Memorandum: Test of revised periphyton nutrient criteria for Otago and Southland Regions

Author: Ton Snelder

Version 2, 6 November 2023

Introduction

Snelder et al. (2022) derived nutrient concentration criteria to achieve the target attribute states for river periphyton set out in the National Policy Statement – Freshwater Management (NPS-FM; NZ Government, 2020). The approach was based on fitting ordinary least squares regression (OLS) models to chlorophyll observations (summarized as the 92nd percentile of the observations and referred to hereafter as Chla92) at 251 monitoring sites distributed across New Zealand. The model explanatory variables comprised nutrient concentrations (summarised as median values of the observations of dissolved inorganic nitrogen (DIN), dissolved reactive phosphorus (DRP), total nitrogen (TN) and total phosphorus (TP)) and other environmental observations at the sites including substrate composition, shade and hydrological indices. These fitted models were subsequently used to defined criteria for DIN, DRP, TN and TP to achieve fixed Chla92 thresholds (50, 120 and 200 mg m^{-2}).

A validation of Snelder et al.'s (2022) criteria for the Otago and Southland Regions concluded that derived criteria were too permissive (i.e., the criteria concentrations are too high; LWP memo to ORC dated 22 February 2023). Validation of the periphyton nutrient concentration criteria derived by Snelder et al. (2022) using data from other regions (i.e., Wellington and the Manawatu-Wanganui regions) also showed them to be too permissive. These findings reduced confidence in the criteria of Snelder et al. (2022).

A study by Snelder and Kilroy (2023) aimed to revise the nutrient criteria based on a regression modelling approach, as used by Snelder et al. (2022), but fitting models using generalised linear models (GLM) instead of the OLS models that were used by Snelder et al. (2022). Briefly, the reason for the change in the modelling approach was because the original OLS models were positively biased. This means that, on average, the models underestimated Chla92 values for a given nutrient concentration. This meant that the models tended to appreciably under-estimate Chla92 at sites with high biomass, which in turn meant the criteria tended to be too permissive. It was anticipated that the revision would produce better nutrient criteria because the GLM models are more able to represent the distribution of the observed site biomass values, particularly for sites with high biomass, and this would reduce overall bias.

The present memo provides a validation analysis of the revised criteria for total nitrogen (TN) and total phosphorus (TP) for sites in the Otago and Southland Regions. The validation focuses on TN and TP because these were the criteria used to estimate reductions in nitrogen and phosphorus loads to achieve periphyton target attribute states in Otago (Snelder and Fraser 2023).

Revised criteria

Snelder and Kilroy (2023) provided details of the methods used to derive revised nutrient criteria for four forms of nutrients (TN, DIN, TP, DRP) to achieve three biomass targets (50, 120 and 200 mg m^{-2}) for 21 River Environment Classification (REC) Source-of-flow classes. Briefly, the GLM models were derived from the data that was used by Snelder et al. (2022) to derive the original criteria. The GLM models were used to predict Chla92 for a wide range of concentrations for each nutrient form at up to 500 individual river locations in each Source-of-flow class. The concentrations at which the predicted biomass was 50, 120 and 200 mg m^{-2} for each location was obtained from these predictions by linear interpolation. The geometric means of the concentrations associated with each biomass target within each Source-of-flow class are the criteria.

For each biomass target, within each Source-of-flow class, Snelder and Kilroy (2023) also obtained the exponentiated standard deviation of the log of the individual nutrient concentrations as a measure of the within-class variability of the concentration criterion. This acknowledges that the derived criterion represents a mean condition for an entire REC class. Using the mean for that segment's class introduces uncertainty because the true criterion for the specific segment will differ from the mean for all segments in the class. The impact of the within-class variation on the validation can be assessed with a Monte Carlo simulation of the validation procedure, which is are explained later.

A detail of the revised criteria derived by Snelder and Kilroy (2023) was that the underlying GLM models tended to over-estimate low Chla92 values (i.e., ≤50 mg m^{-2}). Over-estimation of the low Chla92 values meant that the derived criteria for the lower biomass threshold (i.e., 50 mg m^{-2}) were too stringent (i.e., the concentrations were too low). This issue was also present in the original OLS models and criteria but was slightly more apparent for the revised criteria.

