Update to REC Land Cover categories and review of category membership rules

December 2021

Prepared By:

Caroline Fraser

Ton Snelder

For any information regarding this report please contact:

Ton Snelder

Phone: 027 575 8888
Email: ton@lwp.nz
LWP Ltd
PO Box 70
Lyttelton 8092

LWP Client Report Number:	LWP Client Report 2021-18
Report Date:	December 2021
LWP Project:	2021-18

Quality Assurance Statement

Version	Reviewed By	
Final	Scott Larned	

Table of Contents

Executive Summary v

Glossary vii

1 Introduction 8

2 River Environment Classification (REC) 10

2.1 Overview 10

2.2 Limitations of classifications 10

2.3 The REC Land Cover level 11

3 Literature review 12

4 Data 14

4.1 Land cover data 14

4.2 Water quality monitoring data 14

5 Methods 15

5.1 Update REC Land Cover categories based on LCDB5 15

5.1.1 Original REC Land Cover rules 15

5.1.2 Adjustments to REC Land Cover categories associated with LCDB5 18

5.2 Water quality responses to land cover gradients 19

6 Results 20

6.1 Update REC Land Cover categories 20

6.1.1 Original REC rules 20

6.1.2 Adjustments to REC Land Cover category mapping to LCDB5 22

6.2 Water quality responses to land cover gradients 25

6.2.1 Available data and correlations between land cover gradients 25

6.2.2 Random forest models 30

6.2.3 Linear regression 33

7 Use of land cover categories to describe water quality patterns and alternative approach 37

8 Summary and conclusions 41

8.1 Updating REC Land Cover categories 41

8.2 Review of the REC Land Cover category membership rules 41

8.3 Using REC Land Cover categories to describe water quality variation 43

Acknowledgements 44

References 45

Appendix A LCDB5 land cover descriptions 48

Appendix B REC land cover categories including notations and characteristics 54

Figures

Figure 1: Proportion of monitoring sites and nzsegments in each REC land cover class for alternative versions of LCDB 20
Figure 2: Scatter plot of the R^2 values obtained between with linear regressions of water quality against "Pastoral" land cover, as defined by the original REC classification, and the proposed alternative classification. 23
Figure 3: Proportion of monitoring sites and nzsegments in each REC land cover class for alternative mappings of LCDB v5.0 to the REC land cover classes. 24
Figure 4: Distribution of proportions of upstream land cover for monitoring sites with state data 26
Figure 5: Correlation matrix of continuous REC land cover classes. 27
Figure 6: Correlation matrix, showing correlation coefficients between each pair of water quality variables and continuous land cover classes. 28
Figure 7: scatter plots of proportion upstream land cover against water quality state. Water quality state is log10 transformed median state, with the exception of MCI which is not transformed. Red line is a loess model with a span of 0.5. 29
Figure 8: Partial plots for the eight landcover predictor variables in random forest models of median water quality. 32
Figure 9: Scatter plot of proportion upstream urban against the residuals of the linear model relating proportion upstream pasture to water quality state. Blue lines represent a loess fit of the data (span = 0.5). 34
Figure 10: Comparison of model performance between different water quality models (P-only - Pasture only; U &P - Urban and Pasture; All - All land cover variables; RF - Random forest) 35
Figure 11: Regression coefficients for linear regression of water quality state against upstream proportion of "Urban" and upstream proportion of "Pastoral" continuous REC land cover types. Error bars show 95% confidence intervals for the regression coefficients. 36

Tables

Table 1:River water quality variables, measurement units and numbers of sites that met the state minimum data requirements and total numbers. 15
Table 2: REC land cover classes and mappings to LCDB v1.0 and LCDB v5.0...........17
Table 3: Comparison of nzsegment land cover classes between different LCDB versions. Values shown are percentages of the entire REC network (2 d.p.). Blank cells are land cover pairs that contain no nzsegments...21
Table 4: Comparison of monitoring site land cover classes between different LCDB versions. Values shown are percentages of the entire REC network (2 d.p.). Blank cells are land cover pairs that contain no sites. 21
Table 5:Performance of the water quality models..31
Table 6:Rank order of importance of predictor variables retained in the random forest models for at least one water quality variable. 31

Executive Summary

In this study we updated the REC Land Cover categories to reflect the latest national scale map of land cover of New Zealand (LCDB5). We recommended some changes to the grouping of land cover classes that are defined by LCDB5 prior to applying the REC Land Cover category membership rules. We reviewed literature concerning the relationship between land cover types and water quality and ecological measures and analysed these relationships using data from monitoring programmes in New Zealand. We used these two sources of evidence to consider whether changes to the REC Land Cover category membership rules are justified. We proposed alternative approaches to overcome two issues that arise from the use of REC Land Cover categories in environmental reporting. Finally, we considered the legitimacy of using REC Land Cover categories to describe water quality variation.

For the REC Land Cover category update, we produced a table that assigns all segments of the digital river network (version 2.4) to a Land Cover category based on LCDB5. In making this update, we have not changed the REC Land Cover category membership rules that assign river segments to a "dominant" catchment land cover with special conditions applying to determine dominance by *Urban* and *Pastoral* land cover categories. This is because we did not find strong evidence for changing the original REC Land Cover category membership rules. However, we recommend, and have implemented, two changes to the original REC Land Cover categories associated with the definition of coarsened LCDB5 classes:

- LCDB5 class 41, *Low Producing Grassland* is assigned to coarsened LCDB class *Tussock* (rather than *Pastoral*); and
- LCDB5 class 54, *Broad Leaved Indigenous Hardwood* is assigned to the coarsened LCDB class *Indigenous Forest* (rather than *Scrub*).

We provided both sets of REC Land Cover categories (i.e., based on the original REC Land Cover category membership rules and the recommended changes) with this report as supplementary material.

We reviewed the literature and conducted our own analyses of the relationships between water quality and catchment land cover. Neither the literature, nor our own analyses, revealed threshold responses for any land cover types. In general, relationships between water quality (appropriately transformed) and ecological measures and catchment land cover are approximately linear. However, several studies and our own analyses indicate that the effect of pastoral and urban land cover (i.e., the change in water quality and ecological measures with unit change in catchment pastoral and urban land cover) are similar in magnitude and greater than that of other land cover types. The larger effect of pastoral and urban land cover compared to the other land cover types is consistent with the rules employed by the REC that assigns these as dominant land covers at 25% and 15% occupancy of the upstream catchment. These rules recognise that due to the larger effect size, occupancy of the catchment by upstream pastoral and urban land cover has a disproportionate influence on a range of river characteristics. On balance, we do not consider that there is good evidence for changing the original REC Land Cover membership rules. It is important to emphasise that the REC attempts to discriminate variation in a range of characteristics including hydrology and morphology and that the category membership rules are an attempt to find a suitable compromise across all these characteristics.

REC Land Cover categories are used in environmental reporting to indicate broad scale differences in water quality associated with catchment land cover. This is a legitimate, simple and relevant approach for summarising broad scale variation in water quality patterns for

environmental reporting, but it raises two issues. First, land cover is only one of many factors that affect water quality. REC Land Cover categories leave considerable between-site variation in water quality unexplained. Second, REC Land Cover categories are based on assigning a single dominant land cover to a catchment, which can obscure the contribution of other land cover types to water quality.

We propose that these two issues are best addressed by using multivariable regression modelling to estimate the effect size of different land cover types. This approach provides a more complete picture of the effect of different land cover types on water quality. If environmental reporting needs to indicate the effects of differing land cover types on water quality, effect sizes estimated using regression modelling provides more robust evidence than grouping sites by REC Land Cover categories. However, multivariable regression modelling is more complicated than the use of a single category, and has some analytical complications associated with collinearity between different types of land cover. Therefore, it may not be appropriate in all environmental reporting contexts.

An alternative that retains the simplicity of land cover categories but avoids the need to assign a dominant land cover, is to use a categorical subdivision of a gradient in land cover "anthropogenic modification". The anthropogenic modification gradient can be defined by adding the proportion of catchment area occupied by *Pastoral*, *Urban* and *Exotic Forest* land cover types and subdividing this gradient into categories. The anthropogenic modification categories discriminate catchments on the basis of high and low anthropogenic modification, and these broadly explain variation in the water quality.

Glossary

Term	Meaning in this report
REC Land Cover category	The nine categories defined by the Land Cover level of the REC. One land cover class is assigned to each REC reach based on the dominant land cover of the upstream catchment.
LCDB class	The classes of land cover defined by the LCDB. The number of classes has changed between versions of the LCDB with the current version 5 having 36 classes.
Coarsened LCDB classes	A coarsened version of the LCDB classification that groups some of the original LCDB classes together. The grouping reflects the nine categories defined by the Land Cover level of the REC. However, the REC Land Cover category rules include an additional step that identifies the dominant coarsened LCDB class for the catchment of each segment of the river network.
Continuous LCDB5 land cover variables	Variables that are calculated and available for each segment of the river network that quantify the percentage of the upstream catchment occupied by each LCDB5 class.
Continuous coarsened land cover variables	Variables that are calculated and available for each segment of the river network that quantify the percentage of the upstream catchment occupied by each coarsened LCDB class.
River Environment Classification (REC)	A system that classifies New Zealand's rivers at a range of spatial scales and six hierarchical levels that is used for environmental management purposes.
Digital network (DN)	A digital representation of New Zealand's river networks that is contained within a geographic information system (GIS). The DN comprises approximately 560,000 segments with a mean length of 700 m. Each segment is associated with six attributes that assign the segment to a category at each level of the REC hierarchy.

1 Introduction

The River Environment Classification (REC) is a system that classifies New Zealand's rivers at a range of spatial scales and six hierarchical levels (Snelder and Biggs, 2002). The REC comprises a digital representation of New Zealand's river networks (hereafter digital network; DN) and classification system that are contained within a geographic information system (GIS). The DN in the current version of the REC comprises approximately 560,000 segments with a mean length of 700 m. Each segment of the DN is associated with six categorical variables that indicate category membership at each level of the REC hierarchy. The first four hierarchical levels of the REC are defined by categorical subdivision of the factors: climate, topography, geology and land cover. Each segment of the DN is assigned to a category at each of these levels that indicates the spatially dominant climate, topography, geology and land cover of the upstream catchment. The REC is used as a spatial framework for regional (or larger) scale environmental monitoring and reporting, environmental assessment and management.

The DN was derived from a digital elevation model that was built from the 20 metre contours that appear on Land Information New Zealand's TopoMap50 map sheets. The DN therefore approximates the stream and river features that appear on these maps. The DN has been updated over time with a significant update occurring in 2010 with the release of DN version 2 (DN2.0). The version of the network that is in general use is DN2.4.

REC classes are used in national environmental reporting carried out by the Ministry for the Environment (MFE) and Statistics New Zealand (StatsNZ) to discriminate variation in water quality and ecological indicators (e.g., macroinvertebrate assemblages and periphyton abundance) observed at long term river monitoring sites (e.g., MFE & StatsNZ, 2017, 2019). Typically, site water quality and ecological indicator observations are grouped according to the REC class of the river on which the monitoring site is located. The level of similarity in water quality for sites within classes, and differences between classes, is shown graphically with boxplots, or other types of plots. The variation in water quality explained by REC classes can be formally quantified using analysis of variance (ANOVA) or multivariable equivalents thereof. Analyses of this type link the monitoring sites observations with the REC factors (i.e., climate, topography, geology and land cover) and illustrate broad-scale patterns in water quality and ecological indicators.

Any level of the REC hierarchy can be used to group monitoring sites and discriminate variation (patterns) in river characteristics. For water quality, the Land Cover (fourth) level of the REC is a logical level to use due to the general interest in the relationships between land use and water quality. However, because the REC is hierarchical, classes at lower levels of classification detail (provided by lower hierarchical levels) further subdivide the higher-level classes. This means that at the Land Cover level (fourth) of the REC hierarchy there are many potential classes. The large number of classes at the Land Cover level of the REC can be an impediment to describing water quality variation to non-technical audiences. In addition, because the number of water quality monitoring sites is limited, there is generally only a small number of sites representing each class at lower levels of the REC hierarchy. Poor representation of classes by sites at lower levels of classification detail often means that it is not possible to show that differences between classes are statistically significant. Consequently, for environmental reporting, categories are often used at the REC Land Cover level on their own (i.e., by disregarding the higher three hierarchical levels) to provide a simple model of between site variation in water quality data.

There are three issues that arise from the use of both classes at the REC Land Cover level and the REC Land Cover categories (i.e., ignoring the first three levels of the REC) for environmental reporting. First, REC Land Cover categories are based on an assessment of the dominant land cover in the upstream catchment. The original categorisation was based on the Land Cover Database (LCDB) version 1 (hereafter LCDB1), which was produced from satellite imagery from 1996/1997 (Dymond *et al.*, 2017). Land cover is subject to change over time due to human activities to a much greater degree than the factors that define the first three hierarchical levels of the REC (climate, topography and geology). Therefore, the REC Land Cover categories need to be updated over time so that monitoring sites are assigned to a REC Land Cover category that reflects reasonably current conditions. Second, the assignment of the dominant land cover is based on rules that account for the disproportionate influence of pastoral and urban land cover on multiple river characteristics including water quality, ecology and hydrology. These rules were based on expert judgments that have not been subsequently reviewed, despite significant work on relationships between land cover and water quality and ecological variables since the development of the REC.

The third issue that arises from the use of REC Land Cover categories in environmental reporting is that dominant land cover is a very simple model of water quality patterns. Almost all catchments comprise mixtures of land cover and it is the combined effect of these and other environmental factors (e.g., climate, geology, and topography) that determine water quality. Even the assignment of catchments to a REC Land Cover category is a simplification at two levels. First, the REC categories themselves represent a coarsening of the more detailed land cover classes defined by the LCBD. Second, the representation of the catchment by one "dominant" land cover does not account for areas of other land cover types in the catchment. This simplification can give the impression that observed water quality is attributable to a single land cover, such as pastoral, when in fact there are other types of land cover in the same catchment that are contributing to the observed conditions, such as urban land.

MFE sought assistance and advice regarding the above three issues. Specifically, MFE required

1. an update to the REC Land Cover categories following the recent update of the LCDB to version 5 (hereafter LCDB5), which reflects land cover conditions as of 2018,
2. expert opinion as to whether the rules used defining dominant land cover and assigning the REC Land Cover category are credible or need to be revised, and
3. expert opinion as to whether use of dominant land cover for reporting water quality patterns is justifiable and appropriate.

This report describes the update of REC Land Cover categories based on LCDB5. Changes in the classification of land cover by LCDB5 compared to LCDB1 are described, and changes to the way REC Land Cover categories are defined to best accommodate these changes are recommended. We also report a review of relevant literature and new analyses of the relationships between land cover and water quality and ecological variables. The purpose of the literature review and analyses was to provide evidence for or against revision of the rules used to define dominant land cover and assign the REC Land Cover category. Finally, we discuss the justifiability and appropriateness of using dominant catchment land cover to describe broad scale patterns in water quality as part of environmental reporting. We recommend two alternative approaches to using REC Land Cover categories to describe water quality variation based on land cover and discuss their implications.

2 River Environment Classification (REC)

2.1 Overview

The REC was developed as a spatial framework for management of water resources. Spatial frameworks are tools that assist with organising empirical data; extrapolating data and information to locations with no data; stratifying environmental resources so that management actions can be prioritised; and management expectations and controls can be set that are justifiable, specific and achievable.

The REC groups and classifies each segment of a digital representation of New Zealand's river networks at six hierarchical levels (Snelder and Biggs, 2002). Each of the REC's six hierarchical classification levels is defined by one of six 'controlling factors' (referred to generally as 'factors'). These factors are: Climate; Source-of-Flow; Geology; Land Cover; Network Position; and Valley Landform. Each factor is subdivided into categories that discriminate variation in the physical characteristics of rivers. Each subsequent level of the classification hierarchy subdivides the higher-level class into the categories at that level. Therefore, there is an increasing number of potential classes moving down the REC hierarchy.

The location of each river segment and its class membership at any level of the REC can be mapped so that environmental patterns in rivers can be defined at a range of spatial scales. The mapped patterns of REC categories are expected to be broadly consistent with patterns of physical and biological characteristics that are important for management such as hydrology, hydraulics, water quality and biological communities.

