Derivation of Nutrient Criteria for Periphyton Biomass Objectives

Using regional council monitoring data

April 2022

Prepared By:

Ton Snelder

Cathy Kilroy*

Doug Booker*

*National Institute of Water & Atmospheric Research Ltd

For any information regarding this report please contact:

Ton Snelder

Phone: (03) 377 3755
Email: ton@lwp.nz

LWP Ltd
PO Box 70
Lyttelton 8092
New Zealand

LWP Client Report Number:	
Report Date:	April 2022
LWP Project:	2021-13

Quality Assurance Statement

Version	Reviewed By	
Final	Scott Larned	

Table of Contents

Figures

Tables

Executive Summary

The Ministry for the Environment (MFE) is assisting regional councils with implementation of the National Policy Statement for Freshwater Management (NPS-FM). This report has been prepared for MFE as a contribution to that effort and is related to Section 3.13 of the NPS-FM. Section 3.13 requires councils to set concentration criteria for dissolved inorganic nitrogen (DIN) and dissolved reactive phosphorus (DRP) in rivers to achieve objectives for attributes defined in the National Objectives Framework (NOF), including periphyton.

The objective of this study was to provide updated look-up tables for nutrient criteria to achieve target periphyton attribute states in all hard bottom (i.e., cobble-gravel-bed) streams and rivers in New Zealand, classified into the 21 Source-of-flow classes defined by the River Environment Classification (REC). The tables update those for total nitrogen (TN) and DRP provided in earlier reports and published papers.

The current study aimed to improve on the existing criteria through: (a) the use of new data from regional council periphyton monitoring programmes; (b) the use of periphyton measured as chlorophyll *a* (chlorophyll), which is directly related to the NOF periphyton biomass attribute (rather than data on cover, which needed to be converted to chlorophyll); and (c) application of a range of statistical methods using the larger dataset.

Data available for the study comprised chlorophyll observations at 251 monitoring sites that had at least 20 monthly sample occasions (summarised as the 92nd percentile of the observations and referred to hereafter as Chla92), associated water quality data including nutrient concentrations (typically summarised as median values of the observations), and other environmental observations at the sites including substrate composition and shade. Data also included flow records that were converted to hydrological indices that characterised the flow regime at each site.

The approach we used to derive nutrient concentration criteria started with statistical regression models that sought to explain between-site variation in Chla92 as a function of environmental factors that influence periphyton biomass, including nutrient concentrations, hydrology, and physical habitat. We trialled several statistical methods and determined that the most appropriate was ordinary least-squares regression (OLS). We used OLS r to fit models that explained Chla92 in terms of DIN, TN, DRP and total phosphorus (TP) combined with other environmental variables.

Nutrient concentration criteria for DIN, TN, DRP and TP were then obtained by inverting the fitted OLS models to obtain the concentrations associated with the three NOF periphyton attribute state thresholds: 50 mg m^{-2}, 120 mg m^{-2}, and 200 mg m^{-2}. This process involved using the fitted OLS models to predict the concentration criteria for all segments of the national digital river network (order ≥ 3). The look-up tables were then prepared as the mean of these predicted values in each Source-of-flow class.

Variation in between-site periphyton biomass explained by the statistical models developed in this study was low. For the OLS models that were used to define the criteria, cross-validated R^2 values varied between 0.27 and 0.38. The consequence of low variation explained by the statistical models is that predictions of biomass have large uncertainty. For example, a prediction for Chla92 of 90 mg m^{-2} made with the OLS models has a 95% confidence interval of approximately 15 mg m^{-2} to 300 mg m^{-2} and the 70th, 80th and 90th percentiles of the distribution are 128 mg m^{-2}, 156 mg m^{-2} and 202 mg m^{-2}, respectively. These levels of model performance and associated uncertainties are consistent with other studies in both New

Zealand and internationally. Uncertainties reflect the complexity of the underlying processes and incomplete knowledge of the drivers of periphyton biomass.

The model uncertainty means that a single specific criterion will not ensure that a target level of biomass is not exceeded. Instead, there is a probability distribution that describes the risk that the target level of biomass at a site will be exceeded for a given nutrient concentration. We refer to this probability as under-protection risk. It is important that nutrient criteria describe, as much as possible, the risk of under-protection and allow decision-makers to choose the level of risk that is acceptable. The nutrient criteria derived in this study provide for choice in the level of under-protection risk that might be acceptable. The chosen level of risk of under-protection is not scientifically defined and is a subjective ('normative') choice.

The under-protection risk refers to a river location. Choosing a level of under-protection risk means that, across a domain of interest comprising many sites, a proportion of locations can be expected to have biomass higher than the nominated target despite being compliant with the criteria. The corollary to this is that the objective underlying the criteria is to maintain periphyton biomass at or below the nominated thresholds at a proportion of sites within a domain that is the complement of the under-protection risk. For example, under-protection risks of 30%, 20% and 10% correspond to objectives to maintain biomass below the target level at 70%, 80% or 90% of sites across the domain, respectively.

In this study, we derived look-up tables of nutrient criteria for four forms of nutrient (TN, DIN, TP and DRP) for three target periphyton attribute states (50 mg m^{-2}, 120 mg m^{-2}, and 200 mg m^{-2}.) and for the 21 Source-of-flow classes that encompass all New Zealand's rivers. The look-up tables provide for six levels of under-protection risk, 50%, 30%, 20%, 15%, 10% and 5%. It is noted that the risk of over-protection is the complement of the risk of under-protection (i.e., over-protection risk = 100% - under-protection risk). Because Shade was included as an explanatory variable in all four nutrient OLS models, separate look-up tables of nutrient criteria were derived for shaded and unshaded locations.

The criteria are uncertain in that, for a stated under-protection risk, we do not know what is the precise proportion of sites that will exceed the target biomass threshold when concentrations are held to the criteria. However, validation of the derived criteria using an independent dataset indicates that they perform well. We therefore do not consider that the uncertainties should be a barrier to using the criteria.

The models developed in this study indicate that there is an initially high rate of increase in periphyton biomass with increasing nutrient concentrations for each of the nutrients we considered (TN, DIN, TP and DRP), but periphyton biomass reaches a "ceiling" beyond which there is no further response to increasing nutrient concentrations. Our study indicates that from a practical perspective, in most REC Source-of-flow classes a 92^{nd} percentile chlorophyll biomass of 200 mg m^{-2} would rarely be exceeded because of nutrient enrichment. We refer to the nutrient concentration beyond which there is no further biomass response as the "saturating concentration". The models developed by this study indicated that the saturating concentrations are approximately 1,000 mg m^{-3} for TN and DIN, approximately 50 mg m^{-3} for TP and approximately 25 mg m^{-3} for DRP. The stated saturating concentrations are approximate because they were derived subjectively from simplified graphical representations of the models defined by this study. We are not aware of statistical methods that could be used to more objectively identify the saturating concentrations based on our models. If such methods do exist, any estimate of the saturating concentration will have large uncertainties because our models have considerable unexplained variation.

Because the OLS models used to derive the nutrient criteria are parametric and represent biomass as a log-transformed function of nutrient concentration, some of our derived criteria are higher than the identified saturating concentrations. In addition, for some combinations of nutrient, Source-of-flow class and level of under-protection risk, the model predictions would not exceed the biomass threshold even at the maximum of the range of observed site median nutrient concentrations represented in our dataset. For these combinations, our criteria are appreciably higher than the saturating concentrations of approximately 1,000 mg m^{-3} for TN and DIN, 50 mg m^{-3} for TP and 25 mg m^{-3} for DRP but criteria were limited to the maximum of the relevant observed site median nutrient concentrations.

For all nutrients, the derived criteria typically had high values, relative to saturating concentrations, for Source-of-flow classes that represent rivers with strong physical controls on periphyton biomass (e.g., classes with high flow variability and low temperature). The high values of these criteria should be interpreted cautiously because the sites in our fitting data that represented these physical conditions did not cover a wide range of nutrient concentrations. This means that the models were poorly informed about the nutrient – biomass relationship at sites with strong physical controls on periphyton biomass. It is possible that physical conditions limit maximum biomass in rivers with strong physical controls even when the nutrient concentrations are high. However, because our method involved extrapolation of the model into environmental conditions that are poorly represented by the fitting data, the uncertainty of criteria for Source-of-flow classes that represent rivers with strong physical controls on periphyton biomass is particularly high.

We consider there is high uncertainty about whether biomass can be managed by restricting nutrient concentrations to levels greater than the saturating concentrations. Therefore, where a biomass threshold is exceeded at a site, it is likely that biomass reduction can only be expected if nutrient concentrations are reduced to below the saturating value. Nutrient criteria higher than the saturating values indicate combinations of conditions where periphyton biomass is strongly controlled by non-nutrient factors. Under such conditions, even when nutrient concentrations are greater than saturating levels, the biomass threshold(s) may not be exceeded.

Glossary

The table below defines the terms according to how they are used in this report.

Term	Definition
Biomass	A level of periphyton abundance measured as mg chlorophyll *a* m^{-2}
Biomass thresholds	Levels of periphyton abundance that are considered in this study. Biomass thresholds are measured as specified by the NOF attribute state bands A B, and C as the 92^{nd} percentile value of chlorophyll with boundaries being 50 mg m^{-2}, 120 mg m^{-2}, and 200 mg m^{-2}, respectively.
Chla92	The 92^{nd} percentile value of observations of chlorophyll *a* made at a monitoring site.
$\widehat{Chla}92$	A prediction of the 92^{nd} percentile value of observations of chlorophyll *a* made using a model
Conceptual model	A theoretical model expresses our understanding of the processes that produce differences in biomass between sites.
DIN	Dissolved inorganic nitrogen (NO3-N – NO2-N + NH4-N). One form of the two periphyton nutrients considered by this study.
Disturbance	Components of the conceptual model that are understood to contribute to biomass loss.
DRP	Dissolved reactive phosphorus. One form of the two periphyton nutrients considered by this study.
EC	Electrical conductivity, a measure of the amount dissolved material in the water, dominated by naturally occurring concentrations of the major ions (e.g., calcium, sodium) but also influenced by anthropogenic inputs.
Exceedance probabilities	The probability that 92^{nd} percentile value of chlorophyll will be exceeded based on the predicted probability distribution.
FineSed	The proportion of the bed covered by fine sediment, generally visually assessed using the SAM2 method Clapcott *et al.* (2011)
Independent variable	The variables that are used to explain/predict in a statistical model. Referred to elsewhere as explanatory variables or predictors.
Normal distribution	A type of probability distribution, used in this report to characterise model residuals. A plot of number of cases against value generates a bell-shaped curve with most cases being close to the mean value (= to the median and modal values) and cases higher or lower than the mean with progressively fewer cases in a predictable pattern.
Nutrient criteria	Concentration limits for nitrogen and phosphorus forms considered by this study that will restrict periphyton biomass to a specified threshold.
OLS	Ordinary least squares regression.
PBIAS	Percent bias. A measure of regression model performance that describes the average tendency of the predicted values to be larger or smaller than the observed values.
Probability distribution	A description all the possible values that the predictions of 92^{nd} percentile value of chlorophyll ($\widehat{Chla}92$) represent and the associated likelihoods of these.
QR	Quantile regression.
R^2	Coefficient of determination. A measure of regression model performance that describes the proportion of variation in the response that is explained by the model.
REC	River Environment Classification. A system of classes that discriminates individual segments of New Zealand's rivers and streams into a number of hierarchical levels.
Residual	Or residual error. In a regression, the difference between an observed value and the value predicted by the regression model.

Term	Definition
Resources	Components of the conceptual model that are understood to contribute to biomass accrual (i.e., growth).
Response	The dependent variable in a statistical model (i.e., the variable being explained/predicted by the model). In this study, the response variable was generally Chla92.
RF	Random forest regression.
RMSD	Root mean square deviation. A measure of regression model performance that quantifies the characteristic (i.e., mean) error of the predictions.
Source-of-flow	The second hierarchical level of the REC, which is based on the climate and topography of the upstream catchment. River segments belonging to different Source-of-flow classes are expected to differ significantly with respect to the processes represented by the conceptual model. Consequently, the nutrient criteria pertaining to specific biomass thresholds are expected to differ between Source-of-flow classes.
TN	Total nitrogen. One form of the two periphyton nutrients considered by this study.
TP	Total phosphorus. One from of the two periphyton nutrients considered by this study.
Under-protection risk	The risk, expressed as a percentage, that a randomly chosen location will exceed a specified biomass threshold despite nutrient concentrations being compliant with the specified nutrient criteria. Referred to elsewhere as spatial exceedance.
VIF	Variance inflation factor, a measure of an independent variable's level of collinearity with other independent variables

1 Introduction

The Ministry for the Environment (MFE) is assisting regional councils with implementation of the National Policy Statement for Freshwater Management (NPS-FM). The NPS-FM requires regional councils to set objectives using attributes that are defined in the National Objective Framework (NOF). One of the attributes for streams and rivers is periphyton biomass. Periphyton comprises primarily algae of various types, living on or near to the streambed, but also includes fungi, bacteria and detritus. The algal component of periphyton is generally the predominant primary producer in hard-bottomed (i.e., cobble-gravel-bed) streams. The NOF periphyton attribute state is a measure of the algal component of periphyton (hereafter periphyton biomass) and is measured as mg chlorophyll *a* per square metre of riverbed. The attribute defines four states from A (low biomass, best state) to D (high biomass considered to be below the bottom line)[1].

The periphyton attribute is associated with NPS-FM clause 3.13 (Special provisions for attributes affected by nutrients), which requires that councils must, at a minimum, set river nutrient concentration criteria for dissolved inorganic nitrogen (DIN) and dissolved reactive phosphorus (DRP). The most up-to-date nitrogen and phosphorus criteria to achieve NOF periphyton attribute states was recently published in a peer reviewed paper (Snelder *et al.*, 2019) and has been used to assist MFE's policy development (MFE 2019). The criteria developed by Snelder *et al.* (2019) are risk-based and provide total nitrogen (TN) and DRP concentration criteria for each of 21 River Environment Classification (REC) source-of-flow classes (Snelder and Biggs. 2002) and for the three NOF periphyton biomass attribute state thresholds. The concentration criteria are presented in a series of 'look-up tables' that allow for varying levels of risk that are specified by a 'spatial exceedance criterion'. A spatial exceedance criterion expresses the risk that periphyton biomass at a site is greater than the stated threshold, despite being compliant with the nutrient criteria. To date, lookup tables have been developed for three spatial exceedance criteria 10%, 20% and 30%.

Many regional council monitoring programmes now include measurements of chlorophyll *a*, nutrients and other variables that are directly related to the NOF periphyton biomass attribute measurement unit. The available new data allowed this study to target three aspects of the criteria developed by Snelder *et al.* (2019) for improvement.

1. Snelder *et al.* (2019) developed their criteria using a model derived from data obtained from long-term sampling at 77 National Water Quality Network (NRWQN) sites because this was the best available data at that time. There are now over 350 sites throughout New Zealand at which regional councils have collected periphyton and associated water quality monitoring data that could be used to derive new nutrient criteria.

2. The NRWQN data described periphyton cover (i.e., percentage of the stream bed covered by visible periphyton) rather than biomass (i.e., chlorophyll *a*, which is the measure of periphyton specified in the NPS-FM). Snelder *et al.* (2019) converted the cover data to the equivalent biomass (as chlorophyll *a*) but this conversion was a source of uncertainty in their criteria. Most regional council monitoring programmes

[1] The biomass thresholds of 50, 120 and 200 mg m^{-2}, which separate the NOF bands, refer to the 92^{nd} percentile of periphyton biomass determined from at least three years of data, The 92^{nd} percentile is equivalent to an allowed exceedance of the biomass in one of 12 monthly observations, on average.

now include measurements of periphyton biomass as chlorophyll *a*, which is consistent with the NOF periphyton biomass attribute measurement unit.

3. Snelder *et al.* (2019) used an ordinary least squares (OLS) regression model with some additional analysis to derive their criteria. Quantile regression (QR) is a more direct and less complicated approach that may be easier to understand. QR is commonly applied to the development of nutrient criteria (e.g., Phillips *et al.*, 2018) and can produce the same type of output as the lookup tables for differing spatial exceedance criteria developed by Snelder *et al.* (2019).

In this study, we aimed to produce river nutrient concentration criteria to achieve NOF periphyton attribute states that (1) utilised recent data collected in regional councils periphyton monitoring programmes (2) are based directly on biomass observations (as chlorophyll *a*; hereafter chlorophyll), and (3) that use appropriate methods that are easily understood as possible. The intended purpose of these criteria is to assist with implementation of the NOF periphyton attribute and the tables produced by this study are intended to update those of Snelder *et al.* (2019). We used QR and OLS to model the relationships between biomass and several environmental drivers of biomass accrual and loss including nutrient concentrations, light, temperature and hydrological indices. Where satisfactory models were obtained, we used the models to produce concentration criteria lookup tables for 21 source-of-flow classes and the three NOF attribute band thresholds and for differing levels of risk a site will have periphyton biomass more than the stated threshold.

2 Approach to derivation of criteria

Frequently used approaches to defining nutrient concentration criteria for rivers include: (1) reference state nutrient conditions; (2) identifying nutrient concentrations that represent significant shifts in aquatic ecosystem structure and function; and (3) linking water column nutrient concentrations to measures of autotroph abundance (e.g., periphyton biomass). The aim in this study was to use the third approach to derive criteria for water column concentrations of either dissolved inorganic or total nutrients to achieve objectives for periphyton biomass measured as benthic chlorophyll in hard-bottomed streams (e.g., Biggs, 2000; Dodds *et al.*, 2002; Snelder *et al.*, 2019; Van Nieuwenhuyse and Jones, 1996). The reason for taking this approach is that it will result in criteria that are consistent with the requirements of the NPS-FM to manage periphyton biomass.

The link between nutrient concentration and periphyton abundance can be established based on statistical regression models. These models describe between-site variation in periphyton biomass as a function of environmental variables that affect periphyton, including nutrients, hydrology, and physical habitat. This section outlines this approach in overview as three steps (Figure 1). Greater detail is provided in the following sections.

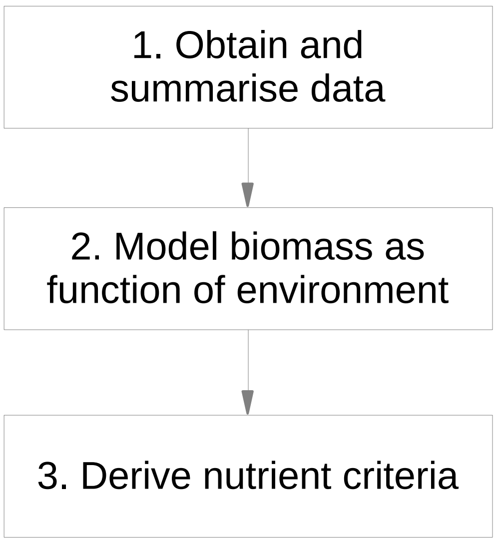

Figure 1. Summary of the approach to deriving nutrient criteria taken by this study.

2.1 Obtain and summarise data

Data comprised monthly observations of periphyton biomass as chlorophyll and several water quality indicators, including concentrations of the nutrients (nitrogen and phosphorus in both dissolved inorganic and total forms), water temperature, electrical conductivity, visual clarity and turbidity. Data on physical condition at each site including substrate size and shade were also included and many sites were associated with a continuous hydrological record.

For each site, the periphyton biomass was represented by the 92nd percentile of the chlorophyll observation time series and water quality variables, and data describing physical conditions, were each summarised as a single statistic (e.g., median)) to represent the characteristic condition for that variable at the site. Continuous hydrological records (as mean daily flows) were used to derive hydrological indicators describing aspects of the hydrological regimes for each site. These summarised variables were used in all subsequent analyses.

2.2 Model biomass as a function of the environment

A conceptual model expressing understanding of the processes that produce differences in biomass between sites guided the development of statistical models that explain those differences. This conceptual model (Figure 2) was proposed by Biggs (1996) and explains periphyton biomass dynamics in streams in terms of counteracting processes that control biomass accrual (e.g., nutrient supply and light) and biomass loss (e.g., from hydrological disturbance and invertebrate grazing). We aimed to define statistical models that:

> (a) explained variation in periphyton biomass using a combination of independent variables that were consistent with the conceptual model (e.g., positive associations with independent variables representing resources and negative associations with variables representing disturbance),
>
> (b) had adequate statistical performance, and
>
> (c) included a nutrient as an independent variable that was positively associated with biomass, to enable development of nutrient criteria in the next step (see below).

The statistical modelling approaches used were quantile regression (QR), ordinary least squares regression (OLS), and random forest regression (RF). Each of these approaches has strengths and weaknesses, which are described in more detail below.

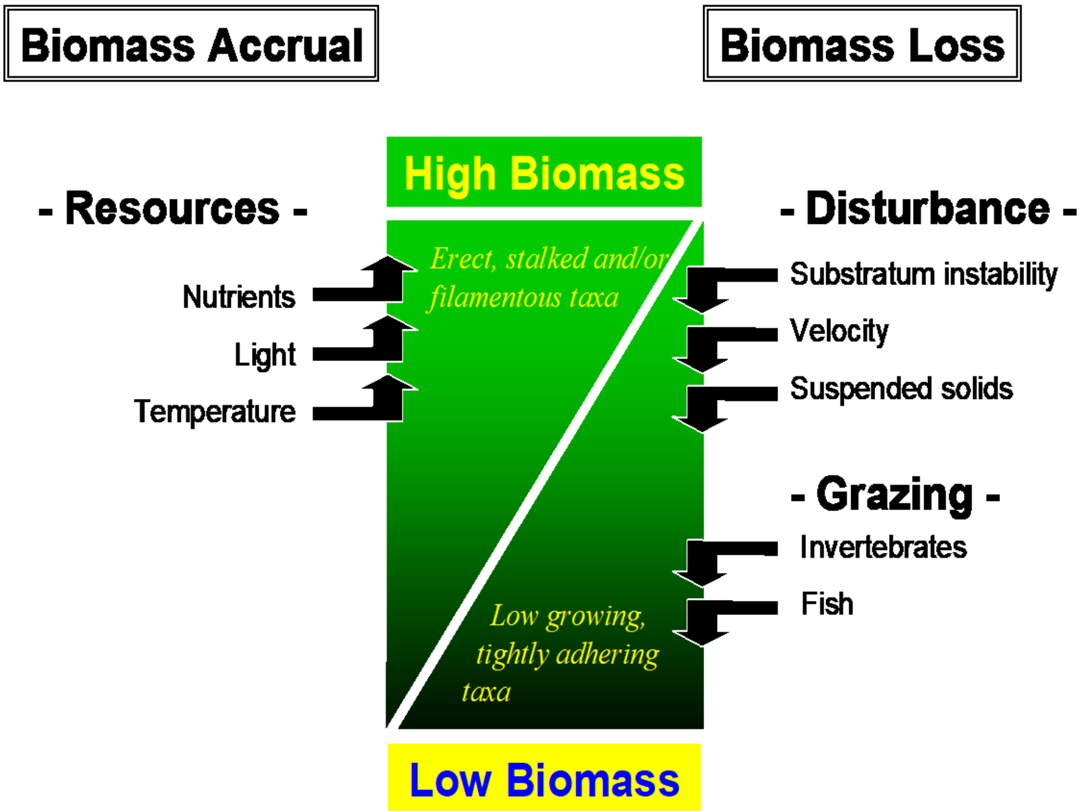

Figure 2. Conceptual model of stream periphyton biomass dynamics proposed by Biggs (1996).

We made separate models for each nutrient form in the water quality dataset: total nitrogen (TN), dissolved inorganic nitrogen (DIN), total phosphorus (TP) and dissolved reactive phosphorus (DRP). We recognise that the influence of nitrogen and phosphorus on periphyton growth are not independent of each other. However, our models included only one nutrient because including both nutrients would not allow the model to be used to derive criteria because the same biomass can be arrived at from differing combinations of the two nutrients. This means the models should be seen as correlative in nature, i.e., establishing the correlation between the characteristic level of nitrogen and phosphorus enrichment and biomass, rather than representing the actual processes of growth and loss through time. We fitted the models to all available sites (i.e., a single global model) to maximise the statistical power that was available from the data. We assumed that the global model would account for regional differences in conditions controlling periphyton biomass because the independent variables describe these differences.

2.3 Derive nutrient criteria

Nutrient criteria were derived using the models from the second step that were judged to perform adequately and were consistent with the conceptual model. Deriving nutrient criteria essentially involves inverting the periphyton biomass models, described above, to predict the nutrient concentration given biomass thresholds of interest, which in this study were NPS-FM attribute state boundaries of 50 mg chlorophyll m^{-2}, 120 mg chlorophyll m^{-2} and 200 mg chlorophyll m^{-2}

The inversion of the periphyton biomass models to derive criteria is shown schematically in Figure 3, in which a response (e.g., biomass) is represented as a function of a stressor (e.g., nutrient concentration) using an OLS model. The black points represent observations from cases with different levels of the stressor and corresponding response. The blue line represents a model of the stressor–response relationship that is fitted to the observations. The inversion of the stressor–response model to derive criteria is indicated by the red horizontal and vertical arrows. The horizontal red arrow represents a target level of biomass of 120 mg chlorophyll m^{-2}. The level of the stressor that will achieve this response is shown by the vertical green line, which defines the criterion as 658 mg m^{-3}.

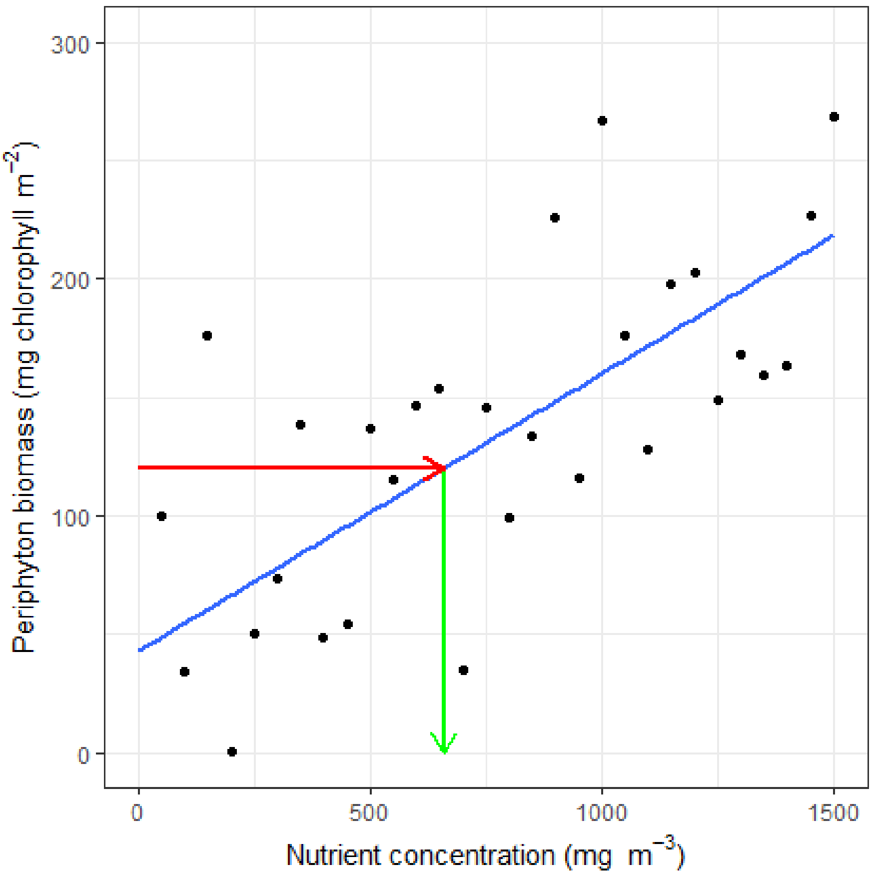

Figure 3. Schematic representation of the derivation of criteria based on a hypothetical stressor-response model. The blue line is an OLS regression fitted to the data. See text for explanation.

While Figure 3 illustrates how an environmental criterion can be derived, the example shown does not account for the uncertainty that always exists in such stressor-response models. In Figure 3, the uncertainty is shown by the scatter of observations around the regression line, due to sampling error and uncontrolled sources of variation. This uncertainty means that there is a risk that the target level of the response will be exceeded even if the stressor is held at or below the derived criterion.

2.3.1 Accounting for under-protection risk

Assuming the model errors are normally distributed, uncertainty in the stressor–response model means 50% of sites will exceed the specified response at the stated criterion. These sites will be under-protected when a criterion has been derived from the model as shown in Figure 3. In addition, 50% of sites will be below the specified response even when the stressor level is higher than the stated criteria. These sites are over-protected when a criterion has been derived from the model as shown in Figure 3.

A subjective decision might be made that the uncertainty is acceptable because the amount by which the 50% of sites that exceed the acceptable response is “small” or ecologically unimportant. However, some stressor–response relationships are less certain than others due to unexplained variation.

Where unexplained variation is large, the assumption that the exceedance is small and acceptable may not be appropriate. A valid criticism of the stressor-response model-based approach to criteria development for periphyton is that relationships between periphyton and in-river nutrient concentrations are generally weak (low variation explained; Dodds *et al.*, 2002; Dodds and Welch, 2000; Lewis *et al.*, 2010; Snelder *et al.*, 2014; Van Nieuwenhuyse and Jones, 1996; Welch *et al.*, 1988). In order to address the issue of a weak stressor-response model, Snelder *et al.* (2019) specifically accounted for under-protection risk in their criteria and called this the spatial exceedance because it indicates the proportion of sites that can be expected to exceed the target level of response even if the stressor is within the stated criteria.

In this study, we anticipated that the stressor-response models would be weak (i.e., the explained variation will be low), and therefore aimed to provide for under-protection risk in the derived criteria. The approach taken in this study is equivalent to that taken by Snelder *et al.* (2019) and is also the discussed in the guidance for deriving nutrient criteria provided by the United States Environmental Protection Agency (USEPA, 2010) and the European Union (Phillips *et al.*, 2018). We note that focussing on under-protection risk indicates a decision to be precautionary with respect to environmental values. For any given level of under-protection risk, there is an associated complementary over-protection risk as shown by the sites below the regression line in Figure 3.

Providing for under-protection risk is shown schematically in Figure 4. The key idea is that a regression model is not simply a single regression line; the model describes the range of values that future, or unobserved, cases will lie. This range is described by a probability distribution that is centred on the regression line in Figure 3. The probability distribution around the line is determined by the model's residual error. If the model residuals are normally distributed, the probability distribution is symmetrical (the regression line represents the mean of the distribution), its width (i.e., the spread of values either side of the regression line) is related to the model variation explained (Figure 4). Therefore, a regression model can be used to predict the entire probability distribution for a specified level of the stressor.

The predicted probability distribution can be used to define response levels that are not exceeded with specified probabilities. For example, the dot-dash, dashed, and dotted lines shown in Figure 4 indicate response levels not exceeded by 70%, 80% and 90% of cases, respectively. We refer to these response levels as those associated with the 70th, 80th and 90th percentiles of the predicted distribution. The probability that the response levels are exceeded is defined by the complement of these percentiles (i.e., 30%, 20% and 10%). Therefore, because the top blue (dotted) line shown in Figure 4 represents the 90th percentile of the distribution, it can be used to define nutrient criteria for which there is a 10% probability the target level of response will be exceeded. We refer to a criterion defined on this basis as having an under-protection risk of 10% (i.e., a probability that the target response is exceeded of 10%).

For the hypothetical stressor-response model discussed here, if the under-protection risk is to be 20%, and the acceptable response is 120 mg chlorophyll m^{-2}, the corresponding criterion is defined by the point at which the red arrow shown in Figure 4 intersects the line representing the 80th percentile of the predicted distribution (dashed blue line in Figure 4). This point is shown in Figure 4 by the green arrow, which indicates a criterion of 266 mg m^{-3}.

Note that this criterion is lower (more stringent) than that defined by the regression line (i.e., the solid blue line shown in Figure 3) because the tolerance of risk of under-protection is lower. The level of risk of under-protection that is used to define criteria is not scientifically defined and is a subjective ('normative') choice. It is noted that the risk of over-protection is the complement of the risk of under-protection (i.e., over-protection risk = 100% - under-protection risk).

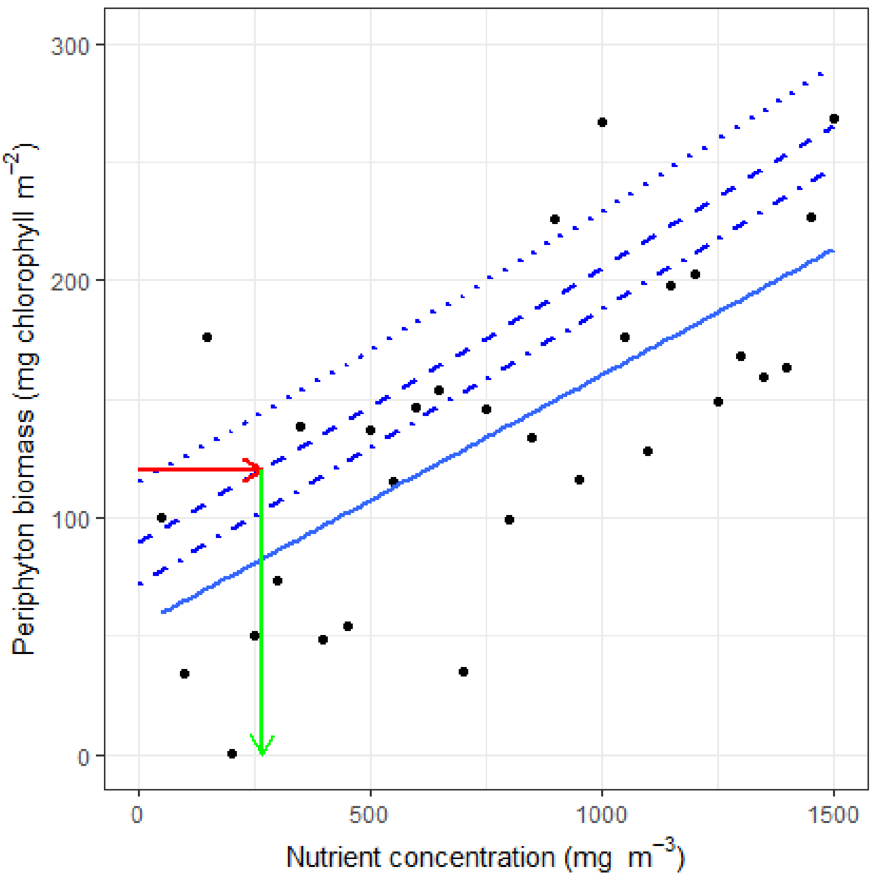

Figure 4. Schematic representation of the derivation of hypothetical criteria showing the fitted OLS regression line (solid blue) and the predicted cumulative probability distribution. The successive blue lines (dot-dash, dashed, and dotted) above the solid blue line represent the 70th, 80th and 90th percentiles of the predicted response distribution. These lines are used to define criteria having under-protection risks of 30%, 20% and 10% respectively. See text for further explanation.

The above example shows how a stressor–response relationship that is fitted to observations using OLS regression can be used to derive criteria with differing levels of under-protection risk. An alternative approach to defining the stressor-response relationship is to use quantile regression (QR). Rather than fitting a model to the conditional mean of the data, QR fits a model to a user-defined conditional quantile of the data (e.g., the 80th percentile; Cade and Noon, 2003). There are several benefits of QR over OLS including being robust when model residuals are heteroscedastic[2] and not normal. Heteroscedastic residuals indicate that there are sources of variation that have not been accounted for in the model and include wedge-shaped plots of stressor versus response (e.g., Figure 5). The wedge-shaped distribution of values indicates that the maximum response is dependent on the value of the stressor, but that there is heterogeneity in the response due to other limiting factors.

[2] Heteroscedastic means that the model residuals change depending on the stressor and response values. Heteroscedasticity is typically identified from a fan-shaped or wedge-shaped plot, in which variance increases as values increase.

A wedge-shaped stressor-response relationship can be quantified by fitting a QR model to some suitable quantile of the data. For example, in Figure 5, the QR model is fitted to the 80th quantile so that 20% of cases lie above the regression line (i.e., a 20% under-protection risk). The regression model for the 80th percentile can be inverted and used to derive criteria in the same way as the OLS. The criteria for a target level of the response of 120 mg chlorophyll m^{-2} is demonstrated in Figure 5 by the point at which the red arrow intersects the regression line. The value of the stressor that is predicted to limit cases with biomass greater than 120 mg chlorophyll m^{-2} to 20% is shown by the green arrow and a criterion of 275 mg m^{-3}.

Note that as for the OLS case, the level of risk of under-protection that is used to define the criteria (the quantile the model is fitted to) is not scientifically defined and is a subjective ('normative') choice.

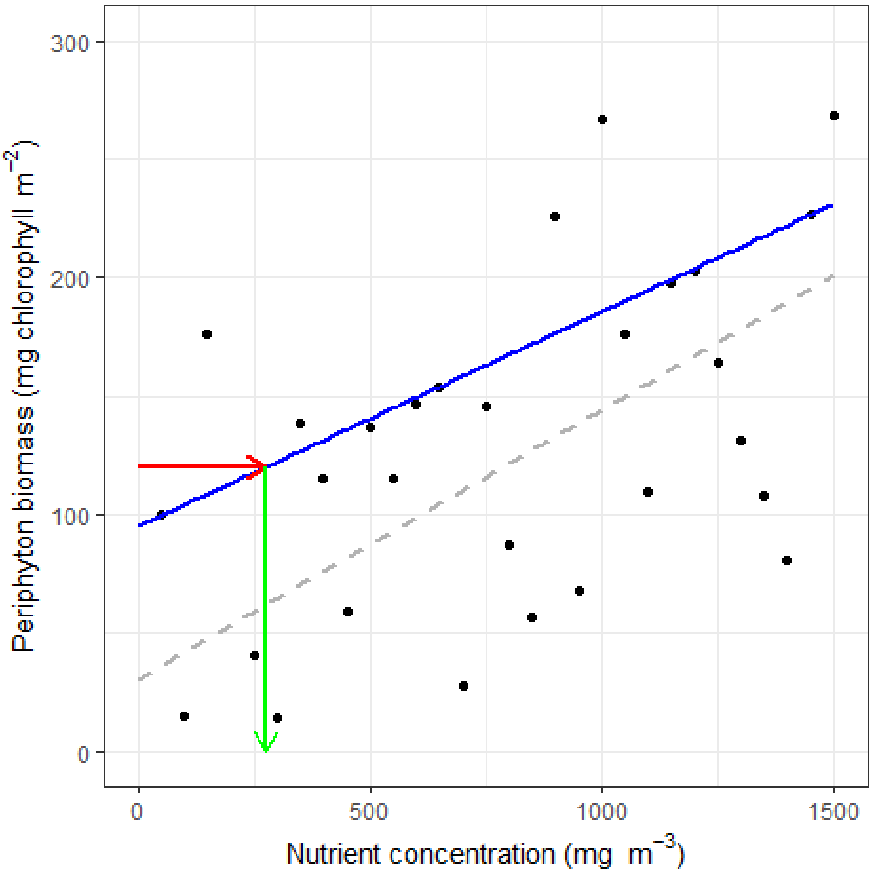

Figure 5. Schematic representation of the derivation of criteria when stressor-response relationship is wedge-shaped. The blue line is a quantile regression fitted to the 80th percentile of the data. The grey dashed line represents an OLS regression. See text for explanation.

2.3.2 Incorporating more than one independent variable

We expected our models to include several environmental variables representing the accrual and loss processes shown in Figure 2, rather than a single independent variable, as illustrated in Figure 3 to Figure 5. Therefore, the relationship between nutrient concentration and periphyton biomass depends on additional environmental variables and differs between locations. If the values of all the environmental variables are known for a location, the models can be inverted and criteria specific to that location can be derived. To provide an easily applied set of criteria, we applied this approach to derive criteria that are specific to rivers grouped into classes.

Following Snelder et al. (2019) river classes were defined by the second level (Source-of-flow) of the REC (Snelder and Biggs, 2002, see section 3.1 for further details). To define nutrient concentration criteria for each Source-of-flow class we calculated the criteria for every segment (order ≥ 3) in the class and then obtained the mean of these values as the class criterion. To calculate the criteria for every segment it was necessary to have the values of the environmental variables that were included in the models. The calculations used estimates of each of the environmental variables that were made for all network segments from spatial models that are described in more detail below.

3 Data acquisition, organisation and processing

The analyses described in subsequent sections depended on water quality, periphyton and hydrological data from regional council SOE river monitoring programmes. In addition, we used existing modelled data that are available for all river segments in New Zealand. In this section we describe the data available from these different sources, followed by an account of selection the preliminary dataset of variables used for data analysis.

3.1 Spatial framework

The spatial framework for our study was the GIS-based digital drainage network, which underlies the REC. The digital network was derived from 1:50,000 scale contour maps and represents New Zealand's rivers as segments bounded by upstream and downstream confluences, each of which is associated with a sub-catchment.

REC Source-of-flow classes were used to broadly stratify New Zealand's rivers into groups with discrete sets of nutrient concentration criteria. The REC Source-of-flow classification subdivides New Zealand's rivers (order ≥ 3) into 21 classes based on differences in catchment climate and topography (Table 1). Source-of-flow classes discriminate differences in the drivers of periphyton including climate (Snelder and Biggs, 2002), hydrological indices (Snelder et al., 2005), physical and chemical characteristics of the water column (Larned *et al.*, 2016) and climate (Snelder *et al.*, 2014). For example, climate categories CD, CW and CX are associated with lower temperatures and solar radiation than climate categories WD, WW and WX. The frequency of high flows varies systematically with climate categories (e.g., CD < CW < CX). In addition, within a climate class there is systematic variation in high flow frequency across topography categories (e.g., Lk < L < GM < M < H). The analysis was therefore expected to produce reasonably similar nutrient concentration targets within a REC Source-of-flow class, and large differences in concentration targets between classes, with the differences understandable in terms of the drivers of periphyton biomass.

Table 1. Defining characteristics, categories, and membership criteria of the River Environment Classifications at each level used in this analysis.

Level	Defining characteristic	Categories	Notation	Category membership criteria
Level 1	Climate	Warm-extremely-wet	WX	Warm: mean annual temperature > 12°C Cool: mean annual temperature < 12°C Extremely Wet: mean annual effective precipitation > 1500 mm Wet: mean annual effective precipitation > 500 and < 1500 mm Dry: mean annual effective precipitation < 500 mm
		Warm-wet	WW	
		Warm-dry	WD	
		Cool-extremely-wet	CX	
		Cool-wet	CW	
		Cool-dry	CD	
Level 2	Topography	Glacial-mountain	GM	GM: M and % permanent ice > 1.5%
		Mountain	M	M: > 50% annual rainfall volume above 1000 m ASL
		Hill	H	H: 50% rainfall volume between 400 - 1000 m ASL
		Low-elevation	L	L: 50% rainfall below 400 m ASL
		Lake	Lk	Lk: Lake influence index > 0.033

3.2 Regional council data

Periphyton biomass has been measured in regional state of the environment monitoring programmes in some regions since about 2008 (e.g., in the Manawatū-Whanganui region). Since the inclusion of periphyton as an attribute in the NOF in the NPS-FM released in 2014, additional regional councils have initiated programmes of regular monitoring of periphyton biomass as chlorophyll and the available data is constantly expanding. The procedure used by all regional councils to collect periphyton samples for chlorophyll analysis is based on the quantitative method 1b described in Biggs and Kilroy (2000), which is the method now prescribed in the National Environmental Monitoring Standards (NEMS) for periphyton. Several councils also collect data from visual assessments of periphyton cover of the stream bed (as percentage cover) by different categories of periphyton (e.g., film, green filaments, cyanobacteria mats), providing high-level descriptions of community composition. Cover data were not used directly in the current analysis but were useful for cross-checking chlorophyll results.