To address the issue of over-prediction of low Chla92 values, Snelder and Kilroy (2023) suggested that an alternative set of criteria for the 50 mg m^{-2} biomass threshold could be derived using quantile regression. This approach was used to derive TN and TP criteria for Otago and Southland sites from the fitting data used by Snelder and Kilroy (2023). These criteria were derived for the same levels of under-protection risk as the revised criteria. However, the quantile regression criteria are spatially uniform (i.e., one value applies to all REC Source-of-flow classes). The alternative set of spatially uniform criteria for TN and TP derived using quantile regression for the 50 mg m^{-2} threshold is provided in Appendix 1.

Validation of the revised criteria for Otago and Southland

Guidance provided by (MFE 2022) suggests that the use of the criteria, for example within a region, should be accompanied by a verification that considers whether they are reasonably consistent with local observations of relationships between periphyton (as Chla92) and nutrient concentrations. There are limited ways to assess confidence in the criteria. However, where a monitoring network for periphyton and nutrients exists within a region, a validation analysis can be performed with the following seven steps.

1. Obtain the median concentration of each nutrient and Chla92 from the observations at each monitoring site.
2. Obtain the REC source-of-flow class and shade status for each site.
3. For a fixed nutrient and level of under-protection risk, obtain the criteria from the lookup tables for the A, B and C bands for each site based on the site's REC source-of-flow class and shade status.
4. For each nutrient and site, and under-protection risk, interpolate the biomass from the criteria by:
    a. treating Chla92 for A, B and C bands of 50, 120 and 200 mg m^{-2} as the variable Y and nutrient criteria for each band as the variable X and assuming biomass is zero when nutrients are zero,
    b. use linear interpolation to estimate the Chla92 (Y values) predicted by the observed site nutrient concentrations,
    c. treating the interpolated Chla92 as a prediction.
5. Calculate, over all sites, the proportion of sites with observed values of Chla92 that exceed the above predicted values. We refer to these sites as the 'exceeding sites'.
6. Repeat this process for each nutrient and level of under-protection risk.
7. Assess whether the nutrient criteria are consistent with the observations by comparing the proportion of exceeding sites with the proportion indicated by the under-protection risk.

MFE (2022) suggests that reasonable agreement (i.e., ± 20%) between the proportion of exceeding sites and level of under-protection risk can be interpreted as evidence that the nutrient criteria are valid for the sites represented by the monitoring network. MFE (2022) notes that perfect agreement should not be expected and that divergence between the proportion of observations that exceed the predictions, and the under-protection risk can be expected to decrease as the sample size increases.

Uncertainty of the validation analysis

The above analysis is uncertain for two reasons. First, the observed values of Chla92 are imprecise (i.e., they are estimates of the population value calculated from the monthly samples). Second, as noted above, there is within-class variability in the estimates of the criteria for each site.

The first component of uncertainty is part of the more general issue that all estimates of attribute states are subject to uncertainty because of sampling error. Recent guidance (Milne *et al.* 2023) has made suggestions for accounting for this uncertainty under subclause (4) of clause 3.10 of the NPS-FM. However, Milne et al. (2023) acknowledge that robust methods for quantifying attribute state uncertainty have not been identified. Milne et al. (2023) acknowledge that standard statistical assumptions (e.g., observations are randomly varying and drawn from the same population), associated with the calculation of confidence intervals, are likely to be violated for typical NPS-FM attributes. For example, observations of chlorophyll have a seasonal component of variation and are, therefore, not entirely random. Attribute states are also assigned to sites using observations collected over time periods of up to five years. Time periods of this duration are likely to include significant changes that are due to long-term trends and inter-annual fluctuations (Snelder *et al.* 2021), which means that the sample does not represent a single population. Therefore, in this study, we ignored the uncertainty associated with observed values of Chla92 and focussed on accounting for the uncertainty associated with the within-class variability in the criteria.