The choice of the factors that define each level of the REC hierarchy and the categories that subdivide each level were guided by scientific knowledge of the causes of patterns in characteristics of rivers at different spatial scales. The higher levels of the REC discriminate large-scale patterns in general characteristics. The lower the level of the classification used, the smaller the scale of the patterns that are defined by REC. Thus, the lower the classification level used, the smaller the scale of the patterns and the more specific the characteristics that can be discriminated.

An important aspect of the first four levels of the REC hierarchy is that the controlling factors are characterising the catchments upstream of each segment in the river network. Hence, category membership for segments at the first four levels of the REC is based on the dominant climate, topography, geology and land cover of the upstream catchment. However, catchments by their very nature are not homogeneous. Therefore, the category assigned to each segment at each of the first four REC levels is a simplified summary of the physical characteristics of the catchment. This implies that the REC, like all other environmental classifications, is an abstract representation of the real-world that cannot represent all the detail of the reality. Nevertheless, REC categories have been shown to discriminate variation in a range of river characteristics including hydrology (Snelder and Booker, 2013; Snelder *et al.*, 2005), water quality (Larned *et al.*, 2016, 2004), and ecology (Larned *et al.*, 2016; Snelder *et al.*, 2004). The REC has been used to discriminate between rivers for applications ranging from regional water quality and quantity planning and regulation (Norton *et al.*, 2010; Norton and Snelder, 2003; Rouse and Norton, 2010; Snelder and Hughey, 2005) to national regulations for river periphyton and sediment (NZ Government, 2017, 2020).

2.2 Limitations of classifications

The use of environmental classifications such as the REC is based on the recognition that, while no two locations are exactly the same, it is not possible to treat every location as a unique entity. Therefore, some level of generalisation is necessary in most environmental management tasks such as monitoring, reporting and policy development (Bailey, 1996; Bryce *et al.*, 1999; McMahon *et al.*, 2001; Omernik, 1995). Classifications are used to group locations that are considered sufficiently alike that they can be treated as the same for a given task. The grouping process is guided by principles that establish how patterns in characteristics will be recognised and how they will be grouped. For example, rivers can be grouped according to size or a single hydrological index such as the mean annual low flow (MALF). A map of a classification of rivers based on these characteristics would show patterns in river size or MALF. Patterns in these single characteristics would be simplified by agreeing on how the compression of detail, or 'graining', of patterns will be performed.

Environmental classifications such as the REC are used to discriminate patterns in multiple ecosystem characteristics. Finding the optimal solution for compressing detail while maximising the utility of a classification becomes increasingly difficult as the number of characteristics of interest increase. Regardless of what principles are used to develop environmental classifications, they cannot represent all the detail of reality. Therefore, environmental classification such as the REC cannot provide optimal discrimination of any individual characteristic and there is no 'correct' solution (Udo de Haes and Klijn, 1994).

2.3 The REC Land Cover level

This study is concerned specifically with the characterisation of catchment land cover to define REC Land Cover categories at the fourth level of the REC hierarchy. Most catchments comprise a mixture of land cover types that the REC Land Cover category summarises according to a deemed 'dominant' land cover.

The determination of the dominant land cover is based on an assessment of land cover within the catchment of each network segment in three steps (Snelder and Biggs, 2002). First, the proportions of the catchment occupied by nine[1] land cover classes[2] derived from the LCDB are obtained by spatial analysis. Second, a rule is applied that generally deems the REC Land Cover category to be whichever of the nine land cover classes occupies the greater proportion of the upstream catchment. Third, an exception is made for the previous rule if *Pastoral* and *Urban* land cover account for greater than 25% and 15% of the upstream catchment, in which case the REC Land Cover class is *Pastoral* or *Urban*, respectively. The rule for the *Urban* land cover category takes precedence over that of *Pastoral* so that segments that have ≥15% *Urban* and ≥25% *Pastoral* are categorised as *Urban*.

The exceptions for *Urban* and *Pastoral* were originally made to account for the disproportionate influence of pastoral and urban land use on multiple river characteristics including water quality, ecology and hydrology. These rules were based on knowledge of the relative effect sizes of different land covers on multiple characteristics of rivers including water quality, ecology, hydrology and geomorphology. The rules represent expert judgments that were made at the time that the REC was developed. In the nearly two decades since, the rules have not been reviewed or revised. The next section of this report reviews recent literature concerning the influence of land cover on the characteristics of rivers with a view to considering whether this provides evidence that a change to the existing rules is warranted.

[1] Snelder and Biggs (2002,2010) variously refer to 7 (table 1.6), 8 (Table 1.9, Appendix 3) or 9 (Table 2.6, Figure 2,10) land cover classes. The differences are related to the inclusion of the "Wetlands" and "Miscellaneous" land cover classes. We have used the 9 level classification in this report, as it provides complete classification of the LCDB terrestrial land cover types.

[2] Note that for clarity we later refer to these nine classes as coarsened LCDB classes because they represent a grouping (coarsening) of the original LCDB classes.

3 Literature review

We reviewed New Zealand and international literature to investigate whether there was evidence that a change to the existing rules for defining REC Land Cover categories is warranted. We examined literature relating to relative effect sizes of different land covers, particularly in terms of river water quality. We were primarily interested in the size of the effect on water quality and ecological metrics in response to varying proportions of catchment land cover types. Differences in effect size may provide evidence for or against singling out of urban and pastoral land covers in the exception rules applied in the REC Land Cover category membership rules. We were also interested in whether studies suggested that responses are continuous and linear or non-linear, or whether there were thresholds (i.e., abrupt inflections or breakpoints) in the land cover-response relationships. Non-linear or threshold responses may provide an obvious and justifiable basis for the threshold levels used in the exceptions of the rules whereas linear responses would mean rules would need to be based on judgements about the levels at which a particular land cover effect is likely to dominate river character (Capon *et al.*, 2015).

A large number of studies and investigations have examined the effects of different types of land cover on water quality (e.g., Khatri and Tyagi, 2015; Larned *et al.*, 2019, and references therein). These studies indicate that the proportion of catchment occupied by agriculture and urban land cover has the largest and most consistent negative effects on water quality. However, the magnitude of these effects, and in particular the relative magnitude of effects between urban and agricultural land covers, vary greatly depending on: the water quality variables of interest; the definition of the land cover classes; the time of the study (due to changes in practices and technologies in both agricultural and urban environments); and the location of the study (i.e., specifics of the natural systems and differences between countries in practices of agriculture and urbanisation). For example, in China, Zhou et al. (2012) found that urban land use exerted a disproportionately large influence on water quality (nutrients and microbiological variables) at multiple scales, whereas agricultural land use was only found to have a significant influence on dissolved oxygen. In contrast, in a study in Kentucky, USA, Coulter *et al.* (2004) found that agricultural land cover generally had a greater contribution to nutrients in rivers compared to urban areas.

Quantifying the effects of land cover on water quality outcomes is confounded by spatial scale and patterns of land cover. For example, Dodds and Oakes (2008) found that the variance explained by land cover decreased as stream order increased and increased when only riparian land cover was used as a predictor. They noted that due to collinearity of the land cover class proportions they were not able to reliably quantify effect sizes for individual land cover categories. Similarly, Lee et al. (2009) found that the spatial distribution of land use (i.e., patch density, shape, edge effects) also contributed significantly to explaining variance in water quality, and Tran et al. (2010) found that land cover associated with riparian buffer zones explained more of the observed variability in water quality outcomes than catchment land cover.

Larned et al (2020) provide a comprehensive summary of the state of the evidence for the effects of land use on freshwater ecosystems in New Zealand. As part of their review, they summarised 35 New Zealand studies that examined the associations between land use/land cover and variables describing state in rivers (Table 1, Larned et al. 2020). Of these 35 studies, 20 represented land cover as continuous variables (% cover), while the remainder only examined categorical land cover types. Of the 20 studies, four focused on only urban land cover, and only four included both agricultural and urban land covers (Ballantine and Davies-

Colley, 2014; Close and Davies-Colley, 1990; Jowett and Richardson, 2003; Larned *et al.*, 2016). In all four of these studies, the effects of the proportion of catchment belonging to land cover types were evaluated based on correlation between continuous land cover variables and water quality or ecological metrics. These studies were consistent in finding that pasture, followed by urban land cover were the land cover variables that explained most of the observed variability.

Larned *et al.*, 2016 was the only study reviewed by Larned et al. (2020) that went further than correlative analyses to evaluate relative effect sizes of the different continuous land cover variables. They used multiple linear regression with water quality variables as dependent variables and continuous percentage upstream urban and pastoral areas as independent variables. As both land cover areas were in the same units (% upstream area), the regression coefficients could be used to summarise relative effect size. The ratio of the urban to pastoral regression coefficients ranged from 0.93 (NO_3N) to 3 (MCI). NH_4N and ECOLI also had large ratios (1.38).

Collier and Hamer (2010) examined the relationship between continuous natural (Indigenous forest plus scrub) land cover and four macroinvertebrate metrics (MCI, ASPM, EPT and %EPT) at 46 sites in Waikato. They applied a LOWESS smoother to scatter plots of the macroinvertebrate metrics against continuous natural land cover. The LOWESS smoother suggested a non-linear response for MCI, %EPT and ASPM, with rapid reductions from 100-80% natural land cover, a plateau from 80-60%, and further decreases from 60-0% natural land cover. However, land cover only explained a small amount of the observed variability and the data only included a limited number of sites that had 60-80% natural land cover, therefore the generality of the observed thresholds is unknown. Tran et al. (2010) also report a threshold response of habitat score (a combined metric of 10 physical habitat characteristics) at 42% percent forest cover (in a 200m riparian buffer), below which habitat score reduced, and above which habitat score was approximately stable. However, a threshold was not reported for percentage forest cover of the entire catchment, and the study was based on a limited (29) number of nested sites, which limits the generality of this result.

D'Amario et al. (2019) identified threshold responses between solute concentrations (TDP, DOC, TDN) and ecological outcomes (a range of structural and functional indicators for diatom, mussels and bacterial communities) for 14-53 sites in Ontario, Canada. They found non-linearity and breakpoints in about half of their analyses. They then compared the distributions of land cover proportions above and below the identified thresholds and reported the land cover at the intersection of the two distributions as land cover breakpoints. Differences in medians of the land cover distributions for sites above and below the ecological response thresholds were generally statistically significant. However, there were large overlaps between the above and below distributions and the land cover breakpoints cannot be interpreted as a land cover proportion at which a threshold ecological response occurs.

Beyond the studies described above, the majority of studies we reviewed indicated that responses of water quality and ecological metrics to land cover gradients are approximately continuous and linear. In a national scale analysis, Larned *et al.* (2016) showed linear responses of seven water quality variables and the macroinvertebrate community index (MCI) to a gradient in the proportion of the catchment occupied by high intensity agriculture. These models explained between 17% and 55% of the variation in site median values of the response variables. This indicates that although land cover is associated with water quality and ecological variables, there are other sources of variation. This is consistent with the conceptual model underlying the REC, which postulates three higher level factors control water quality and ecological variables at scales that are larger than the typical scale of variation in catchment land cover.

In a review of studies that quantified the responses of benthic invertebrate assemblages to gradients in urbanization, Storey *et al.* (2013) reported some studies identified thresholds from 3-18% or more impervious area beyond which there was severe degradation of stream invertebrate communities. However, Storey *et al.* (2013) reported that other studies have reported linear responses with no evidence of an effect threshold. In a large study in nine metropolitan areas of the United States, Cuffney *et al.* (2009) showed strong, linear responses of benthic invertebrate assemblages to urbanisation (i.e., increase in urban land cover) when forest or shrublands were developed. The study found that responses to urban development were difficult to detect when urbanisation occurred in agriculturally dominated catchments because invertebrate assemblages were already degraded. In this study, there was no evidence that assemblages showed any initial resistance to urbanization and therefore there was no evidence for threshold effects.

Data presented in Storey *et al.* (2013) representing urban streams in Auckland and Hamilton indicate that above 10% impervious area the number of sensitive Ephemeroptera, Plecoptera and Trichoptera (EPT) taxa is low. These data show a linear response of EPT taxa to increasing catchment urban land cover, although the size of the dataset was small.

The conclusions from our review of the literature is that pastoral and urban land cover types have larger effect sizes compared to the other land cover types and that there is not clear evidence that indicates threshold responses for any land cover types. In general, relationships between water quality and ecological measures and catchment land cover are approximately linear.

4 Data

4.1 Land cover data

Catchment land cover was derived from the national Land Cover Database 5 (LCDB5) which differentiates 36 categories based on analysis of satellite imagery from the 2018–2019 summer[3].

4.2 Water quality monitoring data

This study uses the monitoring sites and data described by Whitehead, Fraser, Snelder, *et al.*, (2021). The water quality data consisted of measurements of five chemical, one microbiological, one macroinvertebrate and two physical variables from river monitoring sites in council SoE networks and the NRWQN sites (Table 1). Hereafter we refer to these data as 'water quality data'. Detailed methods for processing the water quality data are provided by Whitehead *et al.* (2021). The water quality dataset consisted of data for each of the nine variables, for the 2016–2020 period, at sites for which measurements of a given variable were available in at least 90% of the sampling intervals in that period (i.e., at least 54 of 60 months or 18 of 20 quarters). For the annually sampled MCI, which is generally less variable than physical and chemical water quality variables, we required that data were available for at least four of the five years. In addition, for the MCI data, we shifted the 2016–2020 time period by six months (the five-year period to end of June 2020) to align with water years, in order to prevent splitting summer samples into two calendar years, and only included samples if they

[3] https://lris.scinfo.org.nz/layer/104400-lcdb-v50-land-cover-database-version-50-mainland-new-zealand/

occurred between November and April to align with the draft NEMS guidelines for macroinvertebrates[4].

In the final water quality dataset used for the analyses described below, 976 sites met the minimum data requirements for at least one of the nine water quality variables (Table 1). In addition to sites that met the minimum data requirements, we also collated location information for additional sites that have monitoring data but were not included in the state assessment of Whitehead, Fraser, Snelder, *et al.*, (2021). We used the expanded site list to demonstrate the monitoring network distribution of the REC Land Cover categories. In total there were 1547 monitoring sites used for this purpose.

Table 1:River water quality variables, measurement units and numbers of sites that met the minimum data requirements for use in the analyses.

Variable type	Variable	Abbreviation	Units	Number of monitoring sites
Physical	Visual clarity	CLAR	m	715
	Turbidity	TURB	NTU	834
Chemical	Ammoniacal nitrogen	NH4N	mg l^{-1}	973
	Nitrate + nitrite-nitrogen	NNN	mg l^{-1}	946
	Total nitrogen (unfiltered)	TN	mg l^{-1}	938
	Dissolved reactive phosphorus	DRP	mg l^{-1}	973
	Total phosphorus (unfiltered)	TP	mg l^{-1}	901
Microbiological	*Escherichia coli*	ECOLI	Cfu 100 ml^{-1}	967
Invertebrate	Macroinvertebrate Community Index	MCI	unitless	955

5 Methods

5.1 Update REC Land Cover categories based on LCDB5

5.1.1 Original REC Land Cover rules

The LCDB5 land cover map was used to update land cover data for version 2.4 of the national digital river network using a three-step process. First, the LCDB5 and DN2.4 watershed shapefiles were intersected to identify polygons within each watershed that represented the 34 LCDB5 land cover classes. Second, the area of each LCDB5 class in each watershed was summed. Third, the area of each of LCDB5 class in all watersheds upstream of every segment of the national digital river network was summed. This step was performed using a downstream accumulation algorithm and resulted in a matrix representing every network segment (rows) and the total upstream area belonging to each LCDB5 class (columns). For each network segment, these values were then divided by the total upstream area to derive the proportion of catchment area occupied by each LCDB5 class for every network segment. We refer to these proportions as ‘continuous LCDB5 land cover variables’.

[4] https://www.nems.org.nz/documents/macroinvertebrates/

The original REC Land Cover categories were based on LCDB1. The methodology for determining the REC Land Cover category for each segment of the river network[5] involved the following four steps:

1. Grouping the LCDB classes to define a smaller set of coarsened LCDB classes (see Table 2).
2. Calculating the proportion of the upstream catchment occupied by each coarsened LCDB class for each river network segment. Hereafter, we refer to these proportions as continuous coarsened land cover variables. For each network segment there are as many continuous coarsened land cover variables as there are coarsened LCDB classes.
3. Identifying for each river network segment the dominant land cover as the coarsened LCDB class occupying the greatest percentage of the upstream catchment.
4. The REC Land Cover category for each network segment was deemed to be the dominant land cover unless:
    a. The continuous coarsened LCDB class *Pastoral* exceeded 25%, in which case the category was set to *Pastoral*; or
    b. The continuous coarsened LCDB class *Urban* exceeded 15%, in which case the category was set to *Urban*. (If both *Pastoral* and *Urban* exceeded their respective thresholds, then *Urban* was given precedence).