We obtained periphyton data (as chlorophyll, and percentage cover of the stream bed in some cases) and associated environmental data (i.e., water quality and habitat data from the same sites) from 11 regional councils (Table 2). We requested all available data, starting from the earliest periphyton observation up to the most recent available data, and inspected each dataset for appropriate variables and suitable monthly time series.

For each periphyton monitoring site, we identified the corresponding segment of the digital drainage network based on site location descriptions and geographic coordinates.

*Table 2. Summary of data from periphyton monitoring sites received from regional councils. All sites had monthly time series of periphyton biomass as chlorophyll. Environment Canterbury data are divided into two datasets with different time periods. Only sites with at least 20 observations were included in the analysed dataset (see Section 3.1.1). Ongoing data collection means that sites with shorter records can be included in future analyses. *Data were not obtained because the record is still too short but listed for completeness. Note that three further councils (Gisborne District Council, Tasman District Council, Waikato Regional Council) do not collect monthly data on chlorophyll but may collect periphyton data as cover estimates or as chlorophyll at longer intervals.*

Regional council	No. sites	Mean number observations	Mean duration (months)	Number with flow data	Data start year(s)	Data end year (s)
Auckland Council	12	7	9	9	2020	2021
Bay of Plenty Regional Council	29	49	64	26	2015	2021
Environment Canterbury	38	12	14	38	2020	2021
Environment Canterbury 2	24	33	36	24	2011	2014
Environment Southland	34	45	70	34	2015	2021
Greater Wellington Regional Council	16	42	55	15	2016	2021
Hawkes Bay Regional Council	22	49	81	22	2008 - 2017	2013 - 2021
Horizons Regional Council	67	95	116	48	2008 - 2017	2019
Marlborough District Council	9	39	49	1	2016	2020
Northland Regional Council	38	59	74	27	2014	2021
Otago Regional Council	34	24	26	35	2019	2021
Taranaki regional Council	12	35	40	12	2017	2021
*West Coast Regional Council	10	NA	NA	NA	2020	2021

3.2.1 Chlorophyll data

Time series of chlorophyll (as mg m^{-2}) values at each site were converted to the 92nd percentile, consistent with the definition of the periphyton attribute in the NPS-FM, and hereafter referred to as Chla92.

The periphyton attribute in the NOF requires that the periphyton metric (the 92nd percentile of chlorophyll) is calculated from monthly data collected over at least three years. Because of this open-ended requirement (i.e., no maximum length of record) we used all available monthly periphyton data to calculate the metric.

Data collection from some regional councils has commenced more recently than three years ago. To improve spatial coverage of the data the final dataset used for fitting the models included all sites with at least 20 observations. For details refer to Section 5.1.

3.2.2 Treatment of "missing" chlorophyll data

The time series of chlorophyll included gaps at most sites, with an overall average rate of 20% "missing" data. The rate of missing data varied across regions with highest rates in Southland and Hawkes Bay (approximately 35%).

It is likely that high river flows explained many missed sampling occasions, and that had it been possible to collect a periphyton sample at these times the concentration of chlorophyll would have been low. Therefore, not accounting for such missing low values could lead to a dataset biased towards higher concentrations. However, the amount of periphyton on the riverbed at times of different flow magnitudes is actually unknown. There may also be other unknown reasons for missing a sampling occasion, such as logistics (e.g., site access) and the April 2020 national COVID-19 lockdown.

In the absence of any consistent general rule for dealing with missing data, we chose to calculate the 92nd percentile from the available data only. This decision is unlikely to produce large differences compared to the alternative treatment of setting missing values to zero or some other nominal low value because the 92nd percentile is relatively insensitive to missing data values at the low end of the range.

3.2.3 Water quality data

Data for nine water quality variables were available for most sites (Table 3). The raw data in each regional council dataset were checked for consistency of units before amalgamating all data into a single combined dataset. For example, electrical conductivity was variously reported in mS m^{-1} or µS cm^{-1} and all data were converted to the latter unit. Nutrient data (N and P) were generally reported as mg L^{-1} in the raw data. We converted all these data to mg m^{-3} for the analyses to avoid the use of multiple decimal places, especially for DRP and TP.

For each site with data, we used time series that comprised monthly observations (rather than quarterly or annual data) to ensure that the data were not biased towards a particular season. Summary statistics were calculated for each variable. All nutrient data (TN, DIN, TP, DRP, NH_4-N) were converted to site median concentrations and EC, pH, turbidity and water clarity to site median values. For water temperature we calculated the 95th percentile value at each site (Temp95), to represent the peak summer temperature.

3.2.4 Additional variables describing site physical conditions

Most regional councils also assessed and recorded physical conditions at each site or carried out habitat assessments using methods such as that described by (Edgar *et al.*, 1994).

Measures of shade at each site included densiometer readings, visual assessments of percentage shade, assignments of the amount of shade into categories (e.g., unshaded <20% shade; partial shade 20 – 60% shade; full shade, >60% shade), and riparian shade scores as part of a wider habitat assessment. We converted these different measures to a consistent categorical variable at each site by combining partial shade and shade into a single category. For the few sites where no shade assessment was provided, we manually assigned a shade category based on viewing the Google Earth image of the site at the coordinates provided. The accuracy of this method was confirmed by checking Google Earth images of a subset of sites for which shade data were available.

Data on bed substrate composition were available at many sites as visual assessments in categories ranging from bedrock to sand and silt (e.g., method used by Jowett 1993). The substrate categories used varied across the regional datasets. An estimate of the percentage of the bed covered by fine sediment (sand or smaller) was available for almost all sites and was included as an independent variable.

3.2.5 Variables with incomplete data

The dataset included a range of variables for which data were not available at all sites. These variables could not be used in model development, but we were able to use the incomplete data to test hypotheses that might explain differences in Chla92 between sites.

Several councils provided data on visual estimates of the proportion of the bed covered by different categories of periphyton assessed as described in Kilroy et al. (2013). We used mean and maximum cover by didymo (%) and cyanobacteria (assumed to be *Microcoleus*). Both of these periphyton taxa can potentially cause unexpectedly high chlorophyll densities compared to other periphyton (Kilroy et al. 2009, Hart et al. 2013).

We also used the mean proportion (as a %) of the bed covered by coarse sediment (boulders (> 128 mm across) and/or bedrock), which was available at about 70% of sites.

3.2.6 Hydrological data

The hydrological data sets supplied by regional councils consisted of time series of continuous flows from gauging stations that were at, or close to, the periphyton monitoring sites. We also used data from several NIWA hydrology stations that were close to periphyton monitoring sites. All gauging stations were located on a segment of the digital drainage network based on station location descriptions and geographic coordinates. Flow data were converted to time series of mean daily flows. Measured hydrological data were available for 77% of the periphyton monitoring sites.

The time series of daily mean flows were used to compute 10 hydrological indices that characterize four aspects of the flow regime that characterise disturbance: (1) variation of flows, (2) magnitude and duration of annual extreme flows, (3) frequency and duration of high and low flow pulses (4) rate and frequency of changes of flow (Olden and Poff, 2003; Snelder and Booker, 2013). These indices were derived for two durations: (1) the period of record associated with the periphyton observations and (2) the full period of record available.

Only complete years of data were used to calculate the flow metrics, with "water years" running from 1 October to 30 September, and the year named according to the latter 9 months (e.g., 1 October to 31 December 2000 are placed in the 2001 water year). This definition of water year ensures that summer and autumn low flows (which can be important for periphyton accrual) are not split across years. We searched all records for gaps in the time series longer than 27 days and removed the year of record in which these gaps occurred. Selection of 27 days ensured that in any year there is at least one flow observation in every month.

The magnitude and duration of annual extreme flows was represented by the mean annual maximum and minimum flows over durations of 7 and 30 days (Max7, Max30, MALF7, MALF30) and the standard deviation of flows. These statistics were calculated by first estimating the maximum and minimum of a 7- and 30-day moving average flow in each year of record for each site and the standard deviation of daily flows. Max7, Max30, MALF7, and MALF30 were the mean of these annual values divided by the mean daily flow for the entire record. We divided the standard deviation of the daily flows by the mean daily flow and refer to this as sdQ (note that the standardisation by the mean flows means the sdQ is the coefficient of variation of daily flows). Rivers with high sdQ, Max7 and Max30 have large high flows compared to base flow and tend to have low and sustained periods of low flow (i.e., periods without disturbance; Figure 2). Similarly, rivers with high MALF7 and MALF30 and low sdQ have sustained baseflow and sites with low MALF7 and MALF30 and high sdQ have low and sustained periods of low flow.

The frequency of high flows was represented by the mean number of events per year that exceeded n times the long-term median flow (FREn). We calculated FREn for values of n of 2,3, and 4. If the time interval between an event dropping below the threshold and the next event rising above the threshold was less than 5 days, only a single event was counted. Rivers with high FREn have frequent high flows and therefore high disturbance (short accrual periods; Figure 2).

Rates of change in flow was represented by the number of negative differences in flow between days (i.e., the number of days on which flow was less than that of the previous day; nNeg). The values of nNeg were calculated for each site by first counting the number of days in each year for which the flow reduced on the subsequent day. For each site, nNeg is the mean of these values over all years. Rivers with large values of nNeg have high rates of increasing flow and therefore the rising limbs of their hydrographs are steep. Sites with low values of nNeg have lower rates of increasing flow and therefore gentle rising limbs. In New Zealand, large rivers, spring fed systems and snow melt fed systems tend to have lower nNeg. Rivers with high nNeg are ‘flashy’ and occur in regions that experience stormy high rainfall climate and where the runoff response to rainfall is fast. Rivers with high nNeg have high disturbance (short accrual periods; Figure 2).

The frequency of changes of flow was represented by Reversals. Reversals were calculated for each site as the number of negative and positive changes in flow conditions from one day to the next. Sites with low values of Reversals have infrequent changes in flow conditions and therefore long rising and falling limbs. Rivers with high Reversals have high disturbance (short accrual periods; Figure 2).

3.3 Modelled data

Modelled data that apply to every segment of the digital river network were required in order to estimate nutrient criteria applicable to segments in the 21 Source-of-flow classes (see Section 2.3.2). Some modelled variables were also used as potential independent variables in the initial models.

3.3.1 FENZ variables

Variables describing aspects of the physicochemical and geological environment are available for every segment of the digital river network from the Freshwater Ecosystems of New Zealand

(FENZ) Geodatabase (Leathwick *et al.*, 2010). FENZ variables associated with the REC network include at least 120 variables, developed during the production of a multivariate river classification for the Department of Conservation (Wild et al. 2005). We used only two variables that describe solar radiation from the FENZ dataset (see section 3.4)

3.3.2 Network predictions of hydrological indices

Hydrological indices were obtained for all monitoring sites from predictions made for all segments of the digital river network in previous studies (Booker and Snelder, 2012; Snelder and Booker, 2013). These predicted indices were the same as those derived from hydrological records pertaining to each monitoring site (see Section 3.2.6). The benefit of the predicted hydrological indices over those derived by this study is that they were available for all monitoring sites, not just the 77% of sites with measured hydrological data.

Potentially useful hydrological variables also included predictions of the flow at each segment that would initiate periphyton removal (Haddadchi *et al.*, 2020).

3.3.3 Other predictions of independent variables

In addition to modelled data that were already available (Sections 3.3.1 and 3.3.2) it was necessary to obtain modelled data across the digital network for independent variables used in the periphyton – environment models, for which modelled data were not already available in the FENZ database. This included using predictions generated in recent modelling efforts (e.g., Whitehead 2018), and generating new predictions as part of this project. Further details are provided below (Section 4.6).

3.4 Selection of potential independent variables

The preliminary set of potential independent variables for the regression models is shown in Table 3. The selection was assembled based on both the conceptual model (Figure 2) and knowledge gained from previous studies (e.g., Larned *et al.*, 2015; Matheson *et al.*, 2016; Snelder *et al.*, 2019, 2014; Kilroy et al. 2018).

The water quality and habitat data obtained from regional councils (Section 3.2) were used to represent aspects of the processes that lead to either periphyton accrual or removal (Figure 2). The four nutrient forms (TN, DIN, TP and DRP), temperature (Temp95), visual clarity (Clar) and turbidity (Turb) represent resources (i.e., nutrients, temperature and light). Electrical conductivity (EC) is not explicitly represented in the conceptual model but has been identified as a variable that integrates several higher-level catchment processes (such as catchment enrichment, temperature and flow) (Biggs 1990). EC has also featured as a significant independent variable in previous periphyton – environment models using smaller datasets (e.g., Kilroy *et al.*, 2018, 2020).

Site measurements of Shade (Shade) represent resources (temperature and light). Site measurement of fine sediment cover (FineSed) may differentiate sites in terms of potential disturbance by changes in flow.

From FENZ, we selected solar radiation in June (i.e., winter) and December to represent between-site variation in potential insolation, which is an important factor controlling periphyton biomass (as chlorophyll) accrual (Figure 2). These variables were derived by Wild *et al.* (2005) from a grid of estimated solar radiation that accounted for latitudinal variation in sunlight and interception by cloud cover (Leathwick *et al.*, 2003).

The initial selection of hydrological indices included all those described in Section 3.2.6, calculated from the three alternative sources of the hydrological data (i.e., the record pertaining to the period of periphyton observations, the full hydrological record, and the modelled data).

Periphyton removal flows from Haddadchi et al. (2020) were not included in the initial selection because the predictions were available only for 75% of segments in the river network.

Table 3. Preliminary selection of potential independent variables for the regression models

Abbreviation	Variable	Units	Source
TN	Total nitrogen	mg m^{-3}	Monitoring data
DIN	Dissolved inorganic nitrogen	mg m^{-3}	Monitoring data
TP	Total phosphorus	mg m^{-3}	Monitoring data
DRP	Dissolved inorganic phosphorus	mg m^{-3}	Monitoring data
Temp95	95^{th} percentile of water temperature	degrees C	Monitoring data
Clar	Visual clarity of the water column	m	Monitoring data
Turb	Turbidity of the water column	NTU	Monitoring data
NH4N	Ammoniacal nitrogen	mg m^{-3}	Monitoring data
EC	Electrical conductivity	µS cm^{-1}	Monitoring data
FineSed	Proportion of fine substrate	%	Monitoring data
Shade	Proportion shade	categorical	Monitoring data
SolarRadJune	Solar radiation in June	W m^{-2}	FENZ database
SolarRadDec	Solar radiation in December	W m^{-2}	FENZ database
Max7	Mean annual 7 and 30-day high flow divided by mean flow	Unitless	Three separate sources: daily flow records pertaining to the period of periphyton observations and the full hydrological record, and modelled data.
Max30			
MALF7	Mean annual 7 and 30-day low flow divided by mean flow		
MALF30			
FRE2	Number of events per year that exceeded two, three and four times the long-term median flow	$Year^{-1}$	
FRE3			
FRE4			
nNeg	Number of negative differences in flow between days	$Year^{-1}$	
Reversals	Number of hydrologic reversals	$Year^{-1}$	
sdQ	Standard deviation of daily flows divided by mean flow	Unitless	

3.5 Independent dataset

We used an independent dataset to validate the criteria derived by this study. The independent dataset represented periphyton biomass and nutrient concentration data from river water quality monitoring sites belonging to the National Water Quality Monitoring Network (NRWQN). The NRWQN comprises 77 sites located on 48 of New Zealand's rivers, which broadly represent variation in main-stem rivers across New Zealand (Davies-Colley *et al.*, 2011). Since 1989 a range of water quality variables and a visual assessment of the cover of filamentous and mat forming algae has been carried out monthly at NRWQN sites and flows have been monitored continuously (Davies-Colley *et al.*, 2011; Smith and McBride, 1990).

Data from the NRWQN were used by Snelder *et al.* (2019) to derive the existing periphyton nutrient criteria. Snelder *et al.* (2019) used data for the time-period 1989 to 2010 (22 years) but excluded 11 NRWQN sites for various reasons. The sites AK1, AK2 and GS1 (see Snelder *et al.*, (2014) for site codes) were excluded because they are located on deep rivers with silty beds and do not support conspicuous periphyton biomass. Sites RO2 and RO6 were excluded due to many missing periphyton observations. Sites WH3 and WH4 were excluded because they are dominated by macrophytes. Three sites on large rivers including the Waikato (HM2, HM4 and HM5) and the Clutha (DN4) rivers were excluded due to logistical difficulties in sampling periphyton and artificially fluctuating water levels.

Records were split for some of the 66 retained sites into two portions to account for significant changes that had occurred at the site through the time-period. The two portions of the record were treated as separate sites in the analyses that follow and each portion is referred to as representing a training site. Five sites (HM1, RO4, WA6, WN3, and DN2) were split due to changes in site locations that were required for operational reasons. Five sites (RO3, HV5, WN2, TK2 and DN1) were split due to very obvious changes in mean water quality. Thirteen sites in the South Island (NN3, NN5, GY1, CH1, TK3, TK4, TK6, AX1, AX2, AX3, AX4, DN4 and DN9) were split because they were colonized by the invasive alga *Didymosphenia geminata* (Kilroy et al., 2009). The abundance of *D. geminata* responds to factors that differ from those that promote blooms of other algae in New Zealand rivers (Bothwell et al., 2014). For this reason, the portion of the record prior to the establishment of *D. geminata* was retained and the second portion was discarded. Site TK1 was split due to the pre-commissioning failure of the Opuha dam and the subsequent establishment of *D. geminata*. After excluding some sites and splitting others there was a total of 78 sites comprising either the entire record of the NRWQN site or part thereof, which we hereafter refer to as the training sites.

Monthly observations of periphyton cover in two categories (mats and filaments) had been made by visual assessment at the retained sites since 1989. For details of these observations see (Snelder *et al.*, 2019). The cover estimates were aggregated into a single metric called weighted composite cover (WCC), which is defined as 0.5 x average cover by mats + average cover by filaments (Matheson *et al.*, 2012). The weighting of the two periphyton categories is based on the guideline that filaments and mats are problematic if they exceed 30% and 60% of the visible stream bed (generally < 0.75 m deep) respectively (Ministry for Environment, 2000). For this study we converted the 92nd percentile of observations of WCC at each site to an estimate of equivalent chlorophyll concentration (i.e., equivalent to the Chla92 measure for monitoring sites used in this study). The conversion was made using Equation 1 (the R^2 for the fitted model was 0.61; for further details see Snelder *et al.*, 2019).

$$log_{10}(chlorophyll\ a) = 0.398 + 0.235\left(\sqrt{WCC}\right) \qquad \textit{Equation 1}$$

4 Analysis methods

4.1 Overview of analyses

The analyses were aimed at defining an appropriate statistical model that explained between-site variation in the 92nd percentile of observed chlorophyll (Chla92) as a function of independent variables and then using that model to derive nutrient criteria. These analyses are summarised in Figure 6 as an eight-step process. This process uses the data described in Section 3 and was informed by the conceptual model that is described in Section 2. Details of the eight steps are described in the following sections.

An additional step was to generate a national model of current periphyton state (as Chla92). The purpose of this step was to facilitate discussion of the outcomes of the OLS models generated using observed independent variables.

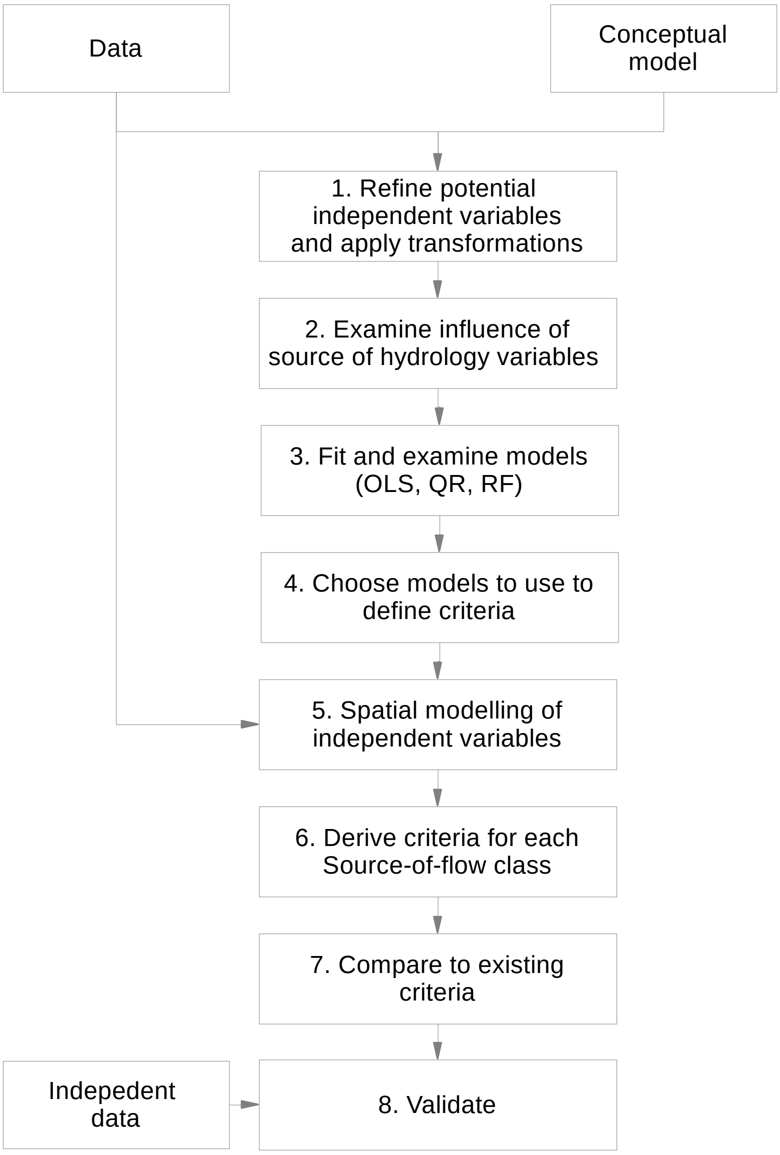

Figure 6. Schematic diagram summarising the analytical process for deriving and validating the criteria.

4.2 Refine selection of potential independent variables and apply transformations

The initial set of potential independent variables listed in Table 1 was reviewed and refined by removing one of any pair of variables that had a Pearson's correlation coefficient of greater than 0.8. When removing variables, we sought to retain variables that had been used in previous studies in New Zealand including those of Biggs (2000); Matheson *et al.* (2016); Snelder *et al.* (2019, 2014). The exception to this was the four nutrient variables (TN, DIN, TP and DRP) because only one of these was ever used as an independent variable in a model.

There were two reasons for this procedure. First, high correlation indicates redundancy in the preliminary selection of potential independent variables, the removal of which simplified the model input data. Second, correlation between independent variables in a regression model means the model may not produce reliable predictions when the value of an independent variable is changed (Zuur *et al.*, 2010). This was particularly important in this study for the independent variables representing nutrient concentrations because the criteria derivation process involved using the inverted model to predict nutrient concentration given the Chla92 threshold of interest.

At this step we also determined transformations to apply to the response and independent variables. Transformations were applied to the response Chla92 to make the data distribution as normal as possible. Transformations were applied to the independent variables to linearise their relationships with the response.

4.3 Examine influence of alternative sources of hydrological indices

The hydrological indices used in our models could be obtained from three alternative sources. The first two sources were based on observed flows at the monitoring sites and pertained to two periods: (set 1) the period of record associated with the periphyton observations and (set 2) the full period of record available. The third source (set 3) was model predictions for all segments of the digital river network made by previous studies (Booker and Snelder, 2012; Snelder and Booker, 2013).

We anticipated that there might be differences in the degree to which indices in sets 1, 2 and 3 can explain between-site variation in Chla92. For example, set 1 (based on the period of record associated with the periphyton observations) might explain more variation than set 2 (full period of record) because the former represents the flow conditions that the periphyton was exposed to. Alternatively, the long-run hydrology represented by set 2 may explain more variation than set 1. This outcome might occur if Chla92 is linked to hydrologically determined geomorphic conditions at sites and these conditions are strongly associated with long-run hydrology (set 2). We also anticipated that the predicted indices might explain variation in Chla92 because national scale models of a range of hydrological indices have been shown to have reasonable accuracy (e.g., (Booker and Snelder, 2012; Booker, 2013). If this were the case, there would be an advantage in that a larger number of sites could be included in our models (i.e., we could include sites that are not associated with measured flows).

We investigated which hydrological data source produced the best models by fitting OLS models, using the approach described in section 4.4.1. We offered the model fitting procedure all the selected potential independent variables and repeated this three times for each nutrient. At each repetition, the source of the hydrological indices was altered, being set 1, 2 or 3. We restricted the fitting dataset to those sites that had all three sets of hydrological data so that all models were comparable.

For each nutrient, we compared model performance across the three models based on the variation explained (i.e., model R^2 value). We also assessed whether the difference in the R^2 values between all pairs of models for each nutrient were statistically significant by comparing the residual sums of squares (RSS) using analysis of variance (ANOVA). In addition, we assessed whether differences in model performance were of practical significance using the root mean square deviation (RMSD), as described in Section 4.4.1 below).

4.4 Fit and examine regression models

4.4.1 OLS regression models

Variable selection has an important role in the model building process. In practice, it is common to have a large number of candidate independent variables available, and they are included in the initial stage of modelling for the consideration of removing potential modelling bias (i.e., increasing the variation explained by the model, Hastie *et al.*, 2001). However, it is undesirable to keep irrelevant independent variables in the final model because this makes it difficult to interpret, may decrease predictive performance and may include redundant and correlated variables. We therefore fitted the OLS models in two steps.

First, 'full' models that included all the independent variables (appropriately transformed, as described in Section 4.2; refer to Section 5.2 for details) but only one of the four nutrients (TN, DIN, TP, DRP), were fitted.

At the second step, standard forward and backward stepwise variable elimination was applied to the saturated models to identify the most parsimonious models. In this procedure, the Akaike information criterion (AIC; Akaike, 1973) was used to apply a penalised log-likelihood method to evaluate the trade-off between the degrees of freedom and fit of the model as independent variables were added or removed (Crawley, 2002). We considered the use of other model variable selection procedures including best subsets and least absolute shrinkage and selection operator. In our experience these alternative procedures do not result in significantly different variable selections and have the disadvantage of being more complicated. Independent variables that were retained in the model after the stepwise variable elimination can be regarded as making significant contributions to model prediction accuracy (Shmueli, 2010).

After fitting the models, we first assessed the effects of collinearity between independent variables in each model using variance inflation factors (VIF; Zuur *et al.*, 2010). Collinearity is the existence of correlation between the independent variables and leads to problems with the reliability of predictions and in interpreting the regression models. When collinearity is low, the signs of the regression coefficients fitted to each independent variable can be reliably interpreted as representing the directions of their relationships with the response. We calculated the VIF values for all independent variables in each model and compared these to the most stringent threshold of 2 indicated by Zuur *et al.* (2010). If all independent variables had VIF values less than 2, we compared the directions of the relationships indicated by the fitted coefficients with the directions indicated by the conceptual model.

After fitting the OLS models, we also considered whether they were consistent with the conceptual model (Figure 2). A positive association between the model response (i.e., Chla92) and the nutrient concentration was necessary for using the model to derive criteria at the subsequent step. We then examined the signs of the other fitted coefficients to confirm that there were also consistent with the conceptual model.

The performance of the OLS models was measured by comparing the observed Chla92 with independent predictions made by a leave-one-out cross validation (LOOCV) of the models and measuring the degree of agreement using three statistics: R^2, bias and the root mean square deviation (RMSD). The R^2 value is the coefficient of determination derived from a regression of the observations against the predictions. The R^2 value shows the proportion of the total variance explained by the regression model but is not a complete description of model performance (Piñeiro *et al.*, 2008). Bias measures the average tendency of the predicted values to be larger or smaller than the observed values. Percent bias (PBIAS) is computed as the sum of the differences between the observations and predictions divided by the sum of the observations (Moriasi *et al.*, 2015). Optimal PBIAS is zero, positive values indicate underestimation bias and negative values indicate overestimation bias (Piñeiro *et al.*, 2008). RMSD is the mean deviation of predicted values with respect to the observed values and quantifies the characteristic (i.e., mean) error of the predictions (Moriasi *et al.*, 2015).

Consistency with the conceptual model was also examined graphically by using the models to predict the response along a gradient in nutrient concentration for each of the 21 Source-of-flow class (see Section 2.3.2). To make these predictions we first extracted the mean values of each of the independent variables in the model for each of the 21 Source-of-flow class classes. We then used the mean value of each independent variable to typify each class and used the model to predict the class response to increments in a gradient in nutrient concentrations that reflected the observed range of each nutrient in our dataset.

Predictions made using the fitted models represent the mean of the probability distribution at each increment of the nutrient concentration. However, based on the assumption that the residuals are normally distributed, the model was used to make predictions for specific percentiles of the prediction probability distribution. We refer to these predictions as $\widehat{Chla}92$, the ‘hat’ symbol indicating that this is a model prediction and Chla92 indicating the prediction is of the 92nd percentile of chlorophyll (mg m^{-2}). The predictions were estimated from the model by setting the response to be prediction intervals. The prediction intervals were set so that the upper confidence limit provided predictions at percentiles of the probability distribution that were consistent with the nominated levels of under-protection risk as follows:

$$Prediction\ Interval = 1 - 2\ \times \frac{Level\ of\ under\text{-}protection\ risk}{100\%} \qquad \textit{Equation 2}$$

For example, the criteria corresponding to a 20% level of under-prediction risk are defined by the 80th percentile of the prediction probability distribution (i.e., 20% of predictions are expected to exceed this value). Following equation 2, the values that define the criteria for the 20% level of under-prediction risk are equivalent to the upper prediction interval of the 60% (i.e., $1 - 2 \times 20/100$) symmetric prediction confidence interval.

The predictions corresponding the upper confidence intervals from the model were back-transformed as follows:

$$\widehat{Chla}92 = CF +\ (Predictions)^4 \qquad \textit{Equation 3}$$

where CF was the correction factor for retransformation bias based on the method of Duan (1983), and *Predictions* are the upper confidence limit produced by the fitted model. Note that a fourth-root transformation was applied to Chla92 and therefore the back-transformation is to raise to the fourth power.

The $\widehat{Chla}92$ values for a given percentile of the probability distribution were plotted against nutrient concentration for each Source-of-flow class. Based on the conceptual model, our expectation was that for a given nutrient concentration, classes with higher resources (i.e., high light and temperature) and lower disturbance (e.g., low FRE3 or sdQ) would have higher $\widehat{Chla}92$ values.

4.4.2 Quantile regression models

Most applications of QR to the development of nutrient criteria are based on simple bivariate models describing the relationship between a nutrient concentration and a trophic state descriptor such as biomass (e.g., Phillips *et al.* 2018). To be useful in this study, we needed to define multivariate QR models that explain variation in Chla92 as a function of the factors represented by the conceptual model. A model that included variables representing resources (i.e., nutrients, light and temperature) and disturbance (i.e., hydrological regime and substrate) could be inverted and used as described in Section 2 to derive criteria for REC classes. Furthermore, to be useful, quantile regression models would be needed for a range of quantiles to define criteria representing different levels of under-protection risk. We were nominally interested in the 10%, 20% and 30% levels of under-protection risk because these have been defined using the method of Snelder *et al.* (2019). These levels of under-protection risk correspond to regression models of 0.7, 0.8 and 0.9 quantiles.

Variable selection and elimination procedures for quantile regression models exist (e.g., Peng and Wang, 2015; Wu and Liu, 2009). However, the dataset we had was relatively small, especially in the context of attempting to estimate models describing extreme percentiles (i.e., 0.8 and 0.9) where a large sample size is required for estimation of the model parameters. We therefore used a simpler approach to assessing whether quantile regression was able to provide useful models. For each nutrient, we took the independent variables that were retained by the corresponding OLS model and used these to fit a QR model.

We inspected the fitted QR models by first determining which of the independent variables were significant terms in the fitted model. We considered an independent variable was significant if its fitted *p*-value was < 0.05. QR models do not have the equivalent of the variation explained (R^2). We also considered whether the fitted model was consistent with the conceptual model (Figure 2). A positive and significant association between the model response (i.e., Chla92) and the nutrient concentration term was necessary for using the model to derive criteria at the subsequent step. Consistency of the other fitted independent variables with the conceptual model was assessed by checking the signs of the coefficients and graphically by predicting the response along gradients in nutrient concentration for each Source-of-flow class (see section 4.4.1).

Using the fitted QR models, we made predictions of the response along a gradient in nutrient concentration for each REC Source-of-flow class. The predictions (i.e., $\widehat{Chla}92$) were plotted against nutrient concentration for each Source-of-flow class and for each regression quantile. As for the OLS models, we expected that for a given nutrient concentration, Source-of-flow classes with higher resources (i.e., high light and temperature) and lower disturbance (e.g., low FRE3 or sdQ) would have higher predicted values (i.e., $\widehat{Chla}92$).

4.4.3 Random forest models

Random Forest (RF) modelling is a machine learning regression and classification method (Cutler *et al.*, 2007). Whereas OLS and QR models quantify the value of the response variable as a linear combination of independent variables, RF models are based on an ensemble of classification and regression trees (CART; Breiman *et al.*, 1984). RF models are therefore not able to be expressed as simple equations, like a linear regression, and the underlying relationships are non-parametric.

RF models are generally used for 'predictive' statistical modelling rather than more traditional 'explanatory' modelling. Predictive statistical models have the objective of predicting a response value (Y) given input values (X) and are based on using a training dataset to 'learn' how to most accurately predict a new observation of Y given X. Explanatory modelling can be considered a process of testing the degree to which observation data is consistent with a causal theoretical model (Shmueli, 2010).

In this study we used RF models in two ways.

1) RF models were used as a non-parametric alternative to OLS and QR models to assess the degree to which a completely data-driven and non-parametric model is consistent with our conceptual model (Figure 2). We considered that significant differences between the structure and/or predictions made by RF compared to the parametric models would suggest that details of our OLS and QR models may be incorrect or missing.

2) RF models were used to make predictions of both Chla92 and several independent variables that had been observed at monitoring sites, but which were not available for other locations in the national river network. The details for these second applications of RF in this study are provided in sections 4.6 and 4.10.

Because RF models are tree-based and non-parametric, they do not predict values of the response that are outside of the range of the observations. However, extrapolation of the models into unsampled environmental conditions would be needed to produce a comprehensive set of criteria. Therefore, we anticipated that RF models would be inappropriate for deriving criteria because coverage of the range of environmental conditions by the fitting data was incomplete compared to the range that exists across the whole country.

RF models are widely used in environmental science (e.g., Messager *et al.*, 2021) and have been used extensively in New Zealand for predicting river and lake water quality based on observations obtained from routine monitoring (e.g., Unwin *et al.*, 2010; Whitehead, 2018). For detailed descriptions of RF models and their diagnostic tools readers are directed to Breiman (2001) and Cutler *et al.* (2007). The brief description below of some aspects of RF models allows the use of, and results obtained from, RF models in this study to be understood.

RF models are free from distributional assumptions and automatically detect and fit non-linear relationships and high-order interactions. RF models achieve high prediction accuracy by basing predictions on an ensemble of single CART models (a forest) (Breiman, 2001). Because they are focussed on predictive statistical modelling, the performance of RF models is generally quantified based on a set of predictions for observations in the training dataset that is independent of the fitting process. During the fitting process, RF model predictions are made for a subset of observations that are excluded from the fitting process; these excluded observations are known as out-of-bag (OOB) observations. To quantify model performance, the predicted response of the OOB observations is compared with the independent predictions made by the model and the degree of agreement is expressed using four statistics described in Section 4.4.1

The relationships between response and independent variables[3] in RF models can be represented by importance measures and partial dependence plots (Breiman, 2001; Cutler *et*

[3] The independent variables used in predictive statistical models such as random forest models are generally called predictors. In this report we refer to them consistently as independent variables for two reasons. First, we are using RF as a non-parametric alternative to OLS and QR and in this context we are considering whether differences in biomass between sites can be explained in a manner that is consistent with the conceptual model. Second, we are using RF models to make predictions of

al., 2007). To assess the importance of each independent variable, the values of the response variable are randomly permuted for the OOB observations, and predictions are obtained from the tree for these modified data. The importance of an independent variable is indicated by the degree to which prediction accuracy decreases when the response variable is randomly permuted. Importance is defined in this study as the loss in model performance (i.e., the increase in the mean square error; MSE) when predictions are made based on the permuted OOB observations compared to those based on the original observations. The differences in MSE between trees fitted with the original and permuted observations are averaged over all trees and normalized by the standard deviation of the differences (Cutler *et al.*, 2007).

A partial dependence plot is a graphical representation of the marginal effect of an independent variable on the response when the values of all other independent variables are held constant (at their respective mean values). The benefit of holding the other independent variables constant is that the partial dependence plot effectively ignores their influence on the response variables. Partial dependence plots do not perfectly represent the effects of each independent variable, particularly if explanatory variables are highly correlated or strongly interacting, but they do provide an approximation of the modelled predictor-response relationships that are useful for model interpretation (Cutler *et al.*, 2007).

RF models include any of the original set of independent variables that are chosen during the model fitting process. However, marginally important independent variables may be redundant (i.e., their removal does not affect model performance) and their inclusion complicates model interpretation. We used cross validation and a backward elimination procedure to remove redundant independent variables from the models (Svetnik *et al.*, 2004). The procedure recursively removes the least important independent variables to identify a 'reduced' model. We chose to reduce the models such that their predictive performance was the smallest error rate (Breiman *et al.*, 1984).

We fitted RF models that described the observed periphyton Chla92 as a function of the potential independent variables. Although RF models do not depend on distributional assumptions, transformation of the response variable to an approximately symmetric distribution improves model performance. We used the same transformation as used in the OLS regression models when fitting the RF models to periphyton Chla92. We fitted four separate models, each model having one nutrient (TN, DIN, TP or DRP) as one of the potential independent variables. We evaluated the four fitted reduced models using the performance statistics: R^2, PBIAS and RMSD and assessed the fitted relationships using independent variable importance scores, partial plots and graphically by predicting the response along gradients in nutrient concentration for each Source-of-flow class (see section 4.4.1).

The predictions for any percentile of prediction probability distribution were estimated as:

$$\widehat{Chla}92 = CF + \ (\mu + Z \ \times RMSD)^4 \qquad \textit{Equation 4}$$

where CF was the correction factor for retransformation bias based on the method of Duan (1983), μ is the prediction returned from the model which represents the mean of the probability distribution, RMSD is the characteristic model error, and *Z* represents increments in the number of standard deviations from the mean for the standard normal distribution. For example, the 80th and 90th percentiles of the probability distribution can be obtained using *Z* values of 0.84 and 1.28, respectively.

several independent variables that are subsequently used to make predictions of periphyton biomass. Referring to independent variables avoids having to refer to predicting predictors, which would be confusing.

4.5 Choose models to use to define criteria

As set out in Section 2.2, to be useful for defining nutrient criteria, models need to explain variation in Chla92 using a combination of independent variables whose relationships with the response are consistent with the conceptual model (including a positive association between the nutrient concentration and Chla92) and have adequate statistical performance. Based on these requirements, at the fourth step, we selected a set of either the OLS or QR regression models (i.e., four models, each using one of TN, DIN, TP and DRP to represent nutrient concentration) from models described above to define nutrient criteria. Our prior expectation was that QR would be a better choice if they were found to be adequate and if not, we expected that OLS could be used based on previous experience (Snelder *et al.*, 2019).

4.6 Spatial modelling of independent variables

At the fifth step in our derivation of nutrient criteria (Figure 6), we used the fitted models to predict nutrient concentrations that would achieve the Chla92 thresholds of interest for REC Source-of-flow classes. For this step we required the values of the independent variables used in the periphyton models for all segments of the digital river network. Some values of the independent variables were already available for all segments of the digital river network including the variables obtained from FENZ and the predicted hydrological indices. However, only the measured data were available for some independent variables that were used to model periphyton Chla92. The variables were: EC, FineSed, Temp95 and the water quality measures Clar, Turb, NH4N (see Table 3 for details).

We used available predictions of Clar, Turb and NH4N made by Whitehead (2018). These predictions were trained on a national dataset of water quality data that comprised between 587 and 882 sites, depending on the variable (Whitehead, 2018). The predictions pertained to median values of all water quality variables (i.e., the same statistic used to define the water quality independent variables in this study) for the five-year period ending 2017. Whitehead (2018) used RF models and quantified model performance using the statistics described in section 4.4.3. These data represent the most up-to-date and spatially comprehensive estimates available for these water quality variables.

We fitted new RF models to the variables EC, FineSed and Temp95 measured at the periphyton monitoring sites. The RF models were fitted using a selection of potential independent variables that are available for all segments of the digital river network (FENZ; Leathwick *et al.*, 2010) and model performance was quantified using the statistics described in section 4.4.3. Where necessary, we transformed the response variable for each model to an approximately symmetric distribution to improve model performance.

We compared and evaluated the performance of the models against the criteria proposed by Moriasi et al. (2015) (Table 4). The fitted models were then used to make predictions for all 120,000 segments of the digital river network with order ≥ 3. Where variables had be transformed as part of the fitting process they were back-transformed to the original variable scale and, where necessary, were corrected for retransformation bias using the smearing estimate (Duan, 1983).

Table 4: Performance ratings for the models used to make spatial predictions of the independent variables in this study. The performance ratings are from Moriasi et al. *(2015).*

Performance Rating	R^2	NSE	PBIAS
Very good	$R^2 \geq 0.70$	NSE > 0.65	|PBIAS| <15
Good	$0.60 < R^2 \leq 0.70$	0.50 < NSE ≤ 0.65	15 ≤ |PBIAS| < 20
Satisfactory	$0.30 < R^2 \leq 0.60$	0.35 < NSE ≤ 0.50	20 ≤ |PBIAS| < 30
Unsatisfactory	$R^2 < 0.30$	NSE ≤ 0.35	|PBIAS| ≥ 30

4.7 Generate criteria for REC classes

4.7.1 OLS regression

Conceptually, the fitted OLS models were used to derive nutrient criteria for each Source-of-flow class as described in Section 2.3. Rather than inverting the models, an equivalent approach to deriving the criteria based on interpolation was used for each of the four nutrients in five stages.

1. For each river segment in each Source-of-flow class, predictions at the upper confidence limit were obtained from the fitted model for increments in a gradient in nutrient concentrations that reflected the observed range of the nutrient in our dataset.
2. Predictions for the upper confidence limit of four prediction intervals were obtained. The prediction intervals were consistent with the six levels of under-protection risk (50%, 30%, 20% 15%, 10% and 5%) based on Equations 2 and 3 (see Section 4.4.1). The derived predictions can be understood as the percentiles of the predicted response distribution shown in Figure 4.
4. For each threshold of interest, an associated nutrient concentration was interpolated from the paired predicted upper confidence limit - nutrient gradient data for the six levels of under-protection risk (50%, 30%, 20% 15%, 10% and 5%).
5. Finally, the mean of the nutrient concentrations derived at stage 4 over all segments in each Source-of-flow class were obtained for each combination of biomass threshold and level of under-protection risk. These mean concentrations are the nutrient criteria for the Source-of-flow classes corresponding to each biomass threshold and level of under-protection risk.

A feature of obtaining criteria for every Source-of-flow class is that the models are sometimes being used to make predictions that are outside of the range of the observations (i.e., the predictions represent extrapolation from the observations). This is because Source-of-flow classes encompass all rivers and streams in New Zealand, but the fitting dataset represents a restricted range of environments. The assumption underlying the derivation of the criteria therefore is that the fitted relationships (i.e., the rate of change in Chla92 with change in independent variables) in the unsampled environmental space are consistent with the environmental space that is represented by the monitoring sites.

