Assessment of *Escherichia coli* Load Reductions to Achieve Draft Freshwater Objectives in the Rivers of Southland Murihiku

To inform the Southland Regional Forum process

August 2021

Prepared By:

Ton Snelder

Caroline Fraser

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:	2020-21
Report Date:	August 2021

Quality Assurance Statement

Version	Reviewed By	
Final	Ned Norton	

Table of Contents

Figures

Tables

Executive Summary

This report describes *Escherichia coli* (*E. coli*) load reductions predicted to be needed in order to achieve options for draft freshwater objectives (FWOs) in rivers in Southland Murihiku. The analysis does not consider how the *E. coli* load reductions would be achieved. The analysis aims to inform the Southland Regional Forum process about the magnitude of the load reductions needed for each option and how these vary across the region.

The study area includes all of Southland except Fiordland and the Offshore Islands. The underlying analysis utilised several models that describe concentrations and loads of *E. coli* in the rivers across the study area. The concentrations and loads were combined with criteria associated with FWOs. Calculations were made of the amounts by which current loads would need to be reduced to allow the draft FWOs to be achieved (i.e., the load reduction required).

The options for FWOs are defined in terms of a band (A, B or C) for all river, lake and estuary receiving environments in the study area. Two sets of FWO options for rivers were nominated and are referred to as the Hauora option and the Proposed Southland Water and Land Plan water quality standards (pSWLP option). The Hauora option was based on hauora principles as reported in Bartlett et al., (2020) and represents the minimum state necessary to support hauora for human health attributes assessed in this report. The pSWLP option represents the possible minimum FWOs based on water quality standards in the pSWLP as reported in detail in Norton and Wilson (2019). In December 2020, the Southland Regional Council and the Te Ao Marama Inc (TAMI) board approved in principle the minimum level of hauora as the draft FWOs for the region. The FWOs will remain draft until they are set through a plan process.

The load reductions required for each FWO option were estimated for individual river segments and these individual results were also aggregated to report on individual 'reporting catchments' and the whole study area. For the whole study area and the Hauora option, the best estimate of the *E. coli* load reductions required were 347 peta *E. coli* yr^{-1}, which represents a best estimate of 90% of the current loads. For the whole study area and the pSWLP option, the *E. coli* load reductions required were estimated to be 311 peta *E. coli* yr^{-1}, which represent 82% of the current loads. These large load reductions reflect the stringency of the FWO options being considered. We note however that an analysis based on meeting the NOF C band state (i.e., the minimum state deemed suitable for primary contact in the national targets laid out in Appendix 3 of the NPSFM) indicate a regional load reduction of 77% of the current *E. coli* load. This result indicates that *E. coli* loads in the Region's rivers far exceed even that minimal accepted state.

This study also provided an estimate that 61% of the large rivers (i.e., those with stream order ≥4) across the whole Southland region (i.e., including Fiordland and the Offshore Islands) are currently suitable for primary contact according to the NPSFM criteria (i.e., in attribute state C or better). This estimate is comparable to a previous estimate of 62% of these large rivers in Southland being suitable for primary contact as reported in MFE (2018). It is noteworthy that the current draft FWOs approved in principle by Environment Southland and the Te Ao Marama Inc (TAMI) board in November 2020, and indeed both sets of FWO options assessed in this report, would see 100% of rivers being suitable for primary contact (i.e., better than C band) within a generation (25-30 years). This is a step further than the regional primary contact targets set by Environment Southland in December 2018 for 66% of rivers to be suitable by 2030 and 80% by 2040, with continued improvement beyond 2040. Those regional targets were Environment Southland's intended contribution to meeting the national targets laid out in the 2017 NPSFM (unchanged in the NPSFM 2020) of 80% being suitable by 2030 and 90% being suitable no later than 2040.

Like all assessments of this type, the predicted load reductions required are subject to considerable uncertainty. The study has quantified these uncertainties and presents the results as best estimates (of the load reduction required) and the 90% confidence interval for these estimates. The broad scale patterns in the estimated *E. coli* load reductions provide a reliable indication of the relative differences between locations. However, there is considerable uncertainty associated with the absolute values of the *E. coli* load reductions and these become larger as the spatial scale over which the reductions are evaluated is reduced. It is unlikely that these uncertainties can be significantly reduced in the short to medium term (i.e., in less than 5 to 10 years) because, among other factors, the modelling is dependent on the collection of long-term water quality monitoring data.

1 Introduction

This report describes an assessment of *Escherichia coli* (*E. coli*) load reductions required to achieve the draft numeric freshwater objectives (FWOs) in the rivers of Southland Murihiku. The purpose is to inform the Southland Regional Forum process, which is making recommendations on how FWOs can be achieved in the Southland Region.

The analysis described in this report does not consider how the *E. coli* load reductions would be achieved; this will be the subject of subsequent studies. The current report therefore only aims to inform the Forum process about the magnitude of the required load reductions, how these vary across the region, and to establish a framework for future scenario testing of methods. The FWOs remain draft until the testing of methods to achieve the reductions has been completed and finalised FWOs are set through a plan process.

