6	1	12	4	7
Hexatomini		6	5	0	11	6	3
Hydropsyche - Orthopsyche	T	6	5	12	11	7	10
Tabanidae		4	7	4	11	6	3
Diaprepocoris		5	5	2	10	4	12
Hydrophilidae		5	5	0	10	5	9
Stenoperla	P	5	5	4	10	4	7
Arachnocolus	E	4	4	0	8	4	5
Cladocera		4	4	0	8	4	22
Culex		4	4	1	8	4	7
Lobodiamesa		4	4	0	8	4	12
Paraleptamphopus		4	4	2	8	4	6
Tepakia	E	4	4	0	8	4	22
Megaleptoperla	P	4	3	2	7	3	2
Zelandoptila	T	3	4	0	7	3	7
Dixidae		3	3	0	6	3	12
Isopoda		3	3	0	6	3	26
Limonia		3	3	0	6	3	2
Procordulia		3	3	0	6	3	14

Taxa	EPT	Appearances	Disappearances	Stays	Total turnovers	# Sites	Mean abundance
Sciomyzidae		3	3	0	6	3	2
Austronella	E	2	3	1	5	2	22
Hydrobiosella	T	3	2	1	5	2	27
Acroperla	P	2	2	0	4	2	3
Aeshna		2	2	0	4	2	2
Ischnura		2	2	0	4	2	4
Oeconesidae	T	2	2	0	4	2	5
Paucispinigera		2	2	0	4	2	5
Rhantus		2	2	0	4	1	9
Siphlaenigma	E	2	2	0	4	2	6
Stictocladius		2	2	2	4	2	3
Polychaeta		1	2	2	3	2	8
Stratiomyidae		2	1	0	3	2	5
Atalophlebioides	E	1	1	0	2	1	1
Chironomidae		1	1	0	2	1	12
Copelatus		1	1	0	2	1	1
Dytiscidae		1	1	0	2	1	1
Hemianax		1	1	0	2	1	1
Molophilus		1	1	0	2	1	2
Mysidae		1	1	0	2	1	1
Nematomorpha		1	1	0	2	1	2
Nothodixa		1	1	0	2	1	4
Spaniocerca	P	1	1	0	2	1	2
Staphylinidae		1	1	0	2	1	1
Taraperla	P	1	1	0	2	1	4
Zelolessica	T	1	1	0	2	1	4
Hydrochorema	T	1	0	0	1	1	1

Appendix G Percent EPT taxa and percent EPT abundance within sites

Figure G-1: Percent EPT taxa and percent EPT abundance over time in each site.