It is not possible to test the validity of this assumption because of the lack of data. However, the representativeness of the monitoring sites of all rivers and streams in New Zealand was assessed to provide some insight into the parts of the environmental space that are poorly represented. A graphical comparison was used to gauge how well the monitoring sites represented environmental variation in stream and rivers at the national scale. Representativeness is the degree to which the distribution of monitored sites over the range of the independent variables used in the models matched the distribution of all river network segments over the range of the same environmental variable. Poor representativeness indicates potentially reduced reliability of the model predictions because certain sets of environmental conditions are not represented in the fitting data.

Histograms of the proportions of monitoring sites numbers over the ranges of the independent variables that were included in the OLS models were visually compared with histograms of the proportions of all segments over the same independent variables. Note that representativeness of monitoring sites is different from model bias, which is defined in Section 4.4.1.

For some combinations of nutrient (i.e., TN, DIN, TP, DRP), Source-of-flow class and level of under-protection risk (i.e., 50%, 30%, 20% 15%, 10% and 5%), we expected that the model predictions would not exceed the biomass threshold even at the maximum of the observed values of the nutrient in our dataset based on criteria derived by Snelder et al. (2019). In these cases, the result returned at stage 4 in the process described above (the interpolation of the nutrient concentration associated with the biomass threshold) was the maximum of the relevant observed nutrient value. Similarly, for some combinations of nutrient, Source-of-flow class and level of under-protection risk, the mean of the nutrient concentrations over all segments (i.e., the result returned at stage 5 in the process described above) was the maximum of the observed values. The interpretation of nutrient criteria that are equal to the maximum of the observed values is therefore that the biomass threshold is not expected to be exceeded even at the given (i.e., maximum observed) nutrient concentration.

4.7.2 Quantile regression

Criteria can be derived from fitted QR models by inverting the fitted models as described in Section 2.3. Conceptually, criteria for the four levels of under-protection risk (50%, 30%, 20% and 10%) could be derived from robust QR models of 0.5, 0.7, 0.8 and 0.9 quantiles. As for the criteria derivation based on OLS models, the mean of the criteria obtained for each segment in each Source-of-flow class could be used to represent the best estimate of the criteria for the class. However, in this study the QR models were not considered to be sufficiently robust for deriving criteria (see Section 5.4.2).

4.8 Comparison with existing criteria

At step seven we compared the derived criteria with those derived by Snelder *et al.* (2019). Snelder *et al.* (2019) derived criteria for the same Source-of-flow classes as this study, for two nutrient forms, TN and DRP and for two levels of under-protection risk,10% and 20%. The two sets of criteria were compared graphically using scatter plots.

4.9 Validation

The derived criteria were validated using the independent NRWQN dataset. The validation was performed by inverting the criteria to predict the biomass (as 92nd percentile of chlorophyll) at the NRWQN sites based on the Source-of-flow class and the observed nutrient concentration. Biomass was predicted for each site using each of the four nutrient forms and each of the four levels of under-protection risk, resulting in 24 sets of predictions. For each set of predictions, the predicted biomass was compared to the observed biomass graphically using scatter plots. Theoretically, 5%, 10%, 15%, 20%, 30% and 50% of the NRWQN sites were expected to have observed biomass that exceeded the predicted biomass when the predictions were made based on the corresponding levels of under-protection risk. We evaluated this expectation by calculating the proportion of sites for which observed biomass exceeded the predicted for the 24 sets of predictions.

4.10 National RF model of current periphyton state

We fitted a final RF model to the site values of Chla92. The purpose of this model was to produce the best possible predictions of current periphyton state from the observation dataset using a model that was not subject to some of the constraints associated with the OLS or QR model. Because this model was not constrained to have a single nutrient or to be a linear combination of the explanatory variables, it might explain more variation in Chla92 than the OLS or QR model. In turn, this might help to identify limitations associated with the OLS or QR models and, therefore, the criteria.

We aimed to use the fitted RF model to make predictions of current periphyton state for all segments of the digital river network (stream order ≥ 3). Therefore, the model had access to all available independent variables including all four nutrients (i.e., TN, DIN, TP and DRP from (Whitehead, 2018). In addition, we offered the modelling fitting process additional potential independent variables from the FENZ Geodatabase (Leathwick *et al.*, 2010), which described catchment characteristics (Table 5).

We used the fitted RF model to predict the current 92nd percentile chlorophyll for every segment of the digital network based on Equation 2 (see Section 4.4.1). We note that predictions will be made for some segments that might not be realised due to shading or fine grained and unstable substrates. The current 92nd percentile chlorophyll for any percentile of the probability distribution could be obtained from the model predictions. For example, the 80th and 90^{h} percentile biomass values (i.e., the biomass exceeded 20% and 10% of time) were obtained from Z values of 0.84 and 1.28, respectively.

Table 5. Additional potential independent variables offered to the national RF model of current Chla92 across river network.

Independent variable type	Variable	Abbreviation	Units
Geography & topography	Geographic coordinate East	Xcoord	Metres
	Geographic coordinate North	Ycoord	Metres
	Stream Order	StreamOrder	Unitless
	Catchment area	usCatArea	km2
	Segment mean elevation	segElev	m ASL
	Percentage of catchment occupied by lakes	usLake	%
	Mean catchment elevation	usElev	m ASL
	Mean catchment slope	usSlope	degrees
	Distance to the coast	DistToCoast	km
	Mean segment slope	SegSlope	degrees
	Distance to furthest headwater segment	DistToHead	km
Climate & flow	Mean segment June air temperature	segTmin	degrees C
	Mean segment January air temperature.	segTwarm	degrees C
	Mean catchment June air temperature	usTmin	degrees C
	Mean catchment January air temperature	usTwarm	degrees C
	Mean annual catchment rainfall	usRain	mm
	Mean catchment coefficient of variation of annual rainfall	usRainvar	mm/yr
	Mean catchment rain days > 10 mm	usRainDays10	days/mo
	Mean catchment rain days > 200 mm	usRainDays20	days/mo
	Mean catchment rain days > 100 mm	usRainDays100	days/mo
	Mean annual catchment potential evapotranspiration	usPET	mm/yr
	Estimated mean flow	MeanFlow	m3/s
Geology	Mean catchment induration (hardness) of regolith	usHard	Ordinal
	Mean catchment phosphorous content of regolith	usPhos	Ordinal
	Mean catchment particle size of regolith	usPsize	Ordinal
	Mean catchment calcium content of regolith	usCalc	Ordinal
Land cover	Proportion of catchment occupied by combination of high producing exotic grassland, short-rotation cropland, orchard, vineyard and other perennial crops (LCDB5 classes 40, 30, 33)	usIntensiveAg	%
	Proportion of catchment in low producing grassland (LCDB5 class 41)	usPastoralLight	%
	Proportion of catchment in native forest (LCDB5 class 69)	usNativeForest	%
	Proportion of catchment in built-up areas, urban parkland, surface mines, dumps and transport infrastructure (LCDB5 classes 1, 2, 6, 5)	usUrban	%
	Proportion of catchment in scrub and shrub cover (LCDB5 classes 50, 51, 52, 54, 55, 56, 58)	usScrub	%
	Proportion of catchment occupied by lake and pond, river and estuarine open water (LCDB5 classes 20, 21, 22)	usWetland	%
	Proportion of catchment in exotic forest (LCDB3 class 71)	usExoticForest	%
	Proportion of catchment occupied in bare or lightly-vegetated cover (LCDB5 classes 10, 12, 14, 15, 16)	usBare	%

5 Results

5.1 Patterns in periphyton biomass and relationships with potential independent variables

The complete regional council dataset included 326 sites with at least one biomass observation, 251 sites with >20 biomass observations and 218 sites with ≥30 biomass observations (Figure 7). Sites in the Auckland region had 8 observations or fewer. Only one site in the Otago region had >30 observations but 19 sites had >20 observations. The 251 sites with >20 biomass observations were retained in the dataset for modelling.

Of the 251 sites that were retained for modelling, two and three were located on segments of the digital network of stream order one and two, respectively. The remaining sites were reasonably evenly distributed across segments of order three to seven (Table 6).

Table 6. Distribution of periphyton monitoring sites retained for modelling across segments of different stream order.

Stream order	Number of sites
1	2
2	3
3	32
4	83
5	63
6	43
7	25

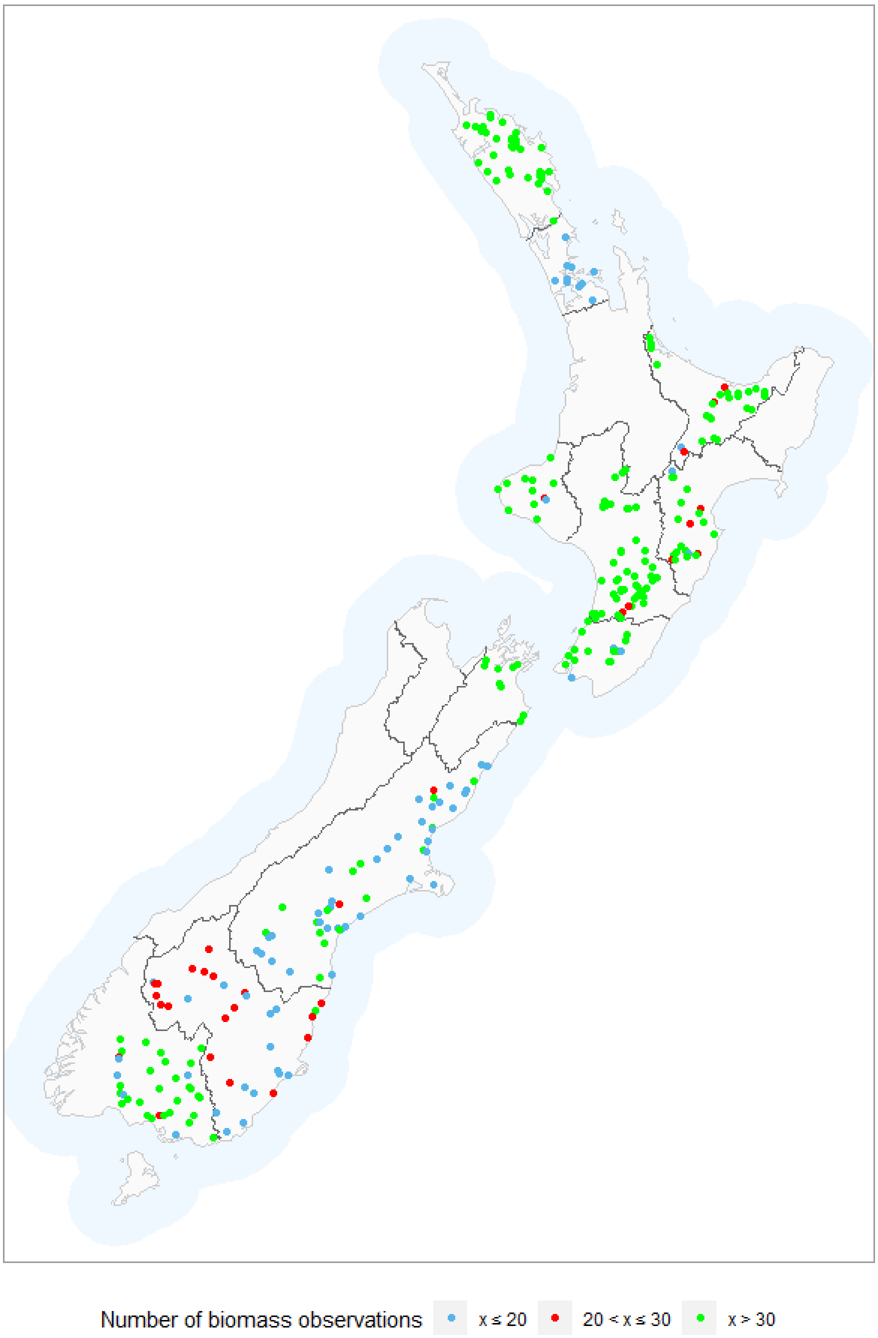

Figure 7. Map of the monitoring sites colour-coded according to number of biomass observations. Note that only sites with >20 observations were retained for modelling.

Mean Chla92 across all sites with >20 observations was 89 mg m^{-2} (Figure 8). Sixty sites had Chla92 >120 mg m^{-2} and 20 sites had Chla92 >200 mg m^{-2}. Site Chla92 was highly variable with no discernible geographic pattern (Figure 9). Site Chla92 >200 mg m^{-2} did not occur at any sites in Southland, Taranaki or Bay of Plenty but occurred in the other regions that were represented by data.

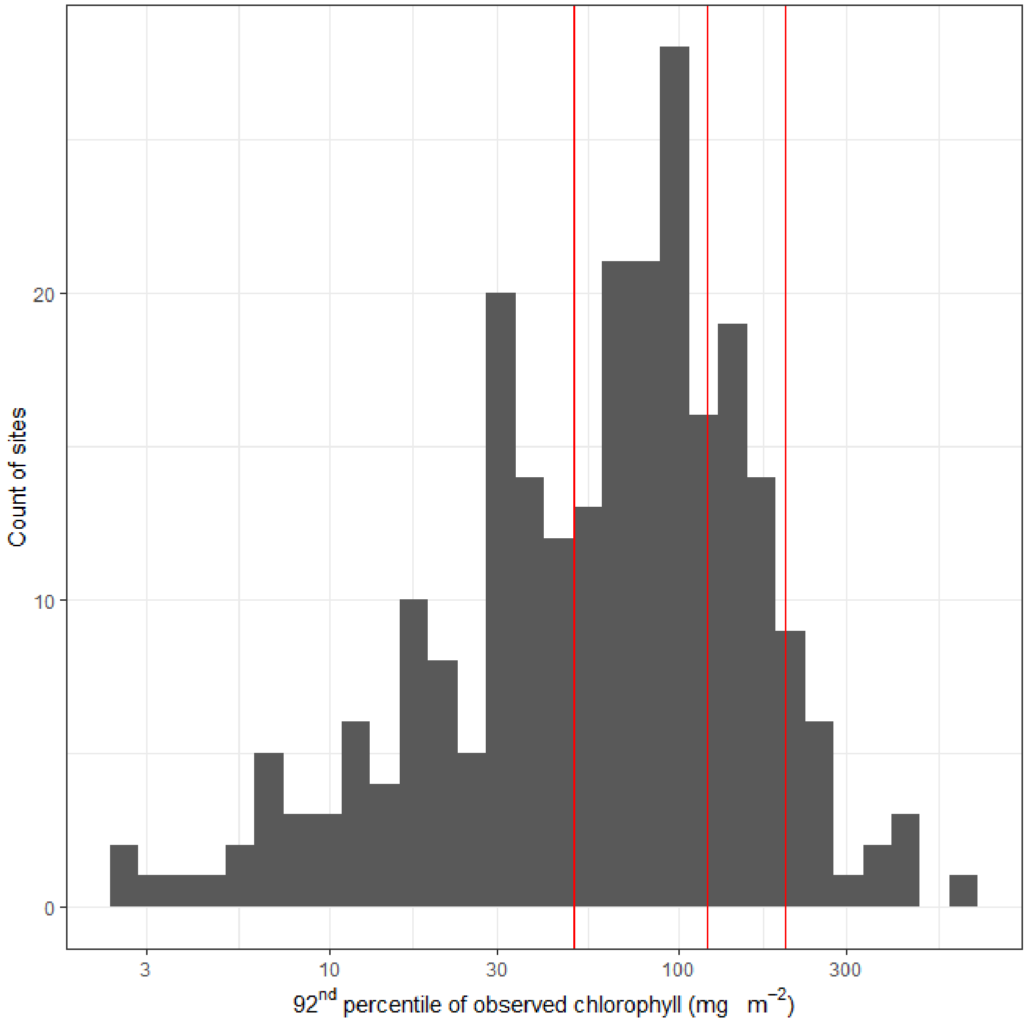

Figure 8. Distribution of 92nd percentile values of chlorophyll at the 251 monitoring sites with > 20 sampling occasions. The vertical red lines indicate biomasses of 50, 120 and 200 mg m^{-2}, which correspond to the thresholds for the NOF periphyton attribute B, C and D bands, respectively.

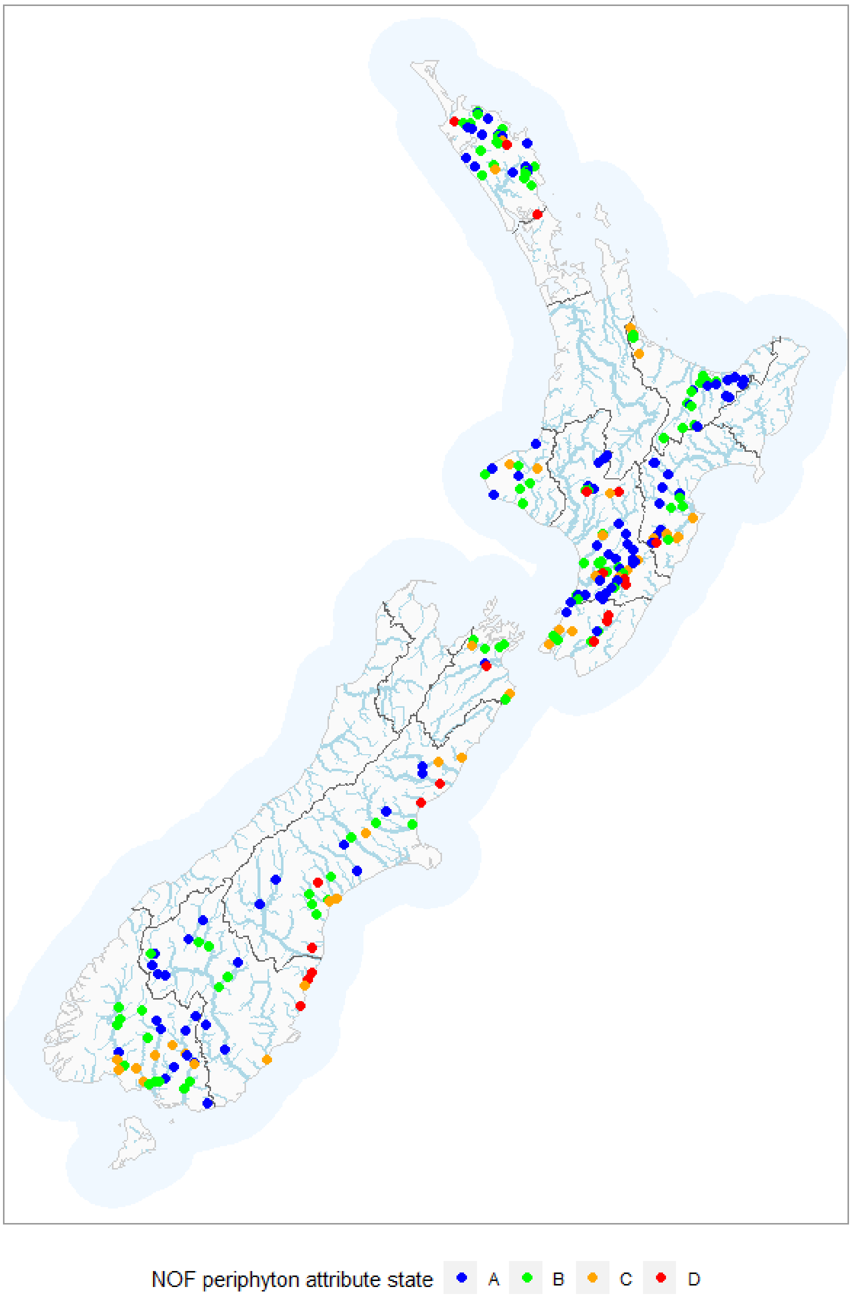

Figure 9. Map of the 251 monitoring sites with > 20 sampling occasions graded according to the NOF periphyton attribute state bands.

The site values of the Chla92 and mean chlorophyll (hereafter ChlaMean) were strongly correlated (Figure 10 A). Based on Snelder *et al.* (2014) we estimated the 92nd percentile of the distribution of chlorophyll observations using the exponential distribution:

$$\widehat{Chla}92 = -log\ (Pr) \times \mu \qquad \textit{Equation 5}$$

where *Pr* (0 ≤ *Pr* < 1) is the probability that biomass is exceeded given the mean (μ>0). $\widehat{Chla}92$ is estimated by setting *Pr* to 0.082 (because 100% - 92% = 8.3%). The consistency of the plot of $\widehat{Chla}92$ versus the observed 92nd percentile of the observations (i.e., Chla92), indicates that the observations within sites approximately followed an exponential distribution for all sites (Figure 10 B). Consistent with this, sites that were graded D tended to have high mean values of chlorophyll and multiple individual observations that exceeded >200 mg m^{-2} (Figure 11). This indicates that all sites had chlorophyll distributions that were approximately exponentially distributed and that the temporal dynamics of biomass are similar at all sites.

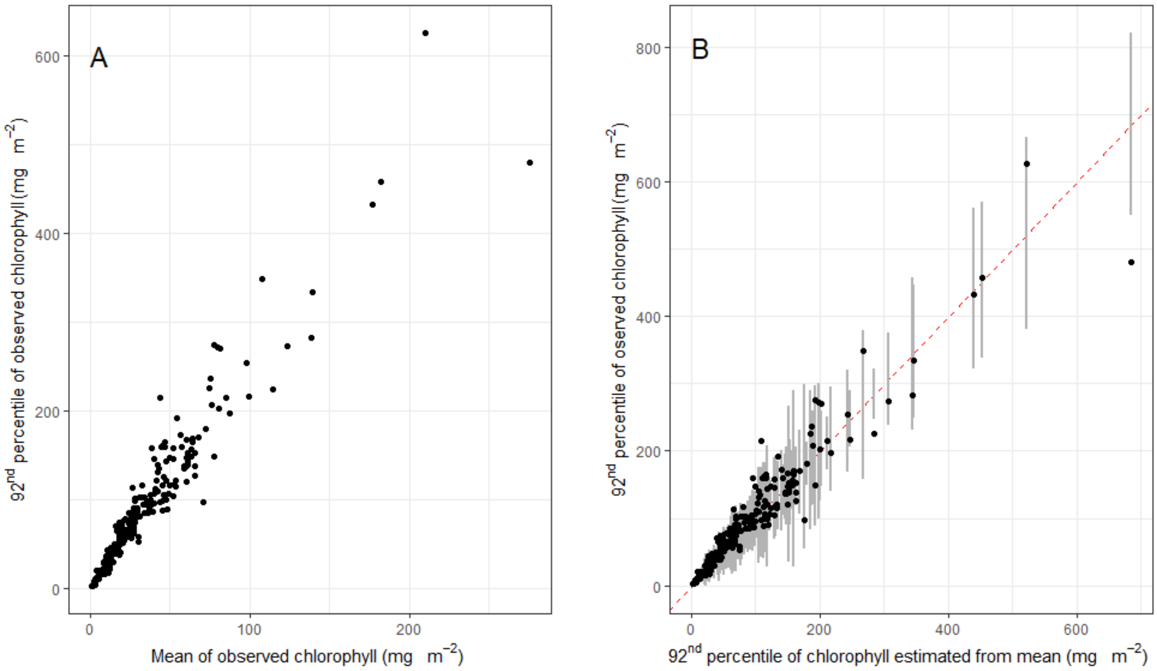

Figure 10. Relationships between the mean and 92nd percentile of chlorophyll at sites with >20 biomass observations. Panel A shows the relationship between the 92nd percentile of observed chlorophyll (Chla92) and the mean of the observed values (ChlaMean). Panel B shows the relationship between the 92nd percentile of observed chlorophyll and the same value calculated from the mean of the observed values based on the exponential distribution. The error bars indicate the 95% confidence interval for the estimated 92nd percentile values. The red dashed line is one to one.

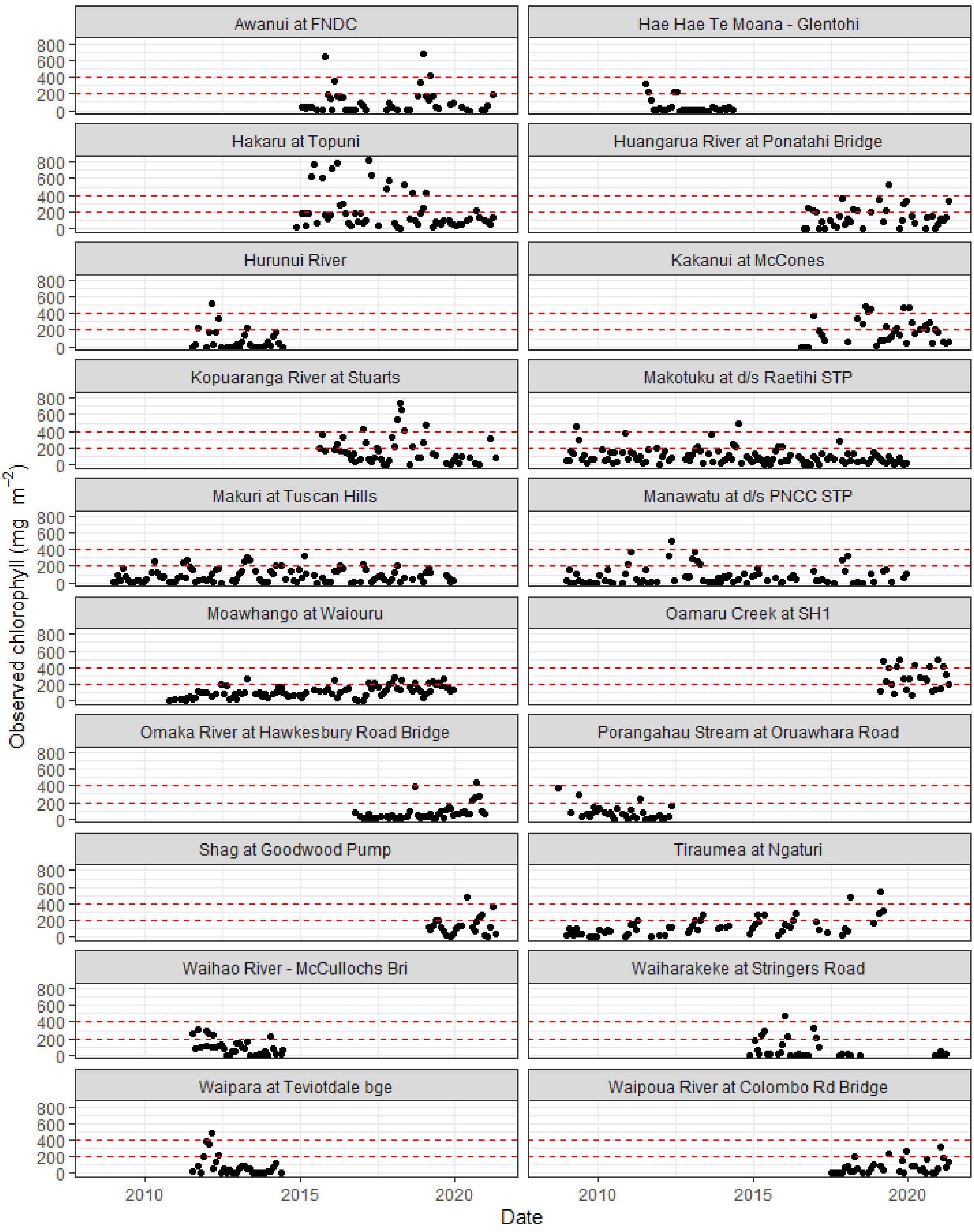

Figure 11. Time series showing biomass observations at the 20 sites at which Chla92 > 200 mg m^{-2}, indicated by the lower red dashed lines. The upper red dashed lines at 400 mg m^{-2} highlight the sites with highest biomass.

When viewed across all sites, relationships between Chla92 and site median values of the nutrients TN and DIN were wedge-shaped and quantile regressions at the 70% quantile were statistically significant ($p < 0.05$, Figure 12). This indicates a limiting relationship between biomass and nitrogen at the national scale but with other factors influencing the Chla92 response. Quantile regressions at the 70% quantile between Chla92 and site median values of the nutrient forms TP and DRP were not statistically significant (Figure 12).

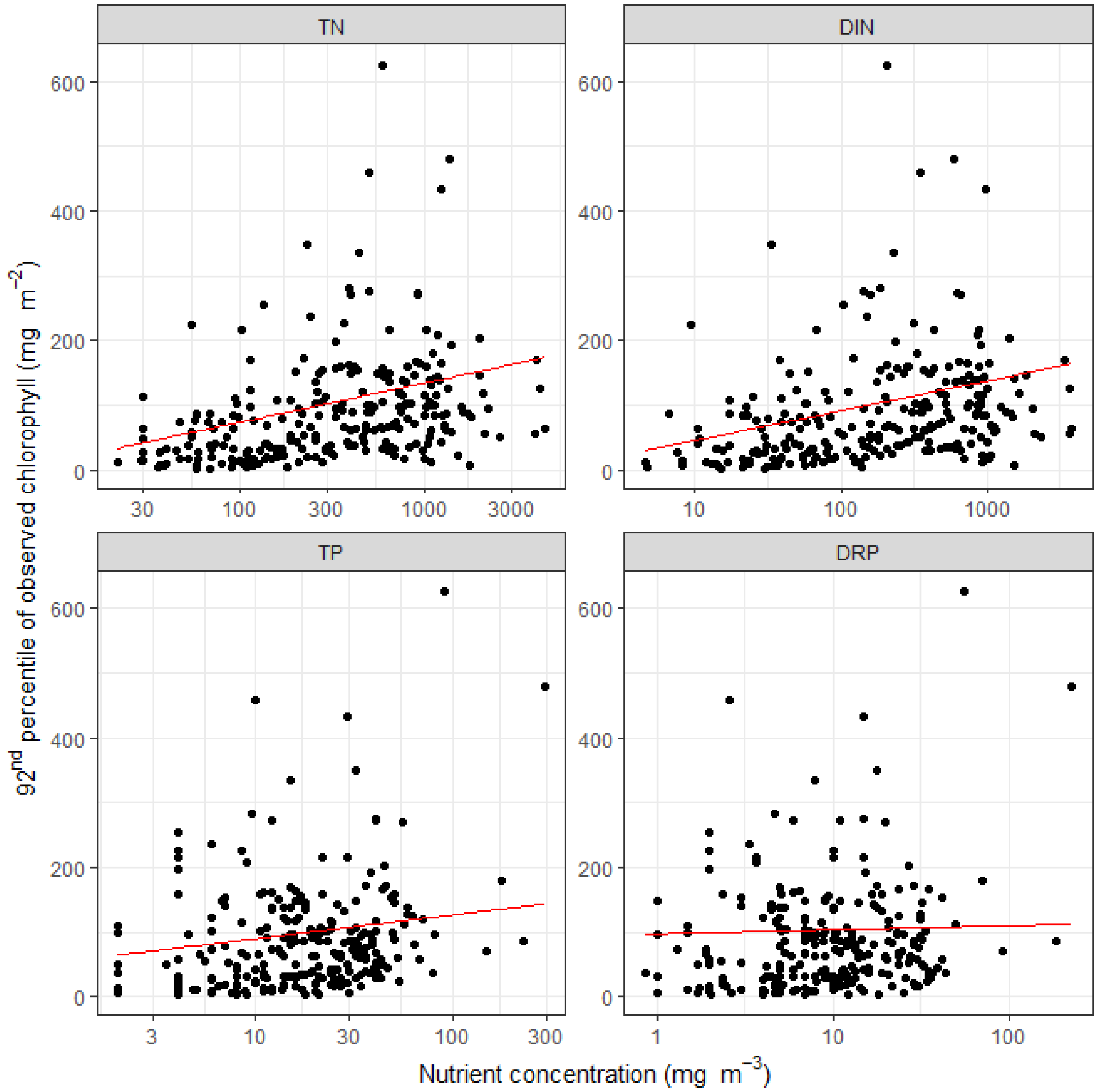

Figure 12. Relationships between Chla92 and nutrient concentrations at the 251 monitoring sites with > 20 sampling occasions. The red lines are quantile regressions fitted to the 70% quantiles. The slopes of these quantile regressions had p-values of 0.03, 0.02, 0.14 and 0.17 for TN, DIN, TP and DRP, respectively.

When viewed across all sites, relationships between the Chla92 and site median values of EC, Temp95, and Clar were wedge-shaped and quantile regressions at the 70% quantile were statistically significant (Figure 13). This indicates a limiting relationship between Chla92 and electrical conductivity, water temperature and water clarity at the national scale but with other factors influencing the Chla92 response at each site. Quantile regressions at the 70% quantile between the Chla92 and site median values of Turb, SolarRadDec and FineSed were not statistically significant (Figure 13).

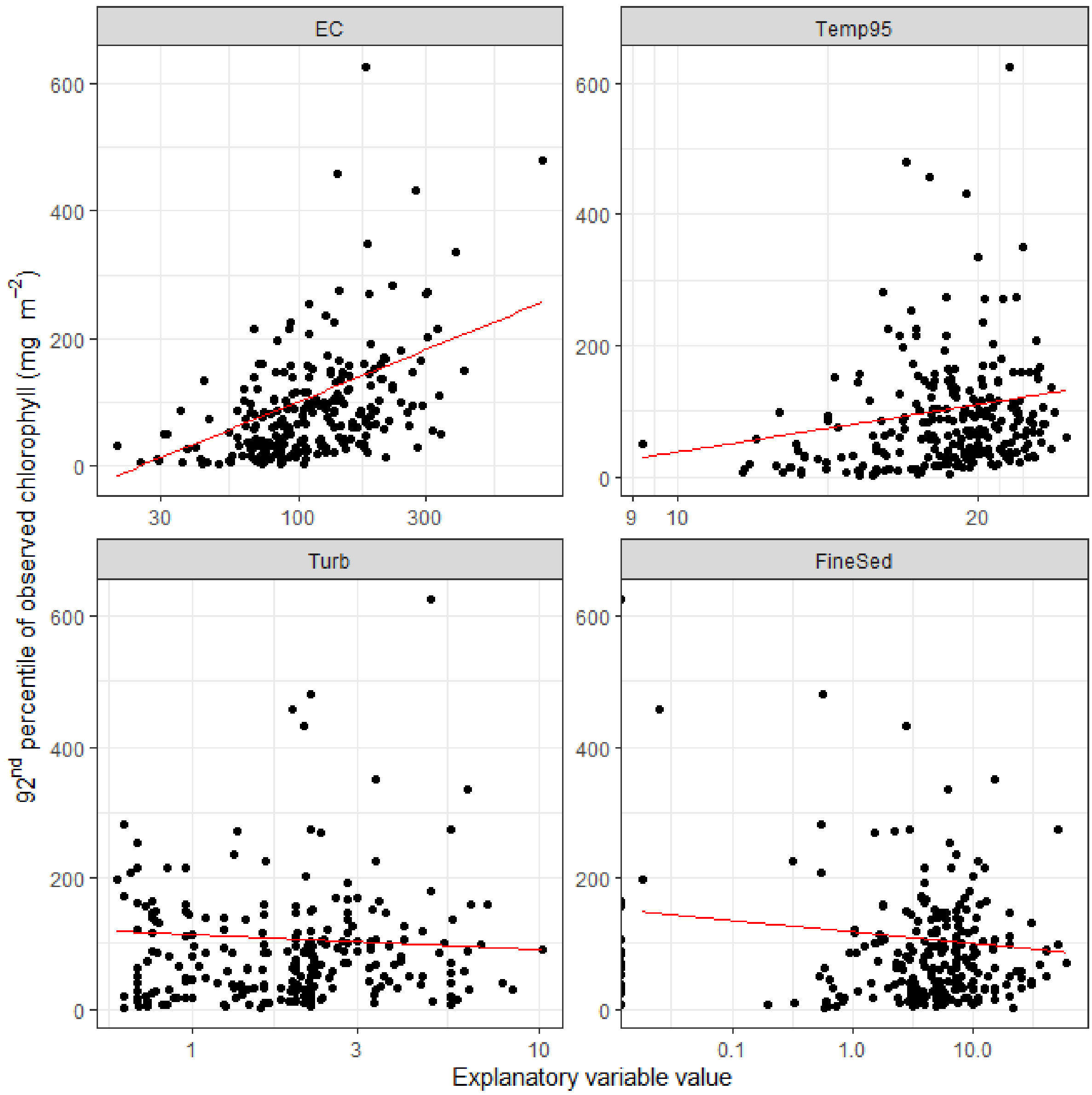

Figure 13. Relationships between Chla92 and potential independent variables at the 251 monitoring sites with > 20 sampling occasions. The red lines are quantile regressions fitted to the 70% quantiles. The slopes of these quantile regressions had p-values of <0.001,0.01, 0.68 and 0.65 for EC, Temp95, Turb, and FineSed, respectively. Note that the variables shown were included in at least one of the models that were chosen for derivation of nutrient criteria.

When viewed across all sites, there were wedge-shaped relationships between Chla92 and two hydrological indices that were derived from daily flows pertaining to the observation period, MALF30_cat and Reversals (Figure 14). Quantile regressions of Chla92 against these two indices at the 70% quantile were statistically significant. This indicates a limiting relationship between Chla92 and these hydrological characteristics at the national scale but with other factors influencing the Chla92 response at each site. Quantile regressions at the 70% quantile between the Chla92 and site median values of FRE3 and nNeg were not statistically significant (Figure 14).

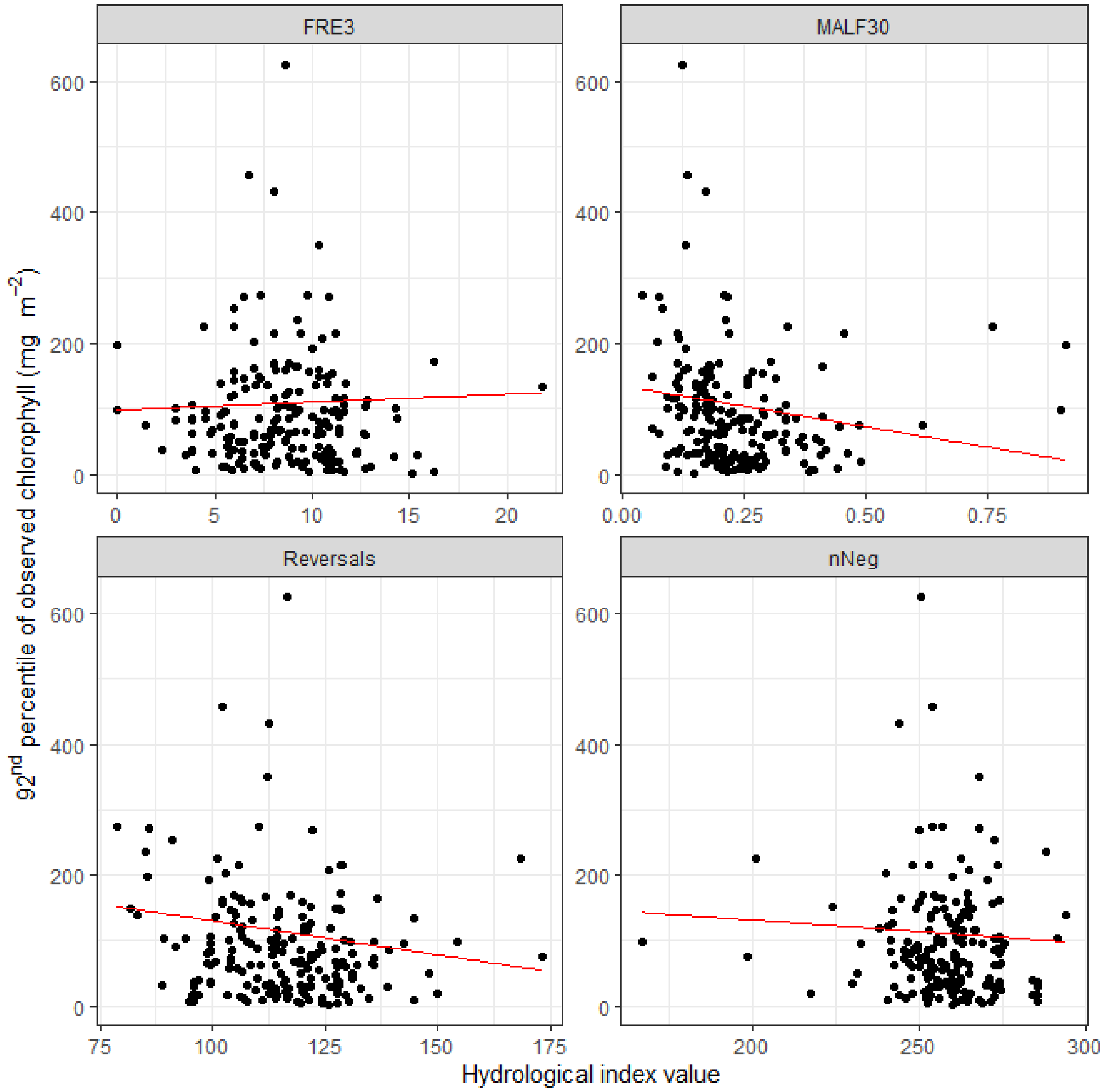

Figure 14. Relationships between Chla92 and hydrological indices that were potential independent variables at the 251 monitoring sites with > 20 sampling occasions. The red lines are quantile regressions fitted to the 70% quantiles. The slopes of these quantile regressions had p-values of 0.58, <0.001, 0.09, and 0.56 for FRE3, MALF30, Reversals, and Nneg, respectively. Note that the variables shown were included in at least one of the models that were chosen for derivation of nutrient criteria.

5.2 Potential independent variables

The nutrients TP and DRP were strongly correlated (0.94) as were TN and DIN (0.98) (Figure 15). However, the nitrogen forms (TN and DIN) were only weakly correlated (between 0.18 and 0.35) with the phosphorus forms (TP and DRP)). Because models were going to be fitted using only one nutrient as a potential independent variable, all nutrients were retained.

From initial potential independent variables (Table 3), we retained electrical conductivity (EC) because it had low correlation (≤ |0.6|) with all other potential independent variables including the four nutrients (Figure 15). We retained FineSed and Shade, which had low correlations (≤ |0.4|) with all other independent variables. Ammoniacal nitrogen (NH4N) had low correlation (≤ |0.4|) with other potential independent variables including the four nutrients and was retained. Because they were correlated > |0.8|, we retained SolarRadDec and discarded SolarRadJune. SolarRadDec had low correlation (≤ |0.6|) with all other potential independent variables. Visual water clarity (Clar) and Turbidity were only weakly correlated with each other and, because they had low correlation (≤ |0.5|) with other potential independent variables, were retained. Water temperature (Temp95) had low correlation (≤ |0.5|) with other potential independent variables and was retained.

We retained the hydrological indices nNeg and Reversals because they had low correlation (< |0.5|) with any other potential independent variables. Because they were correlated > |0.8|, we retained FRE3 and discarded FRE2, and FRE4 (Figure 15). Because they were correlated > |0.8|, we retained sdQ and discarded Max7, Max30, MALF7, and MALF30. The retained potential independent variables offered to the models are shown in Table 7.