The second component of uncertainty is quantified by the within-class standard deviation of the nutrient concentration criteria across river locations. A second validation analysis was undertaken that repeated the first analysis but used this standard deviation in a Monte Carlo simulation to generate 1000 “realisations” of the predicted Chla92 for each site. For each realisation, random errors were added to the criterion for each site and then this “perturbed” criterion was used to produce a realisation of the predicted Chla92. The random error was derived by drawing from a normal distribution with a standard deviation equal to the standard deviation of the log of the individual nutrient concentrations within each class. The 1000 realisations produced by the Monte Carlo analysis were summarised to provide best estimates of the proportion of exceeding sites. The uncertainty of the proportion of exceeding sies was quantified by the 95% confidence interval.

Results

The validation was performed using a dataset pertaining to 64 sites in Otago and Southland (Figure 1). The data for the Otago sites covered the period from February 2019 to March 2022. The data for the Southland sites covered the period from January 2015 to May 2022. Most of these sites were represented in the dataset used by Snelder and Kilroy (2023) to derive the nutriment concentration criteria. However, the period of record for the sites used for this validation analysis was an additional year of observations and therefore the data was semi-independent of the derivation procedure.

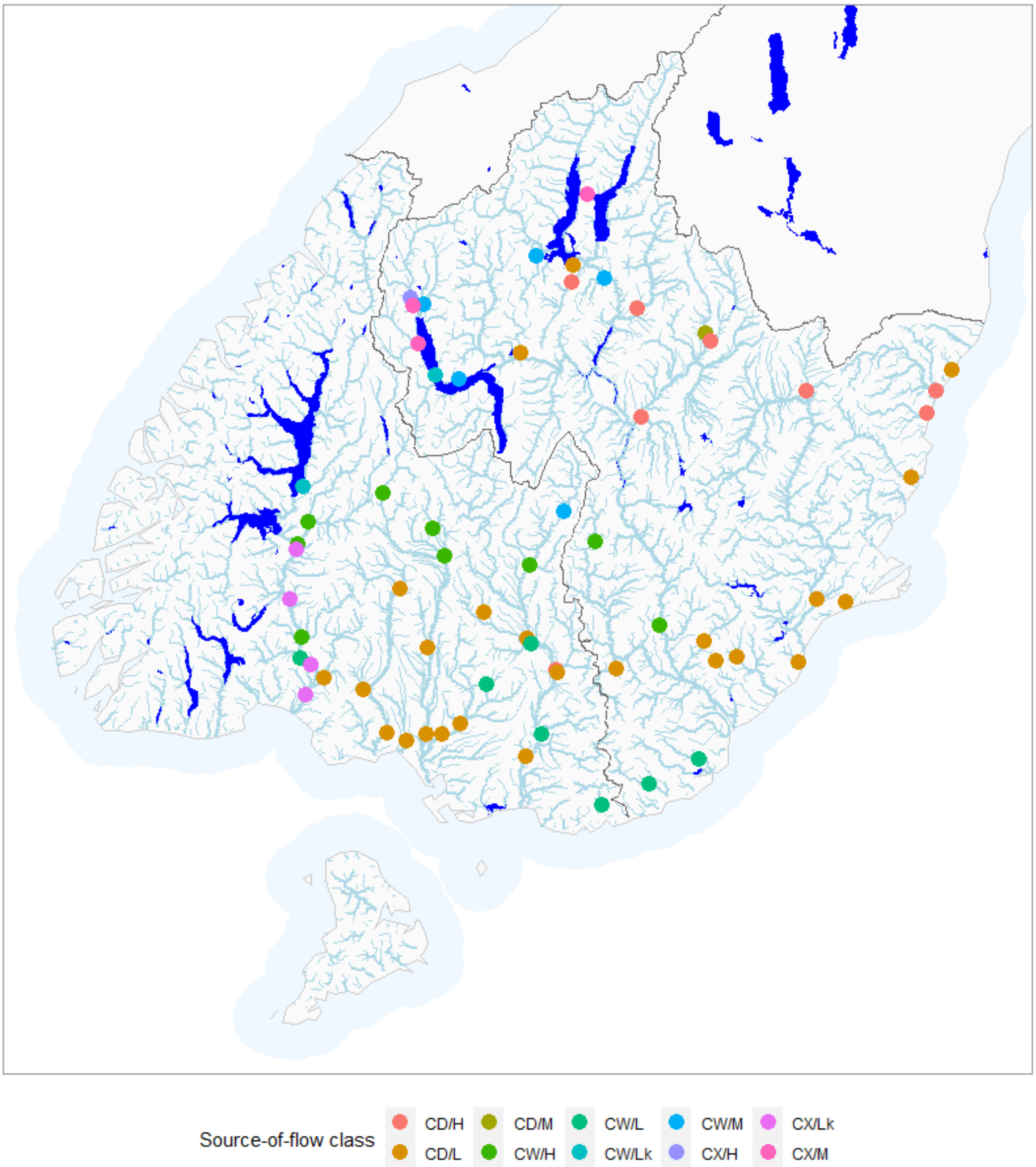

Figure 1. Map of the periphyton monitoring sites in Otago and Southland that were used in the validation process. The Source-of-flow class of each site is identified.

The observed and predicted values of Chla92 at the 64 sites in the region based on the two nutrient forms (TN and TP) are shown as scatter plots in Figure 2. Theoretically, 5%, 10%, 15%, 20%, 25%, 30% and 50% of the sites should have observed biomass that exceeds the predicted biomass when the predictions are made based on the corresponding levels of under-protection risk (i.e., should lie above the red lines on Figure 2).