We repeated the above process with the LCDB5 dataset. Table 2 shows the mapping of LCDB5 to the coarsened LCDB classes. This was done by combining (1) the mapping from LCDB1 to the coarsened LCDB class from Snelder *et al.* (2010) and (2) the mapping from LCDB5 to LCDB1 provided with the LCDB5 dataset[6].

[5] The original version of the digital network used the term “nzreach” to refer to the unique identifier for network segments. Version 2 of the DN used the term “nzsegment”.

[6] “LCDB class correlations.pdf” https://lris.scinfo.org.nz/layer/104400-lcdb-v50-land-cover-database-version-50-mainland-new-zealand/

Table 2: Coarsened LCDB classes and mappings to LCDB1 and LCDB5 classes.

Coarsened LCDB class		**LCDB1 class**	**LCDB 5 class**	**LCDB5 class identifier**
Bare Ground	B	Bare Ground	Landslide	12
			Permanent Snow and Ice	14
			Alpine Grass/Herbfield	15
			Gravel or Rock	16
Indigenous Forest	IF	Indigenous Forest	Indigenous Forest	69
Scrub	S	Scrub	Broadleaved Indigenous Hardwoods	54
			Fernland	50
			Gorse and/or Broom	51
			Manuka and/or Kanuka	52
			Sub Alpine Shrubland	55
			Mixed Exotic Shrubland	56
			Matagouri or Grey Scrub	58
Tussock	T	Tussock Grassland	Tall Tussock Grassland	43
			Depleted Grassland	44
Wetlands	W	Inland Wetland	Herbaceous Freshwater Vegetation	45
			Flaxland	47
		Coastal Wetland	Herbaceous Saline Vegetation	46
Exotic Forest	EF	Planted Forest	Exotic Forest	71
			Forest - Harvested	64
Pastoral	P	Primarily Pastoral	High Producing Exotic Grassland	40
			Low Producing Grassland	41
		Primarily Horticulture	Short-rotation Cropland	30
			Orchard, Vineyard or Other Perennial Crop	33
Urban	U	Urban	Built-up Area (settlement)	1
		Urban Open Space	Urban Parkland/Open Space	2
			Transport Infrastructure	5
		Mines and Dumps[1]	Surface Mine or Dump	6
Miscellaneous	M	Mangrove	Mangrove	70
		Willows and Poplars[2]	Deciduous Hardwoods	68
		Coastal Sand	Sand or Gravel	10

Notes:

(1) Mines and Dumps was a LCDB1 class but was not included in the original Snelder et al. (2010) classification. We assumed that this is most appropriately incorporated into the "Urban" land category.

(2) Snelder *et al.* (Table 2.6, 2010) refer to a land cover Class "Riparian Willows", but this is not a category in LCDB1, so we have assumed that this was the LCDB1 category "Willows and Poplars".

We quantified the impact of updating the REC Land Cover category from LCDB1 to LCDB5 data by examining differences in the distributions of the nine REC Land Cover categories across all network segments and monitoring sites for each LCDB version.

5.1.2 Adjustments to REC Land Cover categories associated with LCDB5

LCDB1 comprised 16 land cover classes (Table 2). The number of LCDB classes have been significantly increased over successive versions of the LCDB; there are 30 classes associated with LCDB5 (Table 2)[7]. The original REC Land Cover category rules were based on grouping the 16 LCDB1 classes to the nine REC land cover categories. The definition of the updated REC Land Cover categories described in the previous section of this report was based on the mapping of the LCDB1 to LCDB5 categories shown in Table 2. We then considered whether the mapping shown in Table 2 is consistent with the original intent of the REC Land Cover categories. We did this by examining the descriptions of the LCDB5 land cover categories (Appendix A) and comparing these with the descriptive characteristics of the REC Land Cover categories provided by Snelder *et al.* (2010) (Appendix B). We developed recommendations for adjustments to the way LCDB5 classes are grouped into coarsened LCDB classes prior to applying the REC Land Cover category rules. These adjustments were associated with the grouping of three LCDB5 classes that represent refinements to the original LCDB1 classes: *Depleted Grassland*; *Low Producing Grassland*; and *Broadleaved Indigenous Hardwoods*.

Depleted Grassland is an LCDB5 class that is mapped to the original LCDB1 class of *Tussock* (Table 2). *Low Producing Grassland* is an LCDB5 class that is mapped to the original LCDB1 class of *Primarily Pastoral*. Therefore, LCDB5 defines a gradient in non-natural grassland[8] as *Depleted Grassland – Low Producing Grassland – High producing grassland*. However, based on the original coarsened LCDB class rules, the mapping of LCDB5 *Depleted Grassland* to the LCDB1 *Tussock* class means *Depleted Grassland* is represented in the *Tussock* coarsened LCDB class, whereas *Low Producing Grassland* is represented in the *Pastoral* coarsened LCDB class. We recommend grouping the LCDB5 classes of *Tall Tussock Grassland*, *Depleted Grassland* and *Low Producing Grassland* and calling this coarsened LCDB class *Tussock* prior to applying the REC Land Cover category rules. This new category would be representative of non-productive to low-production grasslands. This change would allow the *Pastoral* coarsened LCDB class to comprise the LCDB5 classes of *High Producing Grassland* and the two horticultural land cover classes *Short-rotation Cropland* and *Orchard, Vineyard or Other Perennial Crop* and be more representative of high intensity agricultural practices.

Broadleaved Indigenous Hardwoods is an LCDB5 class that is mapped to the original LCDB1 class of *Scrub* (Table 2). This means that based on the original REC rules, *Broadleaved Indigenous Hardwoods* are represented by the coarsened LCDB *Scrub* class. The coarsened LCDB classes *Scrub* and *Indigenous Forest* represent an ecological gradient, from gorse and broom, through successional communities of mixed native and exotic scrub species, to predominantly native bush, and fully established indigenous forest. Consideration of the LCDB5 class descriptions (Appendix A) suggests that the new class *Broadleaved Indigenous Hardwoods* is likely to have ecological, hydrological and water quality effects in common with the LCDB5 class *Indigenous Forest*. We therefore recommend grouping *Broadleaved Indigenous Hardwoods* with *Indigenous Forest* and calling this coarsened LCDB class *Indigenous Forest* prior to applying the REC Land Cover category rules.

We quantified the impact of any potential changes in the mapping between LCDB5 and the coarsened LCDB classes by examining the differences in the distribution of network segments

[7] Six classes (of the original 36 LCDB5 classes) are not associated with coarsened LCDB land cover classes: (0) Not land; (20,21,22) classes associated with open water and (80,81) classes only relevant to the Chatham Islands.

[8] "LCDB classes at Version5.pdf" https://lris.scinfo.org.nz/layer/104400-lcdb-v50-land-cover-database-version-50-mainland-new-zealand/

and monitoring sites between the REC Land Cover categories as defined by the original and the alternative mappings.

5.2 Water quality responses to land cover gradients

The second aim of this project was to review the rules used to assign river segments to REC Land Cover categories. We used the available river water quality data to explore relationships between water quality and upstream land cover and used these results alongside the literature review to consider whether there was a case for revising the REC Land Cover membership rules. We were interested in the relative effect sizes of water quality relationships with continuous land cover variables and whether there were non-linearities or thresholds in any of these relationships. The outcomes of these analyses were used to determine whether the rules for defining REC Land Cover category need to be revised, in particular the exceptions used for *Pastoral* and *Urban* categories. In doing these analyses we were cognisant of the multipurpose nature of the REC and therefore did not have the objective of optimising the discrimination of water quality by REC Land Cover. Rather, we considered whether there was evidence for changing the rules for assigning REC Land Cover categories.

In order to distinguish the individual continuous coarsened land cover variables from the REC Land Cover categories we use the prefix "us" to indicate upstream catchment and the category name, e.g., *usPastoral.* Each continuous coarsened land cover variable is available for each segment of DN2.4 and quantifies the proportion of catchment occupied by each of the coarsened LCDB5 classes.

We used histograms to examine the distribution of monitoring sites along gradients in the continuous coarsened land cover variables (i.e., continuous versions of the land cover types used to define the REC Land Cover categories). The histograms were used to evaluate how well the monitoring sites represent the range in the land cover variables used to define the REC Land Cover categories and therefore how robustly the data can describe the land cover - water quality relationships.

We used scatter plots to examine the shapes/functional forms of relationships between the continuous coarsened land cover variables and the site median values of the water quality variables and MCI scores. We log10 transformed the site median values of all variables, except for MCI, to linearise their relationships with the land cover variables.

We used regression models to describe the water quality relationships with the continuous coarsened land cover variables and used the regression coefficients to quantify the effect sizes. In recent national analyses of river water quality state, the response of water quality to a range of drivers was explored using random forest models (Whitehead, Fraser, and Snelder, 2021) and linear regression models (Whitehead, Fraser, Snelder, *et al.*, 2021).. In this study we repeated these analyses with some modifications to methodologies to allow us to focus on the water quality responses to the nine continuous coarsened land cover variable gradients. For the random forest models we excluded the "land use intensity" predictors described in Table 2.2 of Whitehead, Fraser, and Snelder (2021) and replaced land cover variables with the nine continuous coarsened land cover variables described in this report. For the linear regression models, we extended the linear regression modelling from a focus on "upstream intensive agriculture" to include all nine continuous coarsened land cover variables. We examined the correlation between the nine continuous coarsened land cover variables to aid interpretation of the regressions. We made comparisons of regression coefficients to evaluate effect size of the different coarsened land cover classes.

6 Results

6.1 Update REC Land Cover categories

6.1.1 Original REC rules

Figure 1 shows the differences in the distribution of river network segments and monitoring sites between the nine REC Land Cover categories as defined by LCDB1 and LCDB5. There were differences between the two LCDB versions in REC Land Cover category membership for 19% of network segments and 16% of monitoring sites, respectively.

Table 3 and Table 4 summarise the percentage of digital river network segments and monitoring sites, respectively belonging to pairs defined by the original (rows) and updated (columns) REC Land Cover categories. Note that segments with no change in category between the original and updated versions are represented by the diagonal cells in Table 3 and Table 4. The largest change for in category membership (i.e., non-diagonal cells in Table 3 and Table 4) is for 4.4% of segments that were originally categorised *Pastoral* and were updated to *Tussock*. Other notable changes (between 1-2% of network segments) were: *Exotic Forest* to *Pastoral*; *Indigenous Forest* to *Scrub*; *Tussock* to *Bare*; and *Scrub* to *Indigenous Forest*. For the monitoring sites between 2-3% had changes of: *Indigenous Forest* to *Pastoral*; *Pastoral* to *Tussock*; and *Urban* to *Pastoral*. There was also a change of 1.2% from *Exotic Forest* to *Pastoral*.

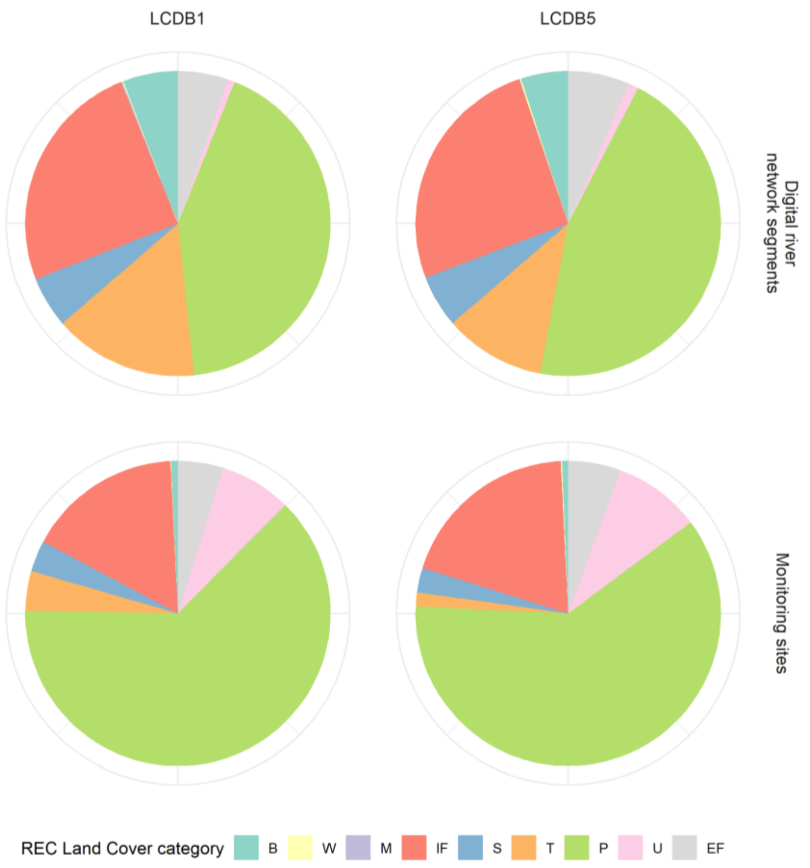

Figure 1: Proportions of monitoring sites and river network segments in each REC Land Cover category for the different versions of the LCDB.

Table 3: Comparison of the proportion of network segments belonging to REC Land Cover categories for assignments based on the original (LCDB1) and updated (LCDB5) versions of LCDB. Values shown are percentages of segments for the entire DN. Blank cells indicate that no segments had membership of the indicated Land Cover categories for the original and updated versions. The diagonal entries indicate the proportion of segments that did not change Land Cover categories between the two LCDB versions.

		REC Land Cover categories based on LCDB5								
		B	**EF**	**IF**	**M**	**P**	**S**	**T**	**U**	**W**
REC Land Cover categories based on LCDB1	**B**	3.94	0.00	0.16	0.00	0.02	0.09	0.79		
	EF	0.02	4.33	0.32	0.01	1.47	0.21	0.11	0.00	0.00
	IF	0.30	0.19	22.05	0.04	0.88	1.23	0.83	0.01	0.01
	M	0.00	0.00	0.00	0.02	0.01	0.00	0.00	0.00	0.01
	P	0.19	0.62	0.79	0.03	38.55	0.69	4.41	0.09	0.02
	S	0.15	0.12	1.10	0.01	0.71	2.74	0.71	0.01	0.01
	T	1.24	0.00	0.39	0.03	0.21	0.40	8.44	0.00	0.00
	U	0.00	0.01	0.03	0.00	0.35	0.02	0.01	0.67	
	W	0.00	0.00	0.02	0.00	0.03	0.03	0.02	0.00	0.06

Table 4: Comparison of the proportion of monitoring sites belonging to REC Land Cover categories for assignments based on the original (LCDB1) and updated (LCDB5) versions of LCDB. Values shown are percentages of monitoring sites. Blank cells indicate that no monitoring sites had membership of the indicated Land Cover categories for the original and updated versions. The diagonal entries indicate the proportion of segments that did not change Land Cover categories between the two LCDB versions.

		REC Land Cover categories based on LCDB5								
		B	**EF**	**IF**	**M**	**P**	**S**	**T**	**U**	**W**
REC Land Cover categories based on LCDB1	**B**	0.52				0.06		0.06		
	EF		3.95	0.32		1.23	0.06			
	IF		0.32	14.68		2.98	0.91	0.13	0.39	
	M									
	P	0.06	0.52	0.97		55.30	0.65	2.65	0.78	
	S	0.06		0.45		0.52	1.49	0.06		
	T	0.06					0.13	1.29		
	U					2.78			6.47	
	W						0.06			0.06

6.1.2 Adjustments to REC Land Cover category mapping to LCDB5

Based on LCDB5, *Low Producing Grassland* occupies approximately 6.6% of the land area of New Zealand. Grouping the LCDB5 classes of *Tall Tussock Grassland*, *Depleted Grassland* and *Low Producing Grassland* and calling this coarsened LCDB class *Tussock* prior to applying the REC Land Cover category rules results in a reduction in the number of segments assigned to the REC Land Cover category *Pastoral* from 45.3% to 37.6% and an increase in number of segments assigned to the REC Land Cover category *Tussock* from 10.7% to 17.7%.