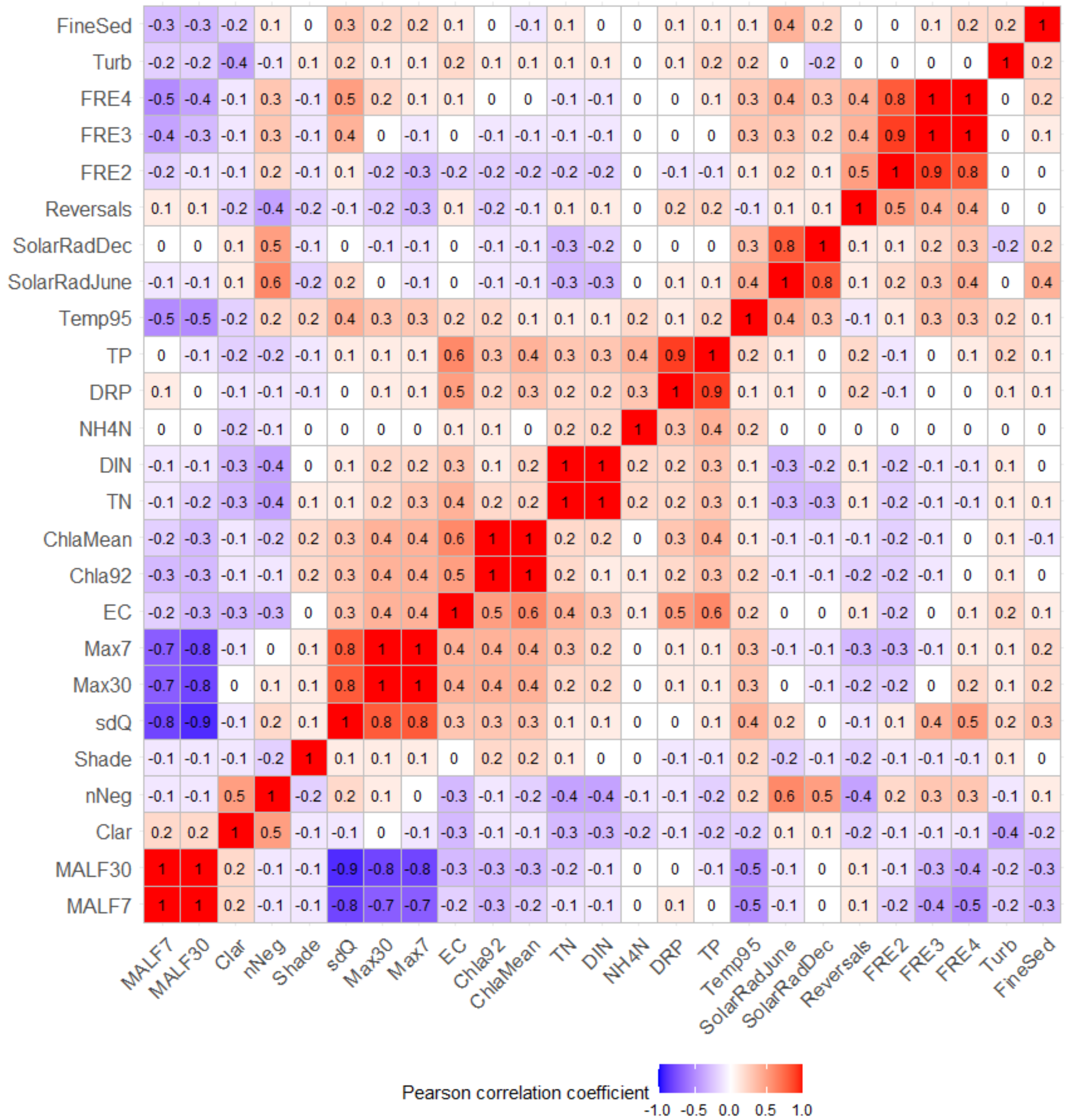

Figure 15. Pearson's correlation coefficients for all pairs of the preliminary selection of independent variables. The dependent variables (ChlaMean, Chla92) are also included, for reference. Each cell in the matrix represents the correlation between a pair of independent variables (shown on the x and y-axes). The colour indicates the strength and direction of the correlation. The variables are arranged in the matrix into groups with high correlation.

Table 7. Independent variables retained for model fitting.

Variable	Abbreviation	Transformation
Nutrient	TN, DIN, TP, DRP	Log (base 10)
Electrical conductivity	EC	
95th percentile of water temperature	Temp95	
Visual clarity of the water column	Clar	
Turbidity of the water column	Turb	
Ammoniacal nitrogen	NH4N	
Proportion of fine substrate	FineSed	
Proportion shade	Shade	
Solar radiation	SolarRadDec	
Number of events per year that exceeded three times the long-term median flow	FRE3	
Number of negative differences in flow between days	nNeg	
Number of hydrologic reversals	Reversals	
Standard deviation of daily flows divided by the mean flow	sdQ	

5.3 Influence of the alternative sources of hydrological indices

For a set of 192 sites that were common to all three datasets (see Section 4.3), the hydrological indices derived from the daily flow records produced the highest R^2 values for the fitted OLS models for all nutrients (Table 8). For TP and DRP the models fitted to the indices derived from the flow record pertaining to the period of periphyton observations had higher R^2 values than the indices derived from the full flow record for each site.

Lower R^2 values for models that were fitted using the modelled hydrological indices were always significantly different to (lower than) models fitted to the indices pertaining to the observation period hydrological indices and the full duration dataset at α=0.05 (Table 9). The models fitted to the modelled hydrological indices were not always significantly different from (lower than) models fitted using indices derived from the full hydrological record at α=0.05 (Table 9).

Table 8. Comparison of R^2 values for OLS models explaining between site periphyton Chla92 as function of selected variables and the nutrients as indicated.

Source of hydrological indices	TN	DIN	TP	DRP
Full record	0.40	0.39	0.35	0.35
Observation record	0.40	0.39	0.38	0.36
Modelled	0.35	0.35	0.33	0.33

Table 9. Significance of differences in variation explained between models fitted using different sources of hydrological indices. Blank cells indicate invalid self-comparisons.

Nutrient	Source data	Full	Observation
TN	Observation	0.40	
	Modelled	0.011	0.003
DIN	Observation	0.18	
	Modelled	0.03	0.007
TP	Observation	0.001	
	Modelled	0.15	0.003
DRP	Observation	0.02	
	Modelled	0.15	0.003

Differences in the performance of the models fitted to the three alternative sources of the hydrological indices as quantified by model uncertainty were small. For example, for TN the model fitted to the indices pertaining to the observation period hydrological indices had a significantly higher R^2 value than the model fitted to modelled hydrological indices (R^2 of 0.40 and 0.35, respectively, Table 8, Table 9). However, the performance differences translated into only minor differences in the confidence in predictions made using the models. For example, the width of the 95% confidence intervals for a site with estimated Chla92 of 90 mg m^{-2} differed very little between models (Table 10).

We judged that differences in model performance achieved with the three alternative sources of hydrological indices were of little to no practical significance. Therefore, the modelled indices were used in final model development so that all 251 sites with >20 periphyton observations could be included.

Table 10. Differences in the width of the 95% confidence intervals for models fitted using different sources of hydrological indices The values shown in each cell are the lower and upper 95% confidence intervals for a site with estimated Chla92 of 90 mg m^{-2}.

Nutrient	Source data		
	Full	Observation	Modelled
TN	16 - 304	16 - 300	15 - 307
DIN	15 - 306	16 - 302	15 - 306
TP	14 - 316	15 - 307	15 - 314
DRP	14 - 316	15 - 307	15 - 314

5.4 Periphyton biomass models

5.4.1 Ordinary least squares regression models

The LOOCV R^2 values for the fitted OLS models ranged between 0.38 and 0.27 (Table 11). Quantile-quantile (Q-Q) plots indicate that the distributions of the regression residuals of the four models were reasonably consistent with the theoretical normal distribution (Figure 16). This indicates that the residual values from a linear regression are reasonably normally distributed. The model may therefore be used to predict the entire probability distribution based on the theoretical normal distribution. This also indicates that the fourth root transformation applied to the model response (Chla92) was appropriate.

Table 11. Performance of the OLS models of periphyton biomass pertaining to each nutrient. Performance was assessed using LOOCV.

Nutrient	N	R^2	PBIAS	RMSD
TN	251	0.38	0.02	0.54
DIN	251	0.38	0.02	0.54
TP	251	0.34	0.00	0.56
DRP	251	0.27	0.07	0.59

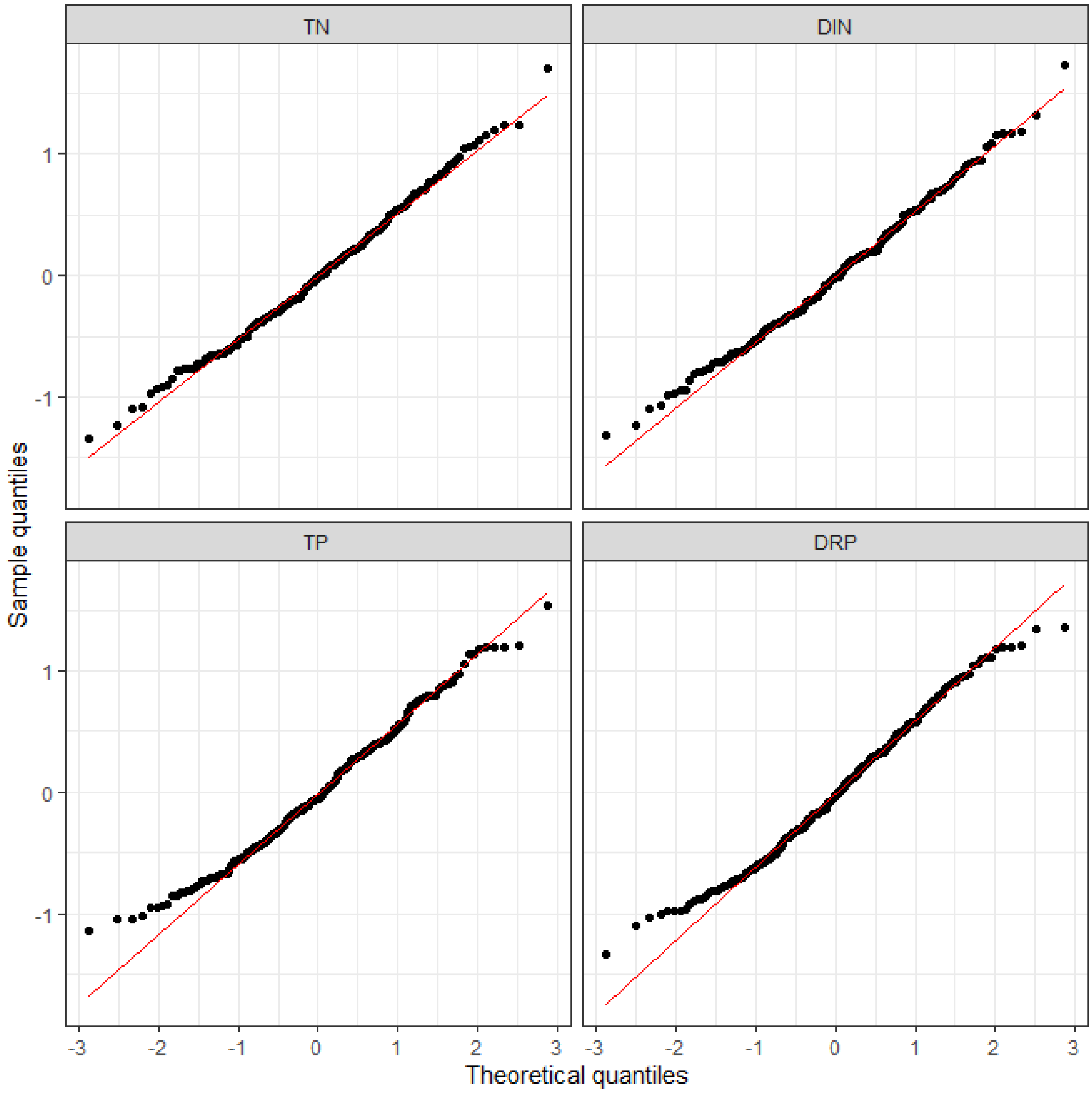

Figure 16. Quantile-quantile plots comparing the OLS regression residuals of the four OLS models to normal distributions.

All variables fitted in all four OLS models had VIF values ≤1.9, indicating low collinearity between independent variables (Table 12). The highest VIF values for the nutrients fitted in each of the four models was 1.6. These low VIF values indicate that the models are robust for inferring the direction of the relationships between the independent variables and biomass and for making predictions about how biomass changes with changing nutrient concentration.

The signs of the coefficients fitted to the OLS model independent variables were consistent with the conceptual model in Section 2.2 (Table 12). For example, for all four models the response (Chla92) was positively associated with nutrient concentration and temperature (Temp95), which represent resource supply (Figure 2). It is noted also that the structure of the models (i.e., the variables that were retained in the models) and the signs of the coefficients, were consistent with previous studies using different datasets and different descriptors of biomass abundance (Snelder *et al.*, 2019, 2014). All four models included the ordered categorical variable Shade with negative coefficients. Because the ordering applied to this variable was unshaded < shaded, the negative coefficient indicates that, all other variables being equal, biomass is higher at unshaded than shaded sites. This is consistent with the conceptual model because it indicates that biomass increases with increasing resource supply.

For all four models, the response was negatively associated with frequency of high flows (FRE3) and changes in flow (Reversals), which represent disturbance (Figure 2). All of the models also included flow variability (sdQ) with positive coefficients (Figure 2). Large values of sdQ indicate sites with large flow ranges and therefore prolonged periods of low flow (without disturbance) that allow for biomass accrual. Note that sdQ was strongly correlated with MALF7 and MALF30 (Figure 15) which are direct measures of the intensity of low flow periods. Three of the four models included electrical conductivity (EC) with a positive coefficient. We assume that EC represents aspects of resource supply; it is weakly correlated with the nitrogen and phosphorus concentrations and may represent additional micro-nutrients.

The models indicated that Chla92 increases at a high rate with increasing nutrient concentration at low biomass, but the rate decreases as concentrations and biomass increase (Figure 17). A Chla92 ceiling (i.e., only very small increases in Chla92 with increase in nutrient concentration) occurred at “saturating” nutrient concentrations of approximately 1000 mg m^{-3} for the TN and DIN models, approximately 50 mg m^{-3} for the TP and approximately 25 mg m^{-3} for the DRP model (assessed subjectively from Figure 17 taking into account the distribution of the data shown by the rug on the nutrient (x) axis and the predicted response). Because the nutrient term in all models is the log transformed concentration, the models predict that biomass continues to increase with increasing nutrient concentration above the ceiling but at gradually reducing rates (Figure 17).

The Chla92 ceiling varied appreciably between Source-of-flow classes (Figure 18). For example, at the saturating DIN concentration of approximately 1000 mg m^{-3} and unshaded locations, the modelled response at the 80th percentile of the predicted probability distribution varied between 50 mg m^{-2} (for the CX/GM Source of flow class) to 220 mg m^{-2} (for the WD/L Source of flow class; Figure 18). Similarly, at the saturating TP concentration of approximately 50 mg m^{-3} the modelled response at the 80th percentile of the predicted probability distribution varied between 71 mg m^{-2} (for the CX/GM Source of flow class) to 205 mg m^{-2} (for the WD/L Source of flow class; Figure 18).

The ceiling in Chla92 occurred at around the 80th percentile of the distribution of the TN and DIN values and the 90th percentile of the TP and DRP values (rug plot shown in Figure 17). The ceiling occurred at between the 70th and 90th percentile of the biomass observations, depending on the Source-of-flow class. Therefore, the modelled relationship between Chla92 and nutrient concentrations was well informed by the data over the steep initial response to increasing nutrient concentration up to the ceiling.

The models indicated that the probability distribution about the mean prediction is wide (Figure 17). For example, for a TN concentration of 1000 mg m^{-3}, the mean of the predicted response is 90 mg chlorophyll m^{-2}, and the 70^{th}, 80^{th} and 90^{th} percentiles of the distribution were 128, 156 and 202 mg chlorophyll m^{-2}, respectively.

Table 12. Fitted coefficients for OLS regression models pertaining to each nutrient variable. For each independent variable the first values are the fitted coefficient (including its sign) and the second value in parentheses is the VIF-value. NA indicates the independent variable was not included in the model.

Nutrient	Intercept	log10(Nutrient)	Temp95	FRE3	Shade	Turb	EC	Reversals	sdQ	nNeg	FineSed
TN	2.34	0.33 (1.4)	0.03 (1.4)	-0.02 (1.9)	-0.16 (1.1)	-0.03 (1.1)	0 (1.4)	-0.01 (1.5)	1.03 (1.7)	NA	NA
DIN	2.42	0.25 (1.2)	0.03 (1.4)	-0.02 (1.9)	-0.17 (1.1)	-0.03 (1.1)	0 (1.4)	-0.01 (1.5)	1.19 (1.6)	NA	NA
TP	5.44	0.21 (1.6)	0.04 (1.5)	-0.02 (2.2)	-0.15 (1.2)	-0.04 (1.1)	0 (1.7)	-0.02 (2)	1.52 (1.7)	-0.01 (1.9)	NA
DRP	8.09	0.32 (1.2)	0.05 (1.4)	-0.03 (2.1)	-0.12 (1.2)	NA	NA	-0.02 (2.1)	2.68 (1.5)	-0.02 (1.7)	-0.01 (1.1)

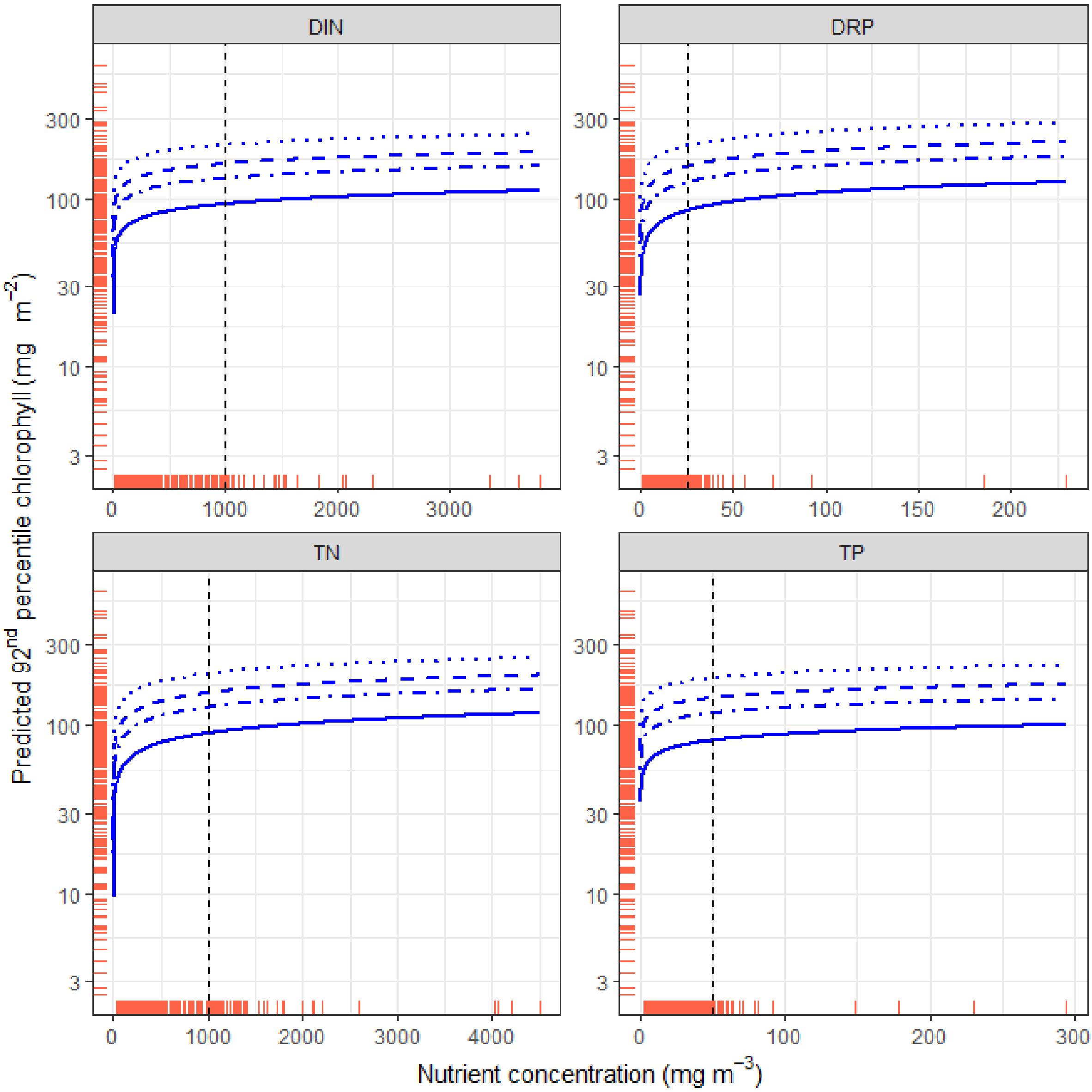

Figure 17. Predicted 92nd percentile chlorophyll based on the OLS models as a function of nutrient concentration for the four nutrients. The predictions represent a site having the mean value of each predictor from the fitting data set. The lower (solid) blue line represents the mean of probability distribution. The successive blue lines indicate the 70th, 80th and 90th percentiles of the predicted probability distribution. The red "rug" indicates the data density on both axes. The black vertical dashed lines indicate the approximate saturating concentrations, which were assessed subjectively taking into account the distribution of the data shown by the rug on the nutrient (x) axis and the predicted response.

Predictions made using the OLS models for each REC Source-of-flow class were broadly consistent with the conceptual model and expectations for each class (Figure 18). Predictions based on all four OLS models (i.e., including TN, DIN, TP and DRP, respectively), indicated that classes with low resource supply (i.e., low temperature) and high hydrological disturbance such as CX/GM, CX/M, CX/H and CX/L had low biomass compared to classes with high resource supply and low disturbance such as WD/L and WD/Lk (Figure 18). In addition, within

Source-of-flow classes, for a given nutrient concentration, predicted biomass was higher when shade status was unshaded compared to shaded (Figure 19).

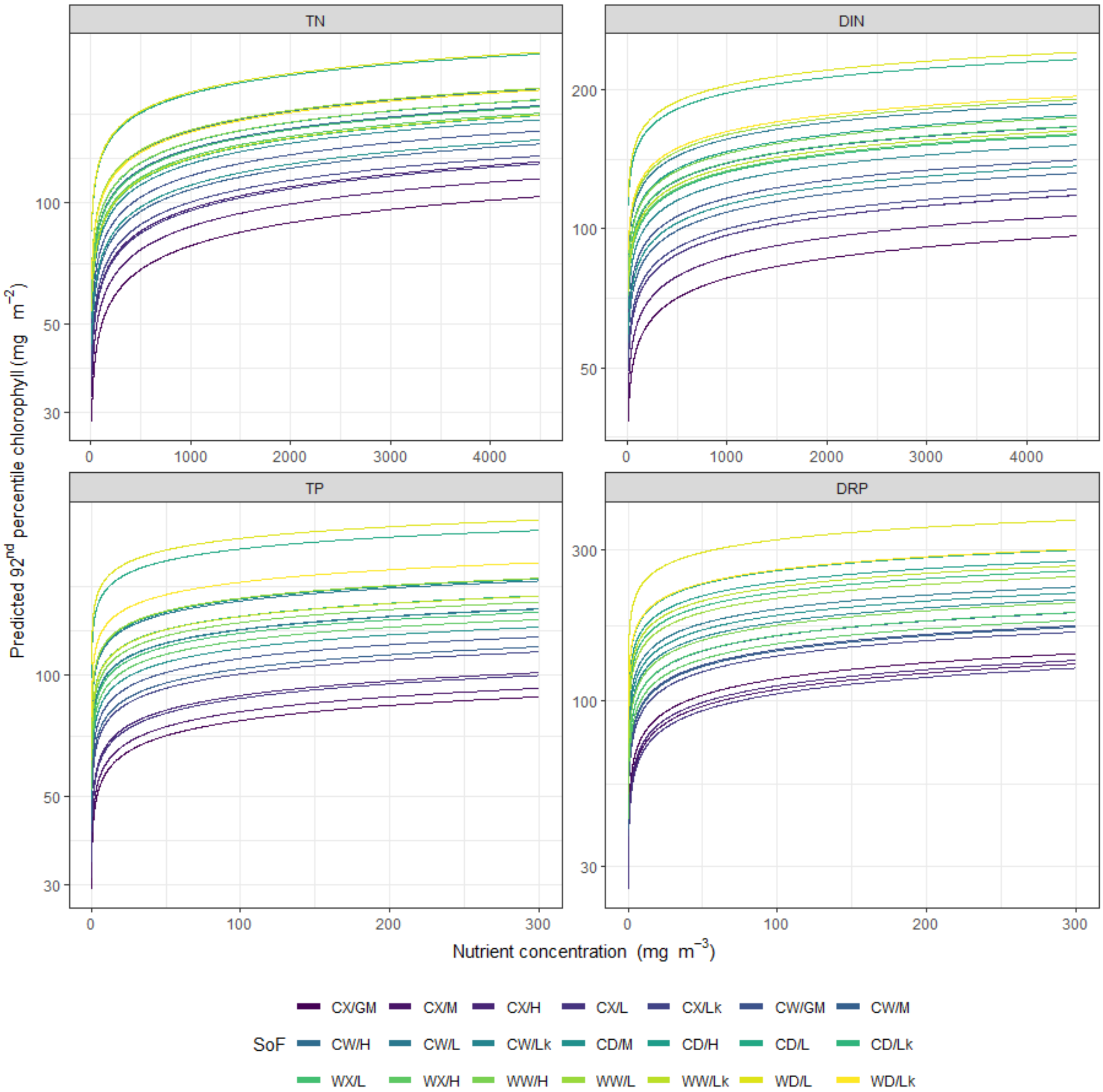

Figure 18. Predicted 92nd percentile chlorophyll at the 80th percentile of the probability distribution as a function of nutrient concentration based on the OLS models for the four nutrients and all Source-of-flow classes. The colour scheme used for REC Source-of-flow indicates an approximate gradient in resources and disturbance as defined by the conceptual model with the darker blue colour indicating low resources and high disturbance and the yellow colours indicating the opposite.

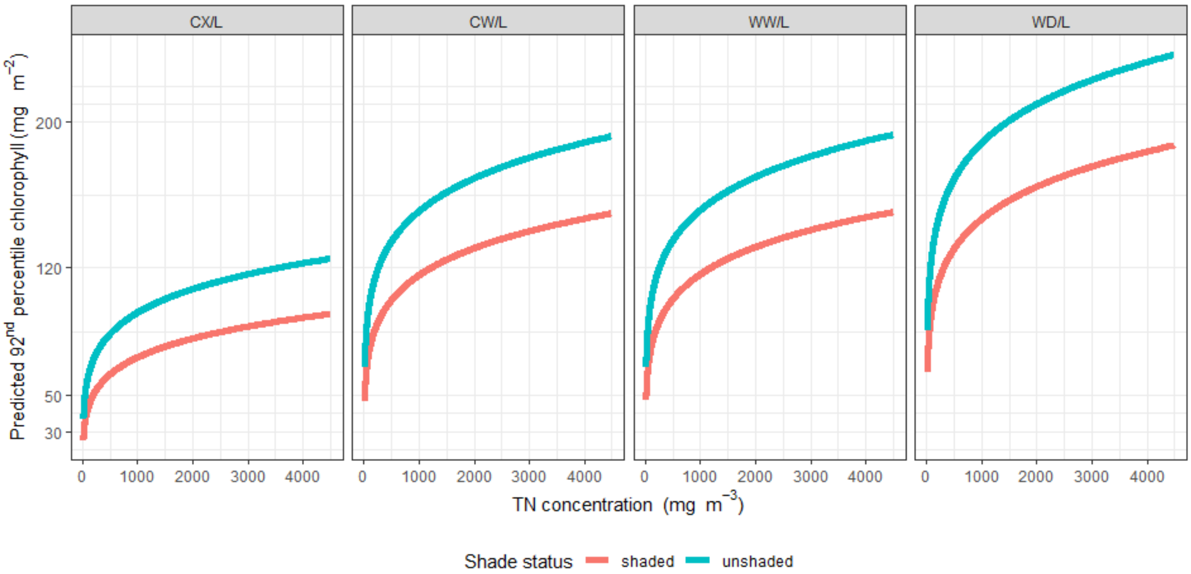

Figure 19. Predicted 92nd percentile chlorophyll at the 80th percentile of the probability distribution for two levels of the shade status variable. These predictions were based on the OLS model that incorporates TN as the nutrient and are for the four Source-of-flow classes indicated.

The OLS models generally underpredicted the highest Chla92 values (i.e., high biomass values had large regression residuals; Figure 20). The underpredicted sites were not associated with low numbers of biomass observations (Figure 20) nor were they associated with any of the incompletely measured site characteristics (Figure 21). This indicates that the independent variables for which we had incomplete data do not explain differences in Chla92 between sites (see Section 3.2.5).

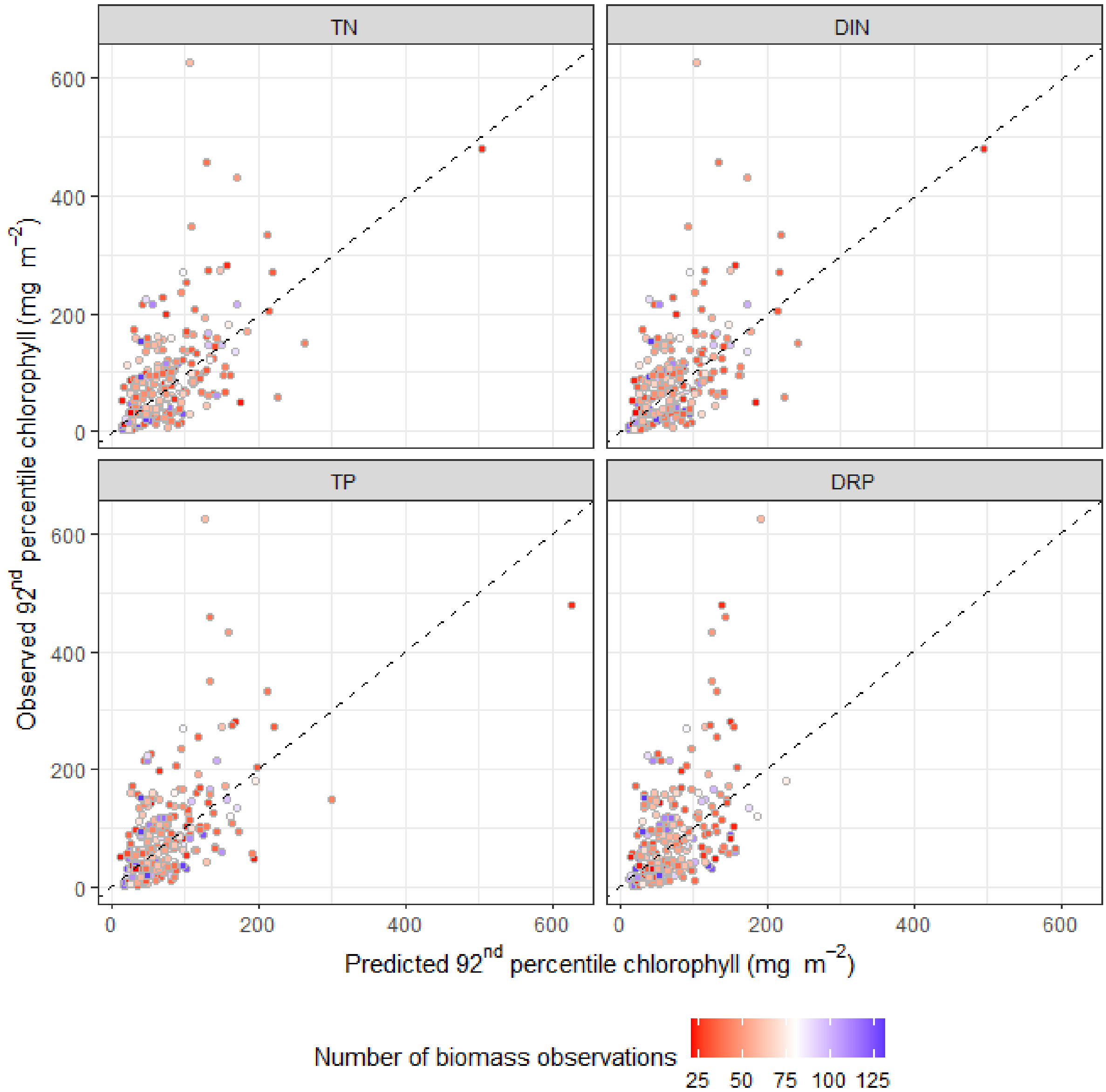

Figure 20. Observed versus predicted values for the OLS regression models. Each point represents a site, and the colour indicates the number of biomass observations. The black dashed line is one to one.

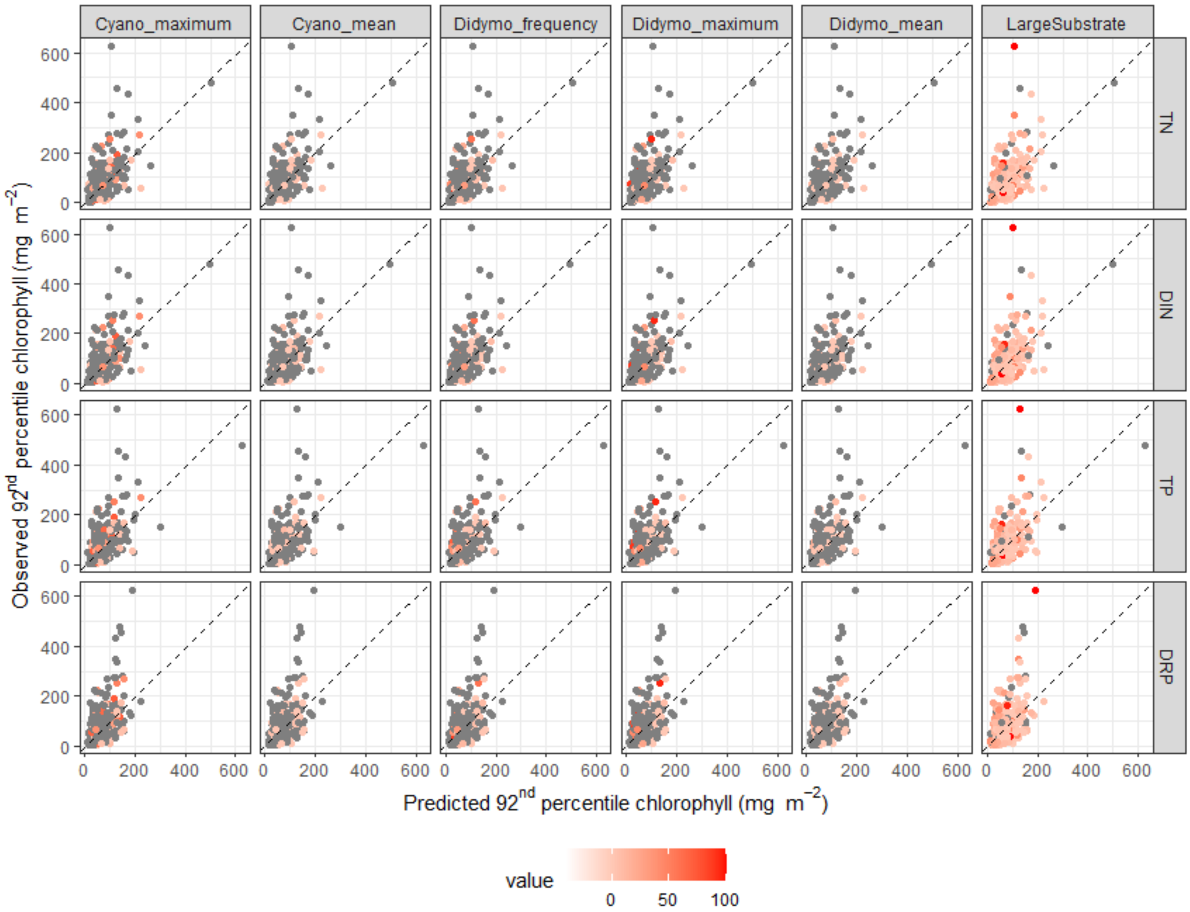

Figure 21. Observed versus predicted values for the OLS regression models showing incompletely measured site characteristics. Each point represents a site, and the colour indicates value of five additional characteristics that were not available for all sites. Grey points indicate the characteristic was not available for the site. The black dashed line is one to one.

5.4.2 Quantile regression models

All or most independent variables that were included in the OLS models (Table 12) tended to be significant model terms for the quantile regression models at the 50% quantile (Table 13). This was expected because the 50% quantile was close to the mean response modelled by the OLS regression and all independent variables offered to the QR were significant terms in the OLS model. However, for the higher quantile models, fewer independent variables were included as significant model terms (Table 13). For example, for the DIN models, four independent variables were included as significant model terms (at α=0.05) in the 50% quantile model but only two variables (Nutrient and EC) were included in the 70% and 80% quantile models and no variables were significant in the 90% quantile model.

Nutrient concentration was included as a significant model term in all TN models and all but the 90% quantile model for DIN. However, the only other model that included nutrient concentration as a significant model term was the 50% quantile TP model. Because inclusion of nutrient concentration in the model is a requirement for using the model to derive criteria, non-inclusion precluded further use of many of the quantile regression models.

All quantile regression models predicted increasing biomass with increasing nutrient concentration which is consistent with the conceptual model (Figure 22). However, the predicted responses were not very consistent with our expectations with respect to differences between Source-of-flow classes. For classes with lower resources (i.e., low light and temperature) and higher disturbance (e.g., high FRE3 or sdQ) the 50% quantile models generally exhibited lower biomass for a given nutrient concentration (Figure 22). But there was a less obvious pattern for classes with higher resources with some high resource, low disturbance classes having low biomass and vice versa. The modelled patterns in biomass across Source-of-flow classes became more inconsistent with expectations for the higher quantile models (Figure 23). For example, for the 80% quantile TP model, biomass decreased with increasing nutrient concentration, which is the opposite of the expectation based on the conceptual model (Figure 2).

The conclusion from the inspection of the quantile regression models is that they were not fit for purpose. These results are likely because fitting an accurate model to the more extreme percentiles (70%, 80% and 90%) requires a larger sample size than the available dataset. In addition, all other things being equal, dataset size requirements increase with the number of significant fitted terms variables. Therefore, the limitations of the size of the dataset were exacerbated by the multiple variables that are associated with the biomass response, as evidenced by the OLS models.

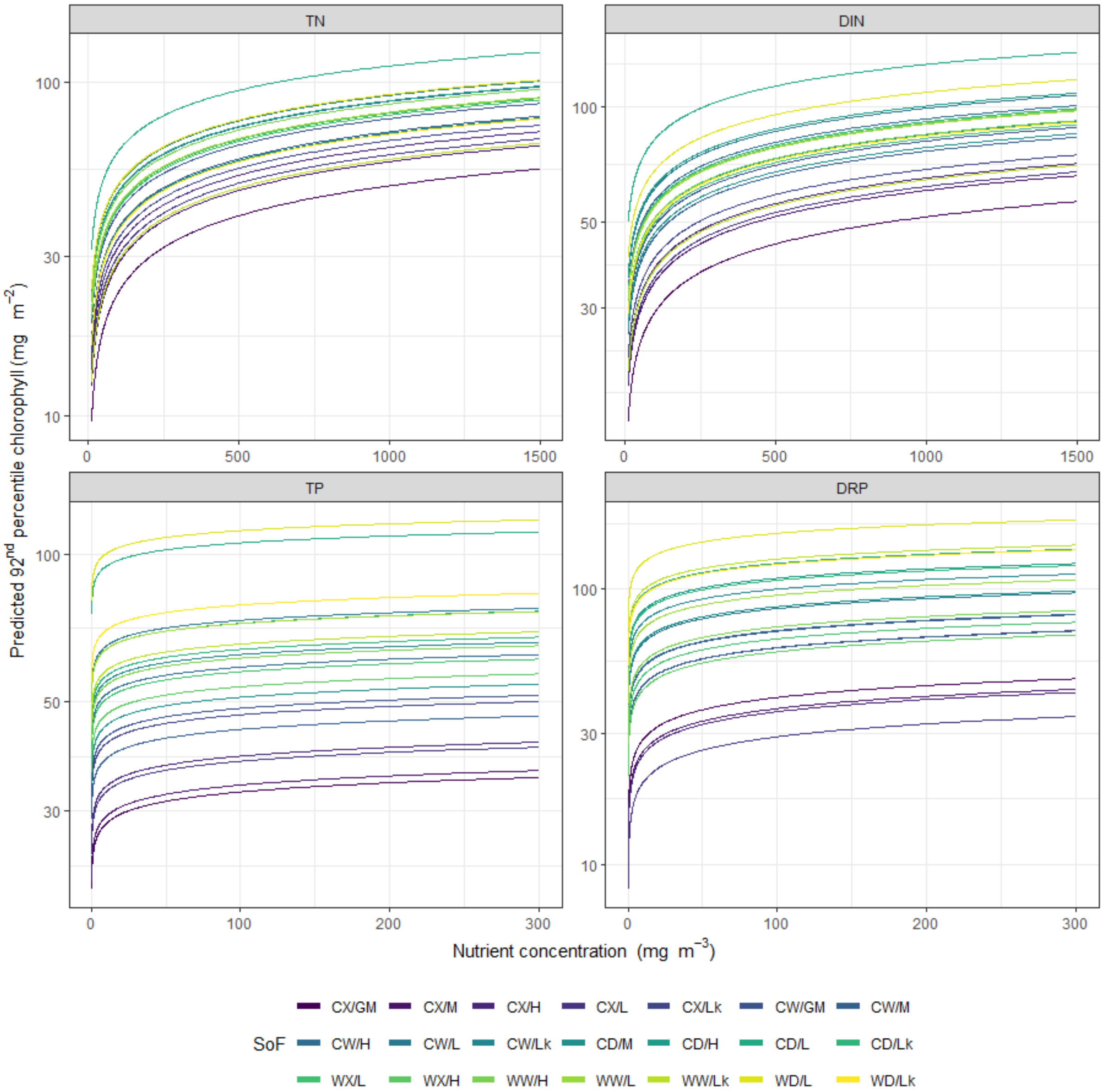

Figure 22. Predicted biomass as a function of nutrient concentration based on the 50% quantile QR models for the four nutrients for all Source-of-flow classes. The colour scheme used for REC Source-of-flow indicates an approximate gradient in resources and disturbance as defined by the conceptual model with the darker blue colour indicating low resources and high disturbance and the yellow colours indicating the opposite.