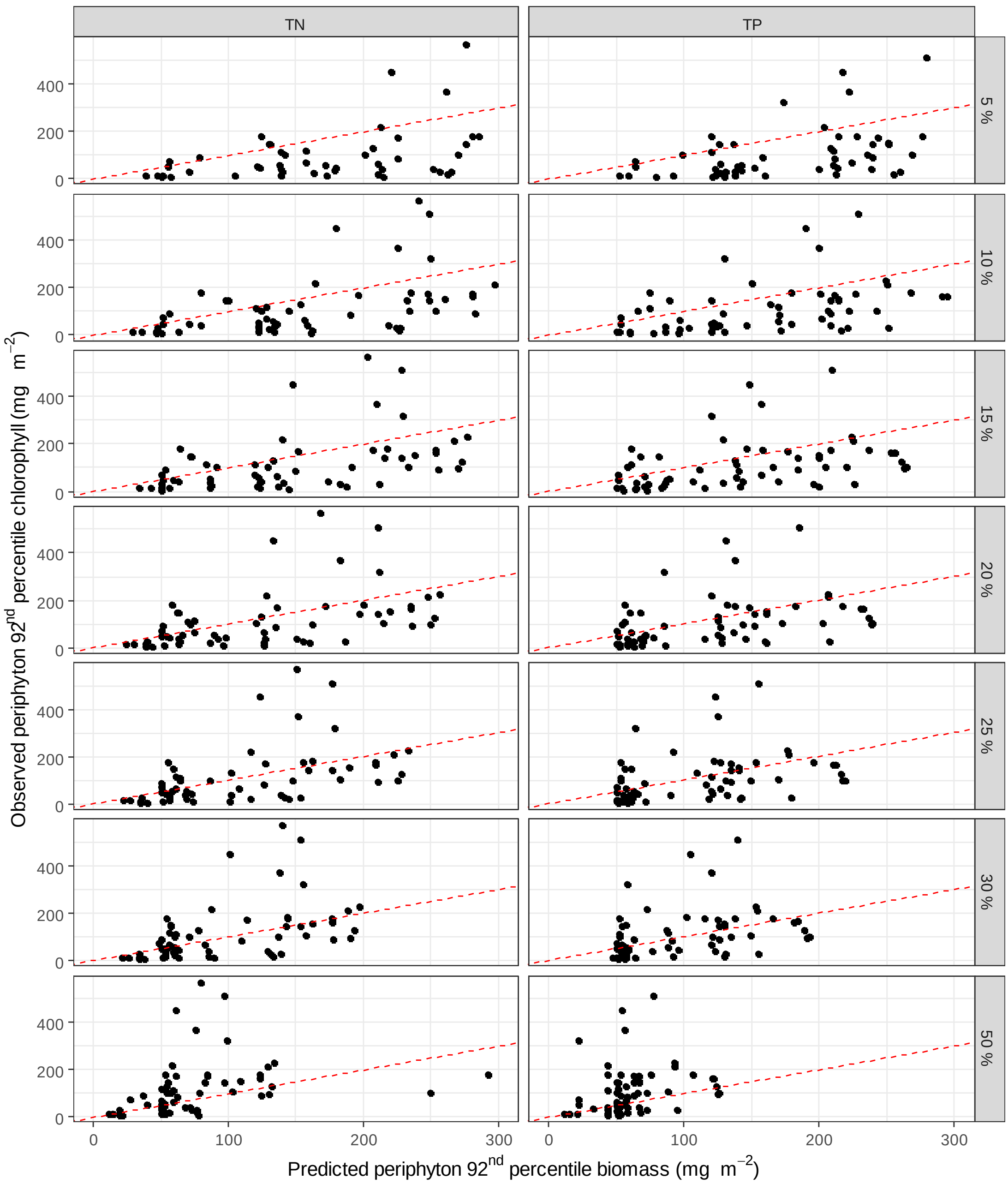

Figure 2. The observed and predicted values of Chla92 at the 64 sites in the Otago and Southland regions where predicted values are derived from the nutrient criteria for under-protection risks of 5, 10,15, 20, 25%, 30% and 50%. Panel labels indicate the under-protection risks and the nutrient form (TN and TP). The dashed red diagonal (one to one) line represents agreement between the predictions and observations. The points lying below the red line indicate 'exceeding sites' (i.e., sites for which the observed biomass was greater than the predicted).

The data shown in Figure 2 indicate that the proportions of sites for which observed Chla92 exceeds predicted Chla92 increases systematically as the under-protection risk increases for both nutrients. Table 1 indicates that the proportion of sites for which observed Chla92 exceeds the predicted is always greater than expected based on the under-protection risk for both nutrients. The column headed "discrepancy" is the difference (for each nutrient) in the under-protection risk and the observed proportion of exceeding sites. Negative values indicate that the criteria are too permissive.

Discrepancies are in the range of -6 to -14%, but these discrepancies are considerably lower than those of the validation of the original criteria (-12 to -36%) reported in LWP memo to ORC dated 22 February 2023 and shown in Table 2.

