Spatial modelling of river water-quality state

Incorporating monitoring data from 2013 to 2017

Prepared for Ministry for the Environment

December 2018

Prepared by:
Amy Whitehead

For any information regarding this report please contact:

Amy Whitehead
Quantitative Freshwater Ecologist

+64-3-343 7854
amy.whitehead@niwa.co.nz

National Institute of Water & Atmospheric Research Ltd
PO Box 8602
Riccarton
Christchurch 8011

Phone +64 3 348 8987

NIWA CLIENT REPORT No: 2018360CH
Report date: December 2018
NIWA Project: MFE18502

Quality Assurance Statement		
	Reviewed by:	Rick Stoffels
	Formatting checked by:	Fenella Falconer
	Approved for release by:	Scott Larned

Contents

Tables

Figures

Executive summary

This report provides model-based predictions of water quality state for each of ~ 590,000 unique river segments that comprise New Zealand's national river network, using data for the period 2013-2017. Two comparable reports were produced in 2010 and 2016 using data for the periods 2003-2007 and 2009-2013, respectively. This report is the second in a series of reports prepared for the Ministry for the Environment on national-scale state and trends in river freshwater quality. The first report provided site-specific river water quality state and trends for several hundred river monitoring sites operated by Regional Councils and NIWA. The river water quality data acquired and processed for the first report were used in the current report.

The predicted water quality values in the current report were generated using Random Forest (RF) models. The RF empirical modelling method predicts the values of response variables using a suite of predictor variables and a dataset of observations (the 'training data'). RF models are an advanced form of regression-tree models. Single regression trees do not identify the optimum tree structure (i.e., the most accurate predictions) and they are sensitive to small changes in the observational data. To overcome these problems, RF modelling employs an ensemble of trees (a forest) and makes predictions based on the average of all trees. RF models have several additional properties that make them suitable for use in situations where the observational data are heterogeneous and the predictor variables are inter-correlated (as is often the case in water quality analyses); they require no assumptions about data distributions to be met, they are minimally affected by multi-collinearity among predictor variables, and they cannot be over-fitted.

An RF model was developed for each of nine water quality variables: visual clarity (CLAR), turbidity (TURB), ammoniacal nitrogen (NH4N), nitrate-nitrogen (NO3N), total nitrogen (TN), dissolved reactive phosphorus (DRP), total phosphorus (TP), *Escherichia coli* (ECOLI), and the macroinvertebrate community index (MCI). The predictor variables consisted of 32 variables for which georeferenced data are stored the River Environment Classification geodatabase. These predictor variables were selected to represent climatic, geological, topographic, land cover, and hydrological conditions in New Zealand rivers and their catchments.

The observational data used in the RF models consisted of site median values from monthly and quarterly measurements (and annual invertebrates for MCI scores) for the period 2013-2017. These data came from 587 - 882 monitoring sites (depending on the variable). The sites are reasonably well-distributed across the North and South Islands, with some gaps in inaccessible areas. To assess the degree to which the monitoring sites used for observational data represent the range of environmental conditions present in New Zealand, we compared histograms of the distributions of predictor variable values for the monitoring sites, with the distributions of the same variables for all river segments in New Zealand. The monitoring sites were reasonably representative, with moderate over-representation of low-elevation, low-gradient catchments with large proportions of intensive agricultural land cover.

The RF models performed well in predicting median water quality state, based on the amount of variation in the observational data explained, the congruence between observed and predicted values, low model bias (tendency to over- or under estimate), and low prediction uncertainty.

The most important predictor variables for the nutrient models and the CLAR, TURB and ECOLI models were the proportions of intensive agricultural and urban land cover in catchments, catchment slope and particle size. For the MCI model, the most important predictor variables were

the proportions of native forest and urban land cover and catchment elevation. Collectively, the models suggest that chemical and microbiological water quality is most severely compromised in low-elevation, low-gradient land under intensive land use.

National maps of predicted median TURB and nutrient and ECOLI concentrations have relatively high values in low-elevation areas on the east coasts of the North and South Island, and in the inland Waikato, Wairarapa Valley, Rangitikei-Manawatu coastal plain, Taranaki Ring Plain, and Auckland Region. Predicted TURB and nutrient and ECOLI concentrations are generally low in major mountain ranges, in large areas of the Department of Conservation estate and in other native forest-dominated areas. Predicted DRP and TP concentrations appear to be elevated in rivers draining phosphorus-rich Tertiary mudstone and volcanic ash on the North Island, suggesting that parent geology affects large scale patterns in river DRP and TP. However, the effects of geology are likely to be inter-correlated with land use and topography. Geographic patterns in predicted CLAR and MCI scores are generally the reverse of the patterns for chemical and microbial water quality, with high values in mountain ranges and Department of Conservation estate, and low values in areas dominated by intensive agriculture and urban land cover.

1 Introduction

River water quality can be characterised by the recent state and trends in variables that are measured at monitoring sites distributed across New Zealand. The sites are monitored as part of the State of Environment (SOE) programmes operated by Regional Councils and unitary authorities and the National River Water Quality Network (NRWQN) operated by NIWA.

This document is the second of two reports updating national river water quality state in New Zealand, using data from the council SOE programmes and the NRWQN. The first report "Water quality state and trends in New Zealand rivers" provided information on water quality state and trends for individual river monitoring sites (Larned et al. 2018). The updated river dataset used in the state analyses of the first report contained measurements for nine physical, chemical, microbiological and macroinvertebrate variables from 587 to 882 sites for the 2013-2017 period. Site-specific water quality state in the first report was based on median variable values for that five-year period.

In the current report, we developed Random Forest (RF) models to relate spatial variation in the same nine water quality variables used in Larned et al. (2018) to a large suite of environmental predictor variables. The predictor variables represent climatic, geological, topographic, land cover, and hydrological conditions in New Zealand rivers and their catchments. The RF models were then used to predict river water quality at un-monitored locations and to produce spatially continuous maps of predicted water quality variation. These maps comprise a high-level statistic that describes predicted patterns in water quality at the national scale.

One of the primary purposes of predictive RF models is to provide large-scale water-quality assessments that are more representative of the true patterns of water quality than assessments based on aggregated data from multiple monitoring sites. The latter approach can lead to conclusions about water quality patterns that are biased by the non-random locations of monitoring sites. Previous analyses have shown that the aggregate network of river water quality monitoring sites in New Zealand is over-represented by sites in catchments dominated by pastoral land cover and under-represented by sites in catchments dominated by native forest (Larned and Unwin 2012). This non-representative distribution of sites can produce biased results when multiple environmental classes are merged. Using RF models to predict water quality state in all river reaches across New Zealand can reduce this problem.

In this report, we provide detailed methods for using RF models to predict river water quality state across the heterogeneous New Zealand environment. The methodological steps include preparation of the water quality state data, selection of environmental predictor variables, assessment of site representativeness, the RF modelling process, and assessments of model performance. In the results section, we present maps of national predictions of river water quality, identify important predictor variables and quantify model performance. In the discussion section, we compare the current RF models with previous models of river water quality state and comment briefly on uncertainty in RF models and alternative modelling methods. We also provided MfE with the model outputs for every river reach in New Zealand as an .RData file.

2 Data

2.1 River water quality data

The monitoring sites and data used in the Stage 1 study to analyse water quality state (Larned et al. 2018) were also used for the current study. The water quality data consisted of measurements of nine physical, chemical, microbiological and invertebrate variables from river monitoring sites in council SOE networks and the NRWQN sites (Table 2-1). Detailed methods for processing the water quality data are given in Larned et al. (2018). The monitoring sites had the following properties: 1) less than 50% of the values for a variable were censored; 2) values for at least 90% of monthly or quarterly sampling dates were available, including imputed values; 3) at least 30 values were distributed over four of the five years from 2013 to 2017. In contrast, the invertebrate data come from annual monitoring and there were no censored data. The sole rule for including invertebrate monitoring sites was that data were available for at least four of the five years from 2013 to 2017.

The Stage 1 study used the original version of the River Environment Classification (REC1; Snelder and Biggs 2002) as a spatial framework to provide environmental context for the analysis. In the current study, we used a recently updated version of the REC, referred to as REC2 (see Section 2.2 for details). All monitoring sites from the Stage 1 study were projected on to the REC2 digital river network, then manually checked. In the final dataset used for RF modelling, between 587 and 882 sites met the inclusion criteria for at least one of the eight water quality variables (Table 2-1).

The geographic distribution of river monitoring sites used for modelling is shown in Figure 2-1. The sites are reasonably well-distributed, although there are gaps in the central North Island and west coast of the South Island. There is a high degree of overlap among the sites used for physical, chemical and microbiological water quality monitoring, as some or all of the corresponding variables are measured at each site in council SOE programmes. There is less overlap among sites used for invertebrate monitoring; several councils operate separate programmes for monitoring physical-chemical water quality and invertebrates, with variable levels of site overlap between programmes.

Table 2-1: River water quality variables, measurement units and site numbers used to develop Random Forest models.

Variable type	Variable	Abbreviation	Units	Number of monitoring sites
Physical	Visual clarity	CLAR	m	587
	Turbidity	TURB	NTU	878
Chemical	Ammoniacal nitrogen	NH4N	mg/m^3	882
	Nitrate-nitrogen	NO3N	mg/m^3	855
	Total nitrogen (unfiltered)	TN	mg/m^3	764
	Dissolved reactive phosphorus	DRP	mg/m^3	877
	Total phosphorus (unfiltered)	TP	mg/m^3	740
Microbiological	*Escherichia coli*	ECOLI	cfu/100 mL	866
Biotic Index	Macroinvertebrate Community Index	MCI	unitless	832

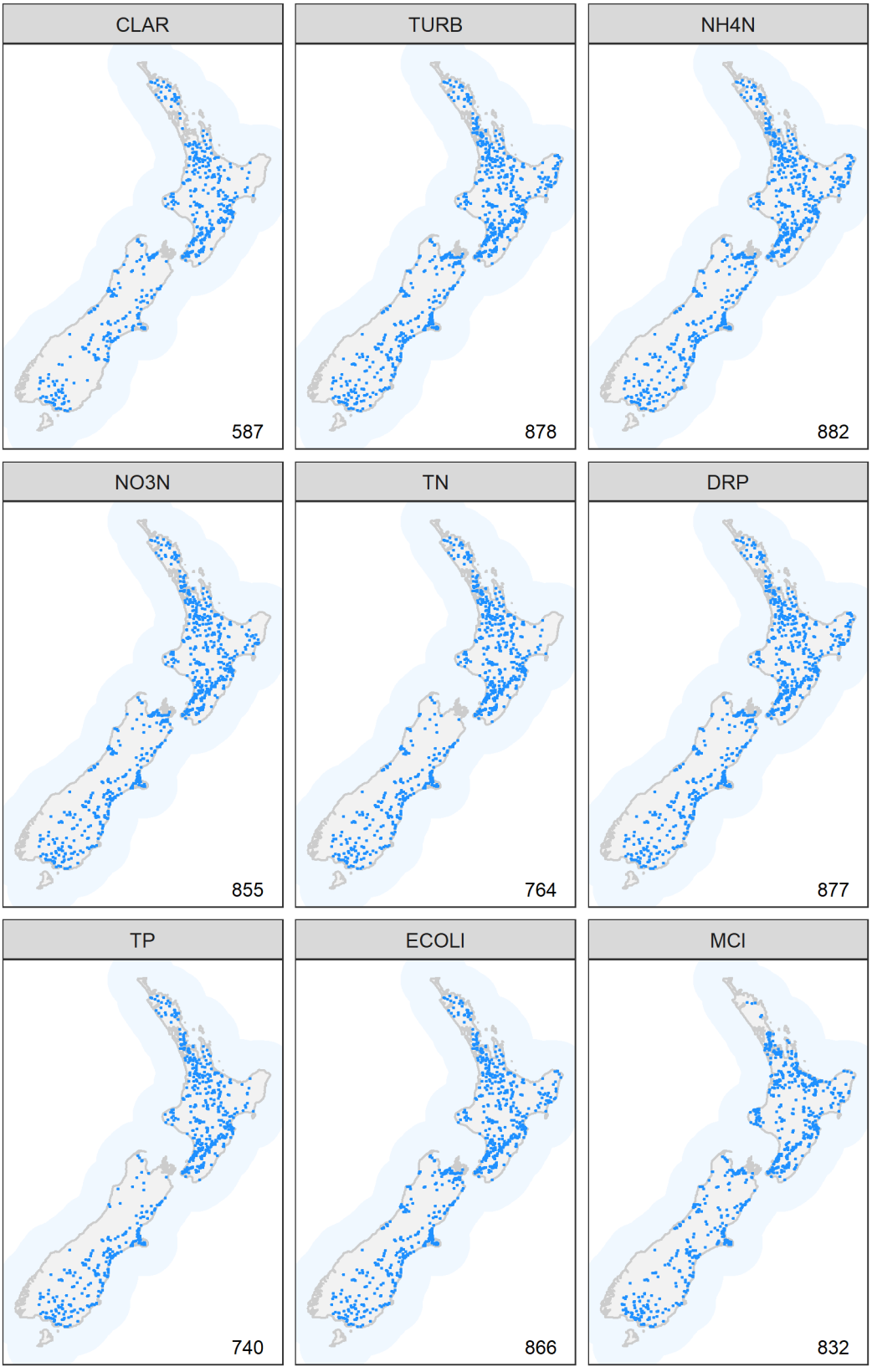

Figure 2-1: Locations of river water quality monitoring sites used for modelling the state of nine water quality variables. The number in the lower right of each panel corresponds to the number of sites included for each variable (Table 2-1).