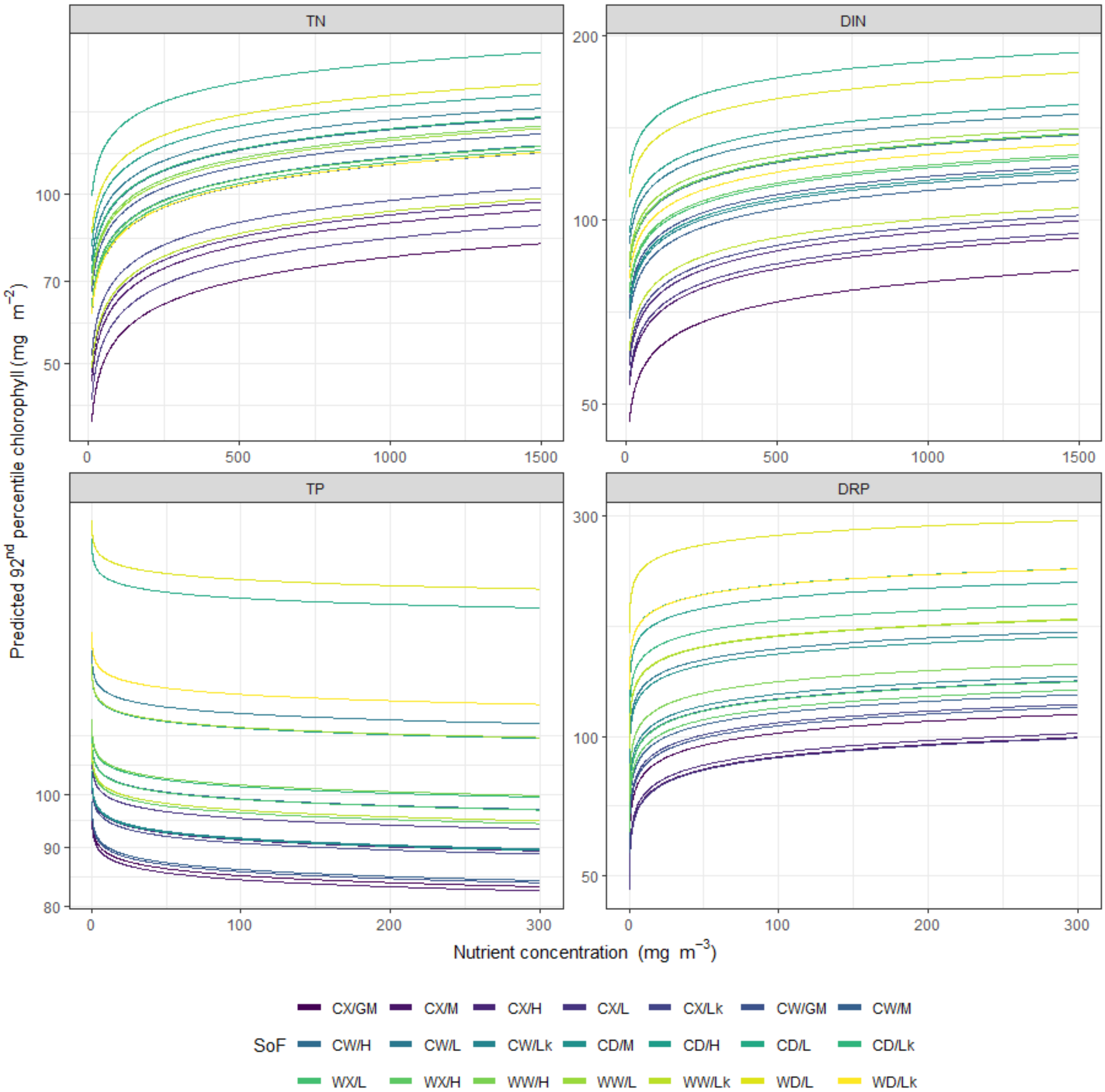

Figure 23. Predicted biomass as a function of nutrient concentration based on the 80% quantile QR models for the four nutrients for all Source-of-flow classes. The colour scheme used for REC Source-of-flow indicates an approximate gradient in resources and disturbance as defined by the conceptual model with the darker blue colour indicating low resources and high disturbance and the yellow colours indicating the opposite.

Table 13. Fitted coefficients for quantile regression models pertaining to each nutrient variable. For each independent variable the first value is the fitted coefficient (including its sign, if negative) and the second value in parentheses is the p-value. NA indicates the independent variable was not included in the model.

Nutrient	Quantile	Intercept	log10(Nutrient)	Temp95	FRE3	Shade	Turb	EC	Reversals	sdQ	nNeg	FineSed
TN	50	2.51 (0)	0.44 (0)	0.02 (0.34)	-0.02 (0.48)	0.15 (0.1)	-0.07 (0.03)	0.003 (0)	-0.01 (0.06)	0.7 (0.35)	NA	NA
	70	3.46 (0)	0.3 (0.01)	0.03 (0.17)	0 (0.86)	0.12 (0.15)	-0.07 (0.02)	0.003 (0)	-0.02 (0.03)	0.58 (0.43)	NA	NA
	80	3.57 (0)	0.23 (0.05)	0.01 (0.66)	-0.02 (0.33)	0.11 (0.22)	-0.06 (0.18)	0.002 (0.007)	-0.01 (0.12)	1.35 (0.17)	NA	NA
	90	2.87 (0.05)	0.19 (0.25)	0.02 (0.47)	-0.03 (0.32)	0.1 (0.34)	-0.03 (0.61)	0.002 (0.081)	-0.01 (0.44)	1.73 (0.15)	NA	NA
DIN	50	2.38 (0.01)	0.35 (0)	0.02 (0.43)	-0.02 (0.32)	0.19 (0.03)	-0.07 (0.03)	0.003 (0)	-0.01 (0.22)	1 (0.21)	NA	NA
	70	3.55 (0)	0.19 (0.03)	0.02 (0.27)	-0.01 (0.71)	0.08 (0.39)	-0.05 (0.07)	0.003 (0)	-0.02 (0.06)	0.9 (0.27)	NA	NA
	80	3.42 (0)	0.18 (0.03)	0.01 (0.49)	-0.02 (0.33)	0.12 (0.24)	-0.05 (0.17)	0.002 (0.011)	-0.01 (0.19)	1.59 (0.15)	NA	NA
	90	3.28 (0.02)	0.16 (0.24)	0.02 (0.35)	-0.02 (0.41)	0.15 (0.11)	-0.05 (0.33)	0.003 (0.053)	-0.01 (0.26)	1.6 (0.17)	NA	NA
TP	50	7.92 (0)	0.09 (0.62)	0.04 (0.13)	-0.01 (0.6)	0.1 (0.29)	-0.04 (0.17)	0.003 (0)	-0.02 (0.01)	1.44 (0.07)	-0.02 (0.08)	NA
	70	7 (0)	0.06 (0.74)	0.02 (0.37)	0 (0.93)	0.08 (0.36)	-0.05 (0.05)	0.003 (0.001)	-0.02 (0.02)	1.2 (0.15)	-0.01 (0.16)	NA
	80	6.82 (0.02)	-0.04 (0.86)	0.01 (0.78)	0 (0.98)	0.05 (0.62)	-0.03 (0.42)	0.003 (0.009)	-0.02 (0.04)	1.33 (0.16)	-0.01 (0.33)	NA
	90	4.99 (0.2)	-0.12 (0.66)	0.03 (0.24)	-0.01 (0.66)	0.07 (0.57)	-0.08 (0.11)	0.003 (0.024)	-0.01 (0.45)	1.79 (0.1)	-0.01 (0.54)	NA
DRP	50	10.47 (0)	0.21 (0.14)	0.07 (0.01)	-0.04 (0.24)	0.09 (0.36)	NA	NA	-0.02 (0.05)	1.83 (0.03)	-0.03 (0)	0 (0.99)
	70	10.62 (0)	0.06 (0.7)	0.03 (0.19)	-0.02 (0.51)	0.02 (0.85)	NA	NA	-0.02 (0.06)	3.07 (0)	-0.03 (0)	-0.01 (0.03)
	80	9.06 (0)	0.16 (0.4)	0.03 (0.23)	-0.03 (0.25)	-0.05 (0.65)	NA	NA	-0.02 (0.07)	3.37 (0)	-0.02 (0.02)	-0.01 (0.04)
	90	12.13 (0)	0.32 (0.09)	0 (0.97)	-0.02 (0.36)	0.04 (0.74)	NA	NA	-0.02 (0.02)	3.54 (0)	-0.03 (0.02)	-0.02 (0)

5.4.3 Random forest models

The R^2 values for the fitted RF models varied from 0.40 to 0.39 for the TN, DIN, TP and DRP models respectively (Table 14). These R^2 values pertain to predictions made for the OOB sample and therefore indicate the models' performance when predicting to sites that were not included in the model fitting process. The RF models therefore performed marginally better than the OLS models by this measure and in terms of RMSD.

Table 14. Performance of the RF models of periphyton biomass pertaining to each nutrient.

Nutrient	N	R^2	PBIAS	RMSD
TN	251	0.40	0.16	0.53
DIN	251	0.39	0.14	0.54
TP	251	0.40	-0.09	0.53
DRP	251	0.40	-0.06	0.53

The TN and DIN models retained all 14 independent variables as significant predictors, but the TP and DRP models retained only seven variables (Table 15). The nutrient was the first and second most important independent variable for the TN and DIN models, respectively, but was not retained when offered to the TP and DRP models.

The directions of the relationships fitted to the RF model independent variables were consistent with the conceptual model (Figure 24). For example, for the TN and DIN models, the response (Chla92) was positively associated with nutrient concentration (Nutrient), electrical conductivity (EC) and temperature (Temp95), which represent resource supply (Figure 2). All models had a negative association between biomass and the hydrological indices FRE3, Reversals and nNeg, which represent disturbance (Figure 2). All models included a positive association with flow variability (sdQ), which represents the intensity of low flow periods (i.e., lack of disturbance; Figure 2). The independent variable Shade was retained in the TN and DIN models with a negative relationship. Because the ordering applied to this variable was unshaded < shaded, the negative relationship indicates that, all other variables being equal, biomass is higher at unshaded than shaded sites. It is noted that the directions of the relationship fitted by the RF models were the same for the OLS models and models developed in other studies (Snelder *et al.*, 2019, 2014).

Table 15. Order of importance of the independent variables fitted by the RF models pertaining to each nutrient. NA indicates the independent variable was not included in the model.

Independent variable	TN	DIN	TP	DRP
EC	2	1	1	1
Nutrient	1	2	NA	NA
sdQ	3	3	2	3
Temp95	4	4	3	2
MALF7	6	5	4	4
FRE3	5	6	5	5
Reversals	7	7	6	6
nNeg	8	8	7	7
SolarRadDec	9	9	NA	NA
FineSed	10	10	NA	NA
Clar	11	11	NA	NA
Turb	12	12	NA	NA
NH4N	13	13	NA	NA
Shade	14	14	NA	NA

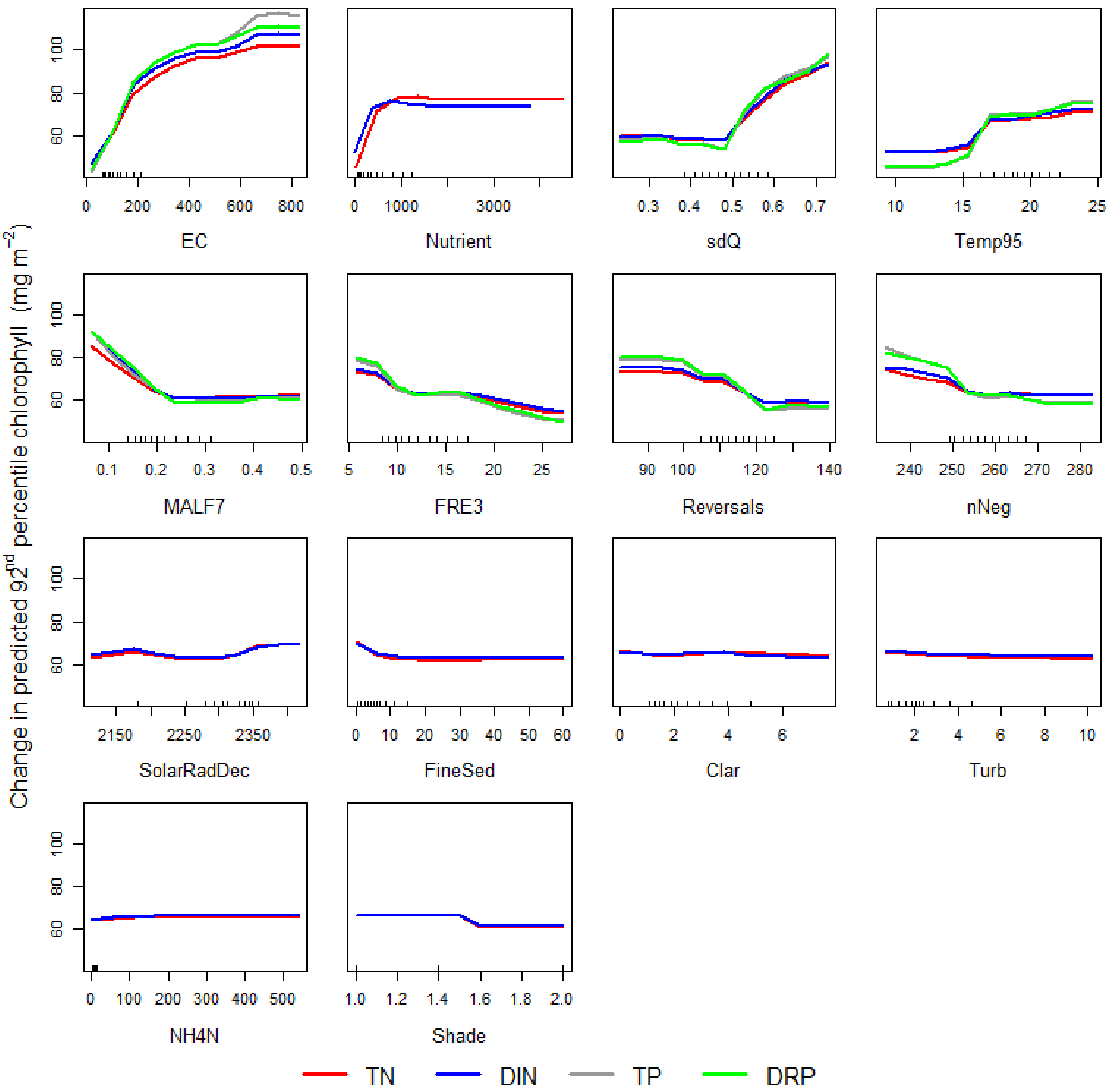

Figure 24. Partial plots for the RF models pertaining to each nutrient. Each panel indicates the marginal contribution to periphyton biomass of an independent variable. The overall order of importance of each independent variable is shown in Table 15. Note that some panels have only 2 or 3 nutrient forms plotted because the predictor was not retained in the models for the other nutrient forms.

Like the OLS models, the RF models generally underpredicted the highest observed Chla92 values (i.e., high values had large regression residuals; Figure 25). The underpredicted sites were not associated with low numbers of biomass observations (Figure 25) nor were they associated with any of the incompletely measured site characteristics (Figure 26). This indicates that the independent variables for which we had incomplete data do not explain differences in Chla92 between sites (see Section 3.2.5).

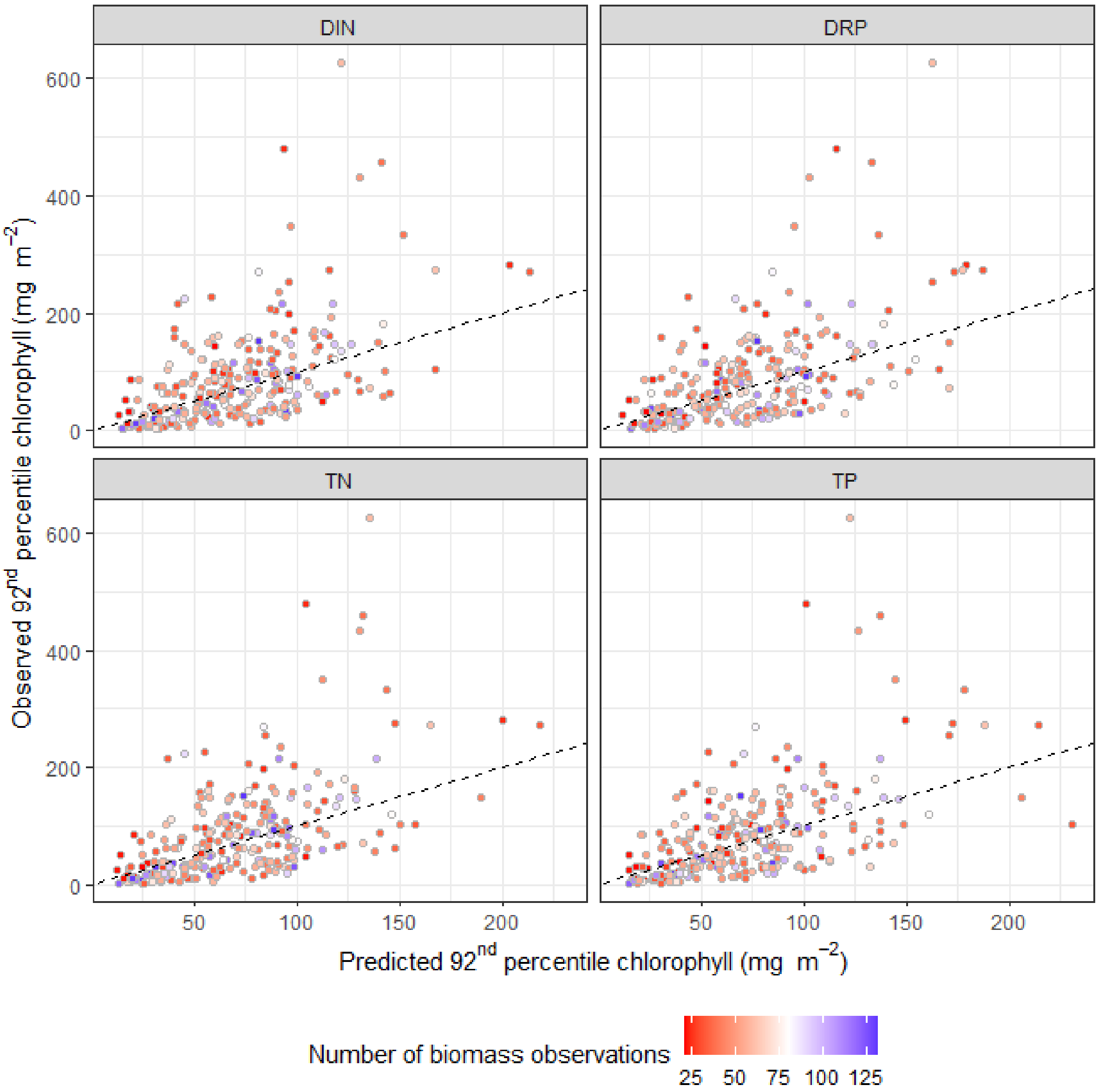

Figure 25. Predicted versus observed values for the RF regression models. Each point represents a site, and the colour indicates the number of biomass observations. The black dashed line is one to one. See Table 14 for regression performance statistics.

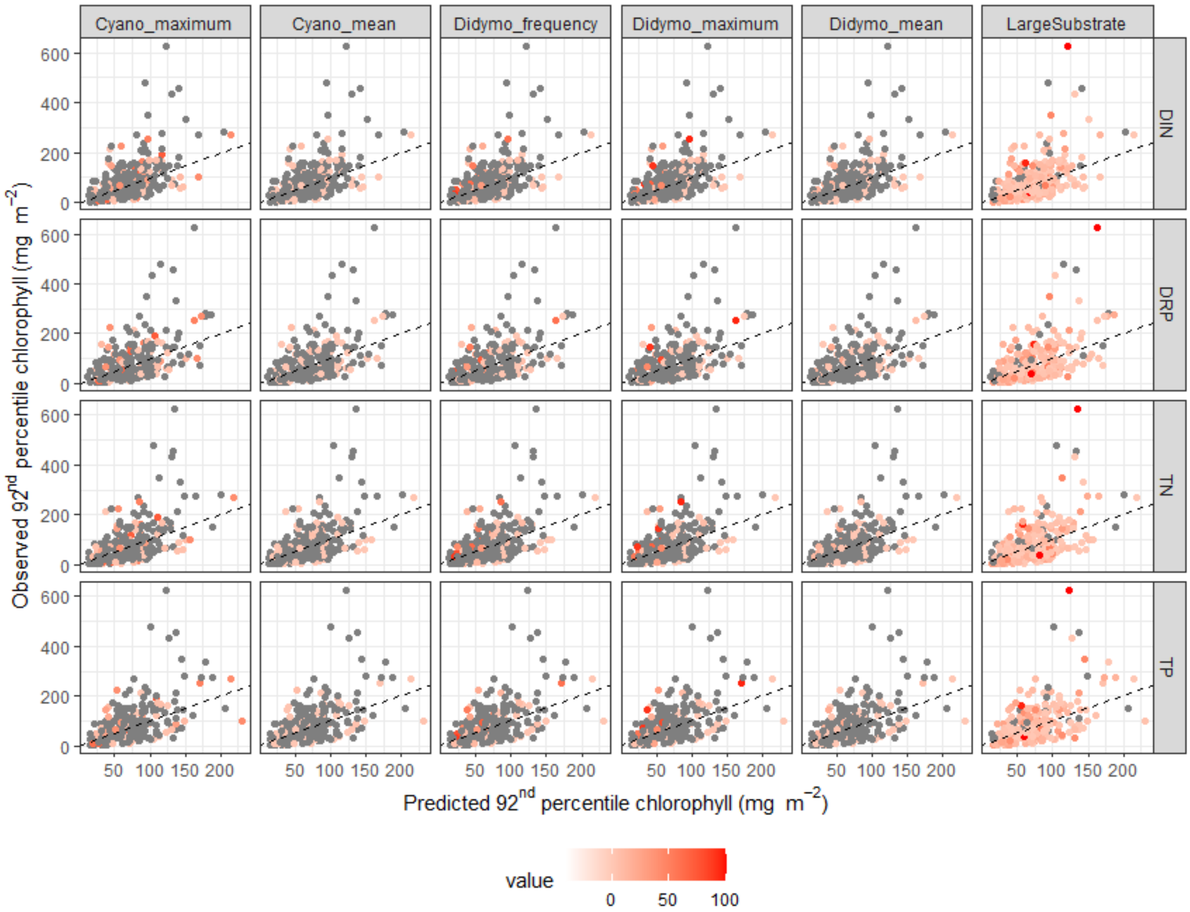

Figure 26. Observed versus predicted values for the RF regression models showing additional characteristics. Each point represents a site, and the colour indicates value of five incompletely measured site characteristics. Grey points indicate the characteristic was not available for the site. The black dashed line is one to one.

Predictions of biomass made using the RF models for each REC Source-of-flow class were broadly consistent with the conceptual model and expectations for each class (Figure 27). Predictions based on the TN and DIN RF models indicated that classes with low resource supply (i.e., low temperature) and high hydrological disturbance such as CX/GM, CX/M, CX/H and CX/L had low biomass compared to classes with high resource supply and low disturbance such as WD/L and WD/Lk (Figure 27).

The predictions of biomass along a gradient in nutrient concentration made using the RF models were not as smooth as the OLS predictions (compare Figure 27 with Figure 22). This is because the RF models are non-parametric. Biomass increased rapidly with increasing TN and DIN concentration until a ceiling occurred at nutrient concentration of approximately 1000 mg m^{-2} after which there was a slightly negative relationship (Figure 24). Note that the TN and DIN concentrations at the ceiling, and the Chla92 ceiling, for the RF models (Figure 24) was very similar to that for the OLS models (Figure 18).

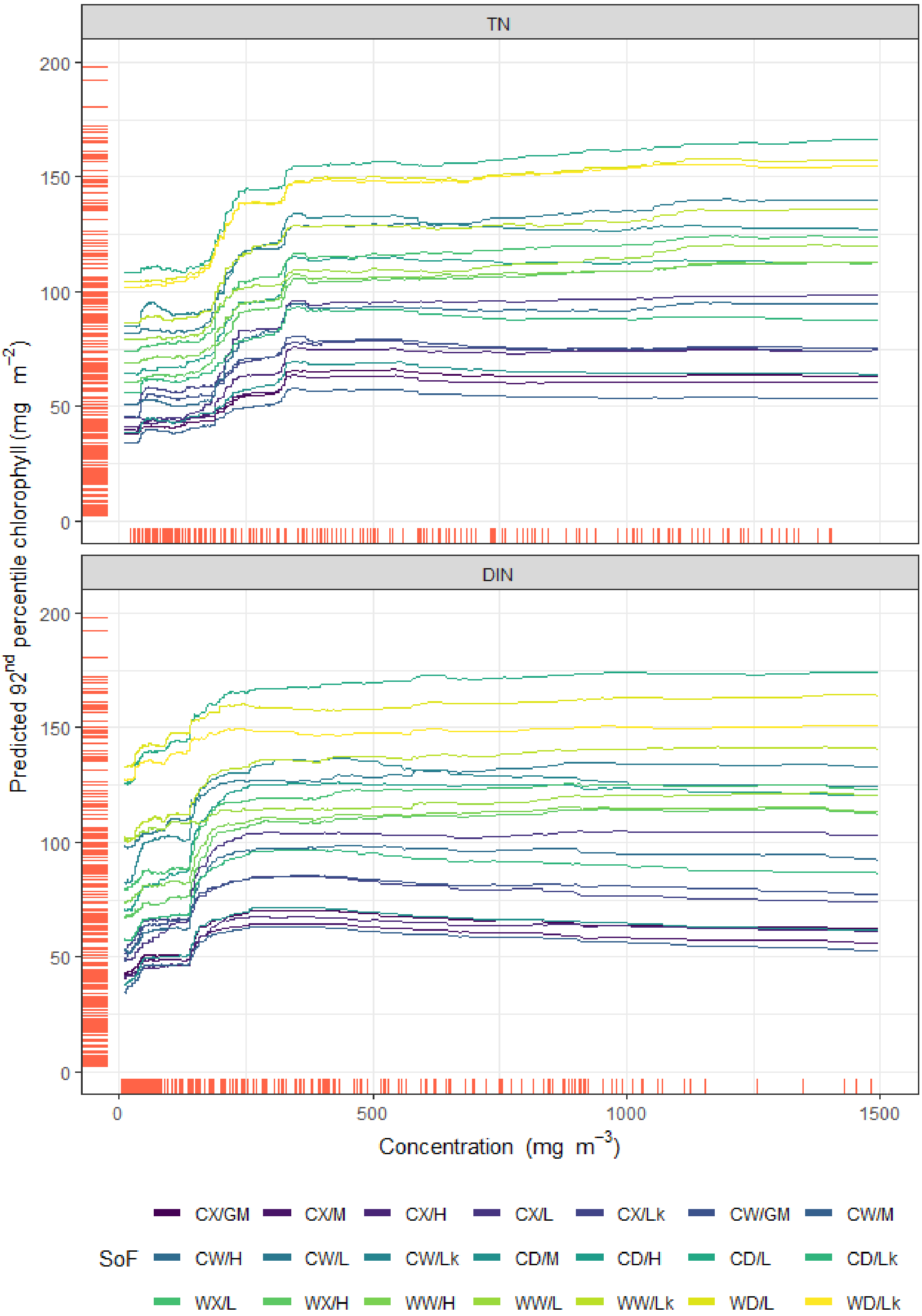

Figure 27. The 80th percentile of the predicted distribution of Chla92 as a function of nutrient concentration based on the RF models for TN and DIN and for all Source-of-flow classes. Note that the RF models that were offered DRP and TP did not include these nutrients and are therefore not shown. The colour scheme used for REC Source-of-flow indicates an approximate gradient in resources and disturbance as defined by the conceptual model with the darker blue colour indicating low resources and high disturbance and the yellow colours indicating the opposite. The red "rug" indicates the data density on both axes. Note that the x-axis has been truncated at 1500 for clarity.

5.5 Choice of models to define criteria

Step 4 of the analysis process outlined in Figure 6 is to choose a set of regression models (i.e., four models, each using one of TN, DIN, TP and DRP to represent nutrient concentration) to derive the nutrient criteria. We selected the four OLS models as the most robust and credible models with which to derive nutrient criteria. The QR models were not robust or credible given that the models for the more extreme percentiles (70%, 80% and 90%) did not include many independent variables as significant terms (Table 13) and the predicted responses were inconsistent with the conceptual model (Figure 23). We had not anticipated using the RF models to derive the nutrient criteria because of limitations in how this type of model can extrapolate predictions for unsampled environmental conditions (see Section 4.4.3).

5.6 Modelled independent variables

There were no existing estimates for digital river network segments for three independent variables that were included in at least one of the fitted OLS models. The variables were: EC, FineSed and Temp95. Spatial models of these independent variables were therefore fitted and used to make predicted values for all network segments that could be used at the subsequent nutrient criteria derivation step.

Prior to fitting the models, distributions of EC and FineSed were made more symmetric by log_{10}- and logit-transformation of the variables (the model responses), respectively. No transformation was necessary for the temperature variable (Temp95).

The RF models of EC and Temp95 had good performance (Table 16), as indicated by the criteria of Moriasi et al. (2015) ($0.60 < R2 \le 0.70$, $15 \le |PBIAS| < 20$). The FineSed model had satisfactory performance ($0.30 < R2 \le 0.60$, $20 \le |PBIAS| < 30$).

At the subsequent nutrient criteria derivation step, predictions of water quality made by Whitehead (2018) and hydrological indices made by Booker and Snelder (2012) and Snelder and Booker (2013) were also used. These models also had at least satisfactory performance based on the criteria of Moriasi et al. (2015).

Table 16. Performance of the RF models used to make spatial predictions of the independent variable EC, FineSed and Temp95. N indicates the number of sites used to fit the model. Transformation indicates the transformation applied to the modelled response, and which are applicable the reported RMSD values in the table.

Modelled variable	N	R^2	NSE	PBIAS	RMSD	Transformation
EC	251	0.63	0.62	0.09	0.15	log10
FineSed	251	0.43	0.43	0.08	0.58	logit
Temp95	251	0.70	0.70	-0.39	0.81	none

Patterns in the predicted values of EC, FineSed and Temp95 are shown in Figure 28. The maps indicate generally higher EC in low elevation locations which may be associated with relatively higher groundwater inputs. Patterns in Temp95 are broadly consistent with the expectation that higher values will be associated with lower elevation and lower latitudes.

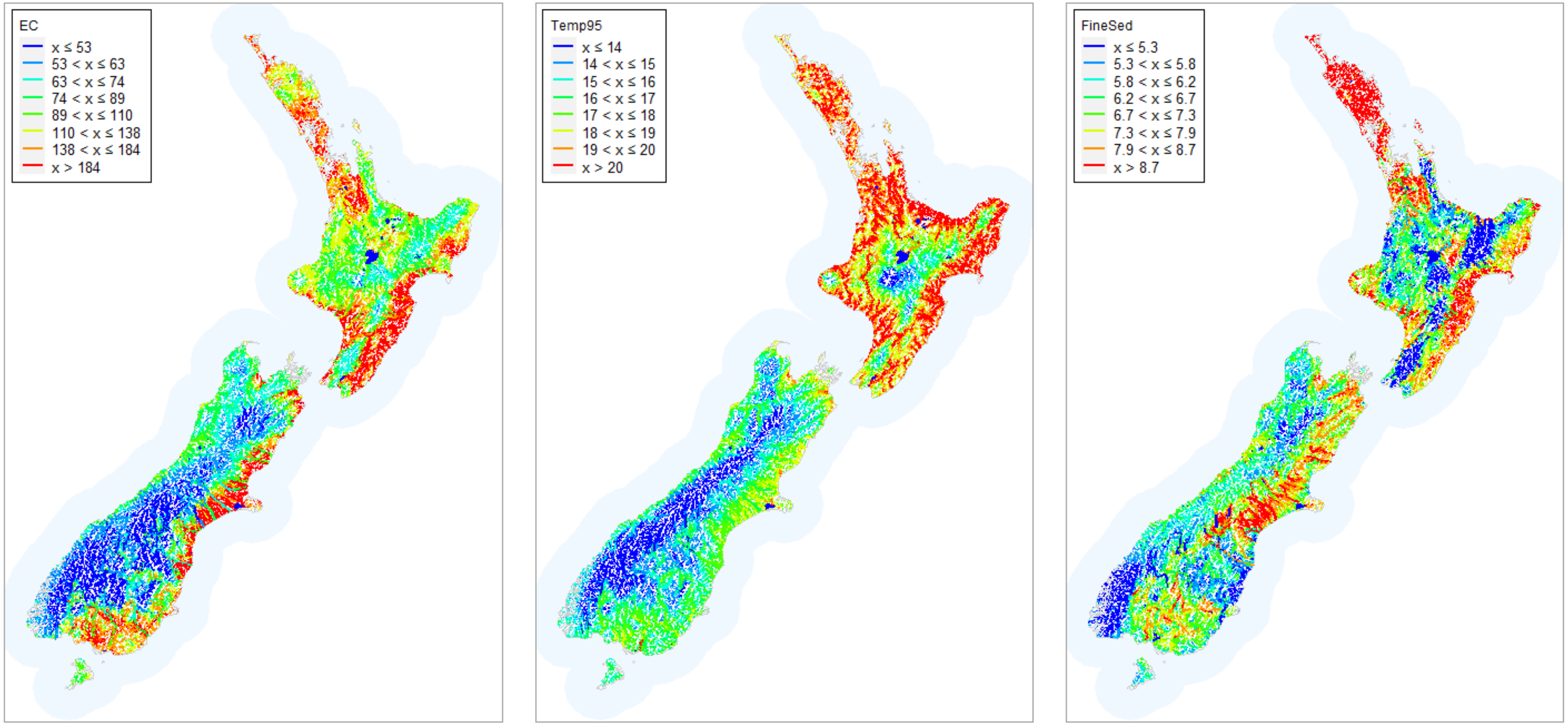

Figure 28. Predicted values for three independent variables (EC, FineSed and Temp95) that were included in the biomass models.

5.7 Nutrient criteria for each Source-of-flow class

The distributions of monitoring sites across the environmental gradients represented by the independent variables retained in the OLS models were generally consistent with the distribution of all network segments nationally across the same gradients (Figure 4). For some environmental gradients, there was moderate over- and under-representation. For example, monitoring sites (represented by the blue histograms in Figure 29) were under-represented in environments characterised by cold temperature (low values of Temp95) and were over-represented in environments characterised by warm temperature (high values of Temp95). Monitoring sites were slightly over-represented in environments characterised by low Reversals and under-represented in environments characterised by low nNeg.

The plots shown in Figure 29 indicate that there are no monitoring sites representing segments with high FRE3 and low solar radiation (SolarRadDec). Model predictions for segments with high FRE3 and low solar radiation are therefore outside of the range of the observations and represent extrapolation from the modelled relationships. The criteria may therefore be less reliable in these environments.

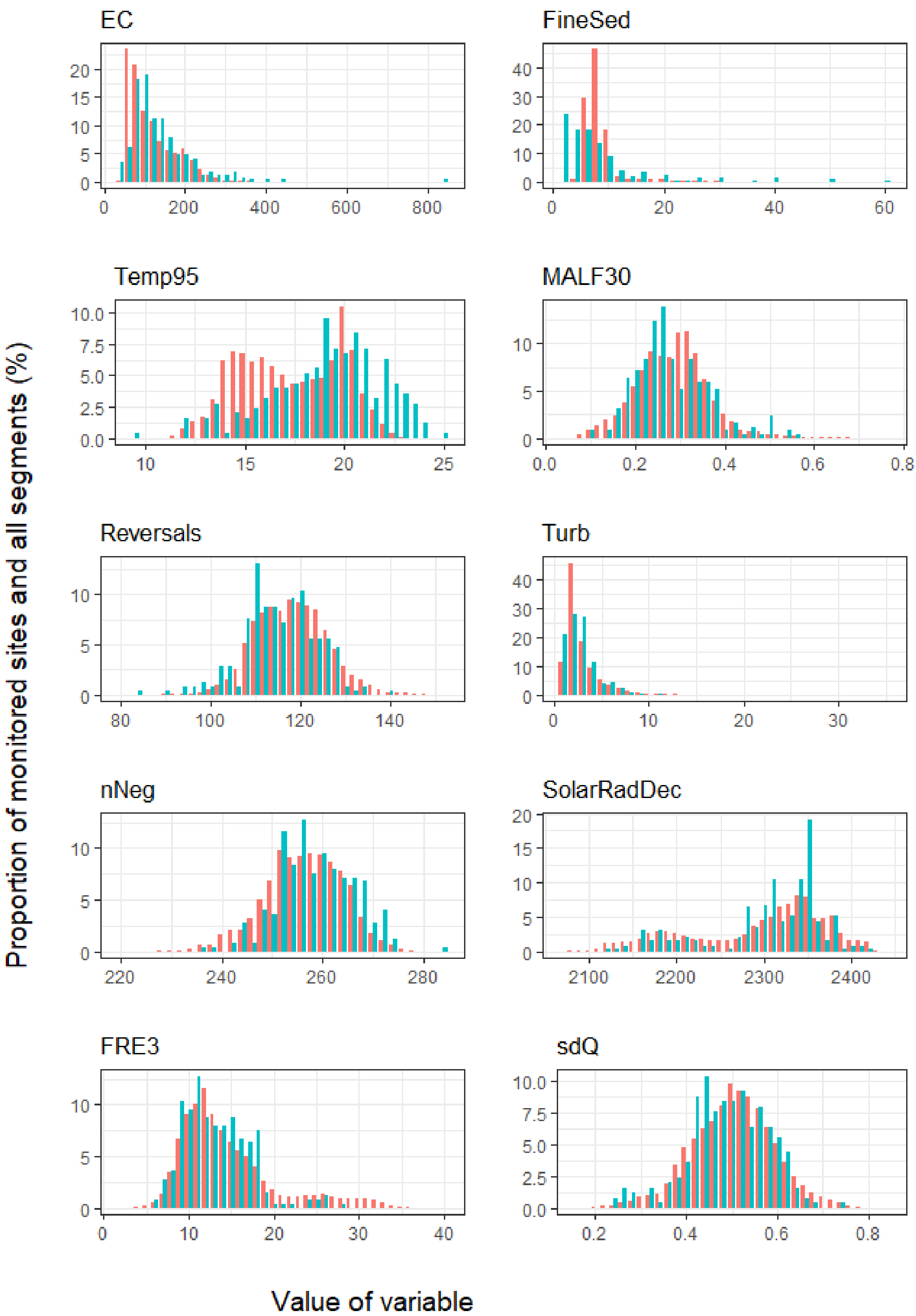

Figure 29. Histograms comparing the distributions of the ten independent variables included in the OLS models for all segments of the river network and the monitoring sites. The network segments are represented by the red histograms and the monitoring sites are represented by the blue histograms. Similarities in the distributions shown in the two histograms in each panel provide an indication of the degree to which environmental variation across the monitoring sites represent environmental variation across all rivers and streams in New Zealand; complete representativeness would be indicated by exact matches between the histograms.

The OLS models were used to predict periphyton biomass across the range in nitrogen and phosphorus concentrations observed in our site data (i.e., up to 4,500, 3,800, 300 and 230 mg m^{-3} for TN, DIN, TP and DRP, respectively). From these predictions, the criteria corresponding to biomass thresholds of 50 mg m^{-2}, 120 mg m^{-2}, and 200 mg m^{-2} were derived for six levels of under-protection risk (50%, 30%, 20% 15%, 10% and 5%) as described in Section 4.7.1. Because Shade was included as an explanatory variable in all four nutrient OLS models, separate look-up tables of nutrient criteria were derived for shaded and unshaded locations.

The criteria for the 20% level of under-protection risk are shown as examples for shaded and unshaded locations and all Source-of-flow classes in Figure 30. The criteria show three patterns that are consistent with the conceptual model. Firstly, criteria are lower for the lower biomass thresholds than the higher thresholds. Secondly, criteria are higher for Source-of-flow classes with low resource supply (i.e., low temperature) and high hydrological disturbance such as CX/GM, CX/M, CX/H and CX/L compared to classes with high resource supply and low disturbance such as WD/L and WD/Lk. Thirdly, within Source-of-flow classes for a given biomass threshold, criteria are higher for shaded locations compared to unshaded (Figure 30, Figure 31).

For each nutrient, the maximum possible value for a criterion is the maximum observed nutrient concentration (i.e., 4,500, 3,800, 300 and 230 mg m^{-3} for TN, DIN, TP and DRP, respectively). Criteria that are equal to these values indicate that on average the model did not predict that the threshold was reached even when the nutrient concentration was maximum.

The complete set of criteria for all thresholds, Source-of-flow classes and levels of under-protection risk for shaded and unshaded locations are provided as tables in Appendix A.

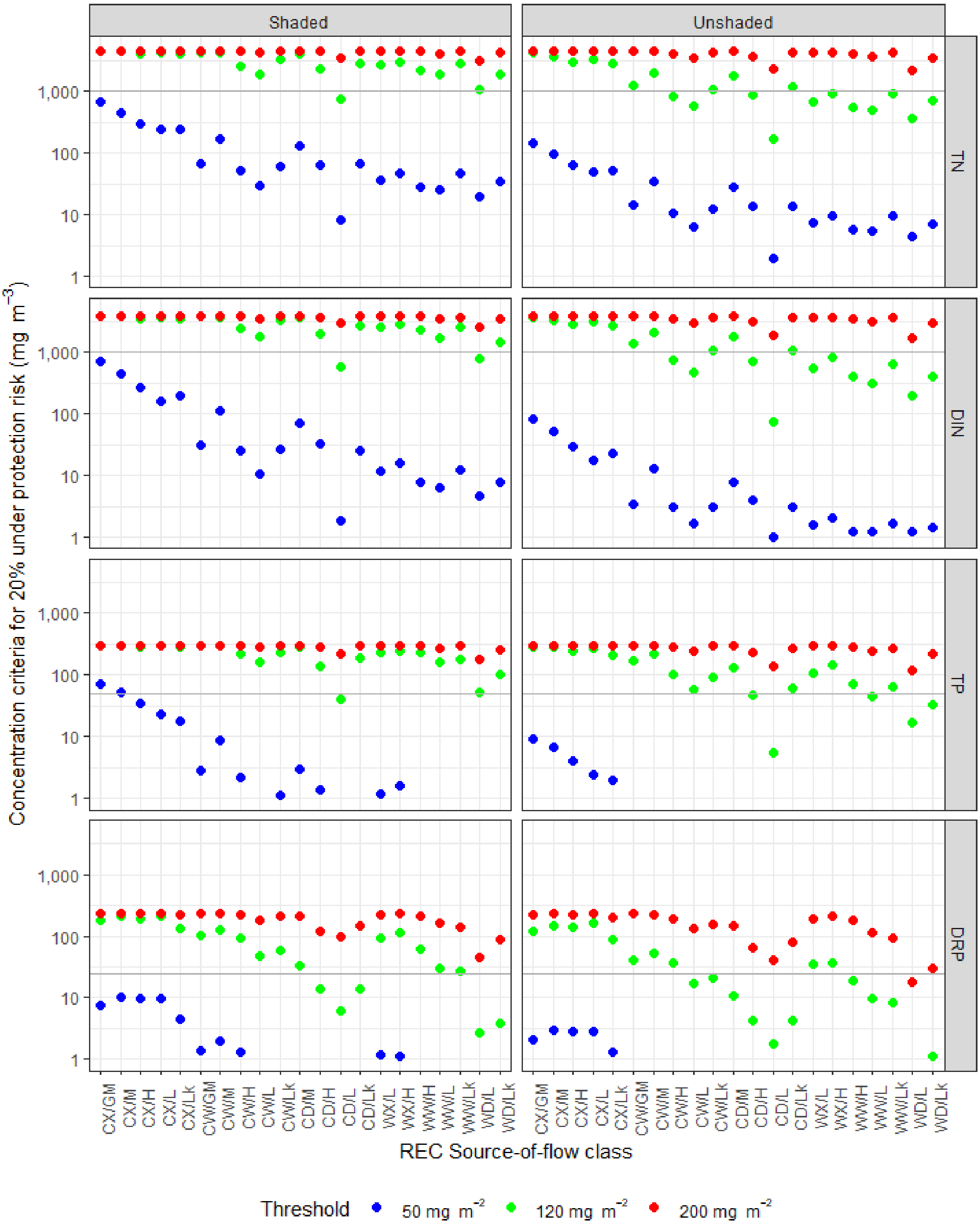

Figure 30. Nutrient criteria for REC Source-of-flow classes and a 20% under protection risk. The grey lines indicate the approximate nutrient concentrations at which the biomass ceiling occurs. For each nutrient, the maximum possible value for a criterion is the maximum observed nutrient concentration (i.e., 4,500, 3,800, 300 and 230 mg m^{-3} for TN, DIN, TP and DRP, respectively).