The reassignment of *Low Producing Grassland* would align the continuous coarsened land cover variables with the continuous land cover predictors used in the river state spatial modelling (Whitehead, Fraser, and Snelder, 2021). The proposed change means the continuous coarsened land cover variable *Pastoral* is consistent with the land cover predictor *usIntensiveAgriculture* and the proposed *Tussock* continuous coarsened land cover variable is equivalent to the land cover predictor *usPastoralLight* used in the river state spatial modelling (Whitehead, Fraser, and Snelder, 2021). The *usIntensiveAgriculture* land cover predictor was used by Whitehead, Fraser, and Snelder (2021) because it has greater explanatory power than a land cover predictor that also includes the LCDB5 class *Low Producing Grassland*.

We used the same linear regression analyses used by Whitehead, Fraser, Snelder, *et al.* (2021) to test the performance of two continuous alternative coarsened land cover variables. The first continuous coarsened land cover variable includes the LCDB5 classes *High Producing Grassland* and the two horticultural land cover classes *Short-rotation Cropland* and *Orchard, Vineyard or Other Perennial Crop*. The second continuous coarsened land cover variable is the same as the first but also includes the LCDB5 class *Low Producing Grassland*. The first variable therefore represents the newly proposed continuous coarsened land cover *Pastoral* variable, and the second variable represents the original continuous coarsened land cover *Pastoral* variable.

The variation explained by the two sets of regressions (R^2 values) are shown in Figure 2. These results indicate that the proposed alternative *Pastoral* continuous coarsened land cover variable explains more of the variation observed in water quality for eight of the nine water quality variables.

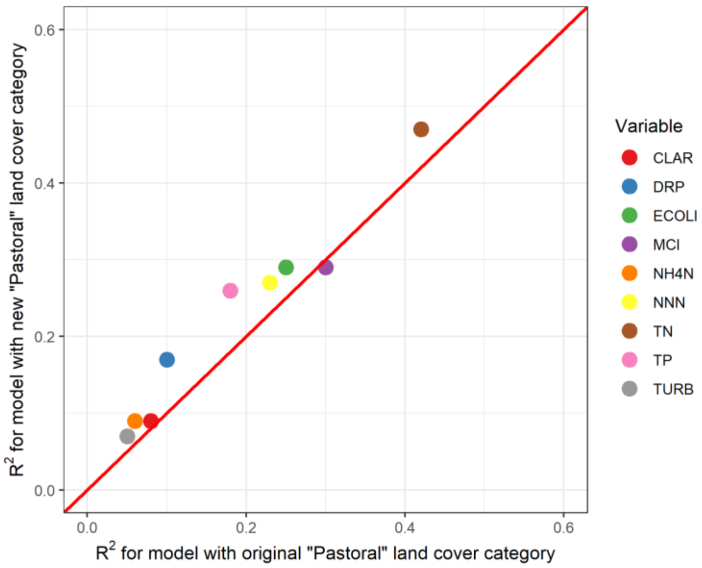

Figure 2: Scatter plot of the R^2 values obtained for linear regressions of the nine water quality variables against two continuous coarsened land cover variables. The x-axis represents a continuous version of the original REC "Pastoral" Land Cover category, and the y-axis represents the proposed alternative REC Land Cover category.

The LCDB5 class *Broadleaved Indigenous Hardwoods* occupies approximately 2.5% of the land area of New Zealand. Grouping this LCDB class with *Indigenous Forest* into the coarsened LCDB class *Indigenous Forest* before applying the REC Land Cover rules results in the number of segments assigned to the REC Land Cover category *Scrub* changing from 5.6% to 4.3% and the number of segments assigned to the REC Land Cover category *Indigenous Forest* from 25.5% to 27.7%.

We also considered whether the LCDB5 class *Gorse and Broom* might be more suitably grouped in with the LCDB classes that define the *Pastoral* continuous land cover variable. This is because the presence of these species might generally indicate low producing pastoral land, and therefore "Gorse and Broom" might be grouped with LCDB classes that define *Pastoral* rather than *Scrub* coarsened LCDB class. However, with the proposal to move *Low Producing Grassland* into the coarsened LCDB class *Tussock*, this change would not be appropriate. We note that the LCDB5 class *Gorse and Broom* category only occupies 0.7% of the land area of New Zealand. Therefore, the grouping of this land cover class is unlikely to have any significant impacts on the overall REC Land Cover classification.

Prior to applying the REC Land Cover category membership rules we applied the following two changes discussed above:

- The LCDB5 *Low Producing Grassland* was grouped with *Tall Tussock Grassland* and *Depleted Grassland* to define the coarsened LCDB class *Tussock*.
- The LCDB5 *Broad Leaved Indigenous Hardwood* was grouped with *Indigenous Forest* to define the coarsened LCDB class *Indigenous Forest*.

We then calculated the dominant land cover class based on the coarsened LCDB classes for each segment of the river network and applied the original rules regarding dominance of *Pastoral* and *Urban* land cover (see Section 5.1.1).

Figure 3 shows the differences in the distribution of river network segments and monitoring sites between the nine REC Land Cover categories as defined by the original and adjusted coarsened LCDB classes. Between the two versions of the coarsened LCDB classes (i.e., original and adjusted), there were differences in REC Land Cover category membership for 10% of all network segments and 9% of monitoring sites.

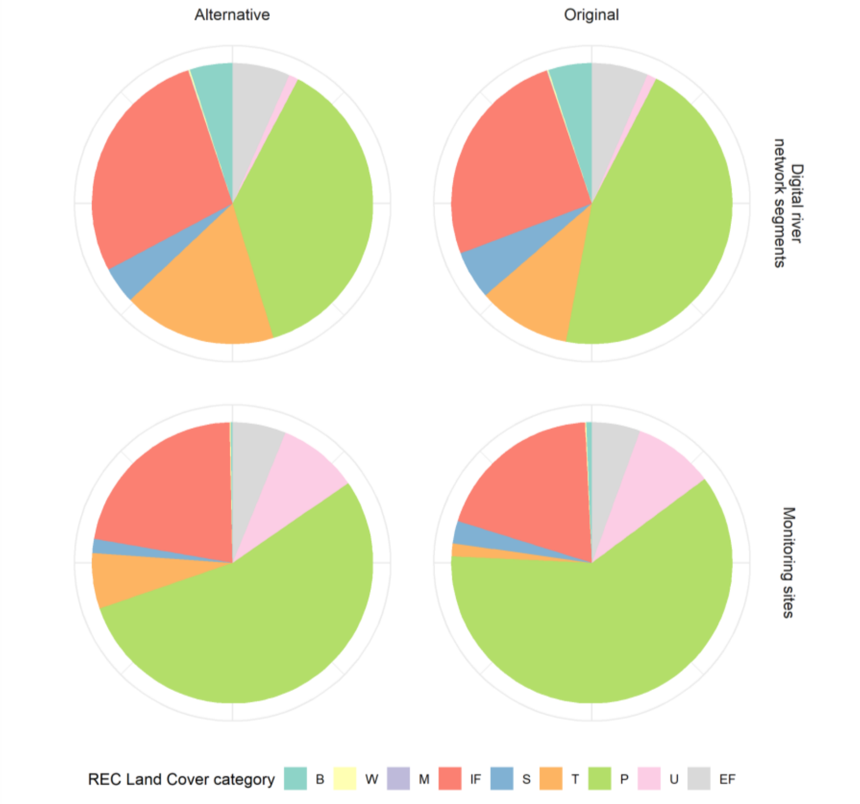

Figure 3: Proportion of monitoring sites and river network segments in each REC Land Cover category for the different versions of continuous coarsened land cover variables.

The proposed new mappings of the LCDB5 class identifiers to coarsened LCDB classes (i.e., and update to Table *2*) is shown in Table *5*.

Table 5. Proposed updated mapping of LCDB5 classes to define the coarsened LCDB5 classes.

Coarsened LCDB5 Classes	LCDB5 class identifiers
Bare	12, 14, 15, 16
IndigenousForest	69, 54
Scrub	50, 51, 52, 55, 56, 58
Tussock	41, 43, 44
Wetlands	45, 46, 47
ExoticForest	71, 64
Pastoral	40,30,33
Urban	1, 2, 5, 6
Miscellaneous	10, 70, 68

6.2 Water quality responses to land cover gradients

6.2.1 Available data and correlations between land cover gradients

The distributions of the continuous coarsened land cover variables (identified by the prefix "us") based on the mapping in Table 5 for the water quality monitoring sites are shown in Figure 4. For most of the land cover variables, the distributions are strongly left skewed (i.e., most sites have low proportions of their catchments occupied by a given land cover class). The exception is for *usPastoral* and *usIndigenousForest*, which have reasonable numbers of sites representing the entire gradient from zero to 100% cover. 99% of sites have less than 5% upstream land occupied by *usMiscellaneous* and *usWetland*. 95% of sites have less than 10% upstream land occupied by *usBare.* Only 8% and 17% of sites had more than 15% of upstream land occupied by *usUrban* or *usTussock*, respectively.

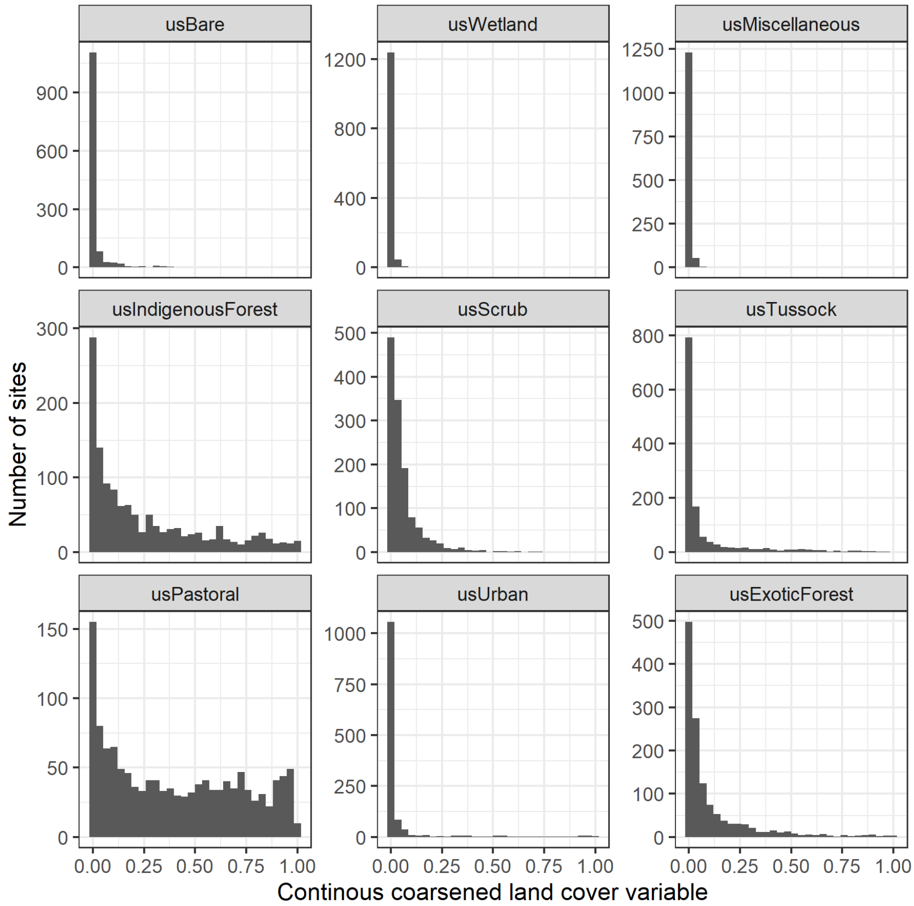

Figure 4: Distribution of the continuous coarsened land cover variables for monitoring sites with water quality data. The plots show the distributions of sites with respect to the nine ccontinuous coarsened land cover variables (i.e., continuous versions of the land cover variables that are used to define the REC Land Cover categories).

The correlation between the continuous coarsened land cover variables for sites with water quality state data is shown in Figure 5. The continuous coarsened land cover variables are not independent of each other, as by definition they sum to one for each site. The largest absolute correlation was between *usPastoral* and *usIndigenousForest*. *usPastoral* also had moderate negative correlations with *usScrub*, *usBare*, *usTussock* and *usExoticForest*. *usUrban*, *usWetland* and *usMiscellaneous* generally had low correlations with other variables, mostly because these are relatively less common classes and are not present upstream (or only as small proportions) for most monitoring sites.

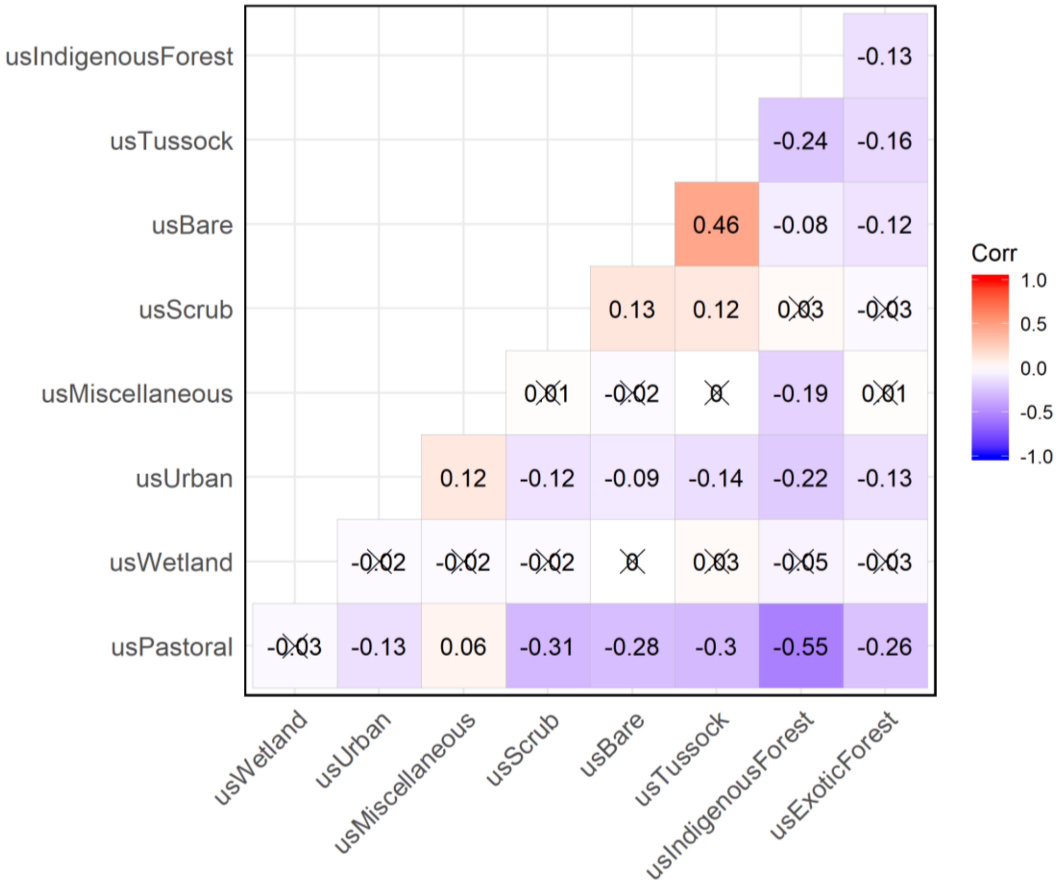

Figure 5: Correlation matrix of continuous coarsened land cover variables at monitoring sites. Cells with 'X' indicate that the correlation was not statistically significant (α = 0.05)

Correlations between the water quality variables and the continuous coarsened land cover variables are shown in Figure 6 and associated scatter plots are shown in Figure 7. Due to their poor representation in the dataset, the classes *usMiscellaneous* and *usWetland* are not shown in the scatter plots. Both the correlation matrix and the scatter plots indicate that the strongest linear relationships between water quality and land cover are for the *usIndigenousForest* and *usPastoral* classes. Increasing proportions of upstream area in the *usIndigenousForest* class are associated with decreasing contaminant concentrations and increasing clarity and MCI, i.e, better water quality. The opposite relationships apply to the *usPastoral* class. The scatter plots shown in Figure 7 indicate relationships are generally linear. Where non-linear responses are indicated by the fitted loess model (red lines shown in Figure 7), there are generally too few data points for these to be considered reliable.