2.2 Predictor data

The digital river network and catchment boundaries used for the REC provided the spatial framework for the RF models of river water quality state. The river network and catchment boundaries were derived from a digital elevation model (DEM) with a spatial resolution of 50 m. The digital network represents New Zealand's rivers as ~ 560,000 segments (bounded by upstream and downstream confluences) and their corresponding catchments. Each segment in the digital network has a unique identifier, the nzsegment number. The links between each nzsegment and its catchment, between adjacent nzsegments and between adjacent catchments facilitate analyses of upstream-downstream connectivity and the accumulation of catchment characteristics in the downstream direction. The digital network has been recently updated to correct errors and to improve its representation of rivers nationally; the REC geodatabase with the updated network is referred to as REC2 (version 2.4).

In addition to the digital network, REC2 contains spatial data layers describing the climate, topography, geology, vegetation, infrastructure and hydrology of New Zealand (https://www.niwa.co.nz/freshwater-and-estuaries/management-tools/river-environment-classification-0). These spatial data are used to link each nzsegment to many attributes that describe the environmental characteristics of the segment and its catchment. Catchment land cover in REC2 is derived from the national Land Cover Database-4 (LCDB4) which differentiates 32 categories based on analysis of satellite imagery from 2012 (lris.scinfo.org.nz). Descriptions of catchment regolith are derived from the Land Resources Inventory (LRI) including interpretations of the LRI categories made by Leathwick et al. (2003). Additional variables for each segment have been derived from national-scale hydrological modelling (e.g., Booker and Snelder 2012).

We selected 32 network attributes from REC2 (Table 2-2) for predictor variables in spatial models of the nine water quality variables listed in Table 2-1. The predictor variables were selected based on their predicted mechanistic or correlative relationships with water quality, and on previous experience with national scale modelling of water (e.g., Unwin et al. 2010; Larned et al. 2016) and invertebrate communities (Clapcott et al. 2013; Leathwick et al. 2011).

Table 2-2: Predictor variables used in random forest models of river water quality variables. *Geological variables are based on regolith, using averages of ordinal values assigned to LRI top-rock categories by Leathwick et al. (2003). The variables usHard and usPsize characterise physical regolith conditions; usPhos and usCalc characterise regolith fertility.

Predictor variable class	Predictor variable description	Abbreviation	Unit
Geography & topography	Catchment area	usArea	m^2
	Segment mean elevation	segElev	m ASL
	Percentage of catchment occupied by lakes	usLakePerc	%
	Mean catchment elevation	usElev	m ASL
	Mean catchment slope	usSlope	degrees
	Distance to the coast	DistToCoast	m
	Mean segment slope	SegSlope	degrees
	Segment sinuosity (segment length divided by the straight line distance between endpoints)	Sinuosity	unitless
	Distance to furthest headwater segment	DistToHead	m
Climate & flow	Mean segment June air temperature	segTmin	degrees C x 10
	Mean segment January air temperature.	segTwarm	degrees C x 10
	Mean catchment June air temperature	usTmin	degrees C x 10
	Mean catchment January air temperature	usTwarm	degrees C x 10
	Mean annual catchment rainfall	usRain	mm
	Mean catchment coefficient of variation of annual rainfall	usRainvar	mm/yr
	Mean catchment rain days > 10 mm	usRainDays10	days/mo
	Mean catchment rain days > 200 mm	usRainDays20	days/mo
	Mean catchment rain days > 100 mm	usRainDays100	days/mo
	Mean annual catchment potential evapotranspiration	usPET	mm/yr
	Estimated mean flow	MeanFlow	m^3/s
Geology*	Mean catchment induration (hardness) of regolith	usHard	Ordinal
	Mean catchment phosphorous content of regolith	usPhos	Ordinal
	Mean catchment particle size of regolith	usPsize	Ordinal
	Mean catchment calcium content of regolith	usCalc	Ordinal
Land cover	Proportion of catchment occupied by combination of high producing exotic grassland, short-rotation cropland, orchard, vineyard and other perennial crops (LCDB4 classes 40, 30, 33)	usIntensiveAg	%
	Proportion of catchment in low producing grassland (LCDB4 class 41)	usPastoralLight	%
	Proportion of catchment in native forest (LCDB4 class 69)	usNativeForest	%
	Proportion of catchment in built-up areas, urban parkland, surface mines, dumps and transport infrastructure (LCDB4 classes 1,2,6,5)	usUrban	%
	Proportion of catchment in scrub and shrub cover (LCDB4 classes 50, 51, 52, 54, 55, 56, 58)	usScrub	%
	Proportion of catchment occupied by lake and pond, river and estuarine open water (LCDB4 classes 20, 21, 22)	usWetland	%
	Proportion of catchment in exotic forest (LCDB3 class 71)	usExoticForest	%
	Proportion of catchment occupied in bare or lightly-vegetated cover (LCDB4 classes 10, 12, 14, 15, 16)	usBare	%

3 Modelling methods

3.1 Random forest models

We modelled median values of each water quality variable as a function of the predictor variables using RF models (Breiman et al. 1984; Breiman 2001; Cutler et al. 2007), with all variables except MCI log-transformed (i.e., the $\log_{10}$ of the median of the untransformed raw data). An RF model is an ensemble of individual classification and regression trees (CART). In a regression context, CART partitions observations (in this case the individual water quality variables) into groups that minimise the sum of squares of the response (i.e., assembles groups that minimise differences between observations) based on a series of binary rules or splits that are constructed from the predictor variables. CART models have several desirable features including requiring no distributional assumptions and the ability to automatically fit non-linear relationships and high order interactions. However, single regression trees have the limitations of not searching for optimal tree structures, and of being sensitive to small changes in input data (Hastie et al. 2001). RF models reduce these limitations by using an ensemble of trees (a forest) and making predictions based on the average of all trees (Breiman 2001). An important feature of RF models is that each tree is grown with a bootstrap sample of the fitting data (i.e., the observation dataset). In addition, a random subset of the predictor variables is made available at each node to define the split. Introducing these random components and then averaging over the forest increases prediction accuracy while retaining the desirable features of CART.

An RF model produces a limiting value of the generalization error (i.e., the model maximises its prediction accuracy for previously unseen data; Breiman 2001). The generalization error converges asymptotically as the number of trees increases, so the model cannot be over-fitted. The number of trees needs to be set high enough to ensure an appropriate level of convergence, and this value depends on the number of variables that can be used at each split. We used default options that included making one third of the total number of predictor variables available for each split, and 500 trees per forest. Some studies report that model performance is improved by including more than ~ 50 trees per forest, but that there is little improvement associated with increasing the number of trees beyond 500 (Cutler et al. 2007). Our models took less than a minute to fit when using the default of 500 trees per forest.

Unlike linear models, RF models cannot be expressed as equations. However, the relationships between predictor and response variables represented by RF models can be represented by importance measures and partial dependence plots (Breiman 2001; Cutler et al. 2007). During the fitting process, RF model predictions are made for each tree for observations that were excluded from the bootstrap sample; these excluded observations are known as out-of-bag (OOB) observations. To assess the importance of a specific predictor variable, the values of the response variable are randomly permuted for the OOB observations, and predictions are obtained from the tree for these modified data. The importance of the predictor variable is indicated by the degree to which prediction accuracy decreases when the response variable is randomly permuted. Importance is defined in this study as the loss in model performance (i.e., the increase in the mean square error; MSE) when predictions are made based on the permuted OOB observations compared to those based on the original observations. The differences in MSE between trees fitted with the original and permuted observations are averaged over all trees and normalized by the standard deviation of the differences (Cutler et al. 2007).

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

RF models can include any of the original set of predictor variables that are chosen during the model fitting process. Inclusion of marginally important and correlated predictor variables does not degrade the performance of the RF models. However, these predictor variables may be redundant (i.e., their removal does not affect model performance) and their inclusion can complicate model interpretation. We used a backward elimination procedure to remove redundant predictor variables from the initial 'saturated' models (i.e., models that included any of the original predictor variables). The procedure first assesses the model error (MSE) using a 10-fold cross validation process. The predictions made to the hold out observations during cross validation are used to estimate the MSE and its standard error. The model's least important predictor variables are then removed in order, with the MSE and its standard error being assessed for each for each successive model. The final, 'reduced' model is defined as the model with the fewest predictor variables whose error is within one standard error of the best model (i.e., the model with the lowest cross validated MSE). This is equivalent to the "one standard error rule" used for cross validation of classification trees (Breiman et al. 1984).

An alternative approach is to choose the model with the smallest error. We used the former procedure as it retains fewer predictor variables than the latter procedure, while achieving an error rate that is not different, within sampling error, from the "best solution". Importance levels for predictor variables were not recalculated at each reduction step to avoid over-fitting (Svetnik et al. 2004).

We note that, because fitting a RF model involves randomly selecting observations and predictor variables throughout the fitting process, successive models fitted to the same data set will exhibit subtle differences in structure and diagnostics such as total explained deviance, MSE, partial dependence plots, and the order of predictor importance. In the current study, the variability in model error between individual fits of the model for each water quality variable were within the reported model performance (see Section 3.2).

All calculations were performed in the R statistical computing environment (R Core Team 2017) using the *randomForest* package (Liaw and Werner 2002) and other specialised packages.

3.2 Model performance

Model performance was assessed by comparing observations with independent predictions (i.e., sites that were not used in fitting the model), which were obtained from the out-of-bag (OOB) samples. We summarised the models using four statistics; regression R^2**,** Nash-Sutcliffe Efficiencies (NSE), bias and root mean square deviation (RMSD).

The regression R^2 value is the coefficient of determination derived from a regression of the observations against the predictions. The R^2 value shows the proportion of the total variance

explained by the regression model (Piñeiro et al. 2008). However, the regression R^2 is not a complete description of model performance.

The NSE (Nash and Sutcliffe 1970) provides a measure of overall model performance by indicating how closely a plot of observed versus predicted values lies to the 1:1 line (i.e., the degree to which two sets of values coincide). NSE values range from $-\infty$ to 1. An NSE of 1 corresponds to a perfect match between predictions and the observed data, an NSE of 0 indicates that the model predictions are as accurate as the mean of the observed data; and an NSE less than 0 indicates that the observed mean is a better predictor than the model.

Model bias measures the average tendency of the predicted values of water quality variables to be larger or smaller than the observed values. Positive values indicate underestimation bias and negative values indicate overestimation bias (Moriasi et al. 2007).

The root mean square deviation (RMSD) is a measure of the characteristic model statistical error or uncertainty. RMSD is mean deviation of predicted values with respect to the observed values (distinct from the standard error of the regression model). RMSD can be used to evaluate the confidence intervals of the predictions.

The relative root mean square error (RSR) is a measure of the characteristic model uncertainty. It is estimated as the mean deviation of predicted values with respect to the observed values (the root mean square deviation), divided by the standard deviation of the observations (Moriasi et al., 2007).

3.3 Representativeness of monitoring sites used in RF models

A graphic comparison was used to gauge how well the monitoring sites used to fit the RF models represented environmental variation at the national scale. Here, representativeness refers to the degree to which the distribution of monitoring sites over the range of an environmental predictor variable matches the distribution of all network segments over the range of the same environmental variable. Poor representativeness can reduce accuracy in model predictions because certain combinations of environmental conditions are not represented in the fitting data.

Histograms of the proportions of monitoring site numbers over the ranges of the 12 most important predictor variables in the RF models (i.e., the predictors with the greatest explanatory power) were visually compared with histograms of the proportions of all network segments over the same predictor variables. Two sets of comparable histograms were derived. The first represented data from all monitoring sites that included at least one water quality variable, excluding MCI (887 sites). The second set of comparable histograms represented the 832 invertebrate monitoring sites that were used for modelling MCI scores. Separate histograms were constructed due to the limited overlap in physical-chemical water quality and invertebrate monitoring sites, as noted in Section 2.1.

Note that representativeness of monitoring sites is different from model bias, which is defined in Section 3.2. Model bias is a measure of systematic error in model predictions (i.e., over- or under-estimation).

3.4 Model predictions

Predictions are made with RF models by "running" new cases down every tree in the fitted forest and averaging the predictions made by each tree (Cutler et al. 2007). The models in this study were fitted to $\log_{10}$-transformed water quality data (except for MCI which used non-transformed data). When these models are back-transformed, the model error term no longer has a mean of zero. Ignoring this

results in retransformation bias (i.e., predictions that systematically underestimate the response). We corrected the retransformation bias using the smearing estimate (*S*) developed by Duan (1983):

$$S = \frac{1}{n}\sum_{i=1}^{n} 10^{\hat{\varepsilon}_i} \quad \text{(Equation 1)},$$

where $\hat{\varepsilon}$ are the residuals of a RF model. The predictions were back-transformed by raising them to the power of 10, then corrected for retransformation bias by multiplying by *S.* The back-transformed and corrected predictions for all river segments in New Zealand were projected on a single national map for each water quality variable.