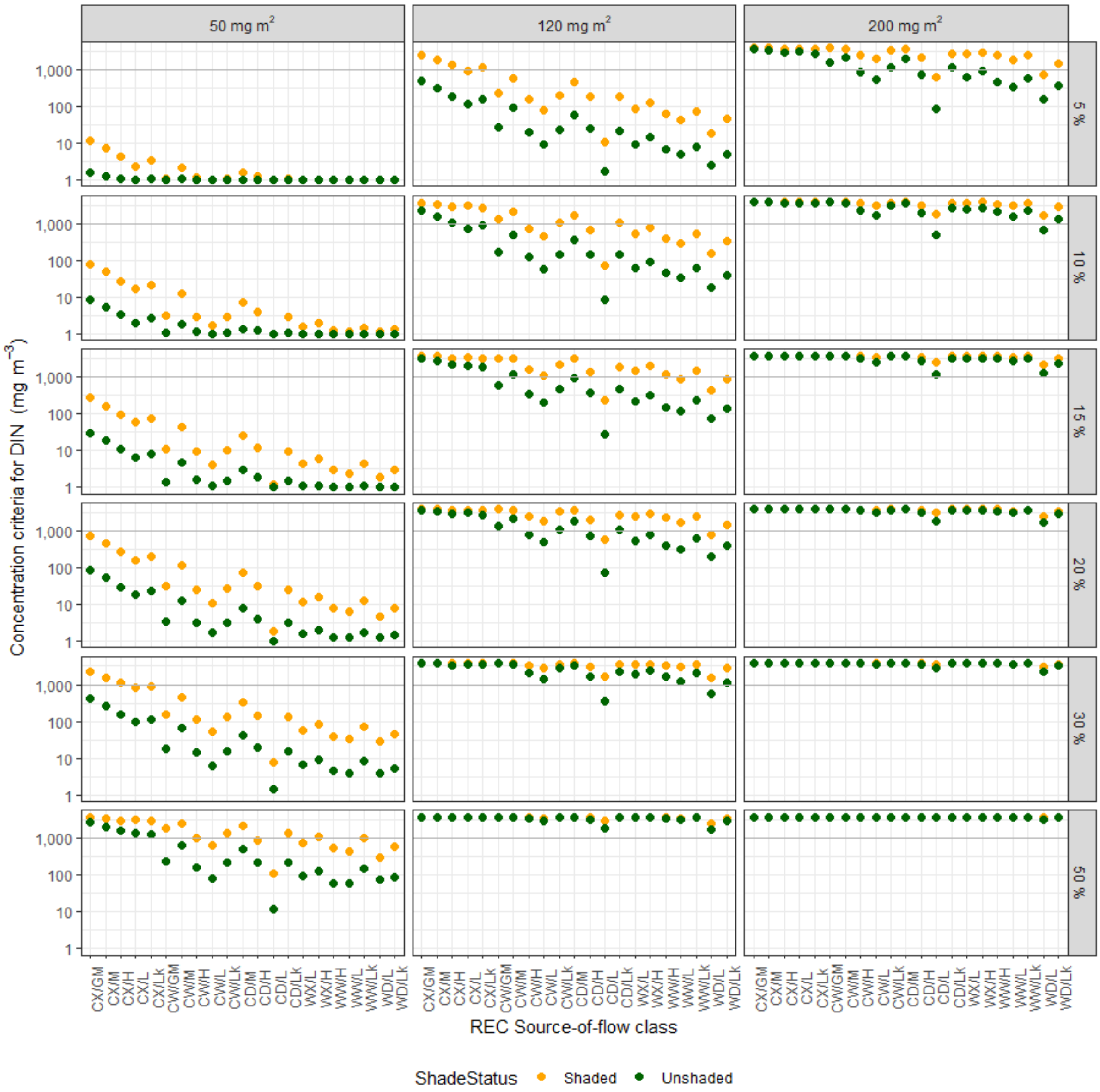

Figure 31. DIN criteria for REC Source-of-flow classes and four levels of under-protection risk and three biomass thresholds for shaded and unshaded sites.The grey lines indicate the approximate nutrient concentrations at which the biomass ceiling occurs. The maximum possible value for a criterion is the maximum observed DIN concentration (3,800 mg m^{-3}).

5.8 Comparison with existing criteria

The criteria derived by this study are compared to those derived by Snelder *et al.* (2019) on Figure 32 and Figure 33. The Snelder *et al.* (2019) criteria did not discriminate between shaded and unshaded locations. Because the NRWQN sites used by Snelder *et al.* (2019) were generally on large rivers, it has been assumed the Snelder *et al.* (2019) criteria apply to unshaded locations. Therefore, the appropriate comparison is with this study's unshaded criteria.

For TN and the of 50 mg m^{-2} threshold and the 10% and 20% levels of under-protection risk, the criteria derived by this study were appreciably lower (all Source-of-flow classes plot below the one-to-one line on Figure 32). In contrast, for TN and the of 200 mg m^{-2} threshold and the 10% and 20% levels of under-protection risk, the criteria derived by this study were appreciably higher (most Source-of-flow classes plot above the one-to-one line on Figure 32). The two sets of criteria showed a greater level of agreement for the 120 mg m^{-2} threshold for both the 10% and 20% levels of under-protection risk (Source-of-flow classes plot close to and either side of the one-to-one line on Figure 32).

Similar patterns in differences in the two sets of criteria are seen for DRP (Figure 33).

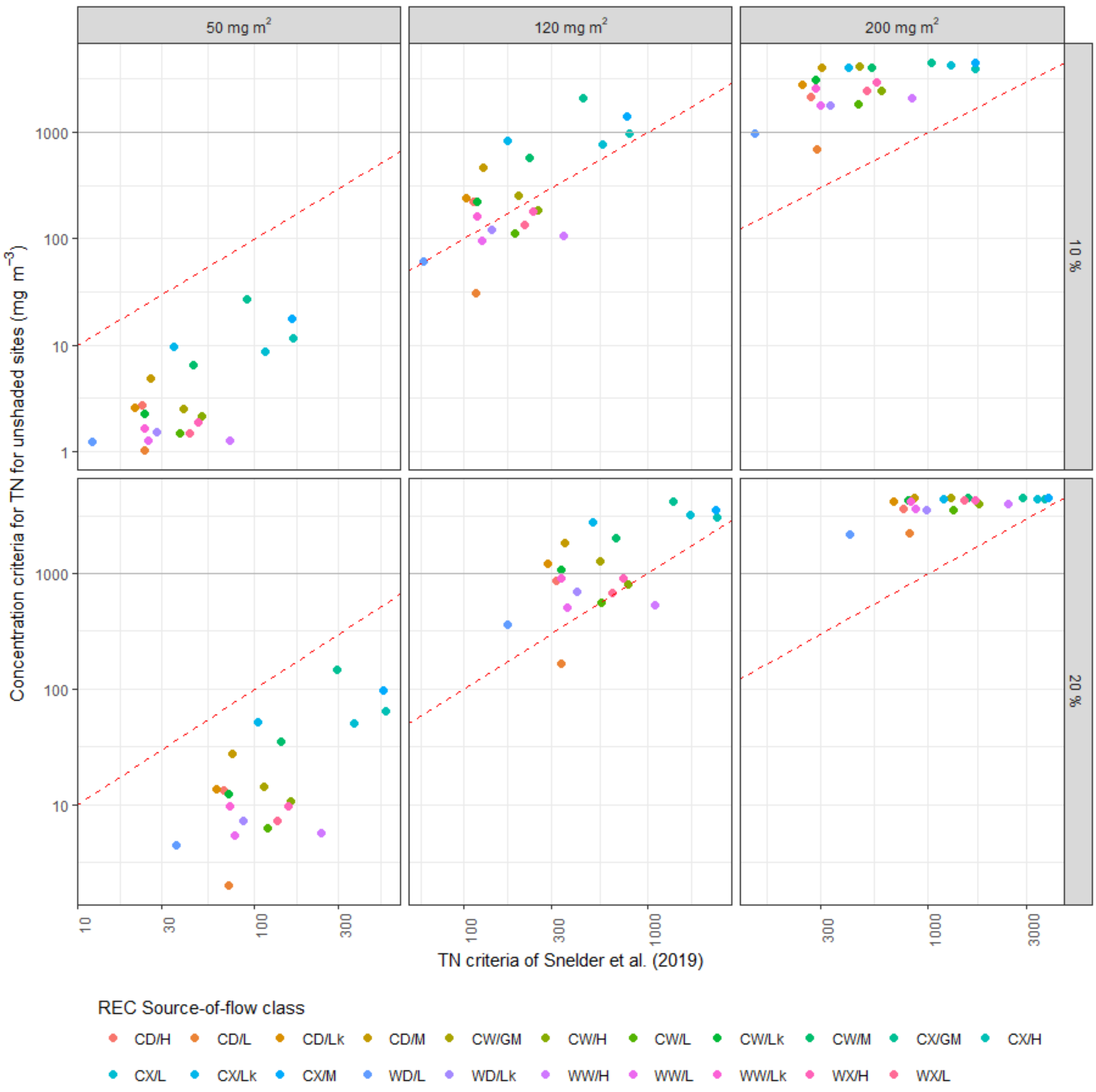

Figure 32. Comparison of TN criteria derived by this study to those of Snelder et al. (2019). The rows are two levels of under protection risk (20% and 10%). The red line is one to one. The grey lines indicate the approximate nutrient concentration at which the biomass ceiling occurs. The maximum possible value for a criterion on the y-axis (this study) is the maximum observed TN concentration (4,500 mg m^{-3}).

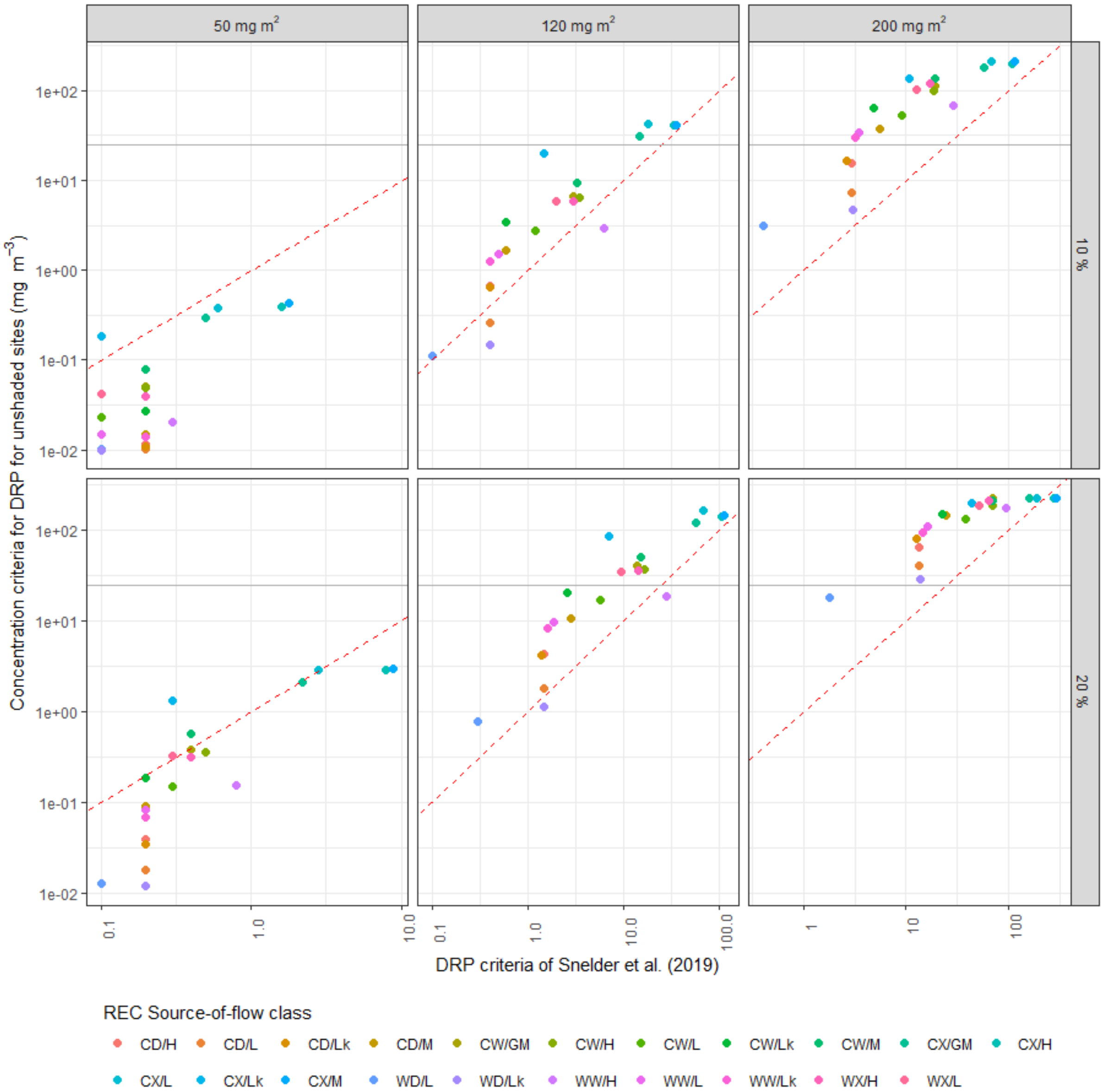

Figure 33. Comparison of DRP criteria derived by this study to those of Snelder et al. (2019). The rows are two levels of under protection risk (20% and 10%). The red line is one to one. The grey lines indicate the approximate nutrient concentration at which the biomass ceiling occurs. The maximum possible value for a criterion on the y-axis (this study) is the maximum observed DRP concentration (230 mg m^{-3}).

5.9 Validation of the criteria

The observed and predicted values of the 92nd percentile periphyton biomass at the 78 sites in the independent dataset based on the four nutrient forms are shown as scatter plots in Figure 34. Theoretically, 5%, 10%, 15%, 20%, 30% and 50% of the independent 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 line on Figure 34).

The data shown in Figure 34 indicate that the proportions of sites for which observed biomass exceeds the predicted increases systematically as the under-protection risk increases for all four nutrient forms. Table 17 indicates that the proportion of sites for which observed biomass exceeds the predicted is approximately as expected according to the level of under-protection risk for all four nutrient forms. Overall, the proportions of sites for which observations exceed predictions are slightly higher than expected based on the level of under-protection risk. Snelder *et al.* (2019) showed that there are large uncertainties associated with both the observations and predictions. Given these uncertainties, it is likely that the independent data are consistent with the derived nutrient targets (i.e., the predictions are within the uncertainty of both the observations and the predictions).

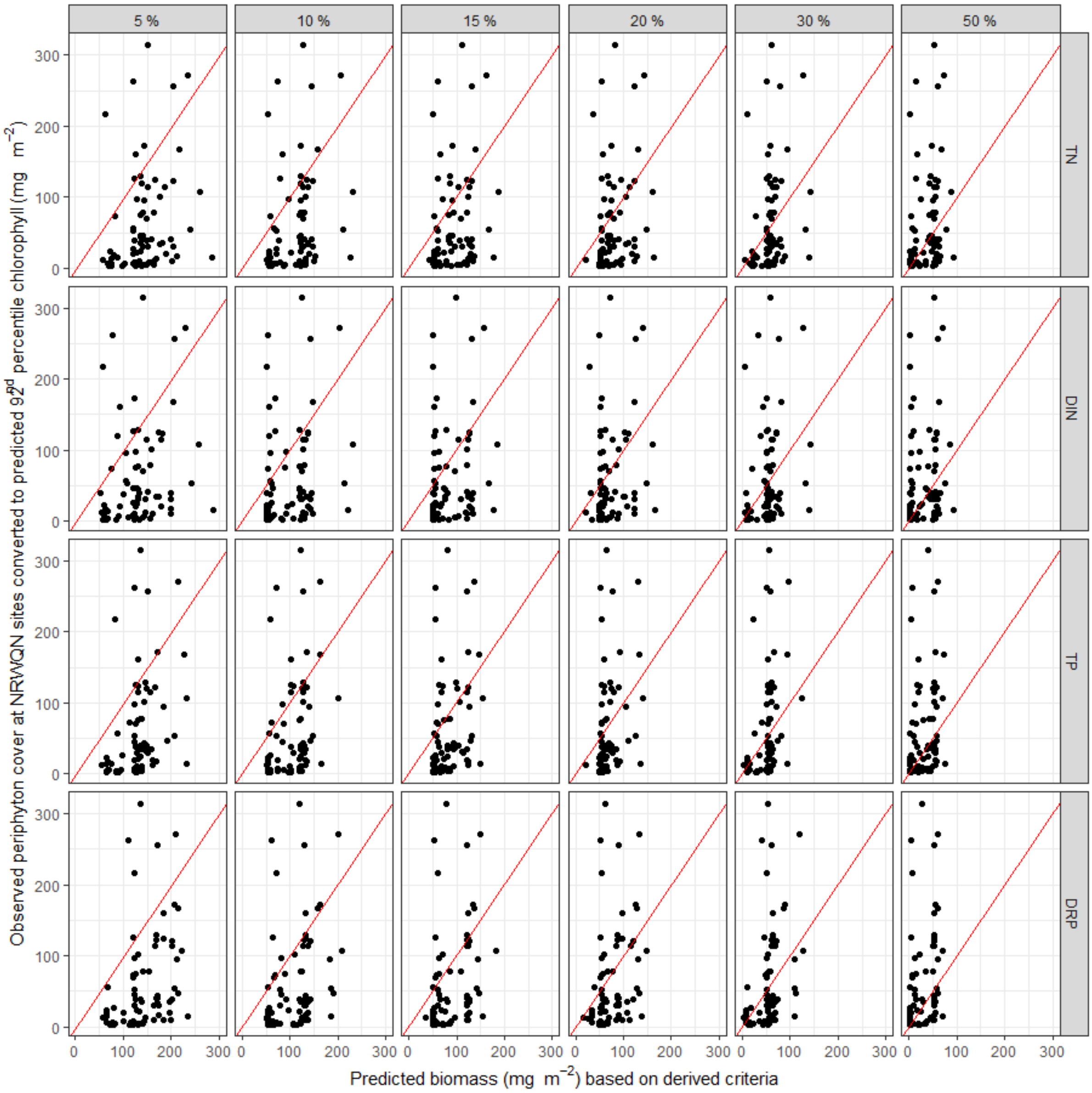

Figure 34. The observed and predicted values of the 92nd percentile periphyton biomass at the 78 NRWQN sites in the independent dataset where predicted values are derived from the nutrient criteria for under-protection risks of 5, 10,15, 20, 30 and 50%. Panel labels indicate the under-protection risks and the nutrient form (TN, DIN, TP and DRP). The red diagonal (one to one) line represents agreement between the predictions and observations. The points lying below the red line indicate sites for which the observed biomass was less than that predicted by the targets and vice versa.

Table 17. Performance of the criteria based on independent dataset. of sites (%) for which observed biomass exceeds the predicted for the four levels of under-protection risk.

Under protection risk (%)	Nutrient form			
	TN	DIN	TP	DRP
5	10	12	10	8
10	15	18	19	17
15	22	23	23	23
20	26	31	29	27
30	35	37	37	38
50	54	63	56	62

5.10 National current periphyton state model

The national current periphyton state model was not subject to the same constraints as the OLS or QR models. It therefore represents that best possible prediction of current periphyton state and might help to identify limitations associated with the OLS or QR models and, therefore, the criteria. The R^2 value for the fitted RF model of national periphyton current state was 0.45. The model had low bias (PBIAS = -0.1%) and an RMSD of 0.51. These performance statistics pertain to predictions made for the OOB sample and therefore indicate the models' performance when predicting to sites that were not included in the model fitting process.

Quantile-quantile (Q-Q) plots indicate that the distributions of the regression residuals of the national periphyton current state model were reasonably consistent with the theoretical normal distribution (Figure 35). This indicates that the model's residual values are reasonably normally distributed. The model may therefore be used to predict the entire probability distribution based on the theoretical normal distribution.

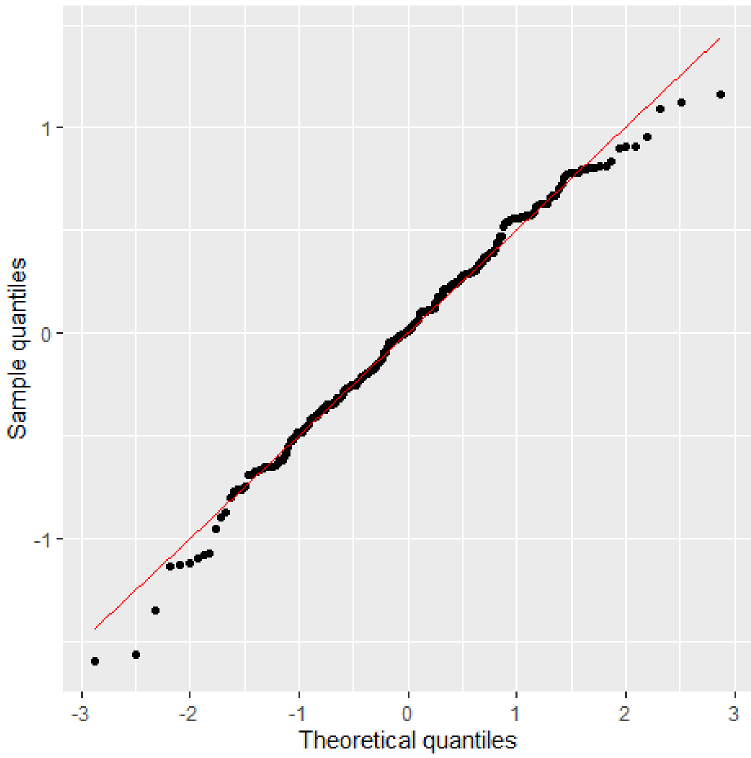

Figure 35. Quantile-quantile plot comparing the RF national periphyton biomass model regression residuals to the normal distribution.

The model retained 26 independent variables, whose relationships with the model response (Chla92) are described by the partial plots shown in (Figure 36). Chla92 was positively associated with concentrations of the nutrients TN and DIN, Temp95, and EC, which is expected because these variables represent resource supply (Figure 2). Chla92 was negatively associated with frequency of high flows (FRE2 and FRE3), which represent disturbance (Figure 2). Biomass was positively associated with Max30 and sdQ, which indicate sustained periods of low flow (i.e., periods without disturbance; Figure 2). Similarly, biomass was negatively associated with MALF30, low values of which indicate sustained periods of low flow (i.e., periods without disturbance; Figure 2).

Several catchment characteristics were included in the national model, and these can be understood as representing gradients in resource supply and disturbance. For example, Chla92 was positively associated with usIntensiveAg. Catchments with high values of usIntensiveAg will tend to have higher temperatures and higher nutrient supply (i.e., high resources; Figure 2). Biomass was negatively associated with usNativeForest, usSlope and usRain. Catchments with high values of usNativeForest, usSlope and usRain will tend to have lower temperatures and nutrient supply (i.e., low resources; Figure 2) and more frequent high flows (i.e., high disturbance; Figure 2).

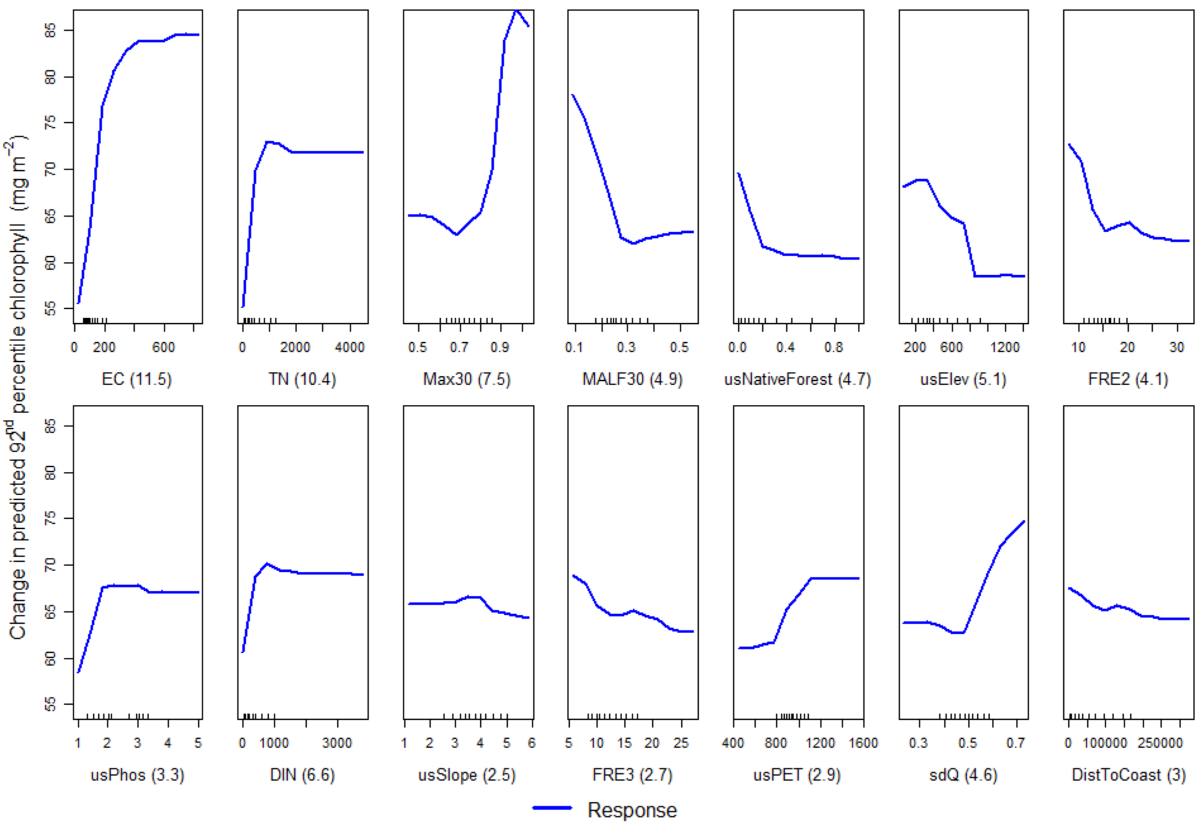

Figure 36. Partial plots for the RF model predicting 92nd percentile chlorophyll using all available predictors. Each panel indicates the marginal contribution of the independent variable to the response. The 14 most important predictors that were retained by the reduced RF are shown.

The RF model generally underpredicted the highest observed biomass values (i.e., high biomass values had large residuals; Figure 37). The underpredicted sites were not associated with low numbers of biomass observations (Figure 37) nor were they associated with any of the incompletely measured site characteristics (Figure 38). This indicates that the independent variables for which we had incomplete data do not explain differences in Chla92 between sites (see Section 3.2.5).

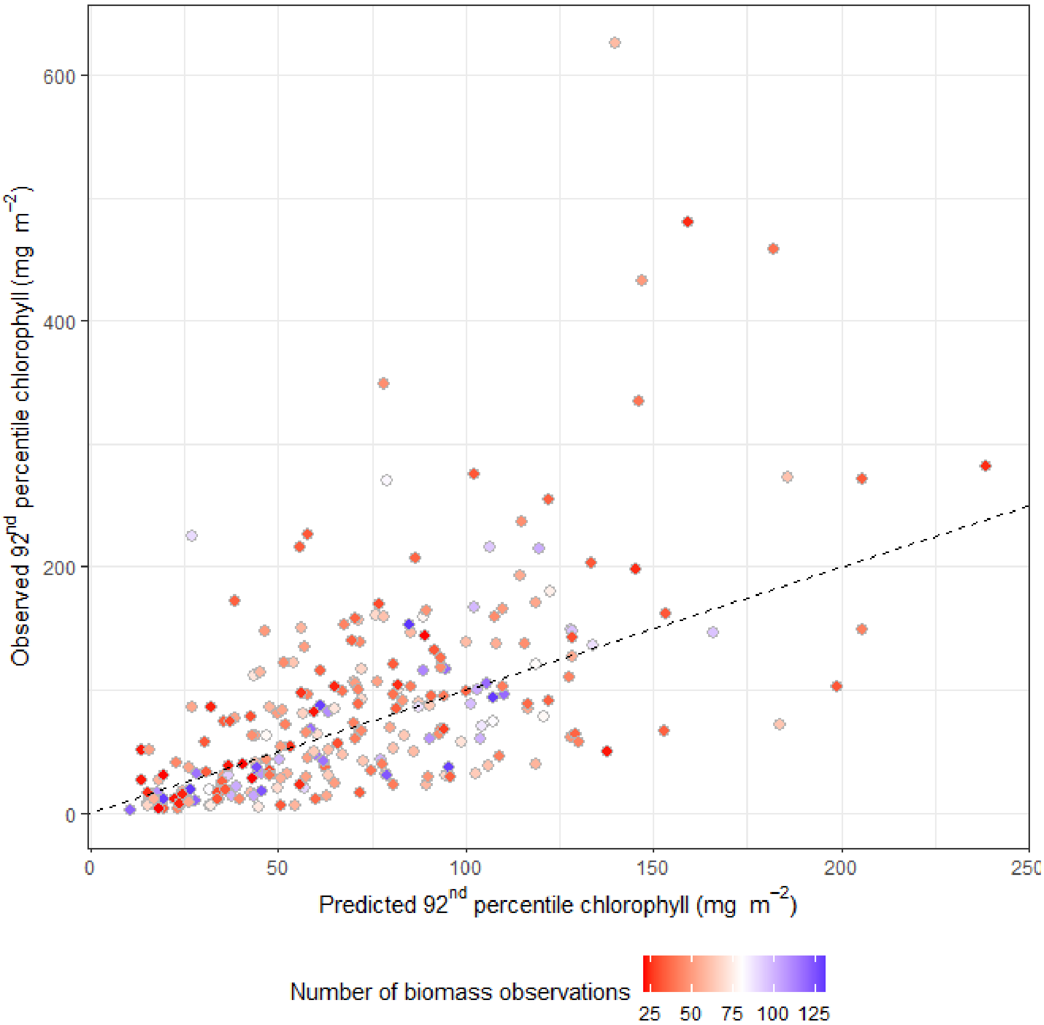

Figure 37. Observed versus predicted values for the national periphyton biomass RF regression model. Each point represents a site, and the colour indicates the number of biomass observations. The black dashed line is one to one.

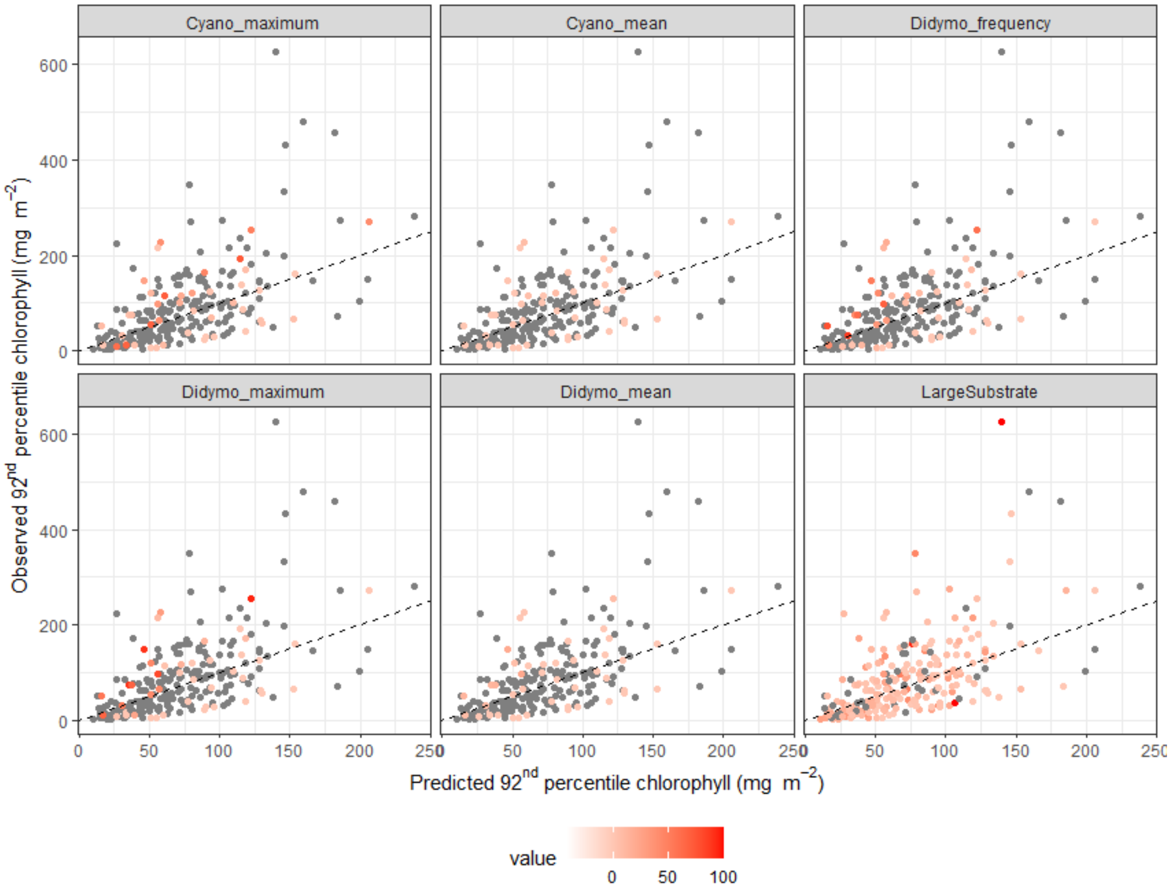

Figure 38. Observed versus predicted values for the national periphyton RF model showing incompletely measured site characteristics. Each point represents a site, and the colour indicates value of five additional characteristics that were not available for all sites. Grey points indicate the characteristic was not available for the site. The black dashed line is one to one.

The fitted RF model was used to make predictions of current 92nd percentile of chlorophyll for all segments of the national digital river network of stream order ≥3. These predictions were then converted to the probability the segment belongs to each of the four NOF attribute states (Figure 39).

The percentage of river segments (stream order ≥3) having different ranges of probability of being in the NOF D attribute state (i.e., 92nd percentile of chlorophyll > 200 mg m^{-2}) are tabulated in Table 18. The predictions indicate that 1%, 5% and 13% of the river network has >30%, 20-30% and 10-20% probability of being in the NOF D attribute state (i.e., below the NOF bottom line), respectively.

Table 18. Predicted percentage of river segments of order ≥3 belonging to ranges of probabilities of being in the NOF D attribute state (below the bottom line).

Probability range	Proportion of network segments (%)
<0.01	47
0.01-0.1	34
0.1-0.2	13
0.2-0.3	5
>0.3	1

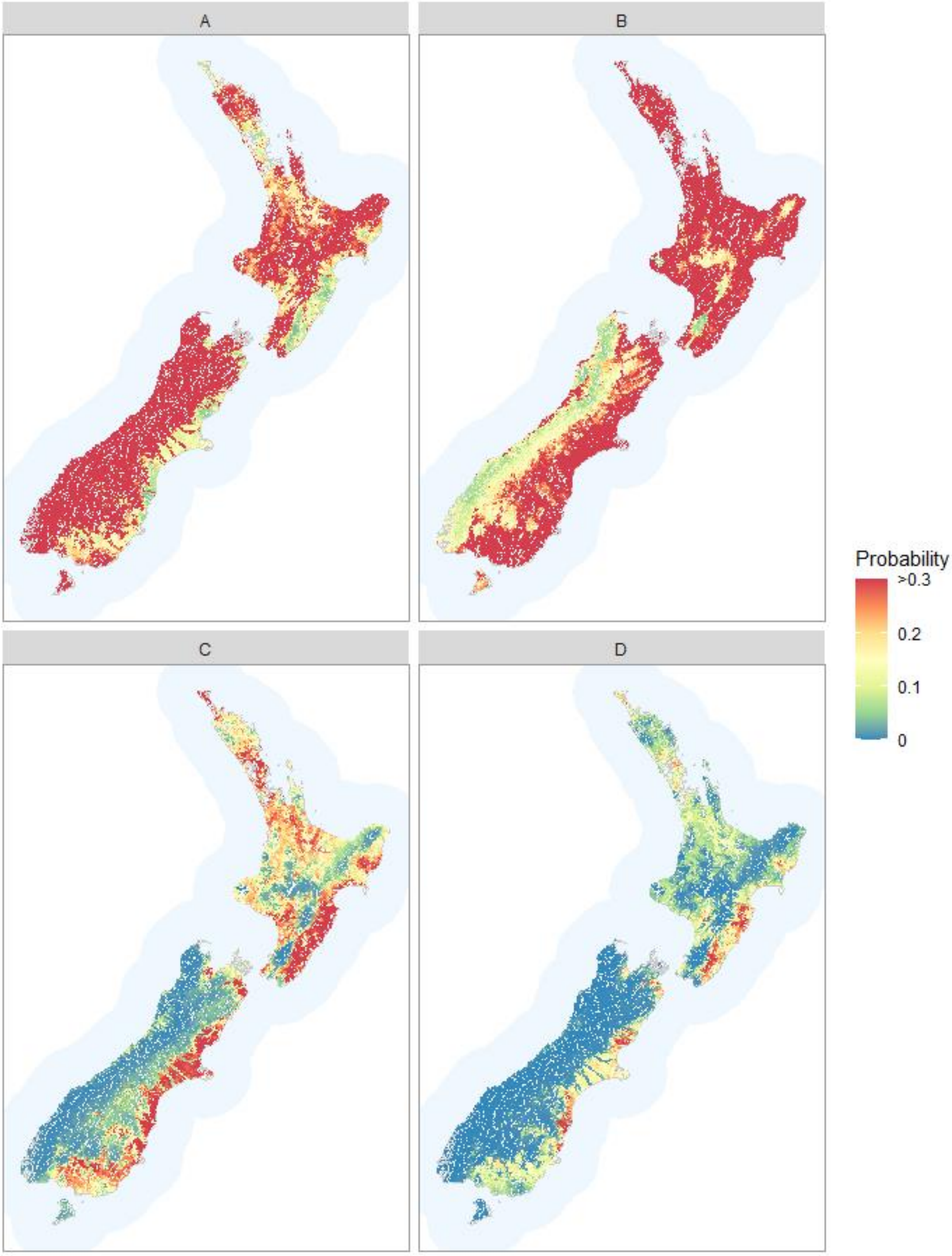

Figure 39. National predictions of current 92nd percentile of chlorophyll for river network segments of stream order ≥3. The maps show the probabilities of network segments belonging to each of the NOF periphyton attribute bands. Note, the colour scale is capped at a probability of 0.3.

6 Discussion

6.1 Approach to defining nutrient criteria

The models used in this study to define the nutrient criteria aimed to explain between-site differences in the characteristic peak periphyton biomass (i.e., the 92nd percentile of monthly samples) as a function of between-site differences in the mean intensity of processes represented by the conceptual model (i.e., biomass accrual and loss). The models therefore do not involve any representation of time. The median nutrient concentration is used as an independent variable in the models because it is an appropriate measure of average difference in nutrient enrichment between sites. The median concentrations incorporate seasonal and flow-driven fluctuations in nutrient concentrations (e.g., high concentrations during winter and low concentrations driven by instream uptake when periphyton biomass is high). Similarly, the hydrological indices are measures of the general difference in levels of disturbance between sites. This approach, as opposed to more temporally detailed approaches (such as using only data leading up to peak biomass in summer) is justified by the observation that periphyton biomass rarely shows consistent seasonality (see Appendix C). It is likely that high biomass is most strongly associated with periods of sustained low or base flow, and these can occur at any time of the year. This indicates that models would not be improved by any discretisation of time. There have been attempts to represent time in regression-based periphyton models, for example, by incorporating antecedent nutrient concentration (e.g., nutrient concentrations averaged over one or more sample occasions prior to a biomass observation) (Kilroy *et al.*, 2018). However, this approach has not been applied to derivation of generalised nutrient criteria because the dynamics are complex and monthly sampling is insufficient to describe the processes involved (see review by Kuczynski, 2019). While there have been some efforts to model periphyton dynamics using detailed mechanistic models (e.g., Suplee *et al.*, 2015), these methods are at early stages of development, are generally applied at a site-specific level, and are not suitable for deriving the nation-wide criteria that this study aimed to produce.

We developed models that incorporated both the dissolved inorganic and total forms of nitrogen and phosphorus. TN and TP often show the best relationships with periphyton biomass (both in New Zealand and internationally, (e.g., Dodds, 2007; Smith and Tran, 2010; Snelder *et al.*, 2014). In this study, the OLS models based on TN and DIN performed similarly but the TP model performed better than the DRP model. In practice, analysts and modelers need to convert between total and dissolved forms of nutrients, depending on the questions being addressed and there are methods to do this. For example, catchment models will generally need to consider the total forms nitrogen and phosphorus in order to conserve mass, but criteria may need to be specified in terms of dissolved inorganic nitrogen and phosphorus. We produced criteria for both dissolved and total forms of both nutrient, which provides a flexible set of tools for users.

Variation in between-site periphyton biomass explained by the statistical models developed in this study was low. For the OLS models that were used to define the criteria, LOOCV R^2 values varied between 0.27 and 0.38 (Table 11). The consequence of low variation explained by the statistical models is that predictions of biomass have large uncertainty. For example, a prediction for the 92nd percentile chlorophyll of 90 mg m^{-2} made with a the OLS models defined in this study has a 95% confidence interval of approximately 15 mg m^{-2} to 300 mg m^{-2} (Table 10) and the 70th, 80th and 90th percentiles of the distribution are 128, 156 and 202 mg m^{-2}, respectively.

The model uncertainty means that there is not a single specific criterion that will ensure that a target level of biomass is not exceeded. Instead, there is a probability distribution that describes the risk that the target level of biomass at a site will be exceeded for a given nutrient concentration. We refer to this probability as under-protection risk. It is important that nutrient criteria describe, as much as possible, the risk of under-protection and allow decision-makers to choose the level of risk that is acceptable. The nutrient criteria derived in this study provide for choice in the level of under-protection risk that is acceptable.

The under-protection risk refers to a specific river location. Choosing a level of under-protection risk means that, across a domain of interest comprising many sites, a proportion of locations can be expected to have biomass higher than the nominated target despite being compliant with the criteria. The corollary to this is that the objective underlying the criteria is to maintain periphyton biomass at or below the nominated thresholds at a proportion of sites within a domain that is the complement of the under-protection risk. For example, under-protection risks of 30%, 20% and 10% correspond to objectives to maintain biomass below the target level at 70%, 80% or 90% of sites across the domain, respectively.

Defining objectives in terms of risk of under-protection acknowledges that there are large uncertainties associated with the statistical (OLS) models used to define criteria by this study at the site-scale. However, because these models explain some variation, they can describe broad-scale variation in the response of benthic chlorophyll to water column nutrient concentrations (Dodds *et al.*, 2002; Snelder *et al.*, 2019, 2014). The large site-scale uncertainties are due to the complexity of processes controlling periphyton growth and loss and to the incomplete representation of these processes by the independent variables in the regression models.

The approach to developing nutrient criteria in this study combines the models' description of broad-scale variation with the assumption that, rather than requiring the target biomass condition at all sites, the management objective is to limit the risk that periphyton biomass will exceed the target level. The approach therefore uses the model in a manner that is appropriate to its precision and encourages management to define management objectives in terms that are tractable.

6.2 Quality of the models and comparison with other studies

The performance of the OLS and RF models defined by this study are similar to that found in past national studies of this type in New Zealand (Kilroy, Snelder, *et al.*, 2020; Kilroy and Stoffels, 2019; Snelder *et al.*, 2019, 2014). The R^2 values are also consistent with results of similar studies overseas. For example, Dodds *et al.* (2002) and Dodds (2006) explained a maximum of 40% of mean biomass (chlorophyll) for observations made at North American sites representing diverse environmental conditions. Other studies have found no meaningful relationship between periphyton abundance and nutrient concentrations (e.g., Lewis and McCutchan, 2010; Welch *et al.*, 1988). A conclusion from this is nutrient-periphyton biomass relationships are complex and our understanding and ability to predict biomass is limited.