Table 1. Validation results for the revised criteria. Proportion of sites (%) for which observed biomass exceeds that predicted for the seven levels of under-protection risk and two forms of nutrient (TN and TP). The discrepancy is the difference between the UPR and the observed proportion of sites exceeding the threshold (%). Negative values indicate that the criteria are too permissive.

Under protection risk (%)	Proportion exceeding (%)		Discrepancy (%)	
	TN	TP	TN	TP
5	17	17	-12	-12
10	17	20	-7	-10
15	22	23	-7	-8
20	27	31	-7	-11
25	31	39	-6	-14
30	38	42	-8	-12
50	58	64	-8	-14

Table 2. Validation results for the original criteria. Proportion of sites (%) for which observed biomass exceeds that predicted for the six levels of under-protection risk and two forms of nutrient (TN and TP). The discrepancy is the difference between the UPR and the observed proportion of sites exceeding the threshold (%). Negative values indicate that the criteria are too permissive.

Under protection risk (%)	Proportion exceeding (%)		Discrepancy (%)	
	TN	TP	TN	TP
5	17	20	-12	-15
10	22	31	-12	-21
15	33	42	-18	-27
20	45	55	-25	-35
30	53	61	-23	-31
50	78	86	-28	-36

Figure 3 summarises the results of the Monte-Carlo procedure and shows the proportion of exceeding sites and the 95% confidence interval for each level of under-protection risk. Figure 3 is consistent with the validation results shown in Table 1; for all levels of under protection risk, the lower confidence limit is always above the associated level of under-protection risk (indicated by horizontal lines). This indicates that the revised criteria are too permissive and the inconsistency between the observation and predictions is larger than what can be attributed to the uncertainty of the criteria.