4 Results

4.1 Model performance

The RF models for most water quality variables performed well, as indicated by the following statistics: $R^2 > 0.5$, NSE > 0.5, and RMSD < 0.5 for most variables (Table 4-1), except NH4N ($R^2 = 0.39$, NSE = 0.39, RMSD = 0.39) and MCI (RMSD = 9.90). Note that the RMSD value for MCI is higher than the other variables because the MCI data were not $\log_{10}$ transformed. Bias in the RF models was low as indicated by the close match between the line representing the regression of the observed versus predicted values (red dashed line in Figure 4-1) and the one-to-one line (blue solid line in Figure 4-1). The close match between the regression and one to one line also indicates that the models are consistent (i.e., that low or high values are not under or over-estimated). Based on NSE values, the TN, MCI, ECOLI and TP models had the best overall performance, the NH4N model had the worst overall performance, and the NO3N, CLAR, TURB and DRP models had intermediate performance.

Table 4-1: 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 R^2 = coefficient of determination, NSE = Nash-Sutcliffe efficiency, RSR = relative root mean square error, 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 and bias are based on non-transformed data.

Water quality variable	Number of sites	Regression R^2	NSE	Bias	RSR	RMSD
CLAR	587	0.59	0.58	0.004	0.65	0.20
TURB	878	0.54	0.53	-0.004	0.68	0.30
NH4N	882	0.39	0.39	-0.007	0.78	0.39
NO3N	855	0.59	0.58	0.002	0.65	0.48
TN	764	0.71	0.70	-0.005	0.54	0.27
DRP	877	0.56	0.56	-0.007	0.66	0.32
TP	740	0.65	0.64	-0.005	0.60	0.27
ECOLI	866	0.66	0.65	-0.007	0.59	0.37
MCI	832	0.68	0.67	0.036	0.57	9.90

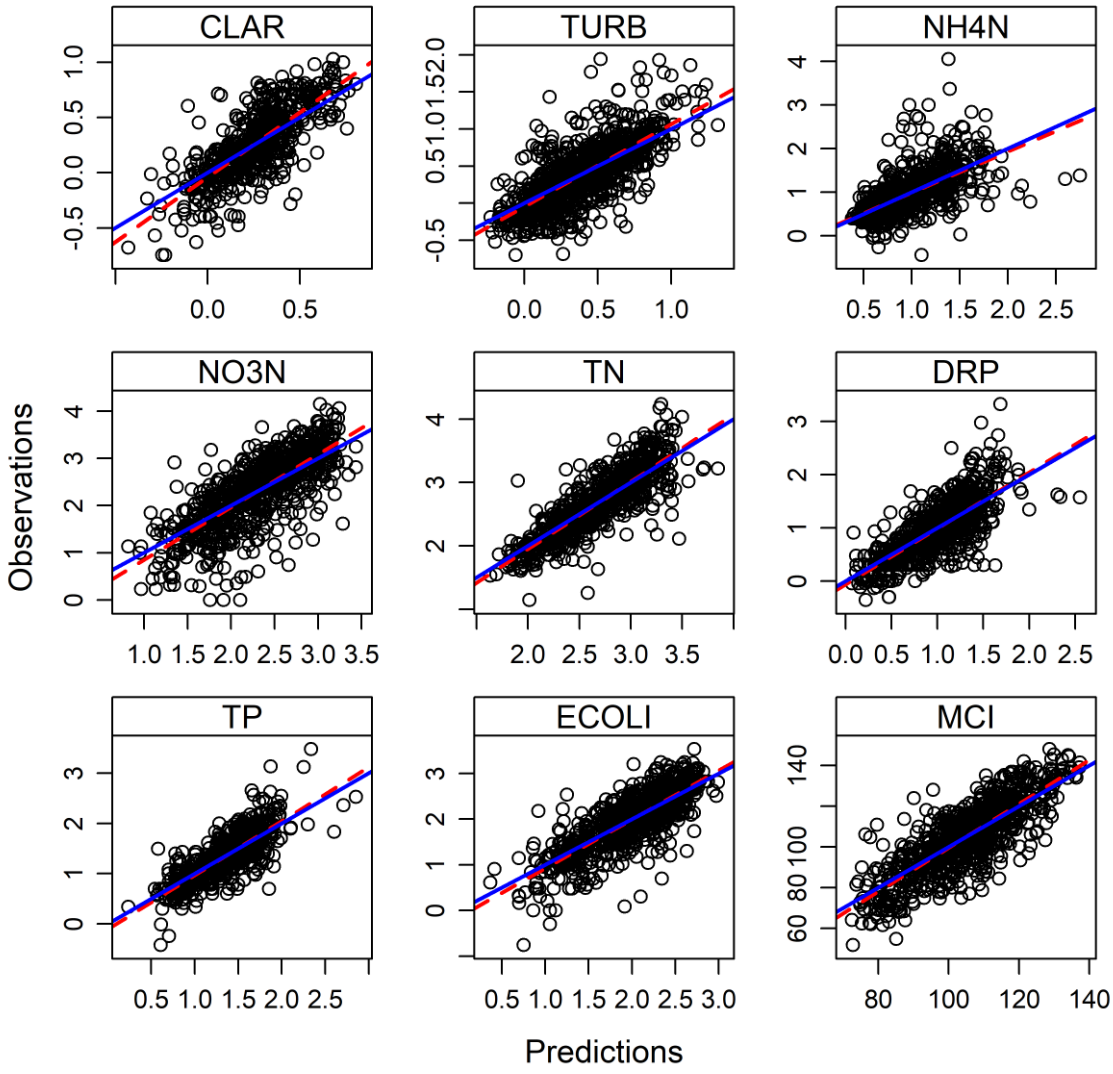

Figure 4-1: Comparison of observed water quality versus values predicted by the random forest models. Note that the observed values are plotted on the Y-axis and predicted values on the X-axis, following Piñeiro et al. (2008). Red dashed line: best fit linear regression of the observed and predicted values. Clue solid line: one-to-one line. Units are the $\log_{10}$ transformed units of all water quality variables except for MCI, which uses non-transformed values.

4.2 Monitoring site representativeness

The distributions of river water quality and MCI monitoring sites across the environmental gradients defined by 12 predictor variables were generally consistent with the distribution of all segments in the river network across the same gradients (Figure 4-2 and Figure 4-3). The predictor variables shown in the histograms were those subsequently found to be most important in the RF models.

There were several cases of moderate over- and under-representation of monitoring sites compared to the river network. Water-quality sites were over-represented in environments characterised by low catchment elevations (usElev) and low catchment slopes (usSlope) (Figure 4-2). Water quality sites were under-represented in catchments with high proportions of native forest land cover (usNativeForest), and catchments with low proportions of intensive agricultural land cover (usIntensiveAg). Invertebrate monitoring sites were over-represented in low gradient rivers (segSlope) close to the coast (distToCoast) and under-represented in catchments with high proportion of native forest land cover (usNativeForest; Figure 4-3).

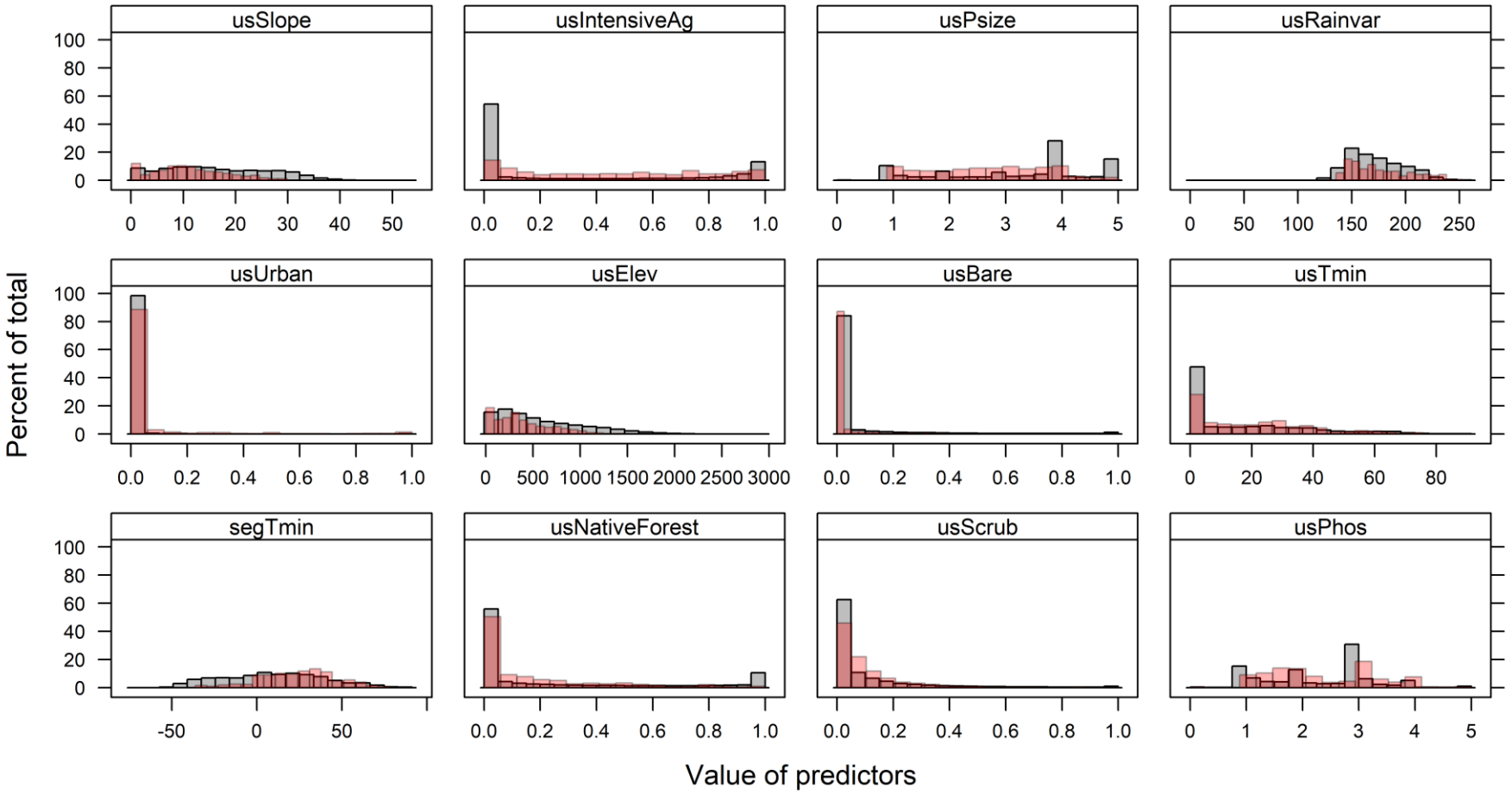

Figure 4-2: The distributions of predictor variables across all segments in the digital river network and at water quality sites (grey and red histograms, respectively). Similarities in the distributions shown in the two histograms in each panel provide an indication of the degree to which environmental variation across the monitoring sites represents environmental variation across the New Zealand river network; complete representativeness would be indicated by exact matches between the histograms. These twelve predictor variables were the most important overall predictors in the water quality RF models (with the exception of MCI – see Figure 4-3) and are ordered from most (top left) to least (bottom right) important (Table 4-2).