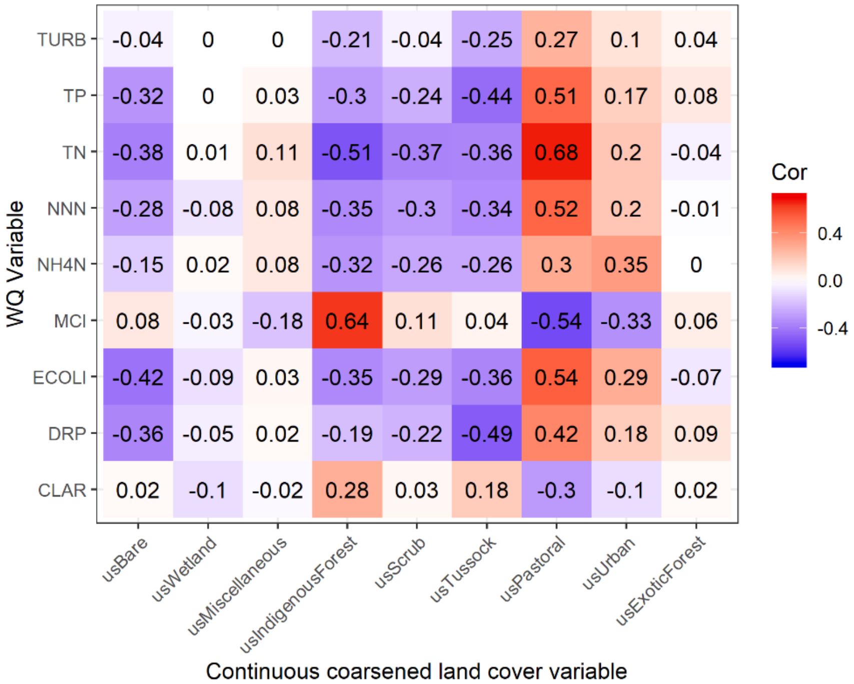

Figure 6: Correlation matrix, showing correlation between each pair of water quality variables observed at the monitoring sites and the corresponding continuous coarsened land cover variables.

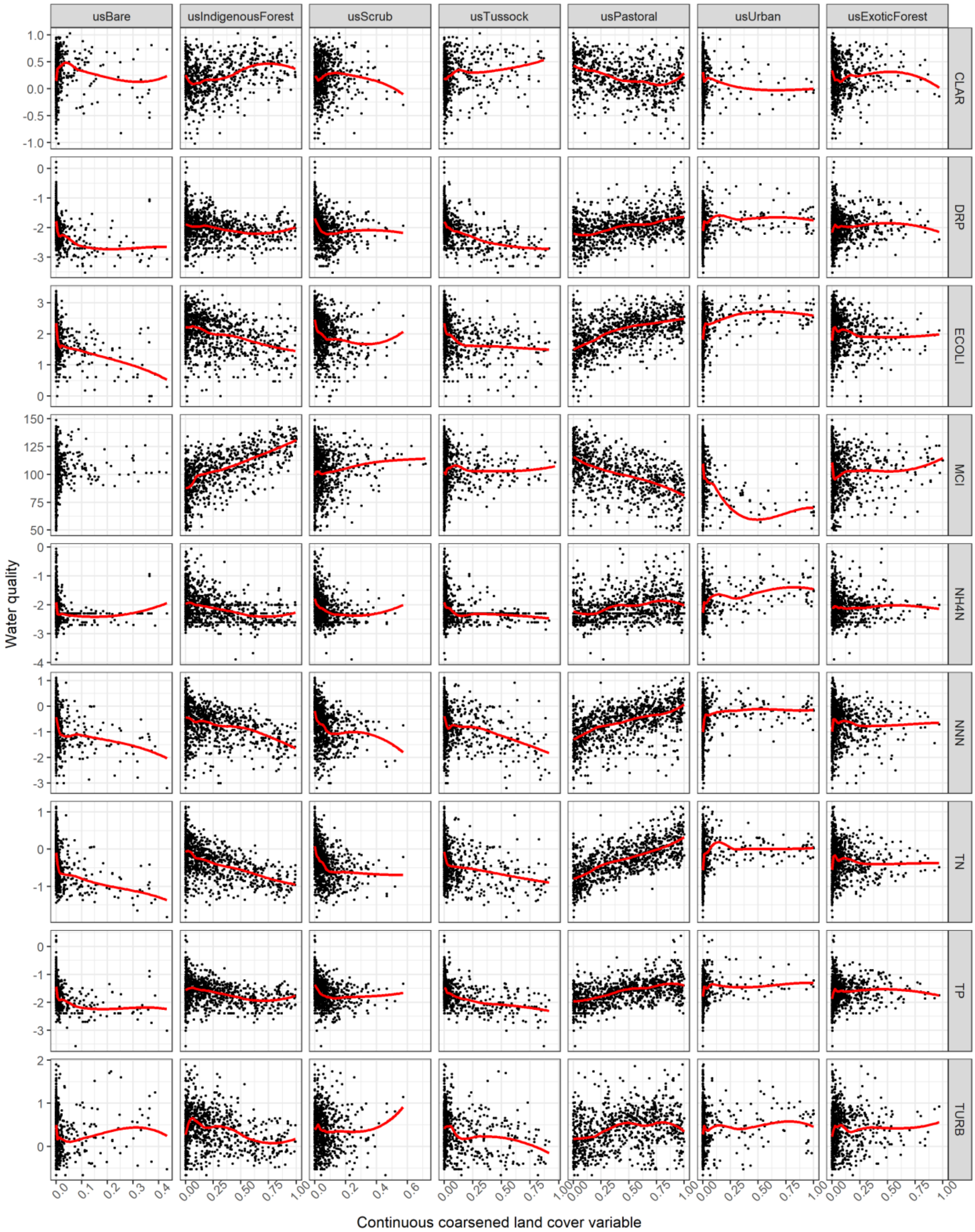

Figure 7: Scatter plots of continuous coarsened land cover variables against the water quality data. The water quality data (i.e., site median values of the observations of the nine water quality indicators) were log10 transformed, with the exception of MCI which is not transformed. The red line is a loess model with a span of 0.5.

6.2.2 Random forest models

The random forest models that used the continuous coarsened land cover variables as predictors performed well for most water quality variables, as indicated by the following statistics: $R^2 > 0.5$, NSE > 0.5, and RMSD < 0.5 for most variables (Table 6). Based on NSE values, the TN, ECOLI and MCI models had the best overall performance, the NH4N model had the worst overall performance, and the TP, NNN, DRP, TURB and CLAR models had intermediate performance. Overall, the random forest model performance and importance was very similar to that of Whitehead, Fraser, Snelder, *et al.* (2021).

The rank importance order of the continuous coarsened land cover variables in the random forest models is shown Table 7. *usPastoral* was included in all models and was in the top three for four variables (ECOLI, TN, NNN, and TP). *usUrban* was in the top 10 predictors for five variables (ECOLI, MCI, NH4N, NNN and TN), but was not included in the models relating to suspended sediment (CLAR, TURB) or phosphorus (TP and DRP). *usBare* was ranked third or lower for ECOLI, TP and DRP. *usIndigenousForest* was the most important variable for the MCI model and ranked fourth for the TN model. *usTussock* was ranked as the third most important variable for the DRP model. *usScrub*, *usExoticForest* and *usWetland* generally had low importance or were not included in the models.

Partial plots show that MCI and CLAR decrease with increasing *usPastoral* and *usUrban* and all other variables increased with decreasing *usPastoral* and *usUrban* (Figure 8). The other continuous land cover variables had the opposite relationship with the response to that of *usPastoral* and *usUrban*.

The partial plots for *usUrban* indicate a non-linear response, with sharp initial increases and then flattening off. The inflection points are between 5-10% cover for ECOLI, TN, NNN and NH4N and around 20% for MCI. However, the distribution of the *usUrban* variable was highly skewed, with more than 90% of sites having less than 10% *usUrban*. This will influence the shape of the response shown in the partial plots. The partial plots show approximately linear responses of MCI, DRP, NH4N and TN to *usPastoral*; the other variables show patterns of higher rates of change for lower *usPastoral* proportions, and lower rates of change after between 30-40% cover.

Table 6:Performance of the water quality models. Performance was determined using independent predictions (i.e., sites that were not used in fitting the models) generated from the out-of-bag observations. Regression R2 = coefficient of determination, NSE = Nash-Sutcliffe efficiency, PBIAS= percent bias, RMSD = root mean square deviation). Units for RMSD and bias are the log10 transformed units of the respective water quality variables except for MCI, for which RMSD is based on non-transformed data.

Variable	Number of sites	Regression R^2	NSE	PBIAS	RMSD
CLAR	715	0.59	0.58	1.38	0.22
DRP	970	0.61	0.61	0.10	0.30
ECOLI	964	0.70	0.69	-0.29	0.33
MCI	955	0.71	0.71	0.05	10.11
NH4N	970	0.51	0.51	0.29	0.34
NNN	943	0.63	0.62	0.16	0.45
TN	935	0.74	0.74	0.66	0.24
TP	898	0.70	0.69	0.12	0.25
TURB	832	0.59	0.58	-0.10	0.30

Table 7:Rank order of importance of predictor variables retained in the random forest models for at least one water quality variable. Blank cells indicate that the predictor was not included in the reduced model. The predictor variables in the first column are listed in descending order of the median of the rank importance over all nine models.

Predictor	CLAR	DRP	ECOLI	MCI	NH4N	NNN	TN	TP	TURB
usPastoral	6	11	2	4	10	1	2	2	6
usUrban			7	8	5	9	8		
usIndigenousForest	16	15	12	1	15	5	4	11	17
usBare	13	1	3	35	13	22	6	3	15
usScrub			11	26		10	9	21	
usExoticForest			13	14		11	17	18	
usTussock		3	15	31	21	17			
usWetland	20		30	34		27			21

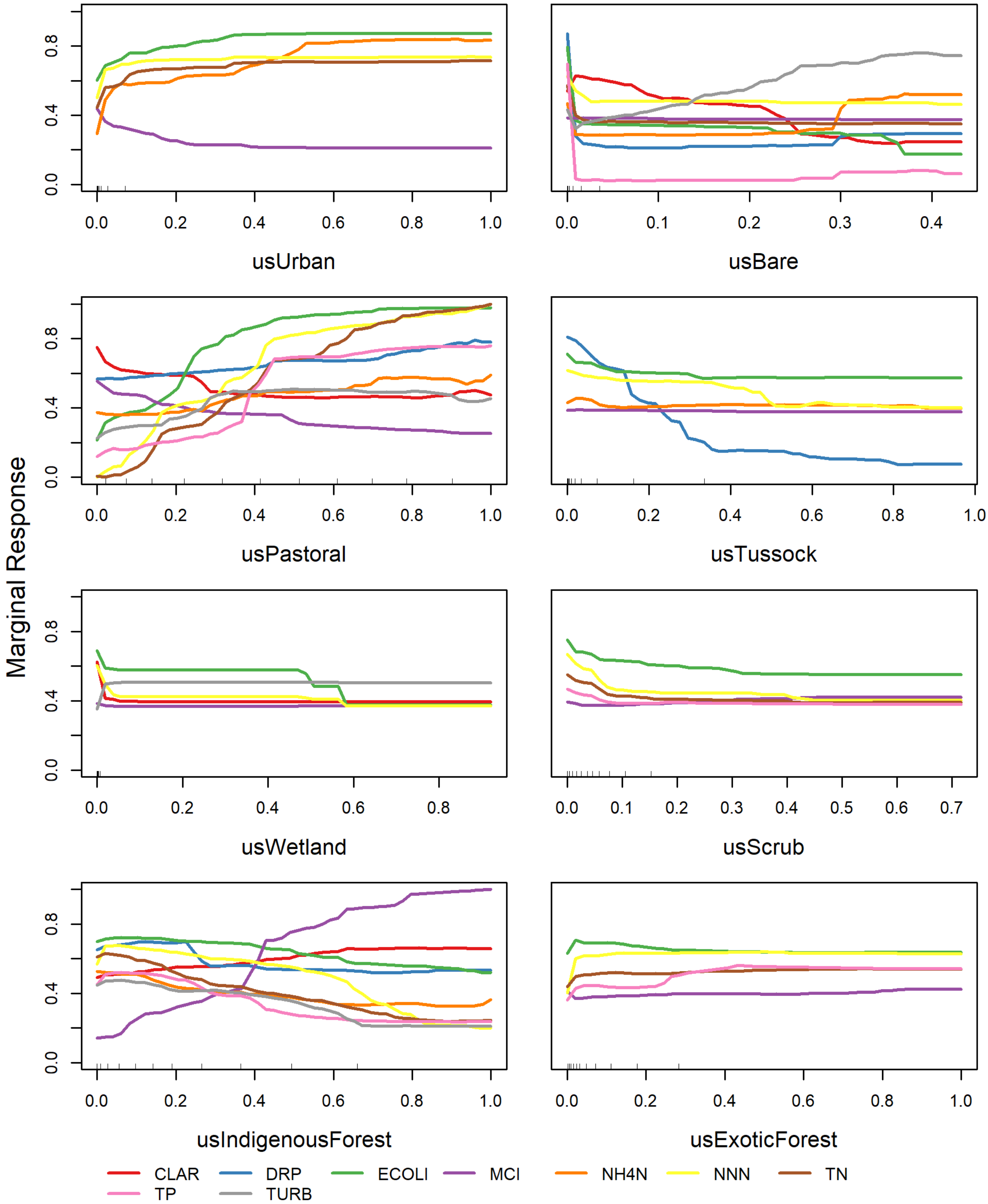

Figure 8: Partial plots for the eight continuous coarsened land cover variables used as predictor variables in random forest models of site median water quality. Each panel corresponds to one predictor. Y-axis scales represent the standardised value of the marginal response for each of the eight modelled response variables. In each case, the original marginal responses over all eight predictors were standardised to have a range between zero and one. Plot amplitude (the range of the marginal response on the Y-axis) is directly related to a predictor variable's importance; amplitude is large for predictor variables with high importance. The grey "rug" indicates deciles of the data on the x-axes of each panel.

6.2.3 Linear regression

The moderate to strong (negative) correlations between the continuous coarsened land cover variables *usPastoral*, *usIndigenousForest* and *usExoticForest* present challenges for isolating the effect size of the continuous coarsened land cover variables based on linear regression analysis. This is because regression coefficients can be interpreted as a mean change in the dependent variable for each unit change in the independent variable when all other variables are held constant. However, when the independent variables are correlated, changes in one variable will be associated with changes in another and regression coefficients are unreliable indicators of the change in the response with unit change in the independent variable.

The Random Forest models indicated that *usUrban* and *usPastoral* variable were, on average across all water quality variables, the two most important land cover predictors. They also indicated that *usExoticForest* was a relatively unimportant predictor for all nine water quality variables. The remaining continuous coarsened land cover variables are highly to moderately negatively correlated with *usPastoral* and *usUrban. usPastoral* and *usUrban* are slightly negatively correlated.

Based on the above observations we refined our approach to using linear regression modelling to investigate the effect size in response to land cover gradients as five steps:

1. Fit linear regressions to each water quality variable using just *usPastoral* as an independent variable. Use scatter plots of the residuals of this model against *usUrban* to determine whether *usUrban* explains additional variation in water quality. Assess whether there are thresholds or non-linear responses.

2. Fit multiple linear regressions to each water quality variable using both *usPastoral* and *usUrban* as independent variables.

3. Fit multiple linear regressions to each water quality variable using all continuous coarsened land cover variables (Table 5) as independent variables. These 'saturated' (i.e., containing all continuous coarsened land cover variables) models provide an estimate of the maximum variability that can be explained by land cover. Regression coefficients are not interpreted due to collinearity of the independent variables.

4. Compare the variance explained (R^2) by each of the models described in steps 1-3, above.

5. Compare the magnitudes of the regression coefficients and their uncertainties from step 3 between *usUrban* and *usPastoral.*

Figure 9 shows scatter plots of *usUrban* against the residuals of the linear models based on *usPastoral*. Note that the scatter plots of *usUrban* against the water quality variables (Figure 7) do not indicate particularly strong relationships with water quality. However, after accounting for the effects of *usPastoral* (Figure 9), there are reasonably linear relationships between *usUrban* and most of the water quality variables.