The models developed in this study were very consistent, both structurally and in terms of performance, with previous national-scale studies using different datasets and descriptors of biomass abundance. For example, Snelder *et al.* (2019, 2014) derived OLS models that included log (base 10) transformed nutrient concentrations (TN and DRP), light and temperature terms and the hydrological indices (FRE3, nNeg, MALF7). The same or similar independent variables were included in the models produced by this study with the directions of the relationships between the independent variables and biomass being the same in all studies and consistent with the conceptual model (Figure 2). Thus, multiple national-scale studies involved checking many independent variables and resulted in retaining the same, or similar, selections. This provides confidence that the independent variables used in this study are widely applicable at the national scale.

Higher performing statistical models than those developed by this study have been derived in some national and regional studies in New Zealand. For example, using monthly data collected over a period of at least a year at 30 river sites throughout New Zealand, Biggs (2000) was able to explain over 70% of the variation in maximum chlorophyll from a combination of mean DIN or DRP concentrations and FRE3. The dataset was limited to a specific river type, primarily in the CW/H REC class. At a regional scale, models explaining between-site variation in Chla92 with $R^2 > 0.75$ were derived for data from the Manawatu-Whanganui region (Kilroy *et al.*, 2018). Better performance of regional models compared to national models could be because of the more limited ranges of conditions that are represented within regions compared to across the whole country

International studies that have identified relatively strong relationships ($R^2 > 0.4$) between chlorophyll and nutrients and other independent variables have generally been based on small datasets (N < 30 sites) associated with gradients of nutrient concentrations starting from a relatively low minimum (e.g., TN < 100 mg m^{-3}) (e.g., Carr *et al.*, 2005; Munn *et al.*, 2010). In addition, variables that are surrogates for nutrient concentrations, such as catchment land use (as percentage cover) sometimes explain considerably more of the variance in chlorophyll than proximal predictors (including nutrients) described in the conceptual model in Figure 2 (e.g., Austin *et al.*, 2015; Taylor *et al.*, 2004). However, models that do not use proximate measures of nutrient supply (e.g., median concentrations) are not useful for deriving nutrient criteria.

The dataset used in the current study is unique in that it has a large spatial coverage (251 sites) and the dependent variable (Chla92) was derived from time series of observations over at least 2 years (and in most cases over 3 years), as well robust data describing all independent variables (including nutrient concentrations). To our knowledge, studies with similar geographic coverage have derived the dependent variable from synoptic data (a single sampling occasion) and relied on inclusion of large-scale variables (e.g., ecotypes) to achieve similar explanatory skill as that seen in the present study (Urrea-Clos *et al.*, 2014).

6.3 Assumptions and limitations of the approach

The key statistical assumptions were:

- that the sample data (i.e., the monitoring sites) are representative of the population (i.e., all rivers in New Zealand), and
- that the residuals of the regression models are normally distributed.

The first assumption is important because obtaining criteria for every Source-of-flow class required using statistical models to make predictions outside the range of the observations (i.e., extrapolation from the observations). This is because Source-of-flow classes encompass all rivers and streams in New Zealand, but the fitting dataset represent a restricted range of environments. We assumed that the fitted relationships were representative of all rivers, but it was not possible to know if this was true. However, some insight into whether the assumption is robust was obtained by considering how representative the monitoring sites were of national scale variation in the explanatory variables included in the models.

Our analysis of representativeness indicated that the criteria derivation process has not involved large extrapolations from the environmental conditions represented by the monitoring sites (Figure 29). Locations that are most poorly represented are rivers with high-frequency high flows (FRE3) and low solar radiation (SolarRadDec), which are characteristic conditions in western and mountainous regions of the country. The poor representation of rivers in the western and mountainous regions of the country is evident from the geographic distribution of monitoring sites (Figure 7).

The analysis of representativeness (Figure 29) only considers the representativeness of the monitoring sites in one-dimension (i.e., with respect to the variable shown on the x-axis), whereas the true representativeness of the sample needs to be considered within the multi-dimensional space defined by all the independent variables. More complex methodologies exist to determine the reliability of the model predictions by considering the degree to which predictions are based on interpolation or extrapolation (Booker and Whitehead, 2018). Generally, the smaller the training set size the greater degree to which model predictions are based on extrapolation and the lower the overall prediction reliability. However, conducting this type of analysis was beyond the scope of the current project.

The assumption that regression model residuals are normally distributed is a condition shared by many types of statistical model. The robustness of the results and conclusions of a study to violations of these assumptions depends on the questions being asked of the analysis. In the current study, we relied on the prediction confidence interval to provide an accurate representation of the probability that a stated biomass threshold will be exceeded. The accuracy of the prediction confidence intervals will be affected by deviations from the normality of the regression residuals. Our regression residuals were reasonably normally distributed (e.g., Figure 16), but some residual bias was evident particularly for sites with high biomass (e.g., Figure 20). The extent to which these violations of the statistical assumptions are a problem has not been formally evaluated. However, we have undertaken several analyses to validate the derived criteria (see Section 6.6), which provides some level of confidence that violation of the assumptions has not led to gross inaccuracies.

The development of the statistical models produced by this study was guided by a conceptual model that summarises our understanding of periphyton biomass dynamics (Figure 2). However, the representation of the conceptual model by the statistical models is subject to two specific limitations; coarse representation of the controlling processes and use of surrogate measures related to periphyton accrual or loss.

Most of the input variables (including the dependent and independent variables) to the statistical model are single values that represent a summary of temporally variable conditions (e.g., 92nd percentile of chlorophyll, median nutrient concentrations, hydrological indices) calculated from time-series data. This means that the processes described by the conceptual model are represented by coarse-grained data that represent only the characteristic condition of the response and the driving processes at each site. In reality, the processes of periphyton accrual and loss described by Figure 2 occur continuously. The independent variables to the statistical models are therefore representing differences in the average intensity of the processes between sites, rather than the processes themselves.

The independent variables are also surrogate measures of the processes they represent. For example, while high flows can ultimately be responsible for removing biomass, the processes on the riverbed that cause that loss are actually drag caused by high water velocities and scour caused by bed sediment movement (Hoyle *et al.*, 2017; Neverman *et al.*, 2018). The hydrological indices used to represent disturbances that lead to the loss of periphyton are therefore only surrogates for these processes. The use of surrogate independent variables is a limitation that will contribute to unexplained variation.

6.4 Under-predicted high periphyton biomass values

Our models tended to under-predict high values of Chla92 (> 200 mg m^{-2}). Informal assessments indicated that the high biomass that the models under-predicted at some sites could not be attributed to aspects of the periphyton community (as determined from visual assessments of cover on the stream bed), the number of biomass observations or the nature of the bed sediment (e.g., Figure 21). Further inspection of the dataset showed that some sites with high biomass were downstream of waste-water treatment plants (e.g., at least four high biomass sites in the Manawatu - Whanganui region were downstream of treatment plants). In some cases, high Chla92 associated with treatment plants may be attributable to elevated concentrations of NH4-N (e.g., Kilroy, Brown, *et al.*, 2020), and to elevated concentrations of other contaminants (e.g., Aristi et al. 2015). However, in this study, all models were offered median NH_4-N as an independent variable, and it was never included as a significant model term.

The distributions of chlorophyll observations at all sites were approximately exponentially distributed (Figure 10, panel A). This indicates that the temporal dynamics of biomass are similar at all sites but that biomasses are generally (i.e., including the mean and maximum biomass) higher at some sites. It may be that we are unable to account for the high biomass at some sites because high Chla92 is caused by multiple different factors at different sites, including unidentified factors. Thus, it is likely that no single model term could better predict high biomass sites at this stage and more research is needed to understand the drivers of high biomass.

6.5 Relationship between biomass and nutrient concentration

All of our models indicated that the increase in biomass with increasing concentration of all nutrients (i.e., TN, DIN, TP and DRP) is initially high but reaches a biomass ceiling beyond which the response to nutrient concentrations is minor (e.g., OLS models - Figure 17, QR models - Figure 22), or there is no further response (RF models - Figure 27). The biomass ceiling depends on site characteristics other than nutrient concentrations with a higher biomass ceiling where resources other than nutrients (e.g., light and temperature) are high and disturbance is low (e.g., Figure 18).

A nutrient concentration beyond which there is no further increase in the biomass response (i.e., the ceiling) could be interpreted as a “saturating concentration”. For the OLS and QR models, the saturating concentrations were assessed subjectively from the plotted model predictions (Figure 17). Saturating concentrations were subjectively assessed to be approximately 1,000 mg m^{-3} for TN and DIN, approximately 50 mg m^{-3} for TP and approximately 25 mg m^{-3} for DRP (Figure 17). For the RF models, the saturating concentrations for TN and DIN were variable across the Source-of-flow classes (Figure 27) and varied between approximately 500 mg m^{-3} and >1,000 mg m^{-3}. For the OLS and QR models, the assessment of the saturating concentration is strongly influenced by the form of the regression model and in particular the log (base 10) transformation of the nutrient term. This means that for the OLS and QR models, modelled biomass does not actually reach a ceiling with increasing nutrient concentration. Therefore, the saturating concentration suggested above indicates where a “practical” ceiling occurs. The RF model is not constrained in the same way as the OLS and QR models. For the RF model there is a biomass ceiling evident in the plot with a slightly negative association between nutrient concentration and biomass at higher nutrient concentrations (Figure 27). However, the exact location of the saturating concentration is subject to uncertainty because the data are limited at this point (see rug on Figure 27).

It is noted that because all the biomass models include multiple explanatory variables, objective definition of the saturating concentration is considerably more complicated than is the case for simple bivariate relationships. In other studies, identification of significant break points (i.e., changes in rates of change) in simple bivariate relationships between chlorophyll and environmental variables (including nutrient concentrations) was subject to high uncertainty (e.g., Dodds *et al.*, 2002; Dodds, Smith, *et al.*, 2006). Given the high dimensionality of the biomass models (as indicated by the numerous significant explanatory variables) any objective identification of the saturating concentration will have considerable uncertainty.

Saturation of periphyton growth occurs as the nutrient concentration increases to a point at which another growth-critical factor becomes limiting. Although there will be absolute upper limits to saturating concentrations of nitrogen and phosphorus for stream periphyton (e.g., when all other resources are non-limiting), in reality, saturating concentrations will vary because of the influence of other limiting factors such as light and temperature. If phosphorus is in short supply, then a growth response beyond a certain concentration of nitrogen (when phosphorus becomes limiting) is not expected, and vice versa for nitrogen (e.g., Lewis et al., 2008). Thus, the ceiling for Chla92 observed in the OLS and RF relationships could represent the point at which peak biomass at some sites stops responding to nitrogen or phosphorus and starts being limited by some other factor, potential candidates being the alternate nutrient (i.e., phosphorus or nitrogen, respectively). Consideration of the ratios of TN:TP and DIN:DRP (calculated from median values at each site) as very approximate indicators of the limiting nutrient (e.g., as used by McDowell et al. 2009) suggested that limitation by the alternative nutrient *might* explain some cases of lower-than-expected Chla92. However, we note that robust assessment of N or P limitation of periphyton biomass requires more detailed analyses, including field experiments. In the context of the models and derived nutrient criteria, the influence of nutrient limitation in influencing periphyton biomass is less important than the effects on biomass of increasing nutrient concentrations (either N or P) up to levels corresponding to the biomass ceiling. The chances of achieving periphyton objectives will be increased by meeting the nutrient criteria for both N and P.

The nutrient concentrations corresponding to the biomass ceiling are generally consistent with concentrations reported in the international literature to be saturating for periphyton growth. For example, in a meta-analysis of nutrient limitation experiments with periphyton responses measured as the ratio of chlorophyll in enriched versus control treatments, an upper saturation concentration for DIN of about 1400 mg m^{-3} was suggested, although responses to increases in DIN started to level off above about 280 mg m^{-3} (Keck and Lepori, 2012). Using a large dataset, Dodds *et al.* (2006) suggested saturation points for TN and TP corresponding to breakpoints in the relationship between log (maximum chlorophyll) and log TN or log TP, of, respectively 367 – 602 and 27 – 62 mg m^{-3}; maximum chlorophyll at the suggested saturation point for TN was, on average, ~150 mg m^{-2}. Keck and Lepori (2012) observed a flattening off (or breakpoint) in the TP versus chlorophyll response relationship TP between 30 and 60 mg m^{-3}. Peak biomass of periphyton in large accumulations (diatoms and green filamentous algae such as *Stigeoclonium* sp., up to 300 mg m^{-2}) was observed to saturate at DRP concentrations > 25 mg m^{-3}; at higher concentrations other factors limited biomass accrual, such as physical instability of the mat (e.g., spontaneous sloughing) or light limitation within the mat (Bothwell, 1989).

Assuming the nutrient and periphyton data used in the current study were representative of nutrient – biomass relationships in New Zealand's rivers, we expect the decreasing responsiveness of biomass to increasing nutrient concentration, and the nutrient concentrations at which the biomass ceiling occurs, are likely to be repeatable and robust findings of this study (Ardon et al., 2020). This is because the region of decreasing responsiveness and the biomass ceiling is well informed by the data. For example, the ceiling in biomass occurs at around the 80th percentile of the distribution of the TN and DIN values and the 90th percentile of the TP and DRP values (rug plot shown in Figure 18).

The RF model of current biomass nationally also indicated that biomass reaches a ceiling in its response to increasing TN and DIN concentrations (Figure 36). The result of this ceiling is that the model predicted a relatively small proportion of the network currently has an appreciable risk of exceeding the 200 mg m^{-2} biomass threshold (Figure 39).

We know from the dataset that a proportion of sites do exceed a biomass threshold of 200 mg m^{-2} by a large margin (Figure 8). However, our model indicates that these large biomass sites are not due to high nutrient concentrations (see Figure 12). We were unable to explain why these sites have such high biomass from the available data (e.g., Figure 21), as discussed above.

6.6 Credibility of the nutrient criteria

Confidence in the nutrient criteria is based in part on their consistency with the conceptual model (Figure 2). Consistent with expectations, concentration criteria are higher in environments with lower resources and higher disturbance. For example, climate categories that have lower temperatures and solar radiation (i.e., CD, CW and CX) have higher concentration criteria than climate categories with higher lower temperatures and solar radiation (i.e., WD, WW and WX). This is consistent with the conceptual model because where non-nutrient resources are lower, we expect biomass accrual to be lower and therefore higher nutrients are required to achieve a given biomass threshold. Concentration criteria are also generally higher in climate categories that have more frequent high flows (e.g., CD < CW < CX). Concentration criteria also increased systematically across topography categories that have more frequent high flows (e.g., Lk < L < GM < M < H). These patterns are consistent with the conceptual model because where disturbance is higher, we expect biomass loss to occur more frequently and therefore higher nutrients are required to achieve a given biomass threshold within the resulting shorter accrual periods. Our models also produced higher concentration criteria for shaded sites than unshaded sites. This is consistent with the conceptual model because where shading reduces light and temperature and where these resources are lower, we expect biomass accrual to be lower.

We validated the derived criteria using the only other national-scale dataset of periphyton abundance that is currently available. These data are based on visual observation of cover at NRWQN sites but were converted to an equivalent measure of the 92nd percentile of chlorophyll based on an empirical model. The validation essentially used the criteria to predict the proportion of under-protected sites in the independent dataset. There were similar proportions of under-protected sites in the independent dataset compared to the under-protection risks indicated for the criteria (Figure 34). This was true for all four nutrient forms and levels of under-protection risk (Table 17). This validation therefore provides confidence that the criteria perform well. It is noted that the validation exercise includes unquantified uncertainties associated with the criteria themselves and with the estimates of Chla92 values for the validation sites.

This study had an objective of exploring the use of quantile regression (QR) instead of OLS to define nutrient criteria. Arguably, criteria derived using QR models have the advantage of being easier to explain because the level of under-protection risk is directly linked to the quantile being modelled (Figure 5). However, we found that QR models could not be used to define criteria for all Source-of-flow classes because fitting an accurate multivariable QR model to the extreme percentiles (i.e., 70%, 80% and 90%) required a larger sample size than was available in this study.

There were some Source-of-flow classes in our dataset with ≥30 sites for which we were able use a QR modelling approach based on simple bivariate relationships (i.e., Chla92 versus nutrient concentration). Some of these models were statistically significant and consistent with the conceptual model (i.e., there were positive relationships between Chla92 and nutrient concentration). For those cases where we were able to use QR modelling, the nutrient criteria we derived were reasonably comparable with those derived using the OLS models (see Appendix B for details). Although limited in scope, the comparability of the criteria derived using the QR models to those based on the OLS is an extra validation exercise that provides confidence in our results.

Because the OLS models are parametric and represent biomass as a log-transformed function of nutrient concentration, some of our derived criteria represent concentrations that are higher than the concentrations associated with the biomass ceiling discussed in Section 6.5. In addition, for some combinations of nutrient, Source-of-flow class and level of under-protection risk the model predictions would not exceed the biomass threshold even at the maximum of the range of observed site median nutrient concentrations represented in our dataset (see examples in Figure 19 for 20% under-protection risk). For these combinations, our criteria are appreciably higher than the saturating concentrations of approximately 1,000 mg m^{-3} for TN and DIN, 50 mg m^{-3} for TP and 25 mg m^{-3} for DRP and, where necessary, were limited to the maximum of the relevant observed site median nutrient concentrations.

For all nutrients, the derived criteria typically had high values, relative to saturating concentrations, for Source-of-flow classes that represent rivers with strong physical controls on periphyton biomass, such as high flow variability, unstable bed sediment, and low temperature (e.g., CX/GM. CX/H, CW/M, CW/H). These river tend to have relatively low nutrient concentrations (Whitehead, 2018) and low biomass (Kilroy *et al.*, 2019). Therefore, the sites in our dataset that represent these types of river did not cover a wide range of nutrient concentrations. This means that the models were poorly informed about the nutrient – biomass relationship at sites with strong physical controls on periphyton biomass and the derived criteria should be interpreted cautiously. In other words, the model was poorly informed about how chlorophyll might respond to high nutrient concentrations at sites where periphyton was strongly controlled by physical factors. It is possible that physical conditions limit maximum biomass in rivers with strong physical controls even when the nutrient concentrations are high. However, because our method involved extrapolation of the model into environmental conditions that are poorly represented by the fitting data, the uncertainty of criteria for Source-of-flow classes that represent rivers with strong physical controls on periphyton biomass is particularly high.

We consider there is high uncertainty about whether biomass can be managed by restricting nutrient concentrations to levels greater than the saturating concentrations. This point was made in one of the few examples we know of in the literature where nutrient controls were demonstrated to achieve reductions in periphyton biomass in a river; in that study it was concluded that: "...nutrient reductions in rivers can be successful in controlling algal biomass but require achievement of concentrations below saturation and likely close to natural background” (Suplee *et al.*, 2012). In other words, where a biomass threshold is exceeded at a site, it is likely that biomass reduction can be expected only if nutrient concentrations are reduced to below the saturating value. Nutrient criteria higher than the saturating values indicate combinations of conditions where periphyton biomass is strongly controlled by non-nutrient factors. Under such conditions, even when nutrient concentrations are greater than saturating levels, the biomass threshold(s) may not be exceeded.

6.7 Criteria uncertainty

The under-protection risk should not be confused with the uncertainty of the nutrient criteria. Snelder *et al.* (2019) used Monte Carlo simulations to quantify the uncertainty associated with the nutrient criteria derived in their study. We have not quantified criteria uncertainty in this study, but we acknowledge that the criteria are uncertain. By this we mean that, for a stated under-protection risk, we do not know exactly what is the true proportion of sites that will exceed the target biomass threshold when concentrations are held to the criteria.

A key reason that we have not attempted to define the uncertainty of the criteria is that our models were unable to explain sites with high biomass. This means that we have an imperfect understanding of the error distribution of our models and cannot accurately represent this error distribution in an analysis such as a Monte Carlo simulation.

The inability of our models to explain sites with high biomass impacts on the definition of the criteria. This is because the model residuals are not perfectly normally distributed (e.g., Figure 16) and therefore there are uncertainties associated with the predicted biomass probability distribution. We consider that these uncertainties are irreducible with current data and knowledge. However, the validation exercise indicates that violation of the assumptions has not led to gross inaccuracies. We therefore do not consider that the uncertainties should be a barrier to using the criteria. Furthermore, we recommend that the choice of level of under-protection risk should reflect the significance of the resources being managed and the consequences of exceeding the target biomass thresholds.

6.8 National predictions

Predictions from an RF model of Chla92 across all segments in the network highlighted that the risk of exceeding the biomass threshold defining the NOF bottom line is low (Figure 39). This is a similar result to that of Kilroy *et al.* (2019). Because the present study’s model was not able to accurately predict chlorophyll observations at the high-biomass sites, and there were limited numbers of high-biomass sites in our dataset, we need to consider the estimate of locations that are below the NOF bottom line as uncertain. However, the results of the national predictions and the identification of the biomass ceiling, indicates that the NOF bottom line (i.e., Chla92 > 200 mg/m^2) is a high level of biomass at any site and is rarely exceeded.

7 Conclusions

In this study we developed look-up tables for nutrient criteria for target periphyton attribute states for the 21 Source-of-flow classes that encompass all New Zealand rivers. The criteria include four levels of under-protection risk that quantify the proportion of sites for which biomass is predicted to exceed the stated target despite compliance with the nutrient criteria. The level of risk of under-protection that is used to when the criteria are applied is not scientifically defined and is a subjective ('normative') choice.

The criteria are uncertain in that, for a stated under-protection risk, we do not know the precise proportion of sites that will exceed the target biomass threshold when concentrations are held to the criteria. However, validation of these derived criteria using an independent dataset indicates that they perform well. We therefore do not consider that the uncertainties should be a barrier to using the criteria.

The models developed in this study indicate that there is an initially high rate of increase in periphyton biomass with increasing nutrient concentrations for each of the nutrients we considered (TN, DIN, TP and DRP), but periphyton biomass reaches a "ceiling" beyond which there is no further response to increasing nutrient concentrations. Our study indicates that from a practical perspective, in most REC Source-of-flow classes a 92nd percentile chlorophyll biomass of 200 mg m^{-2} would rarely be exceeded because of nutrient enrichment. We refer to the nutrient concentration beyond which there is no further biomass response as the "saturating concentration". The models developed by this study indicated that the saturating concentrations are approximately 1,000 mg m^{-3} for TN and DIN, approximately 50 mg m^{-3} for TP and approximately 25 mg m^{-3} for DRP. The stated saturating concentrations are approximate because they were derived subjectively from simplified graphical representations of the models defined by this study. We are not aware of statistical methods that could be used to identify the saturating concentrations based on our models more objectively. If such methods do exist, any estimate of the saturating concentration will have large uncertainties because our models have considerable unexplained variation.

Acknowledgements

The authors thank the field and laboratory staff of regional councils who collect and helped collate the monitoring data used in this study. The authors thank Roger Hodson, Shirley Hayward, Mike Patterson, Sandy Haidekker, Mike Scarsbrook, Adam Canning, Isaac Bain, Karwin Perez and Kohji Muraoka for critical review and feedback during the study process. We thank Scott Larned and David Wood for independent reviews of the report.

References

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

Ardón, M., L.H. Zeglin, R.M. Utz, S.D. Cooper, W.K. Dodds, R.J. Bixby, A.S. Burdett, J. Follstad Shah, N.A. Griffiths, and T.K. Harms, 2021. Experimental Nitrogen and Phosphorus Enrichment Stimulates Multiple Trophic Levels of Algal and Detrital-Based Food Webs: A Global Meta-Analysis from Streams and Rivers. Biological Reviews 96:692–715.

Austin, B.J., N. Hardgrave, E. Inlander, C. Gallipeau, S. Entrekin, and M.A. Evans-White, 2015. Stream Primary Producers Relate Positively to Watershed Natural Gas Measures in North-Central Arkansas Streams. Science of the Total Environment 529:54–64.

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

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

Biggs, B.J. and C. Kilroy, 2000. Stream Periphyton Monitoring Manual. Niwa.

Booker, D.J., 2013. Spatial and Temporal Patterns in the Frequency of Events Exceeding Three Times the Median Flow (FRE3) across New Zealand. Journal of Hydrology (New Zealand):15–39.

Booker, D.J. and T.H. Snelder, 2012. Comparing Methods for Estimating Flow Duration Curves at Ungauged Sites. Journal of Hydrology 434:78–94.

Booker, D.J. and A.L. Whitehead, 2018. Inside or Outside: Quantifying Extrapolation across River Networks. Water Resources Research 54:6983–7003.

Bothwell, M.L., 1989. Phosphorus–Limited Growth Dynamics of Lotic Periphytic Diatom Communities: Areal Biomass and Cellular Growth Rate Responses. Canadian Journal of Fisheries and Aquatic Sciences 46:1293–1301.

Breiman, L., 2001. Random Forests. Machine Learning 45:5–32.

Breiman, L., J.H. Friedman, R. Olshen, and C.J. Stone, 1984. Classification and Regression Trees. Wadsworth, Belmont, California.

Cade, B.S. and B.R. Noon, 2003. A Gentle Introduction to Quantile Regression for Ecologists. Frontiers in Ecology and the Environment 1:412–420.

Carr, G.M., P.A. Chambers, and A. Morin, 2005. Periphyton, Water Quality, and Land Use at Multiple Spatial Scales in Alberta Rivers. Canadian Journal of Fisheries and Aquatic Sciences 62:1309–1319.

Clapcott, J.E., R.G. Young, J.S. Harding, C.D. Matthaei, J.M. Quinn, and R.G. Death, 2011. Sediment Assessment Methods: Protocols and Guidelines for Assessing the Effects of Deposited Fine Sediment on in-Stream Values. Nelson, New Zealand.

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

Cutler, D.R., J.T.C. Edwards, K.H. Beard, A. Cutler, K.T. Hess, J. Gibson, and J.J. Lawler, 2007a. Random Forests for Classification in Ecology. Ecology 88:2783–2792.

Cutler, D.R., J.T.C. Edwards, K.H. Beard, A. Cutler, K.T. Hess, J. Gibson, and J.J. Lawler, 2007b. Random Forests for Classification in Ecology. Ecology 88:2783–2792.

Davies-Colley, R.J., D.G. Smith, R.C. Ward, G.G. Bryers, G.B. McBride, J.M. Quinn, and M.R. Scarsbrook, 2011. Twenty Years of New Zealand's National Rivers Water Quality Network: Benefits of Careful Design and Consistent Operation 1. JAWRA Journal of the American Water Resources Association 47:750–771.

Dodds, W.K., 2006. Eutrophication and Trophic State in Rivers and Streams. Limnology and Oceanography:671–680.

Dodds, W.K., 2007. Trophic State, Eutrophication and Nutrient Criteria in Streams. Trends in Ecology & Evolution 22:669–676.

Dodds, W.K., E. Carney, and R.T. Angelo, 2006. Determining Ecoregional and Reference Conditions for Nutrients, Secchi Depth and Chlorophyll a in Kansas Lakes and Reservoirs. Lake and Reservoir Management 22:151–159.

Dodds, W.K., V.H. Smith, and K. Lohman, 2002. Nitrogen and Phosphorus Relationships to Benthic Algal Biomass in Temperate Streams. Canadian Journal of Fisheries and Aquatic Sciences 59:865–874.

Dodds, W.K., V.H. Smith, and K. Lohman, 2006. Erratum: Nitrogen and Phosphorus Relationships to Benthic Algal Biomass in Temperate Streams. Canadian Journal of Fisheries and Aquatic Sciences 63:1190–1191.

Dodds, W.K.K. and E.B. Welch, 2000. Establishing Nutrient Criteria in Streams. Journal of the North American Benthological Society 19:186–196.

Duan, N., 1983. Smearing Estimate: A Nonparametric Retransformation Method. Journal of the American Statistical Association 78:605–610.

Edgar, N., A. Meredith, P. Empson, and B. Moore, 1994. Regional Ecological Monitoring of Streams in the Waikato (R.E.M.S.). Environmental Quality Section, Environment Waikato.

Francoeur, S.N., B.J. Biggs, R.A. Smith, and R.L. Lowe, 1999. Nutrient Limitation of Algal Biomass Accrual in Streams: Seasonal Patterns and a Comparison of Methods. Journal of the North American Benthological Society 18:242–260.

Haddadchi, A., A. Kuczynski, J.T. Hoyle, C. Kilroy, D.J. Booker and M. Hicks (2020) Periphyton removal flows determined by sediment entrainment thresholds. Ecological Modelling, 434: 16. 10.1016/j.ecolmodel.2020.109263

Hastie, T., R. Tibshirani, and J.H. Friedman, 2001. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer-Verlag, New York.

Hoyle, J.T., C. Kilroy, D.M. Hicks, and L. Brown, 2017. The Influence of Sediment Mobility and Channel Geomorphology on Periphyton Abundance. Freshwater Biology 62:258–273.

Keck, F. and F. Lepori, 2012. Can We Predict Nutrient Limitation in Streams and Rivers? Freshwater Biology 57:1410–1421.

Kilroy, C., L. Brown, L. Carlin, P. Lambert, A. Sinton, J.A. Wech, and C. Howard-Williams, 2020. Nitrogen Stimulation of Periphyton Biomass in Rivers: Differential Effects of Ammonium-N and Nitrate-N. Freshwater Science 39:485–496.

Kilroy, C., M.T. Greenwood, J. Wech, T. Stephens, L. Brown, A. Mathews, M. Patterson, and M. Patterson, 2018. Periphyton - Environment Relationships in the Horizons Region. Analysis of a Seven-Year Dataset. NIWA Client Report, Christchurch, New Zealand.

Kilroy, C., T. Snelder, and R. Stoffels, 2020. Periphyton - Environment Relationships in the Bay of Plenty: Analysis of Data from 2015 to 2019. Prepared for Bay of Plenty Regional Council. NIWA Client Report, NIWA, Christchurch, New Zealand.

Kilroy, C. and R. Stoffels, 2019. Periphyton Growth in Northland Rivers (2019) Current Status, and Development of Relationships for Nutrient Limit-Setting. NIWA Client Report, INIWA, Christchurch, New Zealand.

Kilroy, C., A. Whitehead, S. Howard, and M.T. Greenwood, 2019. Modelling Periphyton in New Zealand Rivers Part 1. An Analysis of Current Data and Development of National Predictions. NIWA Client Report, NIWA, Christchurch, New Zealand.

Kuczynski, A., 2019. Modelling Periphyton in New Zealand Rivers. Part 2. A Review and Prospects for Mechanistic Modelling. NIWA Client Report, NIWA, Christchurch, New Zealand.

Larned, S.T., T. Snelder, M.J. Unwin, and G.B. McBride, 2016. Water Quality in New Zealand Rivers: Current State and Trends. New Zealand Journal of Marine and Freshwater Research 50:389–417.

Larned, S., T. Snelder, M. Unwin, G. McBride, P. Verburg, and H. McMillan, 2015. Analysis of Water Quality in New Zealand Lakes and Rivers. Prepared for the Ministry for the Environment. Wellington: Ministry for the Environment.

Leathwick, J.R., J.M. Overton, and M. McLeod, 2003. An Environmental Domain Analysis of New Zealand, and Its Application to Biodiversity Conservation. Conservation Biology 17:1612–1623.

Leathwick, J., D. West, L. Chadderton, P. Gerbeaux, D. Kelly, H. Robertson, and D. Brown, 2010. Freshwater Ecosystems of New Zealand (FENZ) Geodatabase: Version One User Guide. Department of Conservation, Hamilton, New Zealand.

Lewis Jr, W.M., J.F. Saunders III, and J.H. McCutchan Jr, 2008. Application of a Nutrient-Saturation Concept to the Control of Algal Growth in Lakes. Lake and Reservoir Management 24:41–46.

Lewis, W.M. and J.H. McCutchan, 2010. Ecological Responses to Nutrients in Streams and Rivers of the Colorado Mountains and Foothills. Freshwater Biology 55:1973–1983.

Lewis, J.R., M. William, J.R. Mccutchan, and H. James, 2010. Ecological Responses to Nutrients in Streams and Rivers of the Colorado Mountains and Foothills. Freshwater Biology 55:1973–1983.

Matheson, F., J. Quinn, and C. Hickey, 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. NIWA Client Report, NIWA, Hamilton, New Zealand.

Matheson, F., J. Quinn, and M. Unwin, 2016. Instream Plant and Nutrient Guidelines. Review and Development of an Extended Decision-Making Framework Phase 3. NIWA Client Report, NIWA, Hamilton, New Zealand.

McDowell, R.W., S.T. Larned, and D.J. Houlbrooke, 2009. Nitrogen and Phosphorus in New Zealand Streams and Rivers: Control and Impact of Eutrophication and the Influence of Land Management. New Zealand Journal of Marine and Freshwater Research 43:985–995.

Messager, M.L., B. Lehner, C. Cockburn, N. Lamouroux, H. Pella, T. Snelder, K. Tockner, T. Trautmann, C. Watt, and T. Datry, 2021. Global Prevalence of Non-Perennial Rivers and Streams. Nature 594:391–397.

Ministry for Environment, 2000. New Zealand Periphyton Guideline: Detecting, Monitoring and Managing Enrichment of Streams. Wellington.

Moriasi, D.N., M.W. Gitau, N. Pai, and P. Daggupati, 2015. Hydrologic and Water Quality Models: Performance Measures and Evaluation Criteria. Transactions of the ASABE 58:1763–1785.

Munn, M., J. Frey, and A. Tesoriero, 2010. The Influence of Nutrients and Physical Habitat in Regulating Algal Biomass in Agricultural Streams. Environmental Management 45:603–615.

Neverman, A.J., R.G. Death, I.C. Fuller, R. Singh, and J.N. Procter, 2018. Towards Mechanistic Hydrological Limits: A Literature Synthesis to Improve the Study of Direct Linkages between Sediment Transport and Periphyton Accrual in Gravel-Bed Rivers. Environmental Management 62:740–755.

Olden, J.D. and N.L. Poff, 2003. Redundancy and the Choice of Hydrologic Indices for Characterizing Streamflow Regimes. River Research & Applications 19:101–121.

Peng, B. and L. Wang, 2015. An Iterative Coordinate Descent Algorithm for High-Dimensional Nonconvex Penalized Quantile Regression. Journal of Computational and Graphical Statistics 24:676–694.

Phillips, G., M. Kelly, H. Teixeira, F. Salas, G. Free, W. Leujak, A.L. Solheim, G. Várbíró, and S. Poikane, 2018. Best Practice for Establishing Nutrient Concentrations to Support Good Ecological Status.

Piñeiro, G., S. Perelman, J. Guerschman, and J. Paruelo, 2008. How to Evaluate Models: Observed vs. Predicted or Predicted vs. Observed? Ecological Modelling 216:316–322.

Redfield, A.C., 1958. The Biological Control of Chemical Factors in the Environment. American Scientist 46:230A–221.

Shmueli, G., 2010. To Explain or to Predict? Statistical Science 25:289–310.

Smith, D.G. and G.B. McBride, 1990. New Zealand's National Water Quality Monitoring Network - Design and First Year's Operation. Water Resources Bulletin 26:767–775.

Smith, A.J. and C.P. Tran, 2010. A Weight-of-Evidence Approach to Define Nutrient Criteria Protective of Aquatic Life in Large Rivers. Journal of the North American Benthological Society 29:875–891.

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

Snelder, T., B. Biggs, C. Kilroy, and D. Booker, 2013. National Objective Framework for Periphyton. NIWA Client Report, NIWA, NIWA, Christchurch.

Snelder, T.H., D.J. Booker, J.M. Quinn, and C. Kilroy, 2014. Predicting Periphyton Cover Frequency Distributions across New Zealand's Rivers. JAWRA Journal of the American Water Resources Association 50:111–127.

Snelder, T.H. and D. J. Booker, 2013. Natural Flow Regime Classifications Are Sensitive to Definition Procedures. River Research and Applications 29:822–838.

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

Suplee, M.W., K.F. Flynn, and S.C. Chapra, 2015. Model-Based Nitrogen and Phosphorus (Nutrient) Criteria for Large Temperate Rivers: 2. Criteria Derivation. JAWRA Journal of the American Water Resources Association 51:447–470.

Suplee, M.W., V. Watson, W.K. Dodds, and C. Shirley, 2012. Response of Algal Biomass to Large-Scale Nutrient Controls in the Clark Fork River, Montana, United States 1. JAWRA Journal of the American Water Resources Association 48:1008–1021.

Svetnik, V., A. Liaw, C. Tong, and T. Wang, 2004. Application of Breiman's Random Forest to Modeling Structure-Activity Relationships of Pharmaceutical Molecules. Multiple Classifier Systems:334–343.

Taylor, S.L., S.C. Roberts, C.J. Walsh, and B.E. Hatt, 2004. Catchment Urbanisation and Increased Benthic Algal Biomass in Streams: Linking Mechanisms to Management. Freshwater Biology 49:835–851.

Unwin, M., T. Snelder, D. Booker, D. Ballantine, and J. Lessard, 2010. Predicting Water Quality in New Zealand Rivers from Catchment-Scale Physical, Hydrological and Land Cover Descriptors Using Random Forest Models. NIWA Client Report: CHC2010-0.

Urrea-Clos, G., E. García-Berthou, and S. Sabater, 2014. Factors Explaining the Patterns of Benthic Chlorophyll-a Distribution in a Large Agricultural Iberian Watershed (Guadiana River). Ecological Indicators 36:463–469.

USEPA, 2010. Using Stressor-Response Relationships to Derive Numeric Nutrient Criteria. Technical Report, United States Environmental Protection Agency, Washington D.C.

Van Nieuwenhuyse, E.E. and J.R. Jones, 1996. Phosphorus-Chlorophyll Relationship in Temperate Streams and Its Variation with Stream Catchment Area. Canadian Journal of Fisheries and Aquatic Sciences 53:99–105.

Welch, E.B., J.M. Jacoby, R.R. Horner, and M.R. Seeley, 1988. Nuisance Biomass Levels of Periphytic Algae in Streams. Hydrobiologia 157:161–168.

Whitehead, A., 2018. Spatial Modelling of River Water-Quality State,. Incorporating Monitoring Data from 2013 to 2017. NIWA Client Report, NIWA, Christchurch, New Zealand.

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

Wu, Y. and Y. Liu, 2009. Variable Selection in Quantile Regression. Statistica Sinica:801–817.

Zuur, A.F., E.N. Ieno, and C.S. Elphick, 2010. A Protocol for Data Exploration to Avoid Common Statistical Problems. Methods in Ecology and Evolution 1:3–14.

Appendix A Nutrient concentration criteria for periphyton biomass thresholds derived from the OLS models

Table 19. Criteria derived from the OLS model for TN for three levels of under-protection risk (5%, 10% and 15%) at shaded and unshaded locations. Criteria greater than the saturating concentration of approximately 1,000 mg m^{-3} indicate combinations of conditions where periphyton biomass is strongly controlled by non-nutrient factors. Under such conditions, even when nutrient concentrations exceed saturating levels, the biomass threshold(s) may not be exceeded.

Under-protection risk	SoF	Unshaded 50 mg m^{-2}	Unshaded 120 mg m^{-2}	Unshaded 200 mg m^{-2}	Shaded 50 mg m^{-2}	Shaded 120 mg m^{-2}	Shaded 200 mg m^{-2}
5	CX/GM	6	570	4,176	32	2,335	4,496
5	CX/M	4	382	3,599	21	1,599	4,457
5	CX/H	3	252	3,047	14	1,120	4,050
5	CX/L	2	195	3,127	10	891	4,254
5	CX/Lk	2	219	2,799	11	975	4,055
5	CW/GM	1	66	1,390	3	303	4,342
5	CW/M	2	154	2,135	8	657	4,200
5	CW/H	1	47	885	2	216	2,632
5	CW/L	1	28	625	2	134	2,034
5	CW/Lk	1	56	1,183	3	263	3,400
5	CD/M	1	121	1,934	6	545	4,127
5	CD/H	1	57	910	3	258	2,315
5	CD/L	1	7	182	1	36	806
5	CD/Lk	1	61	1,272	3	283	2,972
5	WX/L	1	33	749	2	159	2,753
5	WX/H	1	45	1,009	2	216	3,116
5	WW/H	1	26	609	1	127	2,383
5	WW/L	1	23	541	1	110	2,006
5	WW/Lk	1	36	868	2	176	2,752
5	WD/L	1	13	312	1	62	1,048
5	WD/Lk	1	27	662	2	132	1,947
10	CX/GM	27	2,089	4,490	130	4,072	4,500
10	CX/M	18	1,416	4,432	86	3,426	4,500
10	CX/H	12	978	3,980	57	2,888	4,436
10	CX/L	9	774	4,215	43	2,938	4,432
10	CX/Lk	10	834	3,998	47	2,595	4,436
10	CW/GM	2	254	4,115	13	1,137	4,500
10	CW/M	6	568	4,090	32	1,867	4,496
10	CW/H	2	183	2,419	9	735	3,975
10	CW/L	1	112	1,833	6	512	3,460
10	CW/Lk	2	221	3,142	11	981	4,306
10	CD/M	5	466	4,006	24	1,682	4,495
10	CD/H	3	223	2,155	12	787	3,576
10	CD/L	1	31	694	2	147	2,136
10	CD/Lk	3	241	2,765	12	1,073	4,134
10	WX/L	1	134	2,476	6	614	4,209
10	WX/H	2	179	2,926	9	819	4,167
10	WW/H	1	105	2,109	5	489	3,953
10	WW/L	1	95	1,786	4	444	3,580
10	WW/Lk	2	162	2,573	7	750	4,143
10	WD/L	1	61	976	3	285	2,051
10	WD/Lk	2	121	1,805	6	570	3,447
15	CX/GM	70	3,498	4,500	327	4,420	4,500
15	CX/M	46	2,639	4,498	217	4,188	4,500
15	CX/H	30	2,113	4,303	145	3,626	4,499
15	CX/L	23	1,894	4,368	112	3,994	4,496
15	CX/Lk	25	1,846	4,327	118	3,599	4,500
15	CW/GM	7	623	4,500	33	2,580	4,500
15	CW/M	17	1,175	4,457	80	3,255	4,500
15	CW/H	5	420	3,512	24	1,587	4,356
15	CW/L	3	278	2,936	14	1,162	3,924
15	CW/Lk	6	545	4,170	28	2,045	4,437
15	CD/M	13	1,057	4,442	62	3,120	4,500
15	CD/H	6	497	3,054	30	1,486	4,128
15	CD/L	1	79	1,480	4	368	2,964
15	CD/Lk	6	600	3,812	31	2,038	4,401
15	WX/L	3	335	3,722	17	1,447	4,491
15	WX/H	5	441	3,732	22	1,946	4,500
15	WW/H	3	263	3,517	13	1,198	4,379
15	WW/L	3	245	3,071	12	1,048	3,937
15	WW/Lk	4	434	3,790	21	1,645	4,362
15	WD/L	2	173	1,654	8	633	2,633
15	WD/Lk	3	323	2,869	15	1,237	3,941

Table 20. Criteria derived from the OLS model for TN for three levels of under-protection risk (20%, 30% and 50%) at shaded and unshaded locations. Criteria greater than the saturating concentration of approximately 1,000 mg m^{-3} indicate combinations of conditions where periphyton biomass is strongly controlled by non-nutrient factors. Under such conditions, even when nutrient concentrations exceed saturating levels, the biomass threshold(s) may not be exceeded.