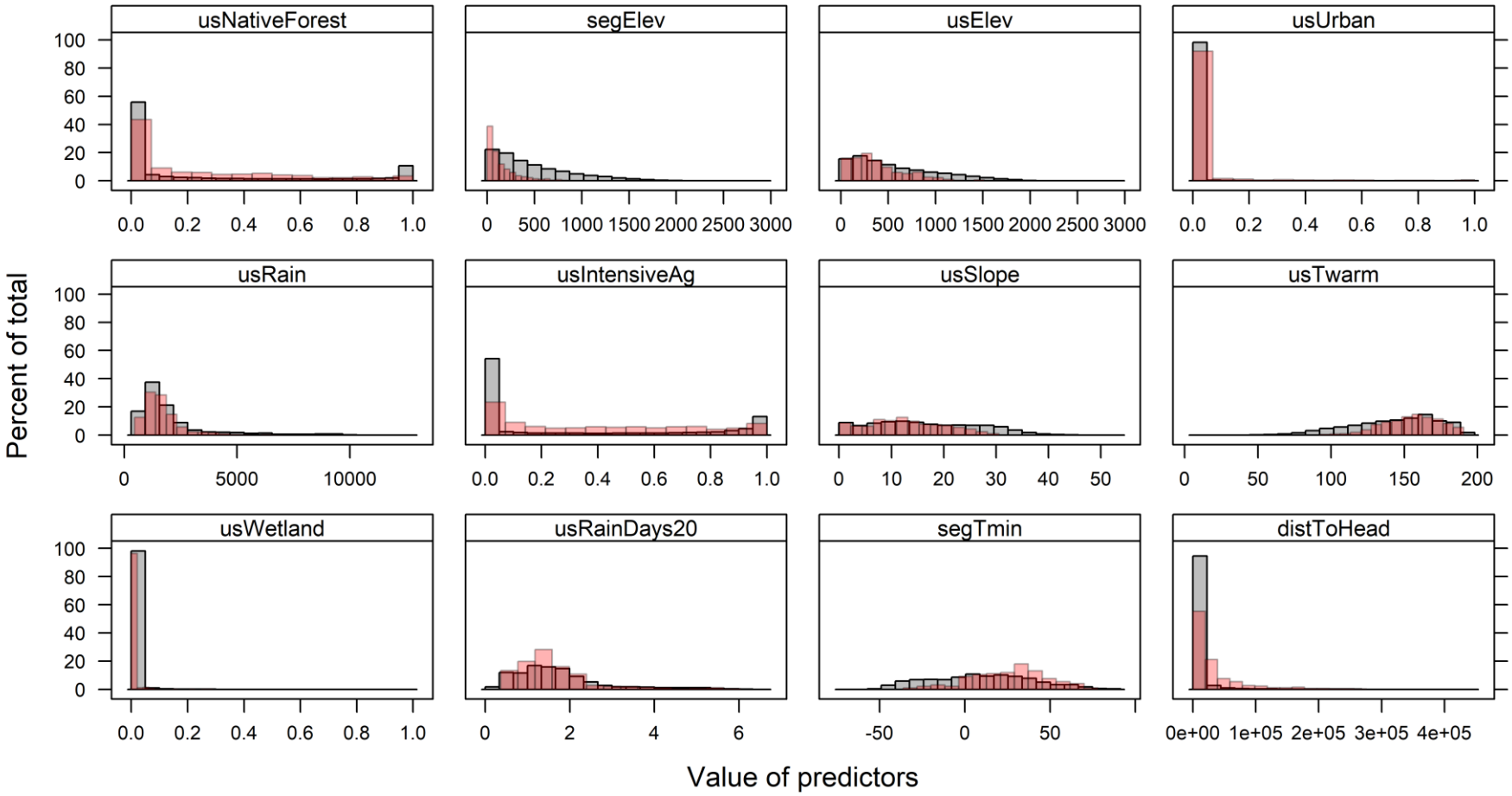

Figure 4-3: Distributions of predictor variables across all segments in the digital river network and at invertebrate sites (grey and red histograms, respectively). Similarities in the distributions shown in the two histograms in each panel provide an indication of the degree to which environmental variation across the monitoring sites represents environmental variation across the New Zealand river network; complete representativeness would be indicated by exact matches between the histograms. These twelve predictor variables were the most important in the RF model of MCI and are ordered from most (top left) to least (bottom right) important (Table 4-2).

4.3 Modelled relationships

The predictor variables with high importance in all RF models reflected strong associations between water quality and land use and catchment topography. The proportion of different landcover types in the upstream catchment were amongst the top six most important predictor variables across all models (Table 4-2). The partial plots indicate that CLAR and MCI decreased with increasing usIntensiveAg and usUrban (ranked 2nd and 3rd, respectively), while nutrients, ECOLI and TURB increased (Figure 4-4 and Figure 4-5). In comparison, CLAR and MCI increased with increasing usNativeForest (the proportion of late-successional native forest, ranked 8th), while the nutrients, ECOLI and TURB decreased. These patterns are consistent with previous correlations between landcover and water quality state (Larned et al. 2004; 2016).

Predictors describing the slope and elevation of the upstream catchment, usSlope and usElev, ranked 1st and 6th overall (Table 4-2).The importance values showed usSlope ranked highly in models of CLAR, NO3N, TN, DRP and TP, while usElev was important in models of Tn, ECOLI and MCI. The partial plots indicated that CLAR and MCI increased with increasing values of usSlope and usElev, while the values of all other water quality variables decreased (Figure 4-4 and Figure 4-5).

The predictors usPsize, usBare and usHard had relatively high importance in the water quality RF models (ranks of 4th, 7th and 14th respectively; Table 4-2). These predictors indicate that the regolith of the catchment is associated with water quality state. The values of most water quality variables decreased with increasing values of usPsize, usBare and usHard, with the exception of TURB which increased with increasing usBare and CLAR which increased with usHard (Figure 4-4). These patterns suggest that water quality generally declines as regolith fertility and erosion potential increases.

Predictors of rainfall usRainvar (variation in mean annual catchment rainfall) and usRain (mean annual catchment rainfall) were the 5th and 11th most important overall predictors, respectively (Table 4-2). Values of CLAR and MCI increased with increasing usRainvar and usRain, and the values of all other water quality variables decreased (Figure 4-4 and Figure 4-5). These results suggest that there is a moderately strong positive association between water quality state and catchment rainfall. The mechanisms that drive this association may include solute dilution and sustained low water temperatures.

Table 4-2: 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	TURB	NH4N	NO3N	TN	DRP	TP	ECOLI	MCI
usSlope	2	16	2	2	2	3	1	8	7
usUrban			3	7	5		25	4	4
usIntensiveAg	4	15	5	1	1		11	1	6
usPsize			6	5	14	2	5		
usRainvar	1	1	4	16	12	9	6	3	17
usElev	7	10	1	6	3	14	8	2	3
usBare	6	7		15	7	7	2	5	
usNativeForest	17	6	16	4	4		10	11	1
usTmin	15	2	8	18	13	6	7	6	18
segTmin	18	4	9	3	10	13	9		11
usRain	10	12	14	11	9	1	3	12	5
usScrub				10	6		27		
usPhos	8	11	10	8	18	12	16	9	
usHard	14	5	12	22	20	10	4		
usWetland						15	26	10	9
usTwarm	13	14	11	13	17	8	12	14	8
distToCoast	19	9	13				21		13
segTwarm	3	3		12	15		23		14
usPET	5	8		19	16	11	14	13	16
distToHead	11	17					24		12
usRainDays20	21	13			19	5	17		10
usPastoralLight				17	11	16	15		
usRainDays10	22	19		21		4	13	16	15
usArea	9	18	15				19	15	21
usRainDays100	16	22			8		28		
meanFlow	12	20					18		20
segSlope				14	22		20		19
segElev	20	21	7	20			22	7	2
usExoticForest				9	21		29		
usCalc							32		22
usLakePerc							30		
sinuosity							31		

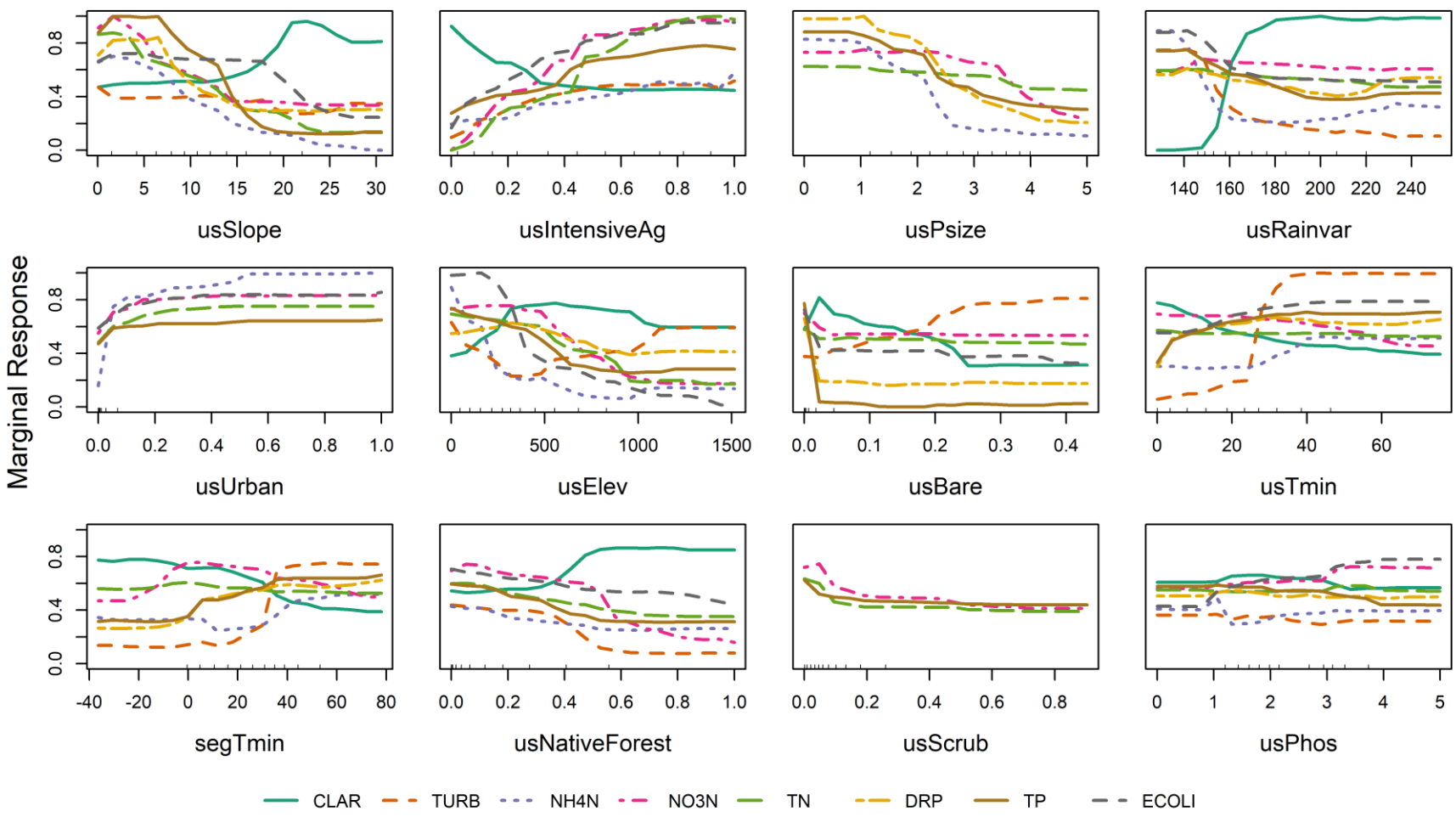

Figure 4-4: Partial plots for the twelve most important predictor variables in random forest models of water quality. Each panel corresponds to one predictor, with predictor variables ordered by overall importance from most (top left) to least (bottom right) important. Y-axis scales represent the standardised value of the marginal response for each of the eight modelled response variables. In each case, the original marginal responses over all twelve 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.

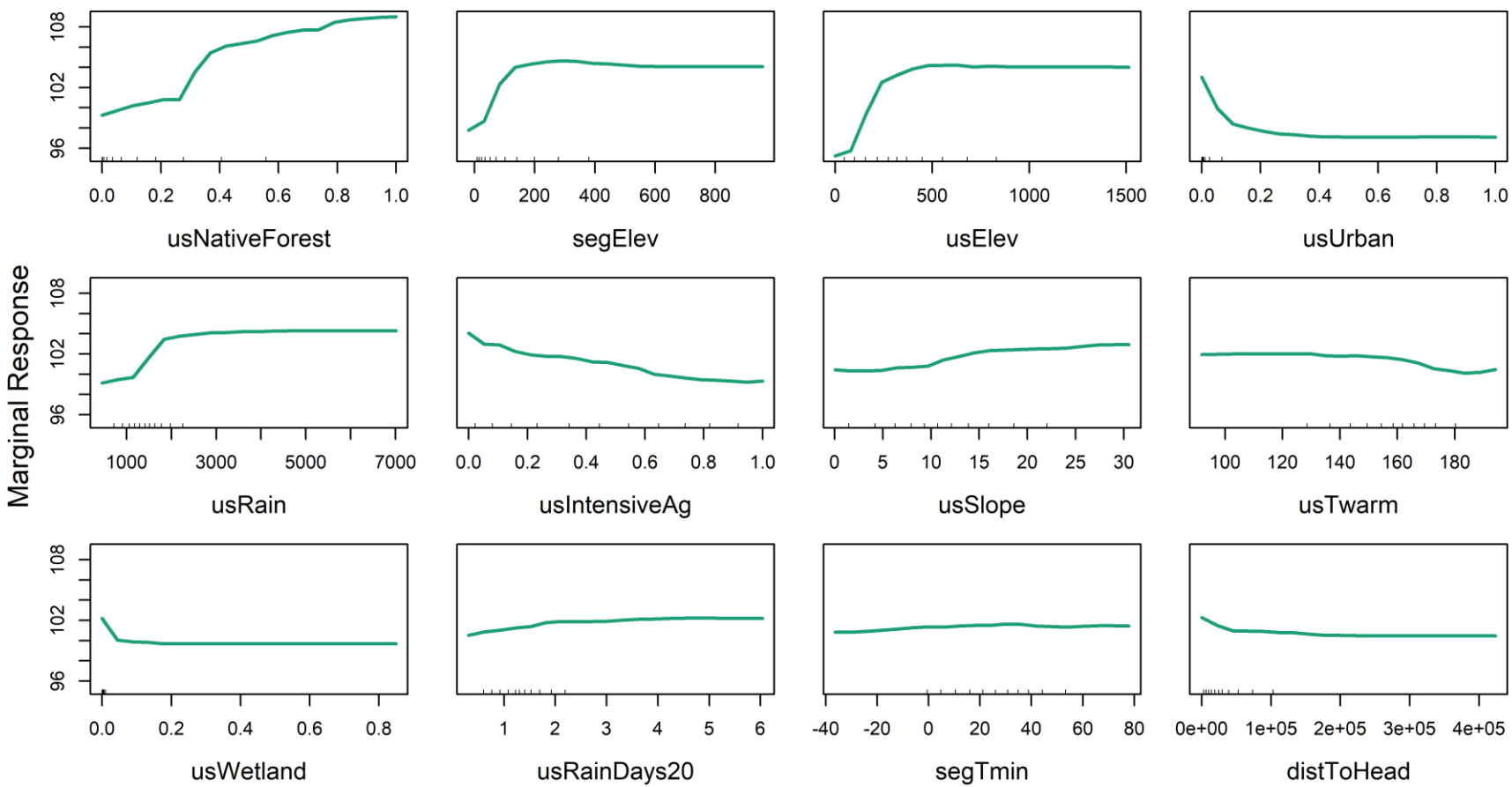

Figure 4-5: Partial plots for the twelve most important predictor variables for the MCI random forest model. Each panel corresponds to one predictor, with predictor variables ordered from most (top left) to least (bottom right) important. The Y-axis scale represents the absolute value of the marginal response (i.e., the scale represents the marginal change in MCI values). The amplitude of each plot (i.e., the range of the marginal response shown on the y-axis) is directly related to a predictor variable's importance.