A summary of the R^2 values from the three linear regression models fitted to each of the nine water quality variables is shown in Figure 10. The regressions that included *usPastoral* and *usUrban* always have higher R^2 values than the regressions that used only *usPastoral*. We note that for each water quality variable, the differences in the R^2 values for models that included *usPastoral* and *usUrban* compared to models that included only *usPastoral* were always statistically significant (ANOVA; $\alpha = 0.05$).

The models that included only *usUrban* and *usPastoral* as independent variables explained a large proportion of the variation explained by saturated models (from 62% to 96% of the maximum explained variation, with a median of 82%). This indicates that of the nine continuous coarsened land cover variables, *usPastoral* and *usUrban* were the most important.

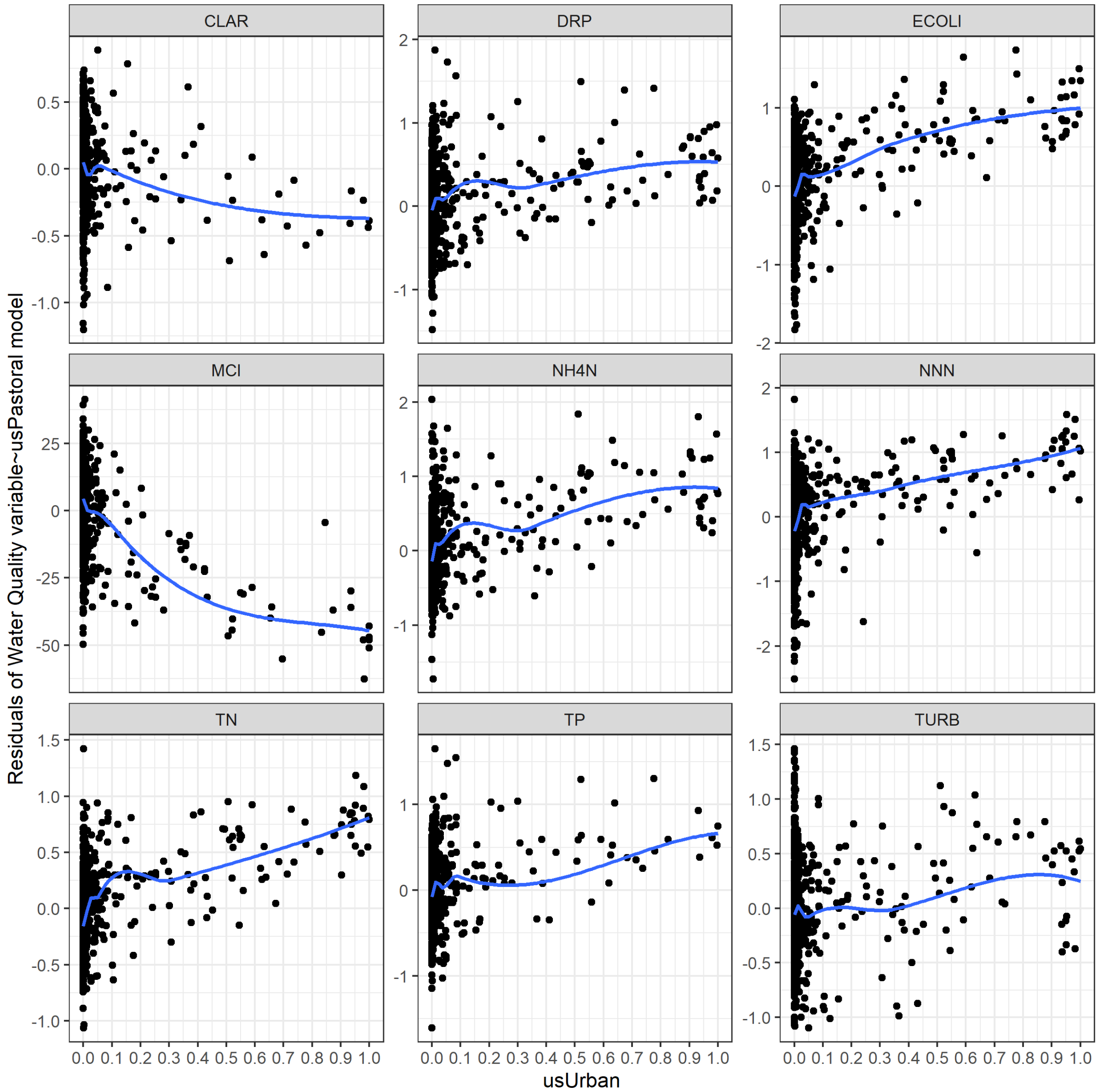

Figure 9: Scatter plot of usUrban versus the residuals of the linear model relating usPastoral to the water quality variables. Blue lines represent a loess fit of the data (span = 0.5).

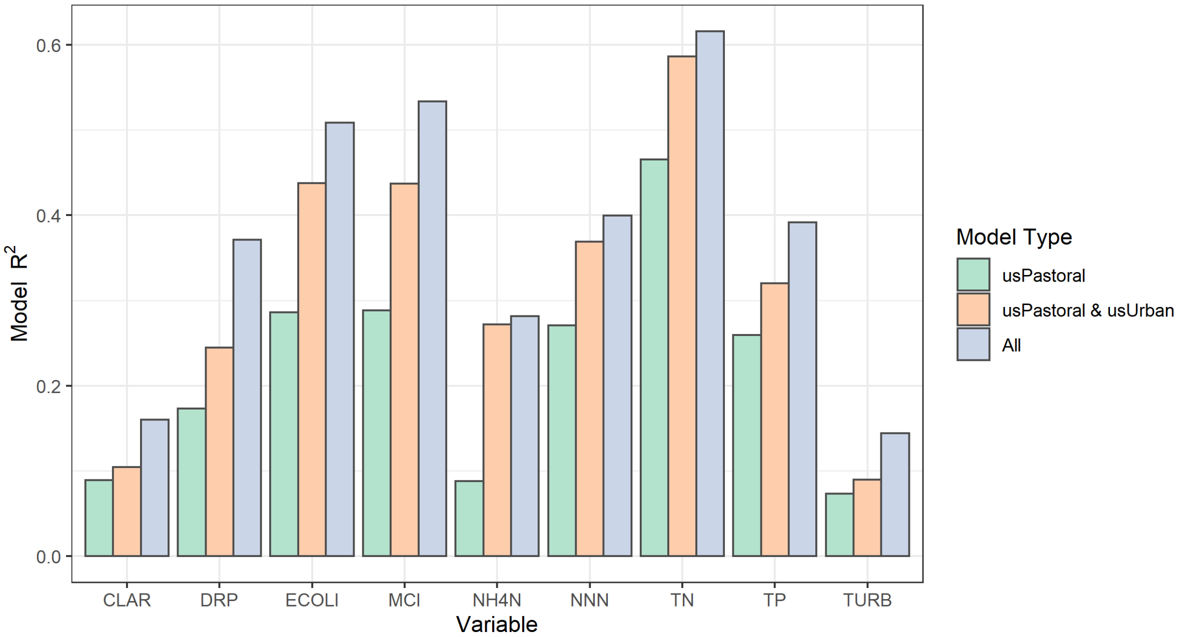

Figure 10: Comparison of model performance between different regression models.

The magnitude of the regression coefficients for the models that used *usPastoral* and *usUrban* as independent variables are shown in Figure 11. Per unit change in *usPastoral* is associated with larger effects for NNN, TN and TURB. For the other water quality variables, the effect of *usUrban* was larger than *usPastoral*. For three water quality variables (TN, NH4N, and MCI) the difference between the regression coefficients did not have overlapping 95% confidence intervals, indicating that the differences in the effects attributable to *usPastoral* and *usUrban* are significant. For the other water quality variables, the differences in the effects attributable to *usPastoral* and *usUrban* were similar. Figure 11 also demonstrates the ratios of the *usUrban* to *usPastoral* regression coefficients and indicates the relative difference in the effect size of *usUrban* compared to *usPasture.* These ratios ranged from 0.82 (TN) to 2.06 (NH4N).

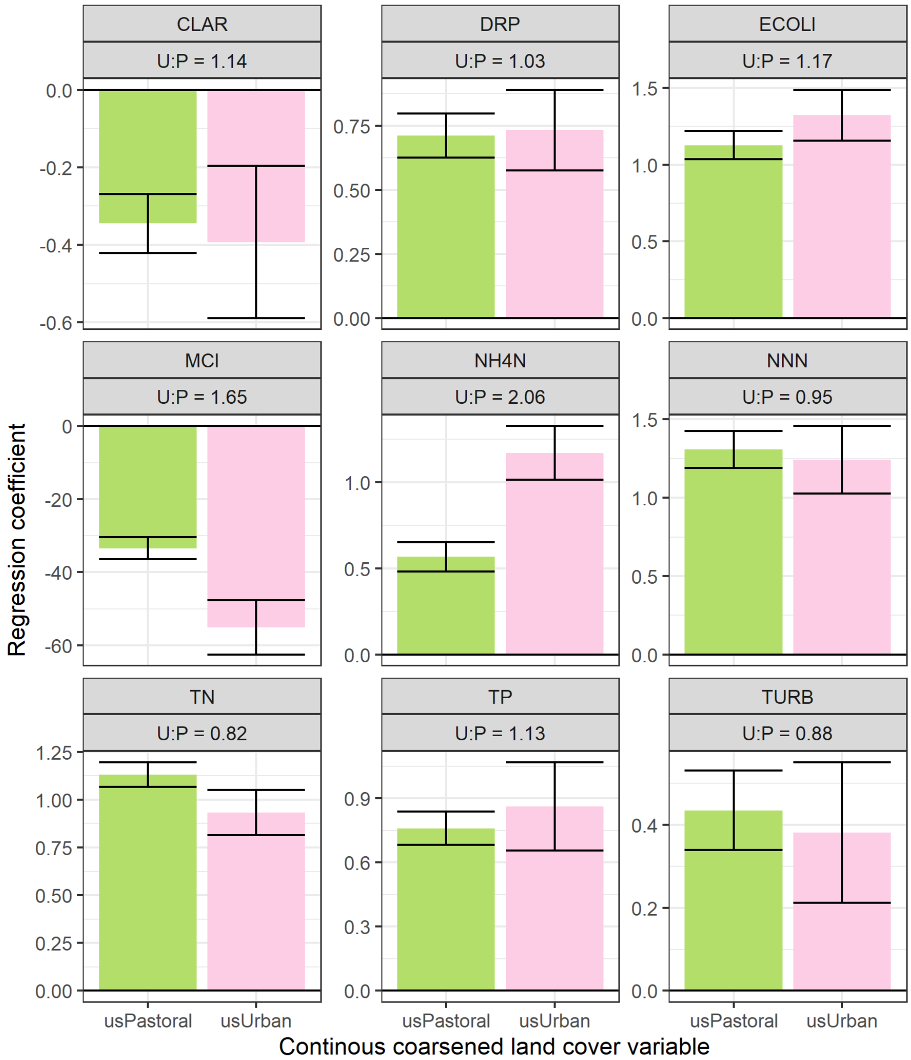

Figure 11: Regression coefficients for multiple linear regressions fitted to each water quality variable using both usPastoral and usUrban as independent variables. Error bars show 95% confidence intervals for the regression coefficients. U:P is the ratio of the usUrban to usPastoral regression coefficients and indicates the relative difference in the effect size of usUrban compared to usPasture.

7 Use of land cover categories to describe water quality patterns and alternative approach

REC Land Cover categories are often used to indicate broad scale differences in water quality associated with different types of catchment land cover. The discrimination of water quality differences based on REC Land Cover categories provides an easily understood method of demonstrating the relationships between land use and water quality variables. REC Land Cover categories have sometimes been further simplified by grouping categories that are interpreted as indicating "Natural" land cover. For example, Larned *et al.* (2018) used box-plots to demonstrate broad-scale variation in water quality variables by defining a *Natural* land cover "super category" by combining *usScrub*, *usIndigenousForest*, *usTussock*, *usBare*, and *usWetland* and retaining the REC Land Cover categories *Pastoral*, *Exotic Forest* and *Urban*. We refer to this classification as Simplified REC Land Cover categories.

The analyses undertaken by this, and other studies, indicate that the continuous coarsened land cover variables (Table 5) partly explain between site variation in water quality variables. Because REC Land Cover categories are categorical versions of the continuous variables, the categories also explain variation in water quality variables. If there are statistically significant differences in water quality between REC Land Cover categories, it is legitimate to indicate that patterns in water quality are approximated by the patterns shown by the REC Land Cover categories. However, there are two important issues that need to be kept in mind when using REC Land Cover categories in this way. First, land cover is only one of many factors that determine water quality. Land cover categories alone are a very simple model of water quality variation, and a considerable amount of between-site variation is unexplained by this model.

The second issue is that catchments generally comprise a mix of land cover types. The REC simplifies catchment land cover into a single 'dominant' land cover category. The dominant land cover designation is a simplification that may obscure the contribution of the remainder of land cover classes in catchments to water quality outcomes. When used in environmental reporting this can lead to the interpretation that a particular land cover category (e.g., urban) is attributable for water quality outcomes when other land cover categories (e.g., pastoral) are also involved.

This problem of obscuring the contribution of other land cover types is particularly relevant for the *Urban* REC Land Cover category. The plot shown in Figure 3 indicates that the *Urban* REC Land Cover category sites made up 7.8% of all sites. For these sites, a significant proportion of the upstream catchment was occupied by pastoral land cover (i.e., *usPastoral;* Figure 12*)*. For example, for *Urban* REC Land Cover category sites in our dataset, the combined proportion of *usUrban and usPastoral* exceeded 0.8 for 75% of sites, the median and maximum of *usPastoral* was 0.31 and 0.84, respectively (Figure 12). In addition, 40% of *Urban* REC Land Cover category sites had *usPastoral* greater than *usUrban,* and for 20% of these sites *usPastoral* was greater than double *usUrban.* Although across all sites the correlation between *usPastoral* and *usUrban* was small (-0.14), within the *Urban* REC Land Cover category the correlation was larger (-0.77). This indicates that the most significant complementary land cover type in catchments assigned to the *Urban* REC Land Cover category was pastoral*.*

In contrast, *55%* of sites were assigned to the *Pastoral* REC Land Cover category*.* Of these sites 26% had zero *usUrban* and 94% of all *Pastoral* sites had less than 5% *usUrban.* The conclusion from this is that there was more likely to be a hidden influence of pastoral land cover at sites in the *Urban* REC Land Cover category than the reverse. Given that the effect size of *usUrban* and *usPastoral* was generally similar in magnitude (Figure 11), the *Urban* REC Land Cover category generally represents sites with large (combined) water quality influences. The reverse is also theoretically true, that is, when urban land cover is a large component of a catchment assigned to the *Pastoral* REC Land Cover category, the water quality outcome is due to the combined influence of the two land cover types. However, for the sites included in our analysis, the ratio of *usUrban:usPastoral* for sites with REC land cover category *Pastoral* range from 0 to 0.36, with a median of 0.003, and 95% of sites had a ratio of less than 0.09.