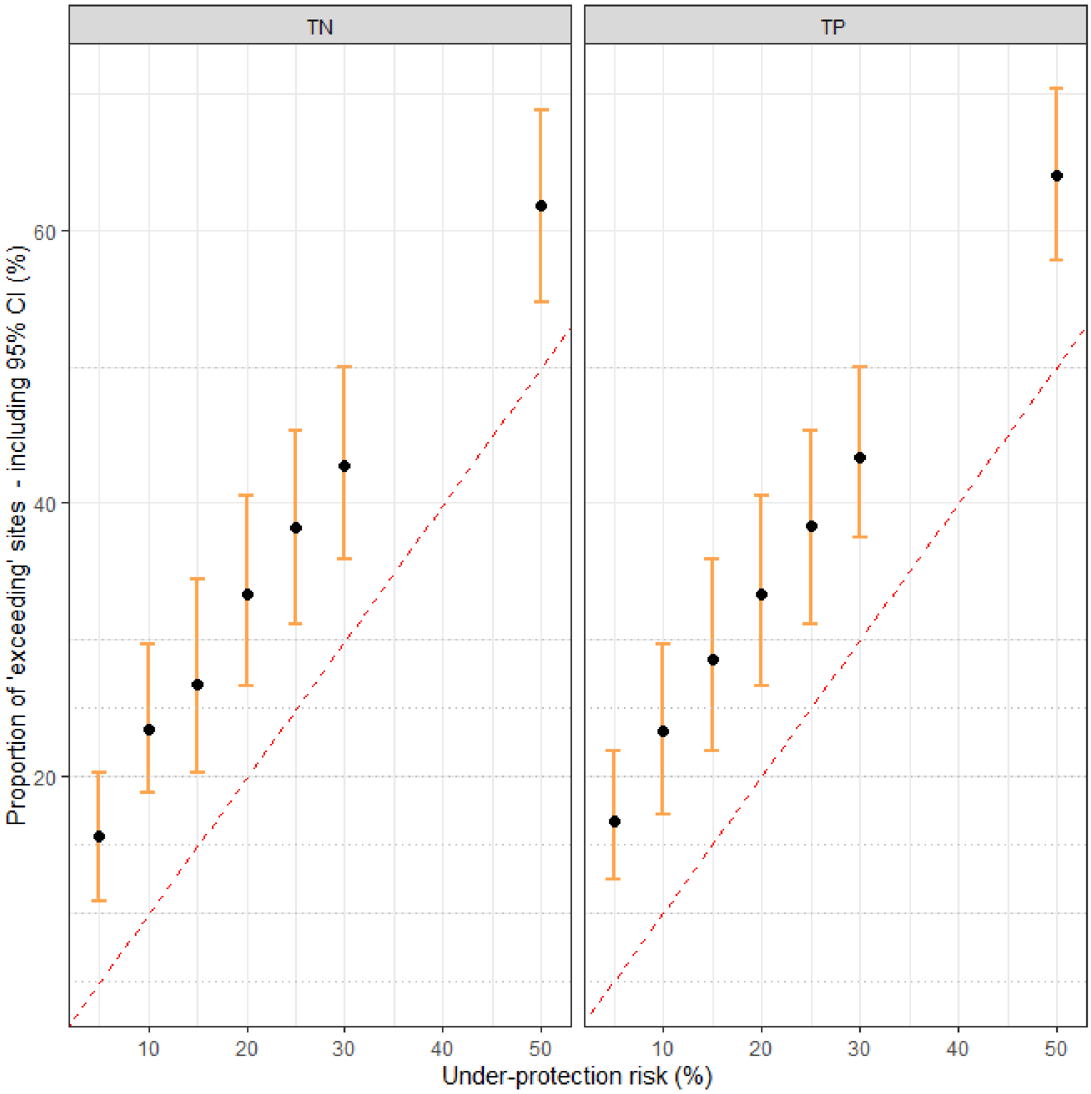

Figure 3. Proportion of "exceeding" sites (i.e., sites that are under-protected) for each level of under-protection risk (x-axis) and the two nutrients. The error bars indicate the 95% confidence interval of the observed "exceeding" sites, which was generated from a Monte Carlo analysis. The dashed red diagonal (one to one) line represents agreement between the proportion of exceeding sites and the under-protection risk.

Conclusions

This validation indicates that the revised criteria are too permissive but are an improvement over the original criteria. In addition, while the criteria are overly permissive, they indicate "reasonable agreement" (i.e., ±20%) between the proportion of exceeding sites and level of under-protection risk.