4.4 Model predictions

The minimum values predicted by the RF models were always somewhat larger than the minimum of the observed values and the maximum predicted values were always somewhat smaller than the maximum observed values (Table 4-3). This is an expected outcome of RF models, which are based on partitioning the data and predictions are derived from the means of observations that are assigned to a particular partition. As a consequence, the predictions for each water quality variable were always within the range of the observations.

Table 4-3: Comparisons of the minimum and maximum observed and predicted values of water quality variables.

Variable and unit	Minimum observed value	Maximum observed value	Minimum predicted value	Maximum predicted value
CLAR (m)	0.18	11	0.27	8
TURB (NTU)	0.20	88	0.40	35
NH4N (mg m^{-3})	0.36	11133	1.61	295
NO3N (mg m^{-3})	1.00	14000	3.89	6041
TN (mg m^{-3})	14.00	17155	32.34	6440
DRP (mg m^{-3})	0.44	2100	0.78	210
TP (mg m^{-3})	0.38	3000	1.20	684
ECOLI (cfu 100ml)	0.17	3400	0.98	1580
MCI (unitless)	51.80	148	60.93	140

The mapped predictions for all five nutrient species (DRP, TP, NH4H, NO3N, TN), TURB and ECOLI have similar coarse-scale spatial patterns, with relatively high values in low-elevation areas on the east coasts of the North and South Island, and in the inland Waikato, Wairarapa Valley, Rangitikei-Manawatu coastal plain, Taranaki Ring Plain, and Auckland Region (Figure 4-7 to Figure 4-13). In contrast, predicted nutrient and ECOLI concentrations are generally low in major mountain ranges (e.g., Southern Alps, Kahurangi, Kaimanawa, and Tararua Ranges), in large areas of the Department of Conservation estate (e.g., Fiordland, Westland, Te Urewera, Egmont, Whanganui and Tongariro National Parks), and in smaller, native forest-dominated areas of Northland and the Coromandel Peninsula.

The low elevation areas characterised by high nutrient and ECOLI concentrations coincide with land used for intensive agriculture and with most of New Zealand's urban centres. High-intensity agricultural and urban land currently account for 60% of the land area below 350 m elevation (Larned et al. 2016). Within these areas, there are some finer scaled differences in predicted water quality. The Canterbury Plains are characterised by high TN and NO3N concentrations, and intermediate TP and DRP concentrations, and the Waikato-Hauraki Plains area is characterised by high concentrations of both nitrogen and phosphorus.

Note that the maps in Figure 4-6 to Figure 4-14 consist of nzsegments of Order 3 and above, and some extensive lowland areas are dominated by low order streams (e.g., eastern Auckland, Tauranga). Steep coastal areas of the Marlborough Sounds, Fiordland, Coromandel and Banks Peninsulas and offshore islands are also dominated by low order streams. The predicted water quality in low order streams in these areas is not shown on the maps in Figure 4-6 to Figure 4-14.

Predicted DRP and TP concentrations are elevated in rivers draining catchments dominated by Tertiary Mudstones (e.g., eastern Wairarapa, and the Aorangi, Puketoi and Ruahine Ranges), and in rivers draining catchments dominated by volcanic andesites, rhyolites and ignimbrites (e.g., central volcanic plateau), as indicated in Figure 4-11 and Figure 4-12. Evidence for phosphorus enrichment due to chemical weathering in these areas comes from several studies of geochemistry and river and lake chemistry (Timperley, 1983; Close and Davies-Colley, 1990, Eden and Parfitt, 1992; McGroddy et al. 2008). The Canterbury coast, Southland and Tasman Bay also showed elevated DRP and TP concentrations that may be associated with anthropogenic sources of phosphorus, such as fertiliser.

Large-scale geographic patterns in predicted MCI scores and CLAR are generally the inverse of those for chemical and microbial water quality (Figure 4-6 and Figure 4-14). Predicted CLAR is relatively high and predicted MCI scores correspond to the excellent and good ecological states (as set out in Stark and Maxted (2008) in mountainous areas, the Department of Conservation estate and other areas dominated by native forest land cover. Predicted CLAR decreases and MCI scores correspond to the fair and poor states in low-elevation alluvial plains and other areas dominated by intensive agriculture and urban land cover. Predicted MCI scores are also fair to poor in some rivers in areas dominated by exotic forest and low-intensity agriculture land cover, such as Central Otago, southwest Canterbury and the Rotorua Lakes-Lake Taupo area.

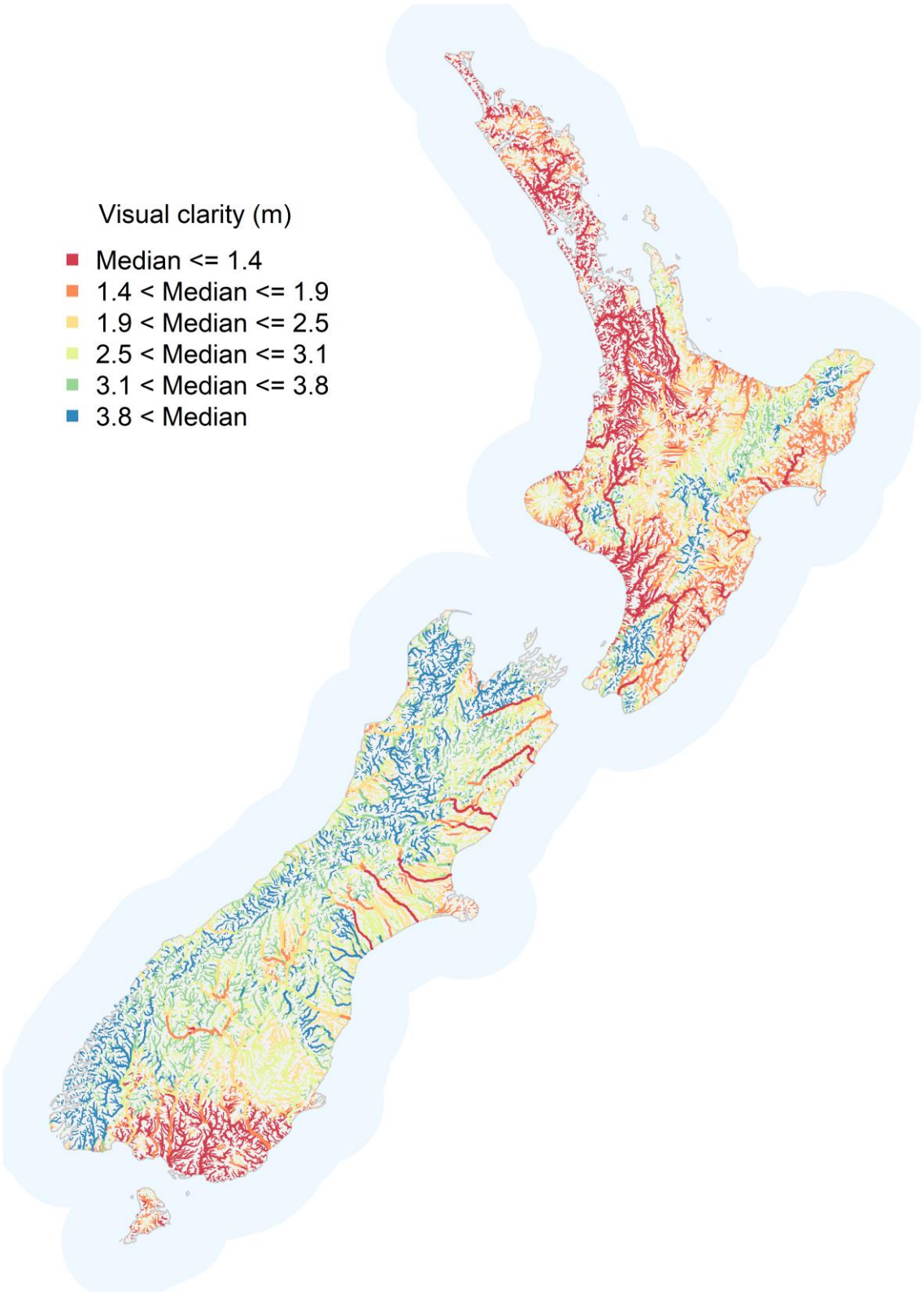

Figure 4-6: Predicted median CLAR in New Zealand rivers. Map shows all nzsegments of Order 3 and higher. Smaller rivers are omitted to make river networks distinguishable.

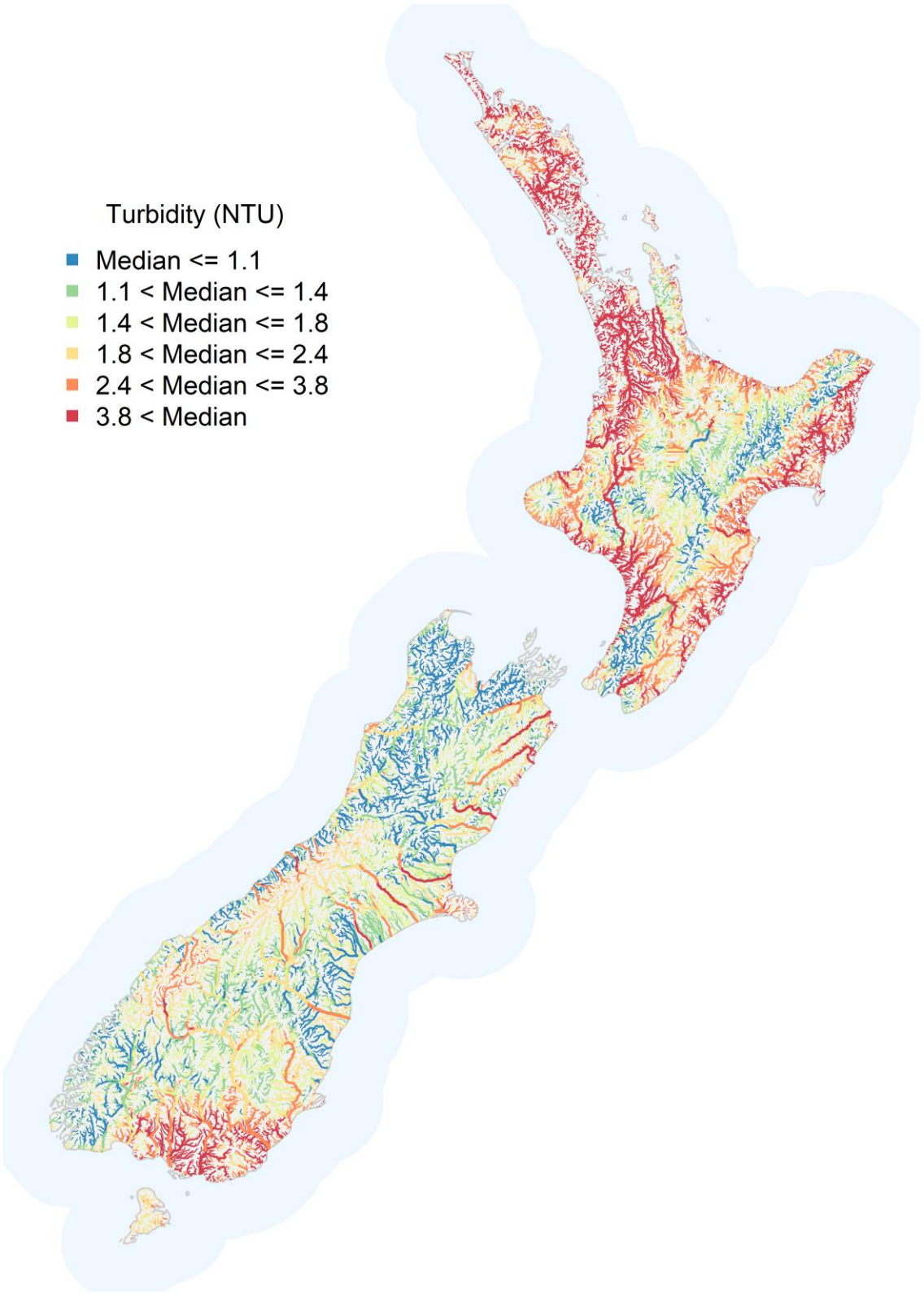

Figure 4-7: Predicted median TURB in New Zealand rivers. Map shows all nzsegments of Order 3 and higher. Smaller rivers are omitted to make river networks distinguishable.