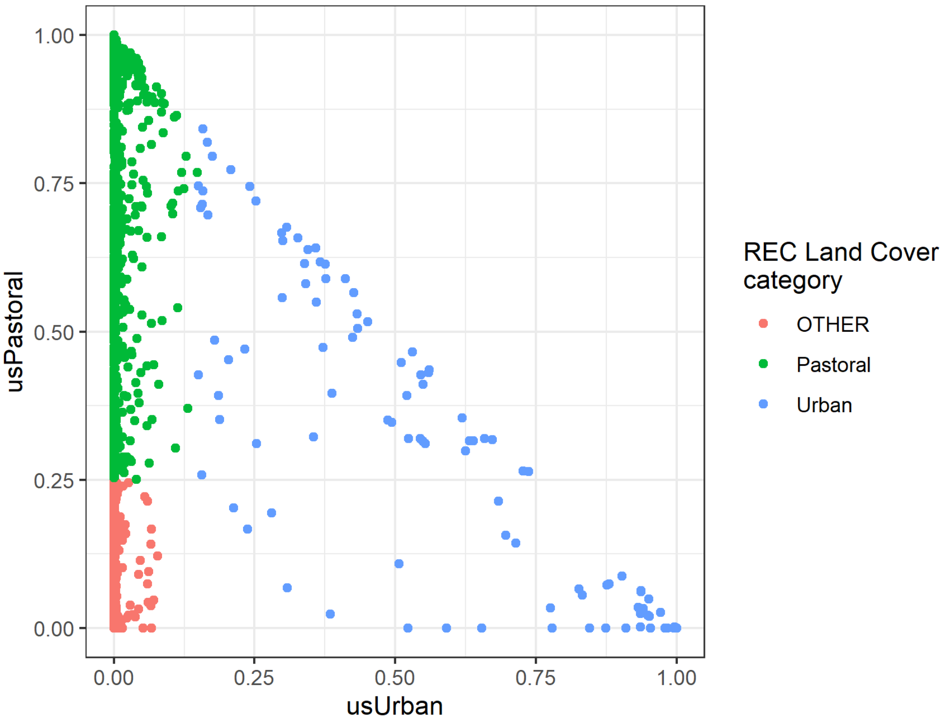

Figure 12. Scatterplot showing usUrban vs usPasture and REC Land cover categories

We propose two alternatives that go some way to addressing the above two issues. Both alternatives have the disadvantage of requiring a higher level of technical understanding from the user, however, at least in some circumstances this may be seen to be outweighed by the advantages. First, the effect of land cover on water quality is best revealed by regression modelling such as demonstrated using both linear and random forest regression models in the previous section of this report. Regression models can be used to indicate the effect size (i.e., mean change in the water quality variable for each unit change in specific land cover types when all other independent variables in the model are held constant). There are complications with this due to collinearity, which we have discussed above. The solution to this is to restrict the independent variables in the model to those that have low correlation (see Section 6.2.3). This approach was taken by Larned *et al.* (2016) to show that water quality variation was strongly associated with high-intensity agricultural and urban land cover types in lowland streams in New Zealand. Similar analyses have also been used in the two most recent national river water quality state analyses prepared for MFE (Larned *et al.*, 2018; Whitehead, Fraser, Snelder, *et al.*, 2021).

If the above approach is too complicated but the assignment of sites to a dominant land cover category is unacceptable, we suggest an alternative categorical approach may be appropriate for using land cover to discriminate water quality variation. The approach is to use categories in the degree of catchment land cover “anthropogenic modification”. Catchment land cover “anthropogenic modification” categories can be derived in two steps. First, for every site and/or network segment, the continuous coarsened land cover variables that represent highly modified land cover types are added to produce a new variable representing an anthropogenic modification gradient. For example, the anthropogenic modification gradient could be defined by adding the following continuous coarsened land cover variables: *usPastoral, usUrban and usExoticForest.* Figure 6 demonstrates that increases in proportions of these variables are associated with poorer water quality outcomes. High values of the anthropogenic modification variable indicate the catchment land cover is subject to high anthropogenic modification and resource use. Low values of the anthropogenic modification variable indicate the catchment land cover is dominated by “natural” or low intensity land cover classes such as *Indigenous Forest* and *Tussock*. Second, the anthropogenic modification gradient is subdivided into a nominated number of categories, for example by subdividing into equal intervals. The categories therefore subdivide the sites or the network into categories representing the least modified catchments at one end of the gradient to the most modified at the other end.

An example of the discrimination of the water quality data used in this study using these anthropogenic modification categories is shown in Figure 13, along with categorisation by the REC land cover categories, and the Simplified REC land cover categories. The example shows four alternative numbers of anthropogenic modification categories (3, 4, 6 and 8 categories). Figure 13 indicates that all water quality variables vary systematically along the anthropogenic modification gradients represented by the categories.

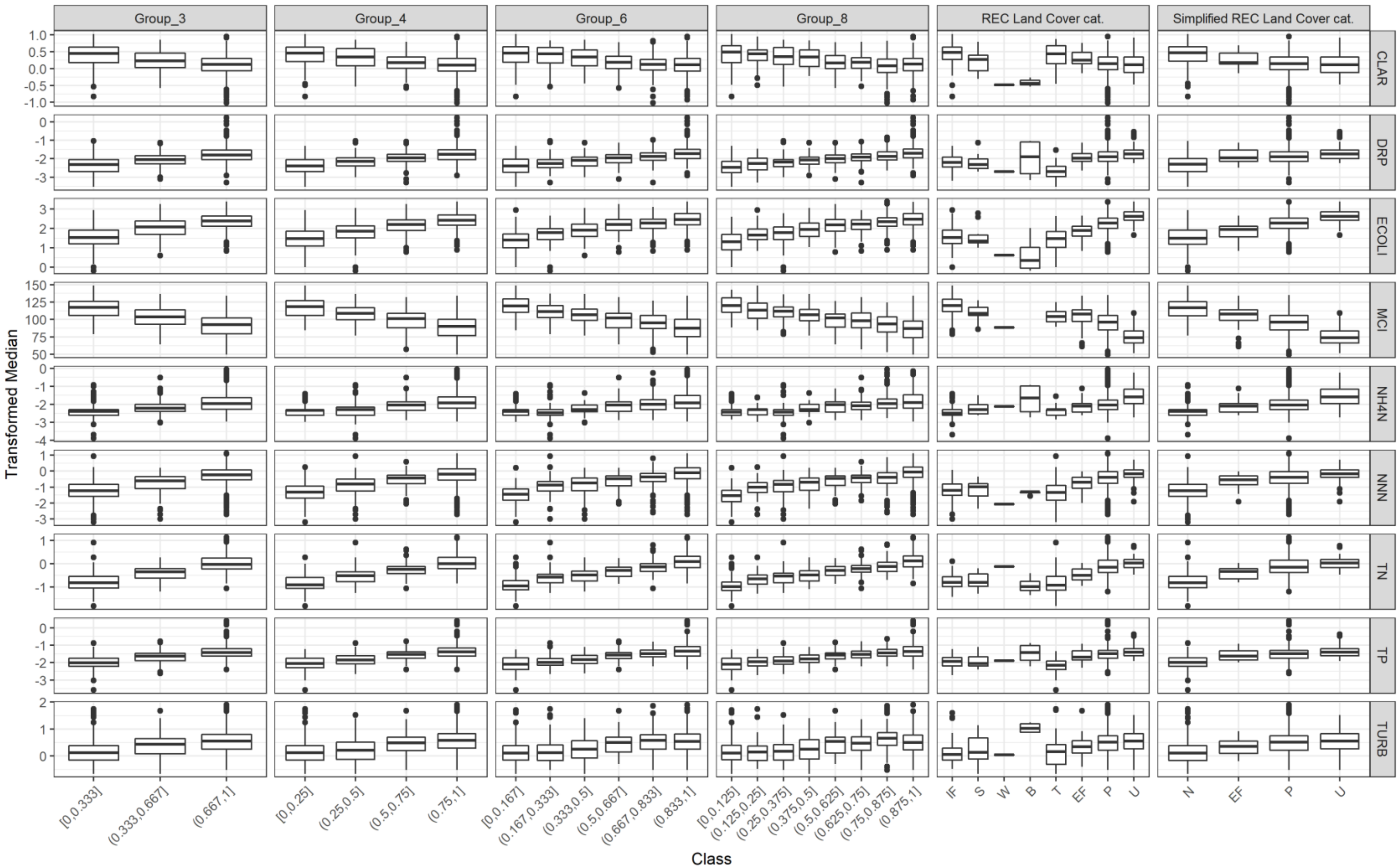

Figure 13: Box and whisker plots of water quality grouped by six alternative land cover classifications. The first four (from left to right) are classifications based on anthropogenic modification categories with the anthropogenic modification gradient subdivided 3,4,6 and 8 categories of equal interval. The two classifications on the right are the REC Land Cover categories (updated as described in section 6.1.2) and the Simplified REC land cover categories.

8 Summary and conclusions

In this study we updated the REC Land Cover categories to reflect the latest national scale map of land cover of New Zealand (LCDB5). We have recommended some changes to the grouping of land cover classes that are defined by LCDB5 prior to applying the REC Land Cover category membership rules. We have reviewed literature concerning the relationship between catchment land cover and water quality and ecological measures and undertaken analyses of these relationships. We have used these two sources of evidence to consider whether changes to the REC Land Cover category membership rules are justified. Finally, we have considered the use of REC Land Cover categories to describe water quality variation. We have proposed alternative approaches to overcome two issues that arise from the use of REC Land Cover categories in environmental reporting. The following sections summarise our study results and findings and draw conclusions from these.

8.1 Updating REC Land Cover categories

This study has produced a table that assigns all segments of the DN (version 2.4) to a REC Land Cover category based on LCDB5. In making this update, we have not changed the REC Land Cover category membership rules which assign segments to a "dominant" catchment land cover with special conditions applying to determine dominance by *Urban* and *Pastoral* land cover categories. This is because we do not consider that there is good evidence for changing the original REC Land Cover category membership rules (see next section). However, we recommend, and have implemented, two changes in the definition of the coarsened LCDB5 classes so that:

- LCDB5 class 41, *Low Producing Grassland* is assigned to coarsened LCDB class *Tussock*; and
- LCDB5 class 54, *Broad Leaved Indigenous Hardwood* is assigned to the coarsened LCDB class *Indigenous Forest*.

We make the first recommendation because we consider that the LCDB5 class *Low Producing Grassland* is more consistent with the *Tussock* REC Land Cover category than the *Pastoral* category that is suggested by the LCDB mapping (Table 2). This change also makes the *Pastoral* REC Land Cover category consistent with *Intensive agriculture* land cover predictor that is used for predictive water quality modelling (e.g., Whitehead, Fraser, and Snelder 2021).

We make the second recommendation because we consider that the LCDB5 class *Broad Leaved Indigenous Hardwood* is more consistent with the *Indigenous Forest* REC Land Cover category than the *Scrub* category that is suggested by the LCDB mapping (Table 2). The implications of these changes in terms of the proportions of network segments that would be assigned to different REC Land Cover categories is detailed in Figure 3 of this report. We provide both sets of REC Land Cover categories with this report as supplementary material.

8.2 Review of the REC Land Cover category membership rules

We reviewed the literature and conducted our own analyses of the relationships between water quality and catchment land cover. We were primarily interested in the size of the effect on water quality and ecological metrics in response to varying proportions of catchment land cover types. We were also interested in whether responses are continuous and linear, non-linear or threshold in nature. Non-linear or threshold responses would provide an obvious and justifiable basis for rules whereas linear responses would mean rules would need to be based on judgements about the levels at which a particular land cover effect is likely to dominate river character.

Neither the literature, nor our own analyses, clearly indicate threshold responses for any land cover types. In general, relationships between water quality (appropriately transformed) and ecological measures and catchment land cover are approximately linear. Several studies and our own analyses indicate that the effect on water quality of pastoral and urban land cover are similar in magnitude (i.e., the change in water quality and ecological measures with unit change in catchment pastoral and urban land cover) and greater than that of other land cover classes. However, several studies and our own analyses indicate that the effect of urban land cover on ecology (as represented by invertebrate assemblages) is greater than that of pastoral land cover (e.g., MCI, Figure 11).

The larger effect of pastoral and urban, compared to the other land cover types is qualitatively consistent with the rules employed by the REC that assigns these as dominant land covers when their areas exceed a fixed proportion of the upstream catchment. These rules recognise that due to the larger effect size, occupancy of the catchment by upstream pastoral and urban land has a disproportionate influence on a range of river characteristics. We consider that the literature and our own analyses support the principle that these two land cover classes are considered dominant when their areas exceed a fixed proportion of the upstream catchment. However, the thresholds of 25% and 15% that are used to assign dominance by pastoral and urban land to a catchment are subjective judgements.

On balance, we do not consider that this study provides sufficient evidence for changing the original dominance rules of 15% and 25% for assigning the *Urban* and *Pastoral* REC Land Cover categories, respectively. The lower threshold for *Urban* is supported by our findings that effect of urban land cover on ecology is greater than for pastoral but there is a less clear distinction for physical and chemical water quality variables. It is important to emphasise that the REC aims to discriminate variation in a range of characteristics including hydrology, water quality, ecology, and morphology and that the category membership rules are an attempt to find a suitable compromise across all these characteristics. The effect sizes of urban versus pastoral land cover on characteristics that were not considered by this study are therefore important considerations. It is well established that large hydrological effects are observed at relatively low levels of catchment urban land cover (Storey *et al.,* 2013) and the physical habitat of urban stream is also generally strongly modified (e.g., Suren *et al.*, 1998). Because the REC is intended to be a classification system that is relevant to a wide range of relevant water management variables (i.e., hydrology, hydraulics, water quality and biological communities), we consider that changing the thresholds for *Urban* and *Pastoral* REC Land Cover categories and altering the relative difference in these thresholds would require analysis of a broader range of response variable than were included in this study including hydrological and morphological/physical habitat characteristics.

The low (15%) threshold and the precedence given to the *Urban* category means that monitoring sites assigned to the *Urban* REC Land Cover category generally have catchments that have a considerable proportion of pastoral land cover (Figure 12). There is the potential with these rules that a site categorised as *Urban* comprises an area of pastoral land cover five times greater than the urban area. The problem with this is water quality and ecological indicators at sites categorised as *Urban* are generally strongly influenced by both types of land cover. There is a risk that the use of the category will lead to simplistic attribution of water quality state to a single (urban) land cover when in reality it is generally the outcome of both urban and pastoral land covers. In the following section we provide two alternative approaches to definition and use of land cover categorisation that reduces the risk of incorrectly attributing water quality to a single land cover class and removes the need to specify land cover thresholds.

8.3 Using REC Land Cover categories to describe water quality variation

REC Land Cover categories are used in environmental reporting to indicate broad scale differences in water quality associated with catchment land cover. We consider that this is a simple and relevant approach to summarising broad scale variation in water quality patterns for environmental reporting but that it raises two issues. First, land cover is only one of many factors that determine water quality. REC Land Cover categories leave considerable between-site variation unexplained. Second, because REC Land Cover categories are based on assigning a single dominant land cover to a catchment, they can obscure the contribution of other land cover types to water quality and may give the impression that observed water quality is attributable to a single land cover, such as pastoral, when in fact there are other types of land cover that are contributing to the observed conditions, such as urban land.

In future water quality analyses and reports, these issues are best addressed by using more advanced methods, such as multivariable regression modelling to estimate the effect size of different land cover types. This approach provides a more complete picture of the effect of land cover on water quality. If environmental reporting needs to establish the association between catchment land cover and water quality, these effect sizes provide more robust evidence than grouping sites by REC Land Cover categories. However, multivariable regression modelling is more complicated than the use of a single category and has some analytical complications associated with collinearity between different types of land cover. Therefore, it may not be appropriate in all environmental reporting contexts, and we propose an alternative land cover categorisation approach when simple demonstrations of water quality – land cover associations are required.

An alternative approach, which retains the simplicity of land cover categories but avoids the need to assign a dominant land cover, is to use a categorical subdivision of an anthropogenic modification gradient defined using land cover classes. The anthropogenic modification gradient could be defined by adding the proportion of catchment area occupied by *Pastoral*, *Urban* and *Exotic Forest* continuous coarsened LCDB land cover variables (i.e., *usPastoral, usUrban and usExoticForest;* Table 2) and subdividing this into categories. The anthropogenic modification categories discriminate catchments on the basis of high and low anthropogenic modification, and these broadly explained variation in the water quality data used in this study (Figure 13). We note that the use of continuous coarsened land cover variables to define the anthropogenic modification gradient is only an example. The gradient could also be defined using alternative combinations of the 34 continuous LCDB5 land cover variables.

Finally, when reporting water quality pertaining to individual sites, we recommend describing catchment land cover using the continuous coarsened LCDB5 land cover variables (Table 5). These variables describe the proportion of the catchment occupied by the same nine land cover categories as the REC. These variables provide a complete picture of catchment land cover, which is appropriate when an enquiry is concerned with an individual site.

Acknowledgements

We thank Scott Larned (NIWA) for reviews of early drafts of this report.

References

Bailey, R.G., 1996. Ecosystem Geography. Springer-Verlag, New York, NY.

Ballantine, D.J. and R.J. Davies-Colley, 2014. Water Quality Trends in New Zealand Rivers: 1989–2009. Environmental Monitoring and Assessment 186:1939–1950.

Bryce, S.A., J.M. Omernik, and D.P. Larsen, 1999. Ecoregions: A Geographic Framework to Guide Risk Characterization and Ecosystem Management. Environmental Practice 1:141–155.