The analysis of Snelder and Kilroy (2023) found that, at the regional level, the revised criteria were variously too stringent or too permissive but also that they were very consistently an improvement on the original criteria. It is noted that a perfect validation may be an unrealistic goal given the inherent uncertainties and potential biases associated with several aspects of these analyses, including the small, and potentially biased, network of periphyton monitoring sites that were used in the analysis. The improved performance of the revised criteria, and the underlying technical explanation for why this was expected, is a sound basis for generally recommending the use of the revised criteria over the original criteria for Otago and Southland.

References

Kelly MG, Phillips G, Teixeira H, Várbíró G, Herrero FS, Willby NJ, Poikane S (2022) 'Establishing ecologically-relevant nutrient thresholds: A tool-kit with guidance on its use' *Science of the Total Environment* **807**, 150977.

MFE (2022) Guidance on look-up tables for setting nutrient targets for periphyton: second edition. ME 1644. Ministry for the Environment., Wellington, New Zealand.

Milne J, Norton N, Franklin P, Wood D, Ausseil O, Young R, Patterson M, Fullard L (2023) Attribute states and uncertainty. Preliminary expert commentary on implementation of clause 3.10(4) of the NPS-FM 2020. 2022364WN. NIWA, Christchurch.

Ministry for Environment (2020) 'National Policy Statement for Freshwater Management 2020.' Available at http://www.mfe.govt.nz/publications/rma/nps-freshwater-management-2014/nps-freshwater-management-jul-14.pdf

Phillips G, Kelly M, Teixeira H, Salas F, Free G, Leujak W, Solheim AL, Várbíró G, Poikane S (2018) Best practice for establishing nutrient concentrations to support good ecological status. EUR 29329 EN,. Publications Office of the European Union, Luxembourg.

Snelder T, Fraser C (2023) Assessment of Nutrient Load Reductions to Achieve Target Attribute States in the Rivers, Lakes and Estuaries of Otago. Using revised periphyton nutrient criteria. LWP Client Report 2023–13. LWP Ltd, Christchurch, New Zealand.

Snelder T, Kilroy C (2023) Revised Nutrient Criteria for Periphyton Biomass Objectives. Updating criteria referred to in Ministry for Environment 2022 guidance. LWP Client Report 2023–08. LWP Ltd, Christchurch, New Zealand.

Snelder T, Kilroy C, Booker D (2022) Derivation of Nutrient Criteria for Periphyton Biomass Objectives. Using regional council monitoring data. LWP Client Report 2021–13. LWP and NIWA Ltd, Christchurch, New Zealand.

Snelder TH, Larned ST, Fraser C, De Malmanche S (2021) 'Effect of climate variability on water quality trends in New Zealand rivers' *Marine and Freshwater Research*.

Appendix 1. Criteria for the 50 mg m^{-2} target state

Plots of observed Chla92 values versus observed site median nutrient values were wedge-shaped (Figure 4). This indicates that there is a limiting relationship between peak biomass (i.e., Chla92) and nutrients at the regional (i.e., Otago and Southland) scale but that other factors influence the Chla92 response (Phillips *et al.* 2018; Kelly *et al.* 2022). Quantile regression models were statistically significant (p < 0.1) for all quantiles for TN and most quantiles for TP (Table 3).

Sites with Chla92 values of 50 mg m^{-2} or less occurred across a wide range of nutrient concentrations and in most Source-of-flow classes (Figure 4). This indicates that there is no obvious landscape scale spatial pattern in the low biomass sites and that, in the absence of variables that can better explain low biomass at these sites, the uniform criteria derived from the quantile regression models are a justifiable approach to defining criteria for the 50 mg m^{-2} biomass target.

Where possible, we derived alternative criteria from all QR models (Table 3) and used these values to replace the criteria pertaining to the 50 mg m^{-2} biomass target for the revised and original criteria.