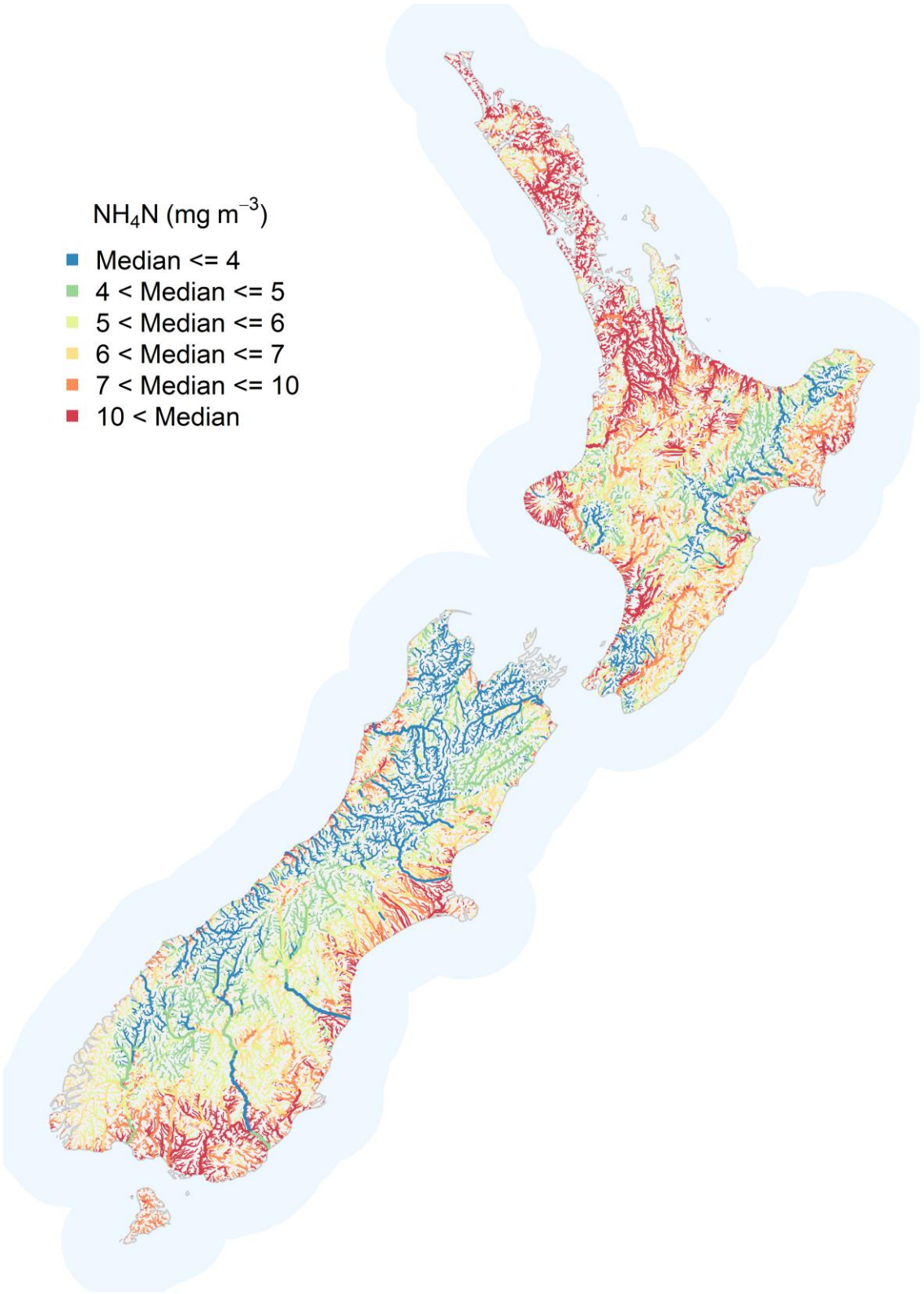

Figure 4-8: Predicted median NH4N in New Zealand rivers. Map shows all nzsegments of Order 3 and higher. Smaller rivers are omitted to make river networks distinguishable.

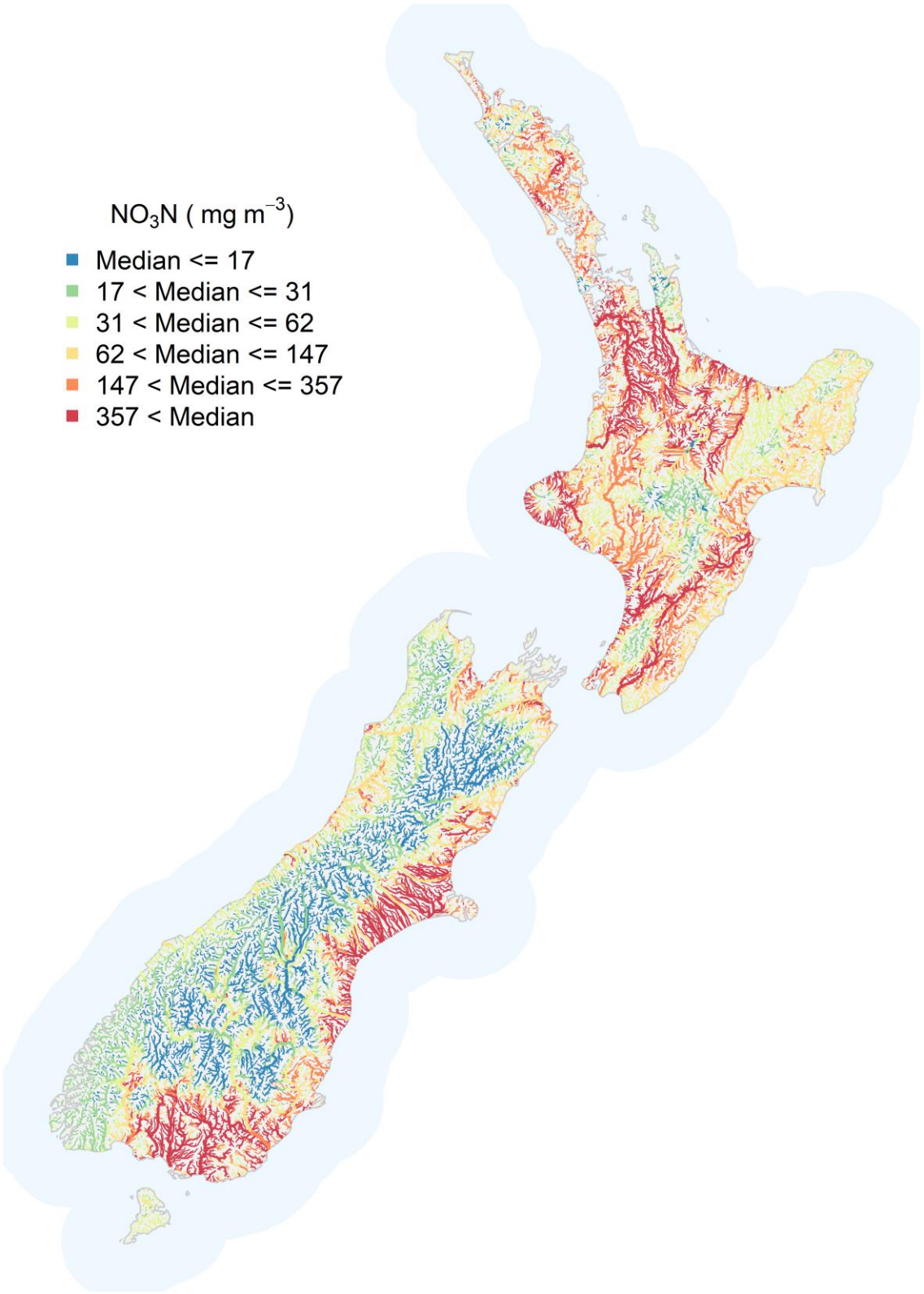

Figure 4-9: Predicted median NO3N concentration in New Zealand rivers. Map shows all nzsegments of Order 3 and higher. Smaller rivers are omitted to make river networks distinguishable.

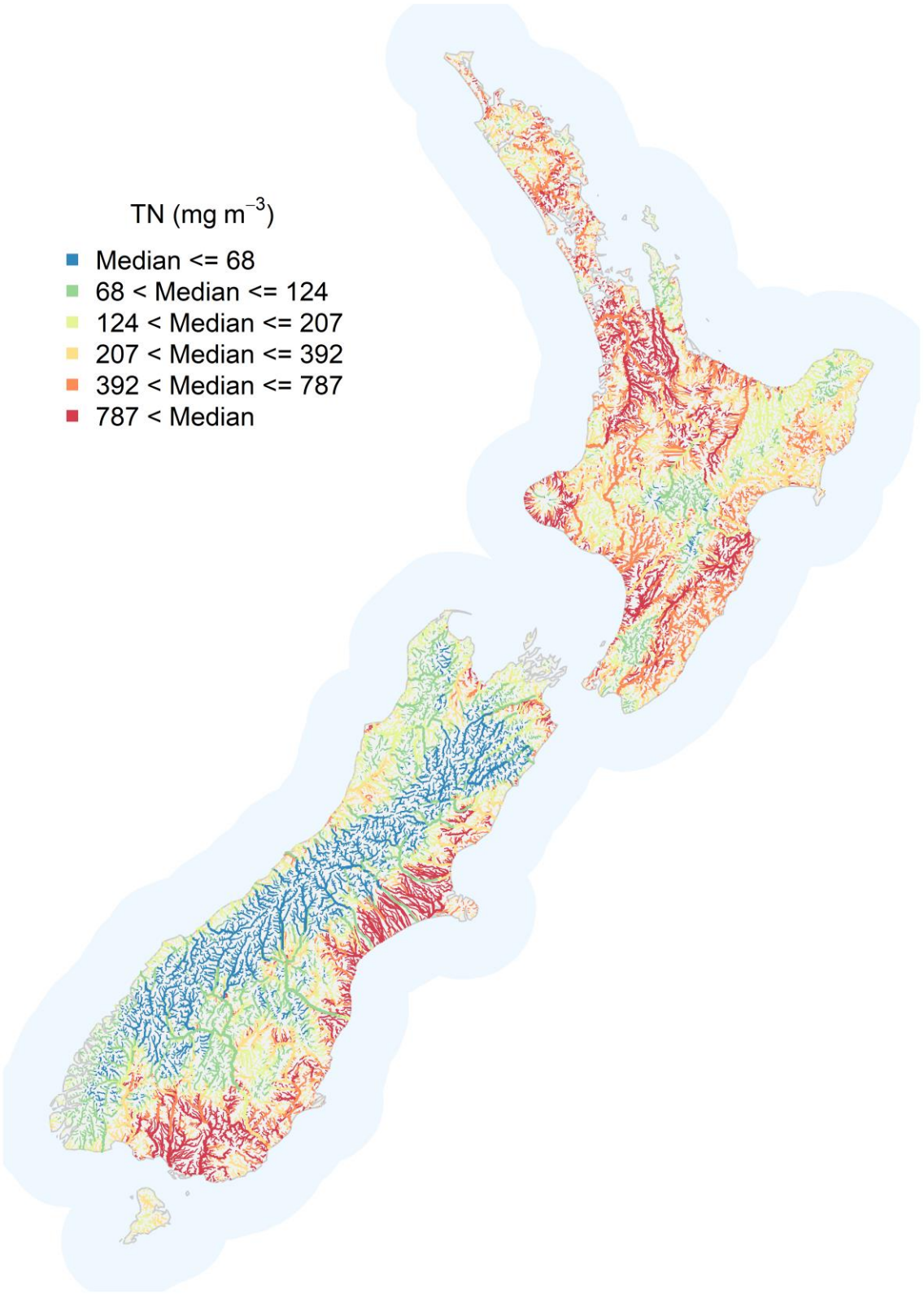

Figure 4-10: Predicted median TN concentration in New Zealand rivers. Map shows all nzsegments of Order 3 and higher. Smaller rivers are omitted to make river networks distinguishable.