Under-protection risk	SoF	Unshaded 50 mg m^{-2}	Unshaded 120 mg m^{-2}	Unshaded 200 mg m^{-2}	Shaded 50 mg m^{-2}	Shaded 120 mg m^{-2}	Shaded 200 mg m^{-2}
20	CX/GM	146	4,153	4,500	678	4,495	4,500
20	CX/M	97	3,554	4,500	451	4,449	4,500
20	CX/H	64	3,026	4,454	302	4,019	4,500
20	CX/L	50	3,205	4,447	237	4,236	4,500
20	CX/Lk	52	2,749	4,457	244	4,033	4,500
20	CW/GM	14	1,269	4,500	68	4,271	4,500
20	CW/M	35	2,017	4,499	165	4,155	4,500
20	CW/H	11	809	4,044	50	2,517	4,459
20	CW/L	6	563	3,527	30	1,917	4,192
20	CW/Lk	12	1,088	4,326	59	3,284	4,490
20	CD/M	27	1,819	4,497	130	4,073	4,500
20	CD/H	13	858	3,668	63	2,239	4,365
20	CD/L	2	164	2,252	8	744	3,529
20	CD/Lk	14	1,202	4,184	66	2,864	4,468
20	WX/L	7	680	4,273	36	2,611	4,500
20	WX/H	10	899	4,250	47	3,016	4,500
20	WW/H	6	538	4,007	27	2,225	4,495
20	WW/L	5	505	3,642	26	1,892	4,132
20	WW/Lk	10	912	4,218	47	2,807	4,433
20	WD/L	4	359	2,154	20	1,063	3,089
20	WD/Lk	7	694	3,527	35	1,901	4,146
30	CX/GM	487	4,472	4,500	2,097	4,500	4,500
30	CX/M	323	4,364	4,500	1,409	4,500	4,500
30	CX/H	217	3,849	4,500	999	4,389	4,500
30	CX/L	170	4,146	4,500	796	4,405	4,500
30	CX/Lk	172	3,883	4,500	796	4,406	4,500
30	CW/GM	47	3,517	4,500	221	4,500	4,500
30	CW/M	115	3,819	4,500	518	4,487	4,500
30	CW/H	35	2,058	4,427	164	3,799	4,499
30	CW/L	21	1,519	4,073	98	3,257	4,428
30	CW/Lk	41	2,666	4,477	194	4,267	4,500
30	CD/M	91	3,730	4,500	423	4,488	4,500
30	CD/H	45	1,878	4,278	206	3,363	4,483
30	CD/L	6	533	3,274	28	1,854	4,158
30	CD/Lk	46	2,433	4,446	218	4,025	4,500
30	WX/L	25	2,002	4,500	119	4,016	4,500
30	WX/H	33	2,549	4,500	154	3,972	4,500
30	WW/H	19	1,670	4,475	91	3,803	4,500
30	WW/L	19	1,443	4,047	89	3,409	4,338
30	WW/Lk	37	2,292	4,414	174	4,123	4,493
30	WD/L	17	871	2,874	79	1,915	3,726
30	WD/Lk	26	1,597	4,066	126	3,254	4,303
50	CX/GM	2,813	4,500	4,500	4,265	4,500	4,500
50	CX/M	1,953	4,500	4,500	3,788	4,500	4,500
50	CX/H	1,530	4,458	4,500	3,249	4,500	4,500
50	CX/L	1,288	4,454	4,500	3,629	4,500	4,500
50	CX/Lk	1,204	4,470	4,500	3,031	4,500	4,500
50	CW/GM	329	4,500	4,500	1,531	4,500	4,500
50	CW/M	724	4,499	4,500	2,319	4,500	4,500
50	CW/H	237	4,059	4,500	962	4,462	4,500
50	CW/L	146	3,537	4,462	675	4,204	4,500
50	CW/Lk	290	4,334	4,500	1,286	4,496	4,500
50	CD/M	619	4,498	4,500	2,109	4,500	4,500
50	CD/H	305	3,691	4,497	998	4,374	4,500
50	CD/L	42	2,280	4,298	200	3,558	4,492
50	CD/Lk	333	4,203	4,500	1,434	4,471	4,500
50	WX/L	180	4,285	4,500	821	4,500	4,500
50	WX/H	228	4,269	4,500	1,074	4,500	4,500
50	WW/H	135	4,013	4,500	637	4,496	4,500
50	WW/L	141	3,658	4,387	624	4,143	4,489
50	WW/Lk	313	4,245	4,500	1,251	4,440	4,500
50	WD/L	155	2,206	3,925	472	3,133	4,416
50	WD/Lk	214	3,550	4,346	889	4,157	4,500

Table 21. Criteria derived from the OLS model for DIN for three levels of under-protection risk (5%, 10% and 20%) at shaded and unshaded locations. Criteria greater than the saturating concentration of approximately 1,000 mg m^{-3} indicate combinations of conditions where periphyton biomass is strongly controlled by non-nutrient factors. Under such conditions, even when nutrient concentrations exceed saturating levels, the biomass threshold(s) may not be exceeded.

Under-protection risk	SoF	Unshaded 50 mg m^{-2}	Unshaded 120 mg m^{-2}	Unshaded 200 mg m^{-2}	Shaded 50 mg m^{-2}	Shaded 120 mg m^{-2}	Shaded 200 mg m^{-2}
5	CX/GM	2	514	3,674	12	2,563	3,800
5	CX/M	1	323	3,380	7	1,782	3,793
5	CX/H	1	186	2,836	4	1,310	3,532
5	CX/L	1	112	3,077	2	918	3,643
5	CX/Lk	1	156	2,759	3	1,128	3,590
5	CW/GM	1	27	1,536	1	233	3,796
5	CW/M	1	93	2,217	2	600	3,715
5	CW/H	1	21	845	1	155	2,572
5	CW/L	1	9	543	1	80	1,941
5	CW/Lk	1	23	1,161	1	201	3,356
5	CD/M	1	58	1,923	2	452	3,652
5	CD/H	1	26	754	1	188	2,070
5	CD/L	1	2	87	1	11	636
5	CD/Lk	1	21	1,150	1	185	2,784
5	WX/L	1	9	611	1	86	2,655
5	WX/H	1	14	938	1	128	2,851
5	WW/H	1	7	466	1	61	2,437
5	WW/L	1	5	346	1	45	1,840
5	WW/Lk	1	8	600	1	75	2,495
5	WD/L	1	3	155	1	19	759
5	WD/Lk	1	5	366	1	45	1,494
10	CX/GM	8	2,254	3,800	77	3,638	3,800
10	CX/M	5	1,508	3,787	48	3,296	3,800
10	CX/H	3	1,067	3,465	27	2,765	3,781
10	CX/L	2	709	3,612	16	3,005	3,781
10	CX/Lk	3	894	3,520	21	2,638	3,797
10	CW/GM	1	170	3,787	3	1,297	3,800
10	CW/M	2	477	3,646	12	2,005	3,800
10	CW/H	1	118	2,322	3	725	3,490
10	CW/L	1	58	1,677	2	461	2,987
10	CW/Lk	1	146	3,150	3	1,008	3,688
10	CD/M	1	347	3,563	7	1,701	3,800
10	CD/H	1	148	1,882	4	671	3,124
10	CD/L	1	8	497	1	71	1,846
10	CD/Lk	1	138	2,533	3	1,012	3,569
10	WX/L	1	62	2,385	2	510	3,688
10	WX/H	1	91	2,672	2	774	3,711
10	WW/H	1	44	2,095	1	377	3,412
10	WW/L	1	34	1,545	1	289	3,050
10	WW/Lk	1	61	2,297	1	529	3,538
10	WD/L	1	18	666	1	151	1,619
10	WD/Lk	1	38	1,311	1	330	2,882
15	CX/GM	30	3,296	3,800	267	3,788	3,800
15	CX/M	19	2,681	3,800	167	3,725	3,800
15	CX/H	11	2,199	3,700	97	3,254	3,800
15	CX/L	6	2,039	3,713	58	3,487	3,800
15	CX/Lk	8	1,911	3,738	75	3,315	3,800
15	CW/GM	1	573	3,800	11	3,279	3,800
15	CW/M	5	1,148	3,793	43	3,304	3,800
15	CW/H	2	337	3,149	9	1,641	3,719
15	CW/L	1	196	2,602	4	1,085	3,335
15	CW/Lk	2	486	3,597	10	2,259	3,790
15	CD/M	3	940	3,788	26	3,134	3,800
15	CD/H	2	375	2,651	12	1,361	3,544
15	CD/L	1	28	1,214	1	242	2,544
15	CD/Lk	2	470	3,307	10	1,894	3,734
15	WX/L	1	214	3,312	4	1,481	3,800
15	WX/H	1	310	3,284	6	2,039	3,800
15	WW/H	1	150	3,114	3	1,186	3,768
15	WW/L	1	120	2,655	2	866	3,336
15	WW/Lk	1	231	3,244	4	1,463	3,721
15	WD/L	1	73	1,244	2	421	2,141
15	WD/Lk	1	143	2,272	3	872	3,320

Table 22. Criteria derived from the OLS model for DIN for three levels of under-protection risk (20%, 30% and 50%) at shaded and unshaded locations. Criteria greater than the saturating concentration of approximately 1,000 mg m^{-3} indicate combinations of conditions where periphyton biomass is strongly controlled by non-nutrient factors. Under such conditions, even when nutrient concentrations exceed saturating levels, the biomass threshold(s) may not be exceeded.

Under-protection risk	SoF	Unshaded 50 mg m^{-2}	Unshaded 120 mg m^{-2}	Unshaded 200 mg m^{-2}	Shaded 50 mg m^{-2}	Shaded 120 mg m^{-2}	Shaded 200 mg m^{-2}
20	CX/GM	83	3,655	3,800	719	3,800	3,800
20	CX/M	52	3,334	3,800	451	3,792	3,800
20	CX/H	30	2,798	3,784	264	3,506	3,800
20	CX/L	18	3,079	3,784	160	3,632	3,800
20	CX/Lk	23	2,689	3,799	200	3,571	3,800
20	CW/GM	3	1,372	3,800	31	3,793	3,800
20	CW/M	13	2,078	3,800	114	3,691	3,800
20	CW/H	3	762	3,510	25	2,470	3,777
20	CW/L	2	483	3,008	11	1,827	3,561
20	CW/Lk	3	1,058	3,702	27	3,295	3,800
20	CD/M	8	1,782	3,800	71	3,618	3,800
20	CD/H	4	705	3,155	32	2,004	3,706
20	CD/L	1	76	1,893	2	580	3,006
20	CD/Lk	3	1,077	3,588	26	2,682	3,782
20	WX/L	2	541	3,701	12	2,556	3,800
20	WX/H	2	810	3,729	16	2,783	3,800
20	WW/H	1	397	3,439	8	2,294	3,798
20	WW/L	1	318	3,075	6	1,732	3,494
20	WW/Lk	2	631	3,622	12	2,529	3,757
20	WD/L	1	200	1,678	5	774	2,550
20	WD/Lk	1	405	2,916	8	1,456	3,492
30	CX/GM	428	3,798	3,800	2,308	3,800	3,800
30	CX/M	267	3,769	3,800	1,552	3,800	3,800
30	CX/H	156	3,377	3,800	1,158	3,759	3,800
30	CX/L	95	3,561	3,800	806	3,759	3,800
30	CX/Lk	115	3,434	3,800	901	3,786	3,800
30	CW/GM	17	3,761	3,800	154	3,800	3,800
30	CW/M	65	3,528	3,800	455	3,800	3,800
30	CW/H	14	1,996	3,751	111	3,373	3,799
30	CW/L	6	1,370	3,440	53	2,853	3,744
30	CW/Lk	15	2,793	3,798	134	3,654	3,800
30	CD/M	40	3,406	3,800	332	3,798	3,800
30	CD/H	19	1,650	3,634	146	2,959	3,794
30	CD/L	1	359	2,760	8	1,608	3,535
30	CD/Lk	15	2,233	3,754	132	3,470	3,800
30	WX/L	7	1,983	3,800	59	3,569	3,800
30	WX/H	9	2,402	3,800	81	3,568	3,800
30	WW/H	4	1,625	3,792	39	3,292	3,800
30	WW/L	4	1,204	3,412	34	2,931	3,662
30	WW/Lk	8	2,041	3,747	72	3,549	3,799
30	WD/L	4	585	2,322	29	1,508	3,100
30	WD/Lk	5	1,140	3,407	45	2,721	3,618
50	CX/GM	2,790	3,800	3,800	3,721	3,800	3,800
50	CX/M	2,039	3,800	3,800	3,507	3,800	3,800
50	CX/H	1,651	3,786	3,800	2,956	3,800	3,800
50	CX/L	1,358	3,787	3,800	3,288	3,800	3,800
50	CX/Lk	1,300	3,799	3,800	2,947	3,800	3,800
50	CW/GM	245	3,800	3,800	1,827	3,800	3,800
50	CW/M	636	3,800	3,800	2,493	3,800	3,800
50	CW/H	164	3,518	3,800	986	3,779	3,800
50	CW/L	83	3,015	3,767	629	3,572	3,800
50	CW/Lk	213	3,713	3,800	1,357	3,800	3,800
50	CD/M	502	3,800	3,800	2,243	3,800	3,800
50	CD/H	218	3,172	3,799	885	3,712	3,800
50	CD/L	12	1,918	3,630	108	3,031	3,795
50	CD/Lk	216	3,599	3,800	1,365	3,786	3,800
50	WX/L	95	3,705	3,800	748	3,800	3,800
50	WX/H	127	3,735	3,800	1,107	3,800	3,800
50	WW/H	61	3,447	3,800	546	3,799	3,800
50	WW/L	58	3,089	3,699	452	3,504	3,791
50	WW/Lk	144	3,644	3,800	987	3,757	3,800
50	WD/L	75	1,734	3,259	302	2,590	3,724
50	WD/Lk	87	2,938	3,652	600	3,499	3,800

Table 23. Criteria derived from the OLS model for TP for three levels of under-protection risk (5%, 10% and 20%) at shaded and unshaded locations. Criteria greater than the saturating concentration of approximately 50 mg m^{-3} indicate combinations of conditions where periphyton biomass is strongly controlled by non-nutrient factors. Under such conditions, even when nutrient concentrations exceed saturating levels, the biomass threshold(s) may not be exceeded.

Under-protection risk	SoF	Unshaded 50 mg m^{-2}	Unshaded 120 mg m^{-2}	Unshaded 200 mg m^{-2}	Shaded 50 mg m^{-2}	Shaded 120 mg m^{-2}	Shaded 200 mg m^{-2}
5	CX/GM	0	53	283	0	184	300
5	CX/M	0	41	281	0	161	300
5	CX/H	0	26	244	0	131	289
5	CX/L	0	17	268	0	118	293
5	CX/Lk	0	14	202	0	82	275
5	CW/GM	0	3	161	0	23	299
5	CW/M	0	8	206	0	48	296
5	CW/H	0	2	101	0	16	221
5	CW/L	0	1	58	0	7	160
5	CW/Lk	0	1	87	0	10	220
5	CD/M	0	3	124	0	22	277
5	CD/H	0	1	46	0	9	132
5	CD/L	0	0	6	0	0	40
5	CD/Lk	0	0	55	0	5	186
5	WX/L	0	1	104	0	11	235
5	WX/H	0	1	142	0	16	245
5	WW/H	0	1	73	0	7	225
5	WW/L	0	0	44	0	4	161
5	WW/Lk	0	0	56	0	4	169
5	WD/L	0	0	12	0	1	49
5	WD/Lk	0	0	27	0	2	99
10	CX/GM	1	186	300	6	277	300
10	CX/M	0	161	300	4	273	300
10	CX/H	0	132	289	2	234	300
10	CX/L	0	120	293	1	261	300
10	CX/Lk	0	81	276	1	188	295
10	CW/GM	0	22	299	0	127	300
10	CW/M	0	46	296	1	181	300
10	CW/H	0	15	219	0	83	281
10	CW/L	0	6	158	0	45	234
10	CW/Lk	0	10	221	0	69	291
10	CD/M	0	21	277	0	103	300
10	CD/H	0	9	132	0	38	221
10	CD/L	0	0	39	0	4	123
10	CD/Lk	0	5	185	0	41	272
10	WX/L	0	10	233	0	80	296
10	WX/H	0	15	243	0	114	299
10	WW/H	0	6	223	0	53	273
10	WW/L	0	4	159	0	32	239
10	WW/Lk	0	5	171	0	43	247
10	WD/L	0	1	50	0	10	108
10	WD/Lk	0	2	99	0	21	208
15	CX/GM	3	256	300	25	298	300
15	CX/M	2	246	300	18	298	300
15	CX/H	1	208	298	11	273	300
15	CX/L	1	234	299	7	285	300
15	CX/Lk	1	157	292	6	245	299
15	CW/GM	0	75	300	1	295	300
15	CW/M	0	122	300	3	276	300
15	CW/H	0	48	267	1	165	294
15	CW/L	0	24	216	0	104	259
15	CW/Lk	0	38	281	0	151	299
15	CD/M	0	64	299	1	222	300
15	CD/H	0	24	195	0	85	265
15	CD/L	0	2	93	0	16	175
15	CD/Lk	0	21	254	0	123	288
15	WX/L	0	41	287	0	180	300
15	WX/H	0	58	290	0	205	300
15	WW/H	0	26	256	0	152	299
15	WW/L	0	16	222	0	96	261
15	WW/Lk	0	23	233	0	117	291
15	WD/L	0	7	91	0	29	147
30	WD/Lk	0	50	49	4	72	59

Table 24. Criteria derived from the OLS model for TP for three levels of under-protection risk (20%, 30% and 50%) at shaded and unshaded locations. Criteria greater than the saturating concentration of approximately 50 mg m^{-3} indicate combinations of conditions where periphyton biomass is strongly controlled by non-nutrient factors. Under such conditions, even when nutrient concentrations exceed saturating levels, the biomass threshold(s) may not be exceeded.

Under-protection risk	SoF	Unshaded 50 mg m^{-2}	Unshaded 120 mg m^{-2}	Unshaded 200 mg m^{-2}	Shaded 50 mg m^{-2}	Shaded 120 mg m^{-2}	Shaded 200 mg m^{-2}
20	CX/GM	9	287	300	72	300	300
20	CX/M	7	284	300	53	300	300
20	CX/H	4	247	300	35	290	300
20	CX/L	2	270	300	23	293	300
20	CX/Lk	2	207	297	18	279	300
20	CW/GM	0	173	300	3	300	300
20	CW/M	1	216	300	9	297	300
20	CW/H	0	104	286	2	223	298
20	CW/L	0	58	242	1	162	277
20	CW/Lk	0	91	295	1	227	300
20	CD/M	0	132	300	3	280	300
20	CD/H	0	48	237	1	138	286
20	CD/L	0	6	139	0	41	215
20	CD/Lk	0	61	277	1	190	297
20	WX/L	0	105	299	1	237	300
20	WX/H	0	144	300	2	246	300
20	WW/H	0	73	285	1	227	300
20	WW/L	0	45	247	0	164	273
20	WW/Lk	0	63	271	1	177	296
20	WD/L	0	17	120	0	53	176
20	WD/Lk	0	33	220	0	103	263
30	CX/GM	58	300	300	188	300	300
30	CX/M	42	300	300	162	300	300
30	CX/H	28	286	300	135	299	300
30	CX/L	18	291	300	126	300	300
30	CX/Lk	13	274	300	81	295	300
30	CW/GM	2	299	300	18	300	300
30	CW/M	6	294	300	42	300	300
30	CW/H	2	209	297	13	277	300
30	CW/L	1	145	273	5	228	293
30	CW/Lk	1	209	300	8	289	300
30	CD/M	2	272	300	19	300	300
30	CD/H	1	124	282	8	215	298
30	CD/L	0	32	204	0	114	266
30	CD/Lk	0	175	295	4	270	300
30	WX/L	1	222	300	9	294	300
30	WX/H	1	235	300	12	297	300
30	WW/H	0	212	300	5	266	300
30	WW/L	0	145	270	3	235	287
30	WW/Lk	0	166	296	5	254	300
30	WD/L	0	47	168	2	105	223
30	WD/Lk	0	91	260	2	202	281
50	CX/GM	236	300	300	294	300	300
50	CX/M	219	300	300	292	300	300
50	CX/H	186	300	300	259	300	300
50	CX/L	206	300	300	277	300	300
50	CX/Lk	132	298	300	224	300	300
50	CW/GM	43	300	300	236	300	300
50	CW/M	81	300	300	248	300	300
50	CW/H	27	288	300	125	299	300
50	CW/L	12	246	297	71	281	300
50	CW/Lk	19	298	300	112	300	300
50	CD/M	40	300	300	171	300	300
50	CD/H	16	246	300	61	289	300
50	CD/L	1	147	283	8	223	298
50	CD/Lk	11	280	300	84	299	300
50	WX/L	20	300	300	132	300	300
50	WX/H	27	300	300	170	300	300
50	WW/H	11	290	300	93	300	300
50	WW/L	9	251	291	60	276	298
50	WW/Lk	15	287	300	85	297	300
50	WD/L	8	129	243	23	183	284
50	WD/Lk	8	226	285	45	266	298

Table 25. Criteria derived from the OLS model for DRP for three levels of under-protection risk (5%, 10% and 15%) at shaded and unshaded locations. Criteria greater than the saturating concentration of approximately 25 mg m^{-3} indicate combinations of conditions where periphyton biomass is strongly controlled by non-nutrient factors. Under such conditions, even when nutrient concentrations exceed saturating levels, the biomass threshold(s) may not be exceeded.

Under-protection risk	SoF	Unshaded 50 mg m^{-2}	Unshaded 120 mg m^{-2}	Unshaded 200 mg m^{-2}	Shaded 50 mg m^{-2}	Shaded 120 mg m^{-2}	Shaded 200 mg m^{-2}
5	CX/GM	0	7	104	0	22	167
5	CX/M	0	10	129	0	31	200
5	CX/H	0	9	126	0	30	186
5	CX/L	0	9	144	0	30	204
5	CX/Lk	0	4	75	0	15	126
5	CW/GM	0	1	34	0	5	91
5	CW/M	0	2	44	0	7	114
5	CW/H	0	1	32	0	5	83
5	CW/L	0	1	15	0	2	43
5	CW/Lk	0	1	18	0	2	52
5	CD/M	0	0	9	0	1	29
5	CD/H	0	0	4	0	0	12
5	CD/L	0	0	2	0	0	5
5	CD/Lk	0	0	4	0	0	12
5	WX/L	0	1	30	0	4	84
5	WX/H	0	1	32	0	4	98
5	WW/H	0	1	16	0	2	53
5	WW/L	0	0	8	0	1	27
5	WW/Lk	0	0	7	0	1	23
5	WD/L	0	0	1	0	0	2
5	WD/Lk	0	0	1	0	0	3
10	CX/GM	0	31	181	1	85	217
10	CX/M	0	41	210	2	104	228
10	CX/H	0	41	196	1	104	220
10	CX/L	0	42	212	1	118	227
10	CX/Lk	0	20	137	1	58	180
10	CW/GM	0	7	114	0	22	223
10	CW/M	0	10	134	0	30	201
10	CW/H	0	6	99	0	21	164
10	CW/L	0	3	53	0	9	107
10	CW/Lk	0	3	65	0	11	123
10	CD/M	0	2	38	0	6	100
10	CD/H	0	1	15	0	2	43
10	CD/L	0	0	7	0	1	23
10	CD/Lk	0	1	16	0	2	50
10	WX/L	0	6	101	0	20	164
10	WX/H	0	6	121	0	20	189
10	WW/H	0	3	69	0	10	150
10	WW/L	0	2	35	0	5	82
10	WW/Lk	0	1	30	0	4	69
10	WD/L	0	0	3	0	0	10
10	WD/Lk	0	0	5	0	1	16
15	CX/GM	1	76	214	3	141	228
15	CX/M	1	92	227	4	173	230
15	CX/H	1	94	217	4	162	227
15	CX/L	1	105	225	4	186	230
15	CX/Lk	1	51	174	2	103	206
15	CW/GM	0	18	214	1	57	230
15	CW/M	0	25	194	1	73	222
15	CW/H	0	17	154	1	53	201
15	CW/L	0	8	97	0	25	152
15	CW/Lk	0	9	114	0	30	173
15	CD/M	0	5	87	0	16	172
15	CD/H	0	2	37	0	6	83
15	CD/L	0	1	19	0	3	57
15	CD/Lk	0	2	43	0	6	103
15	WX/L	0	16	156	0	51	203
15	WX/H	0	16	181	0	54	221
15	WW/H	0	8	138	0	28	189
15	WW/L	0	4	73	0	15	131
15	WW/Lk	0	4	62	0	12	110
15	WD/L	0	0	9	0	1	26
15	WD/Lk	0	0	13	0	2	43

Table 26. Criteria derived from the OLS model for DRP for three levels of under-protection risk (20%, 30% and 50%) at shaded and unshaded locations. Criteria greater than the saturating concentration of approximately 25 mg m^{-3} indicate combinations of conditions where periphyton biomass is strongly controlled by non-nutrient factors. Under such conditions, even when nutrient concentrations exceed saturating levels, the biomass threshold(s) may not be exceeded.

Under-protection risk	SoF	Unshaded 50 mg m^{-2}	Unshaded 120 mg m^{-2}	Unshaded 200 mg m^{-2}	Shaded 50 mg m^{-2}	Shaded 120 mg m^{-2}	Shaded 200 mg m^{-2}
20	CX/GM	2	121	225	7	179	230
20	CX/M	3	148	230	10	208	230
20	CX/H	3	142	226	10	194	229
20	CX/L	3	166	229	10	210	230
20	CX/Lk	1	86	198	5	134	218
20	CW/GM	0	40	229	1	104	230
20	CW/M	1	51	216	2	128	229
20	CW/H	0	37	189	1	93	217
20	CW/L	0	17	135	1	49	178
20	CW/Lk	0	21	153	1	60	206
20	CD/M	0	11	145	0	34	214
20	CD/H	0	4	65	0	14	123
20	CD/L	0	2	40	0	6	99
20	CD/Lk	0	4	81	0	14	151
20	WX/L	0	35	188	1	94	218
20	WX/H	0	36	210	1	111	230
20	WW/H	0	19	178	1	61	207
20	WW/L	0	10	112	0	31	167
20	WW/Lk	0	8	93	0	27	138
20	WD/L	0	1	18	0	3	46
20	WD/Lk	0	1	29	0	4	89
30	CX/GM	8	184	230	29	218	230
30	CX/M	12	212	230	37	228	230
30	CX/H	11	197	229	38	220	230
30	CX/L	12	212	230	39	227	230
30	CX/Lk	5	138	220	18	181	229
30	CW/GM	2	114	230	5	223	230
30	CW/M	2	135	229	8	201	230
30	CW/H	1	99	218	5	163	228
30	CW/L	1	52	181	2	105	207
30	CW/Lk	1	64	210	3	124	229
30	CD/M	0	37	218	1	100	230
30	CD/H	0	15	128	1	42	181
30	CD/L	0	7	104	0	23	159
30	CD/Lk	0	16	159	0	50	209
30	WX/L	1	100	220	5	163	228
30	WX/H	1	119	230	4	188	230
30	WW/H	1	67	209	2	149	228
30	WW/L	0	33	171	1	80	201
30	WW/Lk	0	30	142	1	68	180
30	WD/L	0	3	49	0	10	87
30	WD/Lk	0	4	97	0	15	175
50	CX/GM	71	228	230	135	230	230
50	CX/M	84	230	230	163	230	230
50	CX/H	88	227	230	154	230	230
50	CX/L	100	230	230	180	230	230
50	CX/Lk	45	207	230	95	223	230
50	CW/GM	14	230	230	45	230	230
50	CW/M	19	222	230	58	230	230
50	CW/H	13	199	230	42	221	230
50	CW/L	6	149	220	19	186	227
50	CW/Lk	7	172	230	24	216	230
50	CD/M	3	170	230	12	224	230
50	CD/H	1	80	208	5	140	224
50	CD/L	1	54	194	2	116	221
50	CD/Lk	1	101	222	5	173	228
50	WX/L	12	200	230	39	221	230
50	WX/H	12	219	230	40	230	230
50	WW/H	6	186	230	20	214	230
50	WW/L	3	127	214	11	178	223
50	WW/Lk	3	107	199	9	151	209
50	WD/L	0	24	117	1	57	166
50	WD/Lk	0	40	204	1	114	218

Appendix B Derivation of nutrient concentration criteria for selected Source-of-flow classes using quantile regression

Four Source-of-flow classes in our dataset of 251 sites having at least 20 periphyton observations had ≥30 sites. We examined relationships between the 92nd percentile of biomass observations (Chla92) and nutrient concentration at these sites using scatter plots and fitted quantile regressions to these relationships for the 70%, 80% and 90% quantiles (Figure 40). For some of the nutrients and Source-of-flow classes the relationships between Chla92 and nutrient concentrations were positive (i.e., consistent with the conceptual model; Figure 2).

We assume that relationships between Chla92 and nutrient concentrations that were negative (i.e., inconsistent with the conceptual model) are due to other sources of variation in biomass that are not well constrained by grouping sites into Source-of-flow classes). For example, electrical conductivity (EC) varies appreciably within the Source-of-flow classes and was also generally positively related to biomass (Figure 41).

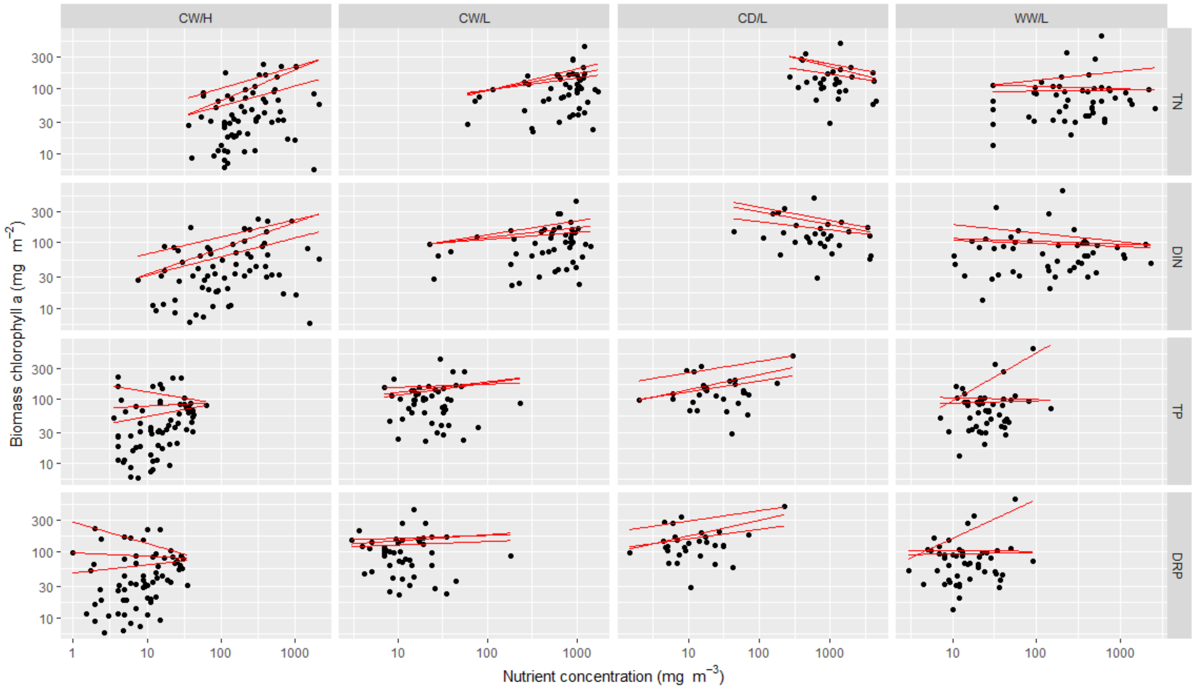

Figure 40. Relationships between the 92nd percentile of biomass observations (Chla92) and nutrient concentrations for Source-of-flow classes having ≥30 sites. The red lines are quantile regressions fitted to the 70%, 80% and 90% quantiles.

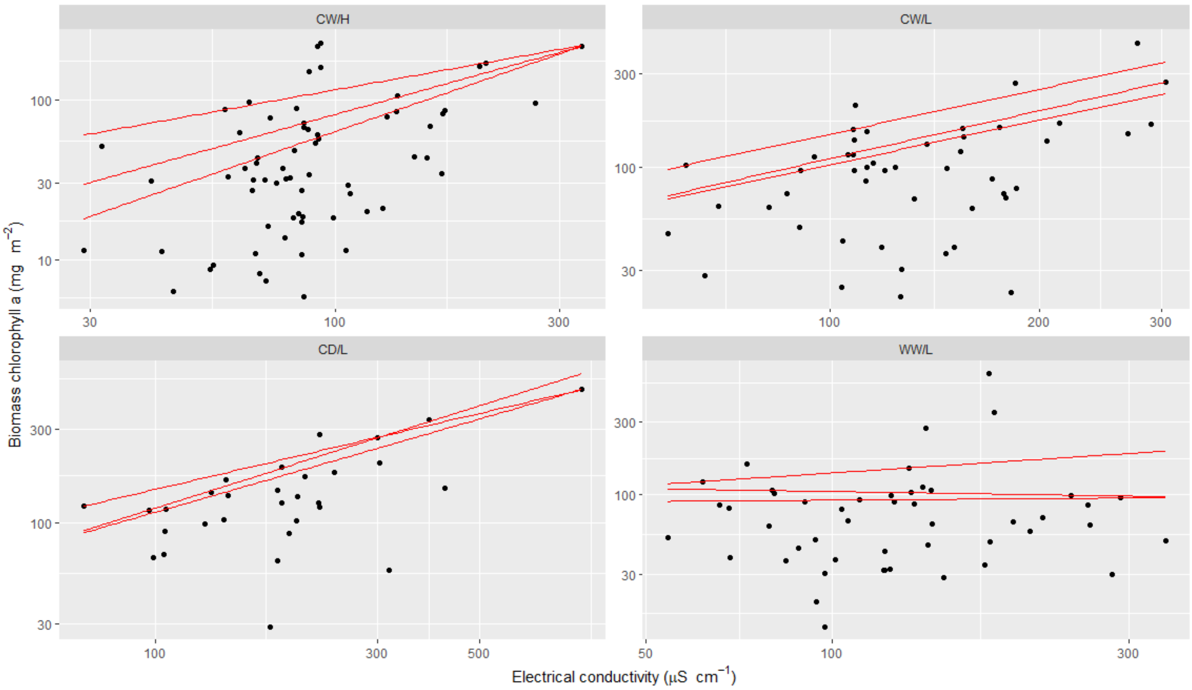

Figure 41. Relationships between the 92nd percentile of biomass observations (Chla92) and electrical conductivity for Source-of-flow classes having ≥30 sites. The red lines are quantile regressions fitted to the 70%, 80% and 90% quantiles.

For each class, we fitted quantile regressions to Chla92 using the combination of nutrient concentration and electrical conductivity (EC) as independent variables. We included EC because it was consistently included as a significant term in the quantile regression models based on all nutrients except DRP (Table 13). Where the relationship between the 92nd percentile of biomass observations (Chla92) and the nutrient concentration was positive (i.e., where the relationship was consistent with the conceptual model; Figure 2), we obtained concentration criteria from the QR models for each nutrient as described in Section 0; Figure 5).

The criteria derived using the QR models as described above are compared to the criteria derived using the OLS models (see Appendix A). Criteria for many combinations of nutrient and under-protection risk could not be derived because relationships between Chla92 and the nutrient concentration was negative. We assume that this is due to other sources of variation in biomass that are not well constrained by grouping sites into Source-of-flow classes. Where criteria were derived based on the QR models, they were reasonably consistent with those derived from the OLS models (points lie close to the one-to-one lines in Figure 42).

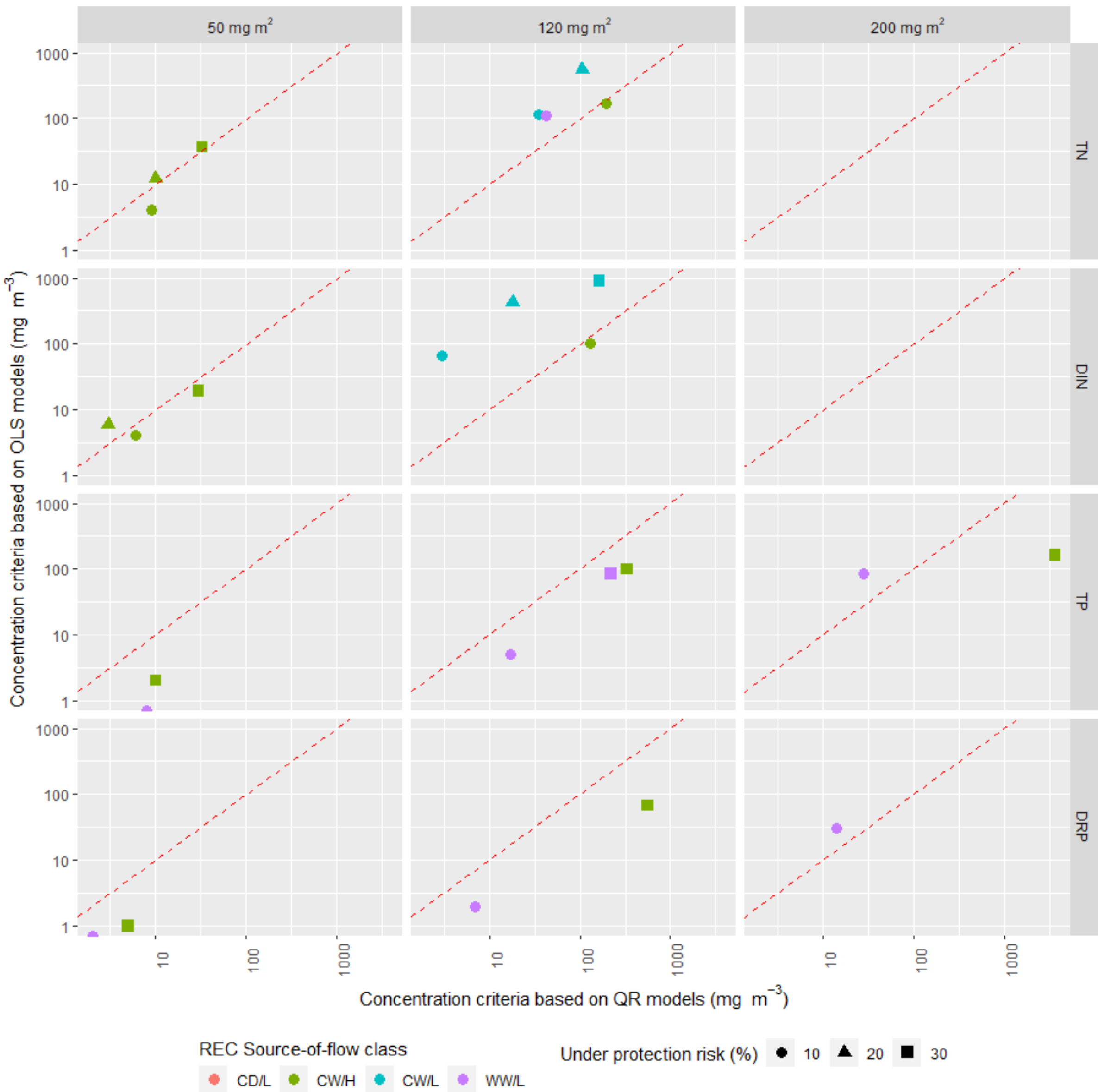

Figure 42. Comparison of concentration criteria for Source-of-flow classes having ≥30 sites derived using OLS and QR models.

Appendix C Temporal patterns in the periphyton biomass observations

The plot shown in Figure 43 shows the temporal patterns of periphyton biomass observations (as chlorophyll) made at the 251 sites with ≥30 biomass observations that were retained for modelling. These plots were made by taking each chlorophyll observation at each site and dividing it by the site mean chlorophyll concentration. These data are plotted by month to show the seasonal pattern of the observations on a standardised scale. The plots indicate some weak seasonality in the observations, but that high biomass (as a proportion of site-mean biomass) can occur at any time of the year.

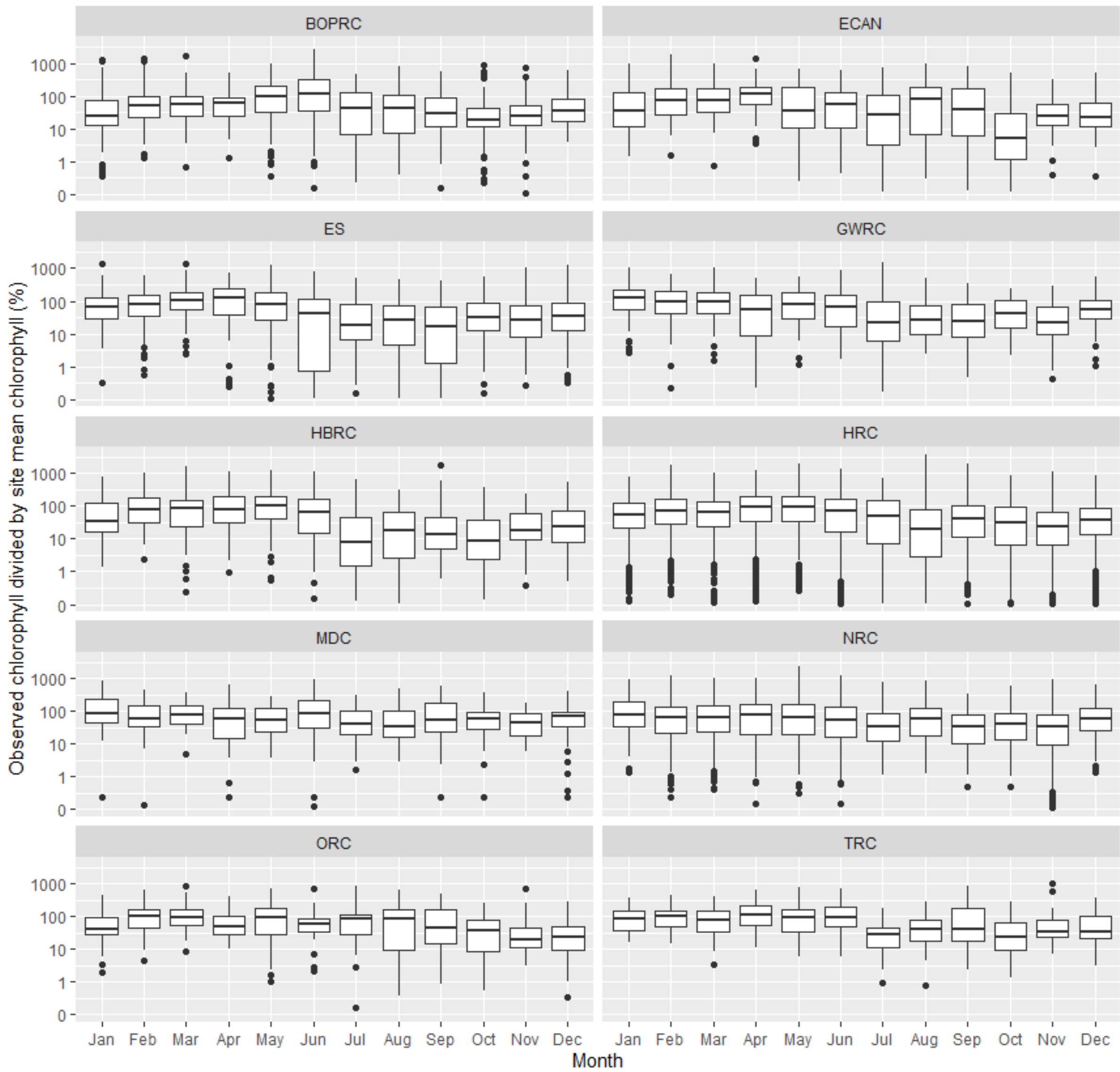

Figure 43. Temporal patterns in the periphyton (as chlorophyll) observations at the 251 sites with ≥30 biomass observations.