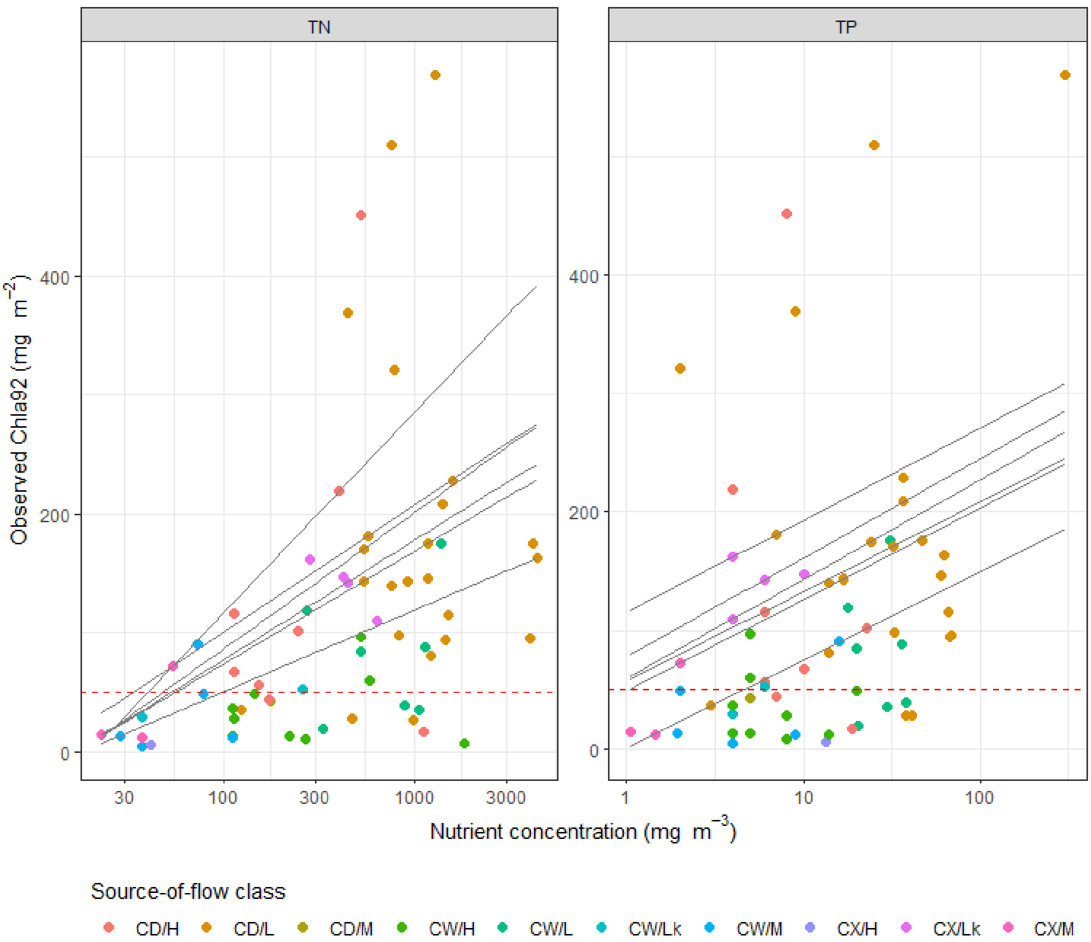

Figure 4. Relationships between Chla92 and median nutrient concentrations at the 251 monitoring sites. The grey lines are quantile regressions fitted to the 0.95, 0.9, 0.85, 0.8, 0.7 and 0.5 quantiles. Not all of these regression lines are statistically significant (see Table 3). The red dashed line indicates a biomass of 50 mg m^{-2}. Points are coloured to indicate the Source-of-flow class of the monitoring site.

Table 3. Criteria derived from the QR models for the 50 mg m-2 Chla92 target state for each nutrient form and level of under-protection risk. The P-value indicates the confidence in the regression coefficient fitted to the nutrient concentration. The criteria have units of mg m-3. NA values indicate that criteria could not be derived from the QR model.

Nutrient	Quantile	Under-protection risk (%)	P value	Criteria
TN	0.5	50	**0**	97.7
	0.7	30	**0**	55.1
	0.75	25	**0**	52.4
	0.8	20	**0**	47.6
	0.85	15	**0.047**	33.4
	0.9	10	**0.044**	39.6
	0.95	5	**0**	29.9
TP	0.5	50	**0**	4.6
	0.7	30	**0.027**	1.1
	0.75	25	**0.078**	1
	0.8	20	0.139	1
	0.85	15	0.357	0.9
	0.9	10	0.493	0.8
	0.95	5	0.451	0.3