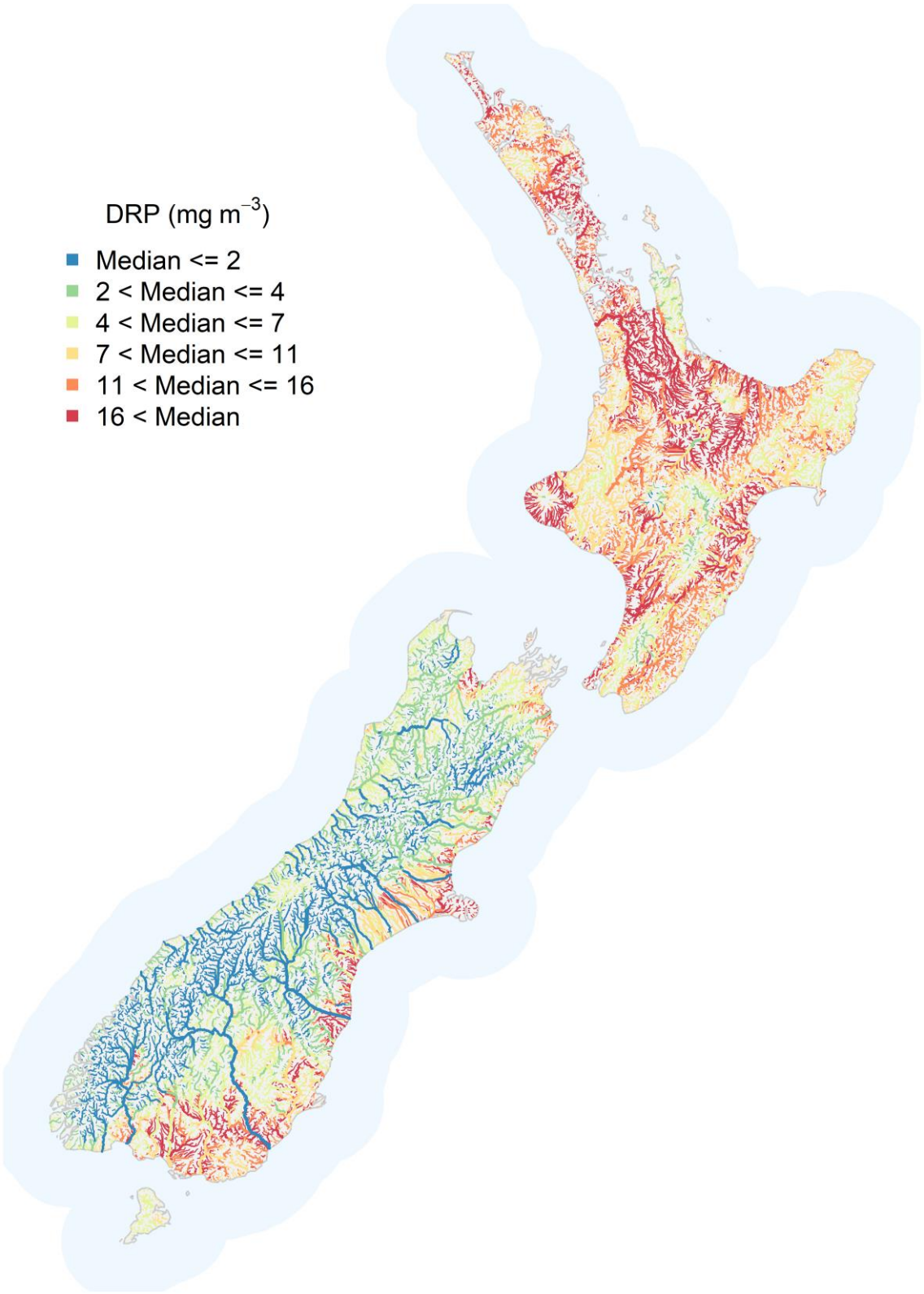

Figure 4-11: Predicted median DRP concentration in New Zealand rivers. Map shows all nzsegments of Order 3 and higher. Smaller rivers are omitted to make river networks distinguishable.

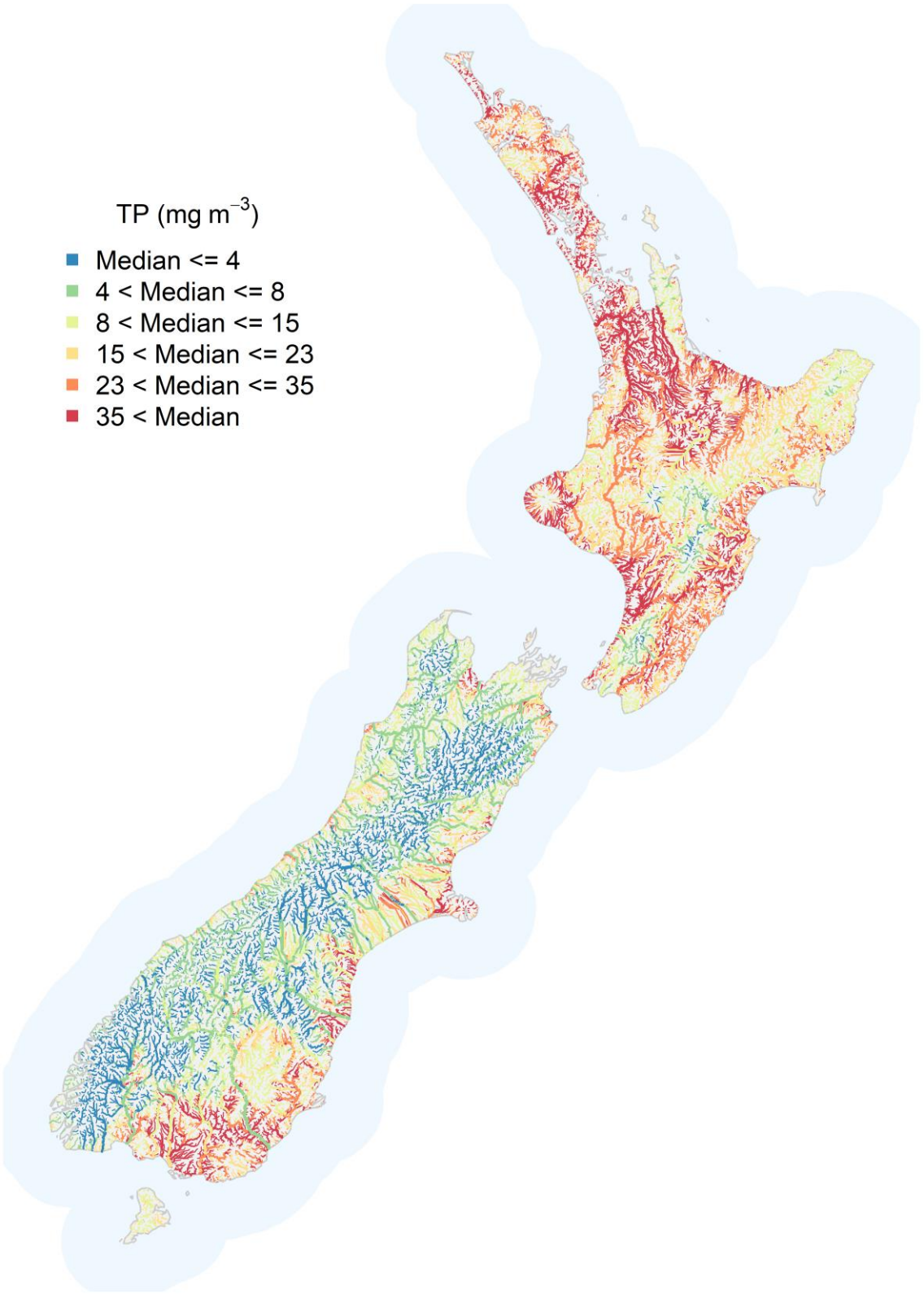

Figure 4-12: Predicted median TP concentration in New Zealand rivers. Map shows all nzsegments of Order 3 and higher. Smaller rivers are omitted to make river networks distinguishable.

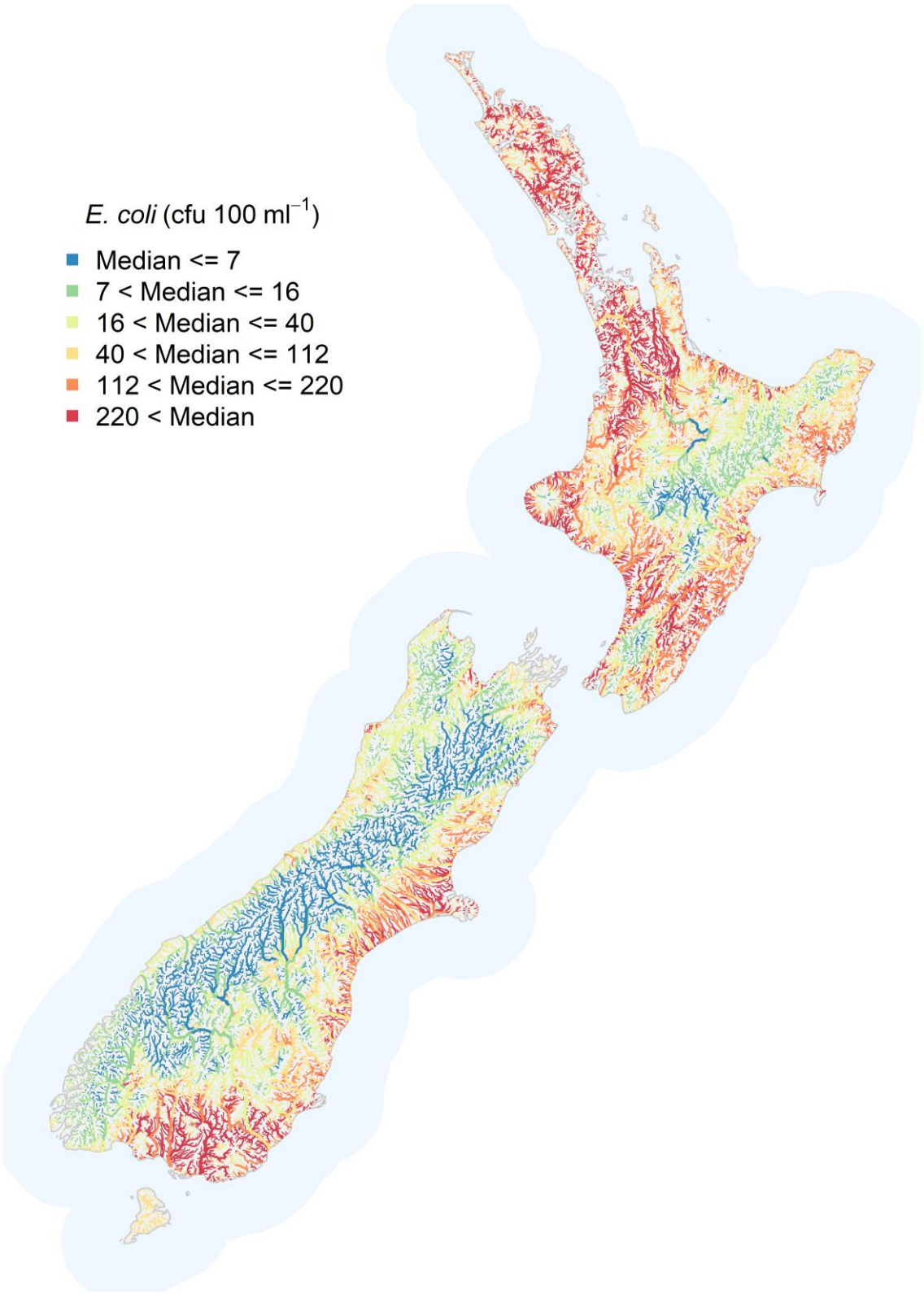

Figure 4-13: Predicted median ECOLI in New Zealand rivers. Map shows all nzsegments of Order 3 and higher. Smaller rivers are omitted to make river networks distinguishable.