Capon, S.J., A.J.J. Lynch, N. Bond, B.C. Chessman, J. Davis, N. Davidson, M. Finlayson, P.A. Gell, D. Hohnberg, and C. Humphrey, 2015. Regime Shifts, Thresholds and Multiple Stable States in Freshwater Ecosystems; a Critical Appraisal of the Evidence. Science of the Total Environment 534:122–130.

Close, M.E. and R.J. Davies-Colley, 1990. Baseflow Water Chemisty in New Zealand Rivers 2. Influence of Environmental Factors. New Zealand Journal of Marine and Freshwater Research 24:343–356.

Collier, K.J. and M. Hamer, 2010. Spatial and Temporal Patterns in the Condition of Waikato Streams Based on the Regional Ecological Monitoring of Streams (REMS) Programme. Environment Waikato, Hamilton, New Zealand.

Coulter, C.B., R.K. Kolka, and J.A. Thompson, 2004. Water Quality in Agricultural, Urban, and Mixed Land Use Watersheds 1. JAWRA Journal of the American Water Resources Association 40:1593–1601.

Cuffney, T.F., G. McMahon, R. Kashuba, J.T. May, and I.R. Waite, 2009. Responses of Benthic Macroinvertebrates to Urbanization in Nine Metropolitan Areas of the Conterminous United States. Third Interagency Conference on Research in the Watersheds., pp. 187–194.

D'Amario, S.C., D.C. Rearick, C. Fasching, S.W. Kembel, E. Porter-Goff, D.E. Spooner, C.J. Williams, H.F. Wilson, and M.A. Xenopoulos, 2019. The Prevalence of Nonlinearity and Detection of Ecological Breakpoints across a Land Use Gradient in Streams. Scientific Reports 9:1–11.

Dodds, W.K. and R.M. Oakes, 2008. Headwater Influences on Downstream Water Quality. Environmental Management 41:367–377.

Dymond, J.R., J.D. Shepherd, P.F. Newsome, and S. Belliss, 2017. Estimating Change in Areas of Indigenous Vegetation Cover in New Zealand from the New Zealand Land Cover Database (LCDB). New Zealand Journal of Ecology 41:56–64.

Jowett, I.G. and J. Richardson, 2003. Fish Communities in New Zealand Rivers and Their Relationship to Environmental Variables. New Zealand Journal of Marine and Freshwater Research 37:347–366.

Khatri, N. and S. Tyagi, 2015. Influences of Natural and Anthropogenic Factors on Surface and Groundwater Quality in Rural and Urban Areas. Frontiers in Life Science 8:23–39.

Larned, S.T., J. Moores, J. Gadd, B. Baillie, and M. Schallenberg, 2020. Evidence for the Effects of Land Use on Freshwater Ecosystems in New Zealand. New Zealand Journal of Marine and Freshwater Research 54:551–591.

Larned, S.T., M.R. Scarsbrook, T. Snelder, N.J. Norton, and B.J.F. Biggs, 2004. Water Quality in Low-Elevation Streams and Rivers of New Zealand. New Zealand Journal of Marine & Freshwater Research 38:347–366.

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

Lee, S.-W., S.-J. Hwang, S.-B. Lee, H.-S. Hwang, and H.-C. Sung, 2009. Landscape Ecological Approach to the Relationships of Land Use Patterns in Watersheds to Water Quality Characteristics. Landscape and Urban Planning 92:80–89.

McMahon, G., S.M. Gregonis, S.W. Waltman, J.M. Omernik, T.D. Thorson, J.A. Freeouf, A.H. Rorick, and K.J. E., 2001. Developing a Spatial Framework of Common Ecological Regions for the Conterminous United States. Environmental Management 28:293–316.

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

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

Norton, N. and T. Snelder, 2003. Options for Numeric Water Quality Objectives and Standards for Rivers and Lakes of Canterbury. Christchurch.

Norton, N., T. Snelder, and H. Rouse, 2010. Technical and Scientific Considerations When Setting Measurable Objectives and Limits for Water Management. Christchurch.

NZ Government, 2017. National Policy Statement for Freshwater Management 2014 (Amended 2017).

NZ Government, 2020. National Policy Statement for Freshwater Management 2020.

Omernik, J.M., 1995. Ecoregions: A Spatial Framework for Environmental Management. W. S. Davis and T. P. Simon (Editors). Biological Assessment and Criteria: Tools for Water Resource Planning and Decision Making. Lewis Publishers, Boca Raton, Florida, pp. 49–62.

Rouse, H.L. and N. Norton, 2010. Managing Scientific Uncertainty for Resource Management Planning in New Zealand. Australasian Journal of Environmental Management 17:66–76.

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

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

Snelder, T.H. and K.F.D. Hughey, 2005. On the Use of an Ecological Classification to Improve Water Resource Planning in New Zealand. Environmental Management 36:741–756.

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., M. Weatherhead, and B.J.F. Biggs, 2010. New Zealand River Environment Classification User Guide. Wellington.

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

Storey, R., G. Brierley, J. Clapcott, C. Kilroy, P. Franklin, C. Moorhouse, and R. Wells, 2013. Ecological Responses to Urban Stormwater Hydrology. Technical Report, Auckland Council, Auckland, New Zealand.

Suren, A., T. Snelder, and M. Scarsbrook, 1998. Urban Stream Habitat Assessment Method (USHA). MFE80502,NIWA Client Report. NIWA, Christchurch, p. 63 + 27 + 41Part A-Urban stream habitat assessment method (USHA) manual Part B(1)-Methodology used to develop USHA. Section 1 Measuring ecological health of urban streams: grading degraded systems. Part B(2)-Methodology used to develop USHA. Section 2 The influcence of habitat variables on urban stream communities: development of an urban stream habitat assessment score.

Tran, C.P., R.W. Bode, A.J. Smith, and G.S. Kleppel, 2010. Land-Use Proximity as a Basis for Assessing Stream Water Quality in New York State (USA). Ecological Indicators 10:727–733.

Udo de Haes, H.A. and F. Klijn, 1994. Environmental Policy and Ecosystem Classification. F. Klijn (Editor). Ecosystem Classification for Environmental Management. Kluwer Academic Publishers., Dordrecht, Netherlands, pp. 1–21.

Whitehead, A., C.E. Fraser, and T.H. Snelder, 2021. Spatial Modelling of River Water-Quality State. Incorporating Monitoring Data from 2016 to 2020. NIWA, Christchurch.

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

Zhou, T., J. Wu, and S. Peng, 2012. Assessing the Effects of Landscape Pattern on River Water Quality at Multiple Scales: A Case Study of the Dongjiang River Watershed, China. Ecological Indicators 23:166–175.

Appendix A LCDB5 land cover descriptions

Class Code	Class Name	Class Description
0	Not land	Used where the shoreline has moved between timesteps and no other appropriate class is available to represent an area outside the coastline.
1	Built-up Area (settlement)	Commercial, industrial or residential buildings, including associated infrastructure and amenities, not resolvable as other classes. Low density 'lifestyle' residential areas are included where hard surfaces, landscaping and gardens dominate other land covers.
2	Urban Parkland/Open Space	Open, mainly grassed or sparsely-treed, amenity, utility and recreation areas. The class includes parks and playing fields, public gardens, cemeteries, golf courses, berms and other vegetated areas usually within or associated with built-up areas.
5	Transport Infrastructure	Artificial surfaces associated with transport such as arterial roads, rail-yards and airport runways. Skid sites and landings associated with forest logging are sometimes also included.
6	Surface Mine or Dump	Bare surfaces arising from open-cast and other surface mining activities, quarries, gravel-pits and areas of solid waste disposal such as refuse dumps, clean-fill dumps and active reclamation sites.
10	Sand or Gravel	Bare surfaces dominated by unconsolidated materials generally finer than coarse gravel (60mm). Typically mapped along sandy seashores and the margins of lagoons and estuaries, lakes and rivers and some areas subject to surficial erosion, soil toxicity and extreme exposure.
12	Landslide	Bare surfaces arising from mass-movement erosion generally in

		mountain-lands and steep hill-country.
14	Permanent Snow and Ice	Areas where ice and snow persists through late summer. Typically occurring above 1800m but also at lower elevations as glaciers.
16	Gravel or Rock	Bare surfaces dominated by unconsolidated or consolidated materials generally coarser than coarse gravel (60mm). Typically mapped along rocky seashores and rivers, sub-alpine and alpine areas, scree slopes and erosion pavements.
15	Alpine Grass/Herbfield	Typically sparse communities above the actual or theoretical treeline dominated by herbaceous cushion, mat, turf, and rosette plants and lichens. Grasses are a minor or infrequent component, whereas stones, boulders and bare rock are usually conspicuous.
20	Lake or Pond	Essentially-permanent, open, fresh-water without emerging vegetation including artificial features such as oxidation ponds, amenity, farm and fire ponds and reservoirs as well as natural lakes, ponds and tarns.
21	River	Flowing open fresh-water generally more than 30m wide and without emerging vegetation. It includes artificial features such as canals and channels as well as natural rivers and streams.
22	Estuarine Open Water	Standing or flowing saline water without emerging vegetation including estuaries, lagoons, and occasionally lakes occurring in saline situations such as inter-dune hollows and coastal depressions.
30	Short-rotation Cropland	Land regularly cultivated for the production of cereal, root, and seed crops, hops, vegetables, strawberries and field nurseries, often including intervening

		grassland, fallow land, and other covers not delineated separately.
33	Orchards, Vineyards or Other Perennial Crops	Land managed for the production of grapes, pip, citrus and stone fruit, nuts, olives, berries, kiwifruit, and other perennial crops. Cultivation for crop renewal is infrequent and irregular but is sometimes practiced for weed control.
40	High Producing Exotic Grassland	Exotic sward grassland of good pastoral quality and vigour reflecting relatively high soil fertility and intensive grazing management. Clover species, ryegrass and cocksfoot dominate with lucerne and plantain locally important, but also including lower-producing grasses exhibiting vigour in areas of good soil moisture and fertility.
41	Low Producing Grassland	Exotic sward grassland and indigenous short tussock grassland of poor pastoral quality reflecting lower soil fertility and extensive grazing management or non-agricultural use. Browntop, sweet vernal, danthonia, fescue and Yorkshire fog dominate, with indigenous short tussocks (hard tussock, blue tussock and silver tussock) common in the eastern South Island and locally elsewhere.
43	Tall Tussock Grassland	Indigenous snow tussocks in mainly alpine mountain-lands and red tussock in the central North Island and locally in poorly-drained valley floors, terraces and basins of both islands.
44	Depleted Grassland	Areas, of mainly former short tussock grassland in the drier eastern South Island high country, degraded by over-grazing, fire, rabbits and weed invasion among which Hieracium species are conspicuous. Short tussocks usually occur, as do exotic grasses, but bare ground is more prominent.

45	Herbaceous Freshwater Vegetation	Herbaceous wetland communities occurring in freshwater habitats where the water table is above or just below the substrate surface for most of the year. The class includes rush, sedge, restiad, and sphagnum communities and other wetland species, but not flax nor willows which are mapped as Flaxland and Deciduous Hardwoods respectively.
46	Herbaceous Saline Vegetation	Herbaceous wetland communities occurring in saline habitats subject to tidal inundation or saltwater intrusion. Commonly includes club rush, wire rush and glasswort, but not mangrove which is mapped separately.
47	Flaxland	Areas dominated by New Zealand flax usually swamp flax (harakeke) in damp sites but occasionally mountain flax (wharariki) on cliffs and mountain slopes.
50	Fernland	Bracken fern, umbrella fern, or ring fern, commonly on sites with low fertility and a history of burning. Manuka, gorse, and/or other shrubs are often a component of these communities and will succeed Fernland if left undisturbed.
51	Gorse and/or Broom	Scrub communities dominated by gorse or Scotch broom generally occurring on sites of low fertility, often with a history of fire, and insufficient grazing pressure to control spread. Left undisturbed, this class can be transitional to Broadleaved Indigenous Hardwoods.
52	Manuka and/or Kanuka	Scrub dominated by mānuka and/or kānuka, typically as a successional community in a reversion toward forest. Mānuka has a wider ecological tolerance and distribution than kānuka with the latter somewhat concentrated in the north with particular

		prominence on the volcanic soils of the central volcanic plateau.
54	Broadleaved Indigenous Hardwoods	Lowland scrub communities dominated by indigenous mixed broadleaved shrubs such as wineberry, mahoe, five-finger, Pittosporum spp, fuchsia, tutu, titoki and tree ferns. This class is usually indicative of advanced succession toward indigenous forest.
55	Sub Alpine Shrubland	Highland scrub dominated by indigenous low-growing shrubs including species of Hebe, Dracophyllum, Olearia, and Cassinia. Predominantly occurring above the actual or theoretical treeline, this class is also recorded where temperature inversions have created cooler micro-climates at lower elevations e.g. the ‘frost flats’ of the central North Island.
56	Mixed Exotic Shrubland	Communities of introduced shrubs and climbers such as boxthorn, hawthorn, elderberry, blackberry, sweet brier, buddleja, and old man’s beard.
58	Matagouri or Grey Scrub	Scrub and shrubland comprising small-leaved, often divaricating shrubs such as matagouri, Coprosma spp, Muehlenbeckia spp., Casinnia spp., and Parsonsia spp. These, from a distance, often have a grey appearance.
80	Peat Shrubland (Chatham Is)	Low-growing shrubland communities usually dominated by Dracophyllum spp. in association with Cyathodes spp. and ground ferns. Mapped only on the Chatham Islands.
81	Dune Shrubland (Chatham Is)	Low-growing shrubland communities dominated by Leucopogon spp., Pimelia arenaria and Coprosma spp., in association with sedges and scattered herbs and grasses.

		Mapped only on the Chatham Islands.
70	Mangrove	Shrubs or small trees of the New Zealand mangrove (Avicennia marina subspecies australascia) growing in harbours, estuaries, tidal creeks and rivers north of Kawhia on the west coast and Ohiwa on the east coast.
64	Forest - Harvested	Predominantly bare ground arising from the harvesting of exotic forest or, less commonly, the clearing of indigenous forest. Replanting of exotic forest (or conversion to a new land use) is not evident and nor is the future use of land cleared of indigenous forest.
68	Deciduous Hardwoods	Exotic deciduous woodlands, predominantly of willows or poplars but also of oak, elm, ash or other species. Commonly alongside inland water (or as part of wetlands), or as erosion-control, shelter and amenity plantings.
69	Indigenous Forest	Tall forest dominated by indigenous conifer, broadleaved or beech species.
71	Exotic Forest	Planted or naturalised forest predominantly of radiata pine but including other pine species, Douglas fir, cypress, larch, acacia and eucalypts. Production forestry is the main land use in this class with minor areas devoted to mass-movement erosion-control and other areas of naturalised (wildling) establishment.

Appendix B REC land cover categories including notations and characteristics

Land-Cover category	Notation	Characteristics of river environment
Bare ground	B	The Bare Ground category tends to occur over large areas only in mountainous catchments. The hydrological and water chemistry characteristics of this class tend to accentuate the characteristics of the Mountain Source-of-Flow category. Runoff response is rapid, low nutrient concentration and suspended sediment tends to be high
Indigenous forest, Scrub, Tussock (Natural land cover)	IF S T	Flood peaks are attenuated by vegetation, and low flows are generally more sustained than Pastoral or Bare Ground Land-Cover categories. Nutrient concentrations tend to be low. Suspended sediment concentrations tend to be low resulting in high water clarity.
Pastoral	P	Flood peaks tend to be higher and recede faster. Low flows are generally more extreme relative to catchments with natural land cover. Nutrient concentrations are high relative to natural Land-Cover categories. Erosion rates tend to be high, resulting in low water clarity and fine substrates (silts and mud) compared to natural land cover.
Exotic forestry	EF	Flow regime dependent on the age of the forest. Mature forests display a regime relatively similar to that found in native forest; recently logged forests display a regime similar to pastoral sites. Variable nutrient and suspended sediment concentrations depending on the cutting cycle of the forest. Nutrients for mature forests are typically lower than for rivers with a Pastoral Land-Cover category.
Urban	U	Flood peaks are very 'peaky' and recessions return quickly to base flow. Base flows are very low. High concentration of many contaminants. High suspended sediment load during development and typically low afterward. Fine substrates (silts and mud) high relative to natural Land-Cover categories.

.