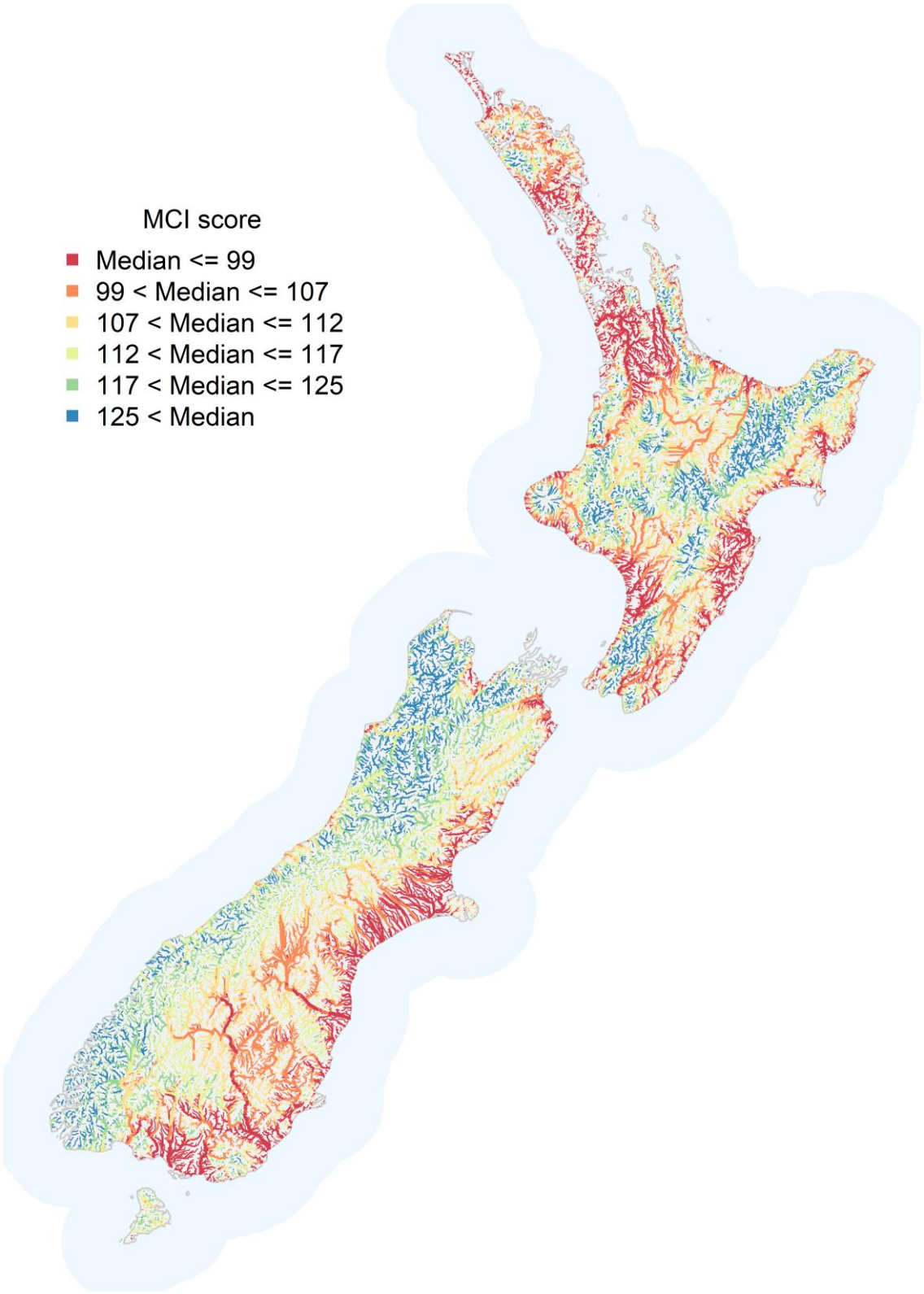

Figure 4-14: Predicted median MCI scores in New Zealand rivers. Map shows all nzsegments of Order 3 and higher. Smaller rivers are omitted to make river networks distinguishable.

5 Discussion

5.1 Comparison with previous studies

The river water quality models presented in this study update previous modelling work carried out by Unwin et al. (2010) and Larned et al. (2016). The models in the previous two reports were based on data from 1996-2007 and 2009-2013, respectively, while the current models are based on data from 2013-2018. The results of the current study are generally consistent with those of Unwin et al. (2010) and Larned et al. (2016), with the structures of the models (as indicated by the relative importance of predictor variables and directions of partial plots) also similar. In addition, the performance of the models in the present study (as indicated by percent variance explained) was generally comparable to model performance in Unwin et al. (2010) and Larned et al. (2016).

Improvements in the modelling methodology and predictor variables between the 2010 and 2016 studies (see Larned et al. 2016) increased the performance of the RF models. In the current study, we used the same modelling procedures and most of the predictor variables as Larned et al. (2016). However, we generated new landcover predictors using 2012 landcover data (LCDB4), improving on the 2008 landcover data (LCDB3) used in the two previous reports. This spatial layer represents the most current landcover data available at a national scale.

In a previous study, Clapcott et al. (2013) fitted an RF model to site median MCI scores and reported a cross-validated R^2 of 0.63. Clapcott et al. (2013) also reported a cross-validated $R^2$of 0.64 with an alternative technique, boosted regression trees. The equivalent R^2 statistic for the MCI model in the current study was 0.69 (Table 4-1). The small improvement in performance in the current study may reflect the longer-term dataset and the use of the REC2 river network.

5.2 Model uncertainty

In this study, we modelled broad scale patterns in water quality using catchment characteristics and segment scale descriptors as predictor variables. Because the processes determining water quality at any location are complex, some unexplained variation in our models is to be expected. Predictions made for individual locations are associated with uncertainties characterised by model RMSD (Table 4-1). However, the level of model bias for each water quality variable was low, which indicates that the predicted patterns reflect broad scale relative differences between locations.

The 95% confidence intervals for values predicted by our models for individual segments can be obtained using the following equations. Equation 1 should be used for calculating the intervals for the MCI predictions. Equation 2 should be used for calculating the intervals for the other water quality variables for which the variables were $\log_{10}$ transformed prior to model fitting and the prediction uncertainty (RMSD) values have been reported in the $\log_{10}$ transformed space.

$$95\%\ CI = x \pm 1.96 \times RMSD \quad (1)$$

$$95\%\ CI = 10^{[\log_{10}(x) \pm 1.96 \times RMSD]} \quad (2)$$

where x is the estimated value in the original units, RMSD is the model error and 1.96 is the standard normal deviate or Z-score for probability ($0.025 \leq Z \geq 0.975$). The prediction confidence intervals for the $\log_{10}$-transformed variables, when expressed in the original units of the variables, are asymmetric and their values vary in proportion to the predicted water quality value. For example, if we let x be a

predicted value for CLAR of 0.1 m, the lower and upper 95% confidence intervals are 0.04 and 0.25 m, respectively, whereas if x is 1.0 m the lower and upper 95% confidence intervals are 0.4 and 2.5 m, respectively.

RF model performance differed among water quality variables (Table 4-1). This variation may be attributed to differences in the biophysical processes that control those variables. Some biophysical processes may be poorly represented by our catchment-averaged spatial predictor variables. For example, concentrations of dissolved and total nitrogen and phosphorus in rivers are influenced to differing degrees by adsorption-desorption processes, deposition and suspension, and biological assimilation, transformation and removal; these mechanisms are not explicitly represented in the RF models. The absence of predictors that account for these and other processes means that some level of unexplained variation is inevitable.

5.3 Alternative modelling methods

The RF method that we used to develop river water quality models is well-suited to data from monitoring sites that represent a wide range of environmental conditions. However, it is not the only method available. Alternative statistical models include generalised additive models (GAMs; Hastie et al. 2001), artificial neural networks (e.g., Joy and Death 2001), and boosted regression trees (e.g., Leathwick et al. 2005). We did not employ these alternatives, but it is possible that some water quality applications would be better served by models developed by one of the alternative methods. In particular, if it is important to identify areas with potentially extreme water quality values, models such as GAMS that can extrapolate beyond the range of the fitting data would be useful, although such predictions may lead to spurious results . In addition, models that incorporate biophysical processes (e.g., CLUES; Alexander et al. 2002) are available; in some circumstances, process models are better suited to inform environmental policy. We considered RF models to be the best tool for predicting river water quality state for national scale reporting for 3 general reasons:

1. Spatial data that correspond to land cover and other environmental characteristics are widely available in New Zealand. These data are suitable for investigating associations between water quality and environmental characteristics, and empirical models are appropriate tools for identifying those associations. In contrast, process models require measurements or estimates of catchment processes (e.g., erosion, contaminant transport and transformation) and these data are in far shorter supply. In addition, process models are generally more time consuming and complex to calibrate than purely empirical models.

2. RF model predictions can be mapped at scales ranging from single network segments to the entire county. These maps provide a useful description of spatial patterns in water quality for environmental reporting purposes.

3. Among empirical modelling methods that generate associations between water quality and environmental characteristics, RF models have several advantages: they are minimally affected by multi-collinearity among predictor variables, they cannot be over-fitted, and they are unaffected by variation in data distributions. RF models cannot predict beyond the range of the observations, which may limit their utility in some applications. In the present study, limiting model predictions to the range of observations was a positive attribute as it ensured that those predictions were conservative.

6 Acknowledgements

Thanks to Dan Elder, Annabelle Ellis and Sheree De Malmanche (Ministry for the Environment) for support of this project and for feedback, enthusiasm, and patience. Thanks also to Abi Loughnan and the LAWA team for data updates and the many regional council staff who provided additional data and information about monitoring programmes. Thanks to Ton Snelder for provision of analytical code and statistical advice. Finally, thanks to Fenella Falconer, Rick Stoffels and Scott Larned for formatting and reviewing the report.

7 References

Alexander, R.B., Elliott, A.H., Shankar, U., McBride, G.B. (2002) Estimating the sources and transport of nutrients in the Waikato River basin, New Zealand. *Water Resources Research* 38: 1268.

Booker, D., Snelder, T. (2012) Comparing methods for estimating flow duration curves at ungauged sites. *Journal of Hydrology* 434: 78-94.

Breiman, L. (2001) Random forests. *Machine Learning* 45: 5–32.

Breiman, L., Friedman, J.H., Olshen, R., Stone, C.J. 91984) Classification and Regression Trees. Wadsworth, Belmont, California.

Clapcott, J.E., E. Goodwin, E., Snelder, T. (2013) Predictive models of benthic macroinvertebrate metrics. *Cawthron Institute Report* 2301.

Close, M.E., Davies-Colley, R.J. (1990) Baseflow water chemistry in New Zealand rivers 2. Influence of environmental factors. *New Zealand Journal of Marine and Freshwater Research* 24: 343-356.

Cutler, D.R., Edwards, J.T.C., Beard, K.H., Cutler, A., Hess, K.T., Gibson J., Lawler, J.J. (2007) Random forests for classification in ecology. *Ecology* 88: 2783–2792.

Donath, F.M., Daughney, C.J., Morgenstern, U., Cameron, S.G., Toews, M.W. (2015) Hydrochemical interpretation of groundwater-surface water interactions at catchment and local scales, Lake Rotorua catchment, *New Zealand. Journal of Hydrology (NZ)* 54: 11-32.

Duan, N. (1983) Smearing estimate: a nonparametric retransformation method. *Journal of the American Statistical Association* 78: 605–610.

Eden, D.N., Parfitt, R.L. (1992) Amounts and sources of phosphate in hill country rocks, south-eastern North Island, *New Zealand. Australian Journal of Soil Research* 30: 357-370.

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

Joy, M.K., Death, R. G. (2001) Control of freshwater fish and crayfish community structure in Taranaki, New Zealand: dams, diadromy or habitat structure? *Freshwater Biology* 46: 417–429.

Larned, S.T., Scarsbrook M.R., Snelder. T., Norton, N.J., Biggs, B.J.F. (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., Unwin M.J. (2012) Representativeness and statistical power of the New Zealand river monitoring network. New Zealand Ministry for the Environment. *NIWA Report* CHC2012-079.

Larned, S.T., Snelder, T., Unwin, M., McBride, G. B. (2016) Water quality in New Zealand rivers: current state and trends. *New Zealand Journal of Marine and Freshwater Research* 50: 389–417.

Larned, S.T., Whitehead, A.L., Snelder, T.H., Fraser, C., Yang, J. (2018) Water quality state and trends in New Zealand rivers. *NIWA Client Report 2018347CH*. NIWA, Christchurch.

Leathwick, J.R., Overton, J.M., McLeod, M. (2003) An environmental domain analysis of New Zealand, and its application to biodiversity conservation. *Conservation Biology* 17: 1612–1623.

Leathwick, J.R., Rowe, D., Richardson, J., Elith, J., Hastie, T. (2005) Predicting the distributions of New Zealand's freshwater diadromous fish. *Freshwater Biology* 50: 2034–2052.

Leathwick, J., Snelder, T., Chadderton, E., Elith, J., Julian, K., Ferrier, S. (2011) Use of generalised dissimilarity modelling to improve the biological discrimination of river and stream classifications. *Freshwater Biology* 56: 21–38.

Liaw, A., Wiener, M. (2002) Classification and Regression by randomForest. *R News* 2: 18–22.

McDowell, R., Snelder, T., Booker, D., Cox, N., Wilcock, R. (2012) Establishment of reference or baseline conditions of chemical indicators in New Zealand streams and rivers relative to present conditions. *Marine and Freshwater Research* 64: 387–400.

McGroddy, M.E., Baisden, W.T., Hedin L.O. (2008) Stoichiometry of hydrological C, N, and P losses across climate and geology: An environmental matrix approach across New Zealand primary forests. *Global Biogeochemical Cycles* 22, GB102.

Moriasi, D.N., Arnold, J.G., Van Liew, M.W., Bingner, R.L., Harmel, R.D., Veith, T.L. (2007) Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE* 50: 885–900.

Nash, J.E. Sutcliffe, J.V. (1970) River flow forecasting through conceptual models Part I—A discussion of principles. *Journal of Hydrology* 10: 282–290.

Piñeiro, G., Perelman, S., Guerschman, J., Paruelo, J. (2008) How to evaluate models: observed vs. predicted or predicted vs. observed? *Ecological Modelling* 216: 316–322.

R Core Team (2017) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Stark, J D., Maxted, J. (2007) A user guide for the MCI. New Zealand Ministry for the Environment. *Cawthron Report* 1166.

Svetnik, V., Liaw, A., Tong, C., T. Wang, T. (2004) Application of Breiman's Random Forest to modeling structure-activity relationships of pharmaceutical molecules. *Lecture Notes in Computer Science* 3077: 334-343.

Timperley, M.H. (1983) Phosphorus in spring waters of the Taupo volcanic zone, North Island, New Zealand. *Chemical Geology* 38: 287-306.

Unwin, M., Snelder, T., Booker, D. Ballantine, D., Lessard, J. (2010) Predicting water quality in New Zealand rivers from catchment-scale physical, hydrological and land cover descriptors using Random Forest models. *NIWA Client Report* CHC2010-037.