The analysis methodology is based on similar studies that assessed national-scale studies of nitrogen load reduction requirements (MFE, 2019; Snelder *et al.*, 2020) and regional scale nutrient reductions (nitrogen and phosphorus) in the Southland region (Snelder, 2020). However, the current analysis involved some modifications to methods used by these earlier studies to represent the Southland region in greater detail, and to assess load reductions for *E. coli* rather than nutrients. To keep the current report simple, the methods are described only in broad terms and the reader is referred to MFE (2019) and Snelder *et al.* (2020) for the details of the methodology. The exceptions to this are descriptions of details of the method where these pertain to modifications made for the current study.

2 Methods

2.1 Overview

Conceptually, this study represents *E. coli* loads being generated in catchments and transported to downstream river and stream receiving environments by the drainage network (Figure 1). The loads of *E. coli* (i.e., *E. coli* organisms per year) arriving at each receiving environment determine the distribution of *E. coli* concentrations through time and therefore the risk to human health (MFE and MoH, 2003). Acceptable risks to human health are defined by levels of four statistics (i.e., criteria) that describe the distribution of *E. coli* values at a site. These statistics are the annual median and 95th percentile concentrations (Median, Q95), and the proportion of samples for which concentration thresholds of 260 and 540 *E. coli* 100mL^{-1} are exceeded (G260, G540). These statistics are used because they are the basis for the *E. coli* attribute states in the National Objectives Framework (NOF) appended to the National Policy Statement – Freshwater (NPS-FM; NZ Government, 2017, 2020). Where one or more of the values of the four statistics exceed a defined criterion, there is a requirement to reduce the current load of *E. coli*. The four statistics are also the basis for national targets to increase the proportions of large rivers that are suitable for primary contact (i.e., that are C band state or better), as set out in Appendix 3 of the NPS-FM (NZ Government, 2017, 2020).

This study's calculations were based on a spatial framework that represents the drainage network (i.e., streams and rivers) and associated catchments. Calculations were performed for every segment of the network, which represent river receiving environments.

The calculation of load reductions required were based on statistical models fitted to *E. coli* data obtained from river state of environment (SOE) monitoring in the Southland region. The analyses were undertaken in five steps (Figure 1). At step 1, observations made at each SOE site were analysed to calculate the four *E. coli* statistics and to calculate annual loads of *E. coli* (i.e., number of organisms) per year. In addition, linear regression models were used to relate the observed values of the *E. coli* statistics at each SOE site with the associated *E. coli* loads (expressed as yields by dividing by the catchment area of each SOE site). At step 2, the four statistics and the loads (expressed as yields) were used as training data in spatial models. The spatial models were used to predict the value for the four statistics and the *E. coli* loads for every segment of the network (i.e., every stream and river receiving environment) within the study area.

The criteria to achieve two sets of options for FWOs in river receiving environments are defined in terms of four statistics representing *E. coli* concentrations. At step 3 of the analysis, compliance was assessed by comparing these criteria with the associated predicted value for each river segment. In addition, the linear models relating *E. coli* statistics and loads defined at step 1 were used to calculate the maximum allowable load (MAL), which is the load that will ensure the four *E. coli* statistics do not exceed their associated criteria at each segment. The local excess load for each segment was then calculated as the current load minus the MAL. The local excess load is the amount by which the current load at each segment would need to be reduced to achieve the FWO.

At step 4 of the analysis, the load reduction required at every point in the drainage network was calculated as the maximum of the local load reduction at that and all upstream receiving environments. The load reduction required differs from the local excess load in that it considers the excess load of all upstream receiving environments. Thus, a point in the network may have a local excess load of zero but, if it is situated downstream of receiving environments that have local excess loads, it will have a load reduction required that reflects those upstream local excess loads. The load reduction required can be expressed in absolute terms as a load of organisms per year (*E. coli* yr^{-1}), as a yield (organisms per catchment area per year; *E. coli* ha^{-1} yr^{-1}) and as a proportion of the current load of *E. coli* (%).

At step 5 critical point catchments were identified by first identifying critical points in each sea-draining catchment in the study area. For every point in the drainage network there is a critical point, which is the downstream segment that has the highest ratio of current load to MAL. The catchment upstream of the critical point is a critical point catchment and has a load reduction required, which is the local excess load at the critical point. The critical catchment load reduction required is expressed as a yield (i.e., number of *E. coli* organisms per catchment area; *E. coli* ha^{-1} yr^{-1}) or as a percentage of current *E. coli* load (%). The critical catchment load reduction required indicates the spatially averaged reduction rate that would be required over the entire area of the critical point catchment to reduce the load sufficiently to allow FWO to be achieved across the entire catchment. Sea-draining catchments can have one critical point (the most downstream receiving environment) or multiple critical points, which include the most downstream receiving environment and other sub-catchments. Critical catchments can have a catchment load reduction required of zero when the current load is less than the MAL or have positive values when the current load exceeds the MAL.

It should be kept in mind that a critical catchment load reduction required represents a load reduction for the whole critical catchment. If the catchment includes areas of non-productive land, and the methods for load reduction are restricted to mitigation actions associated with pastoral land use, the required load reduction from productive land would need to be higher than the reported value.

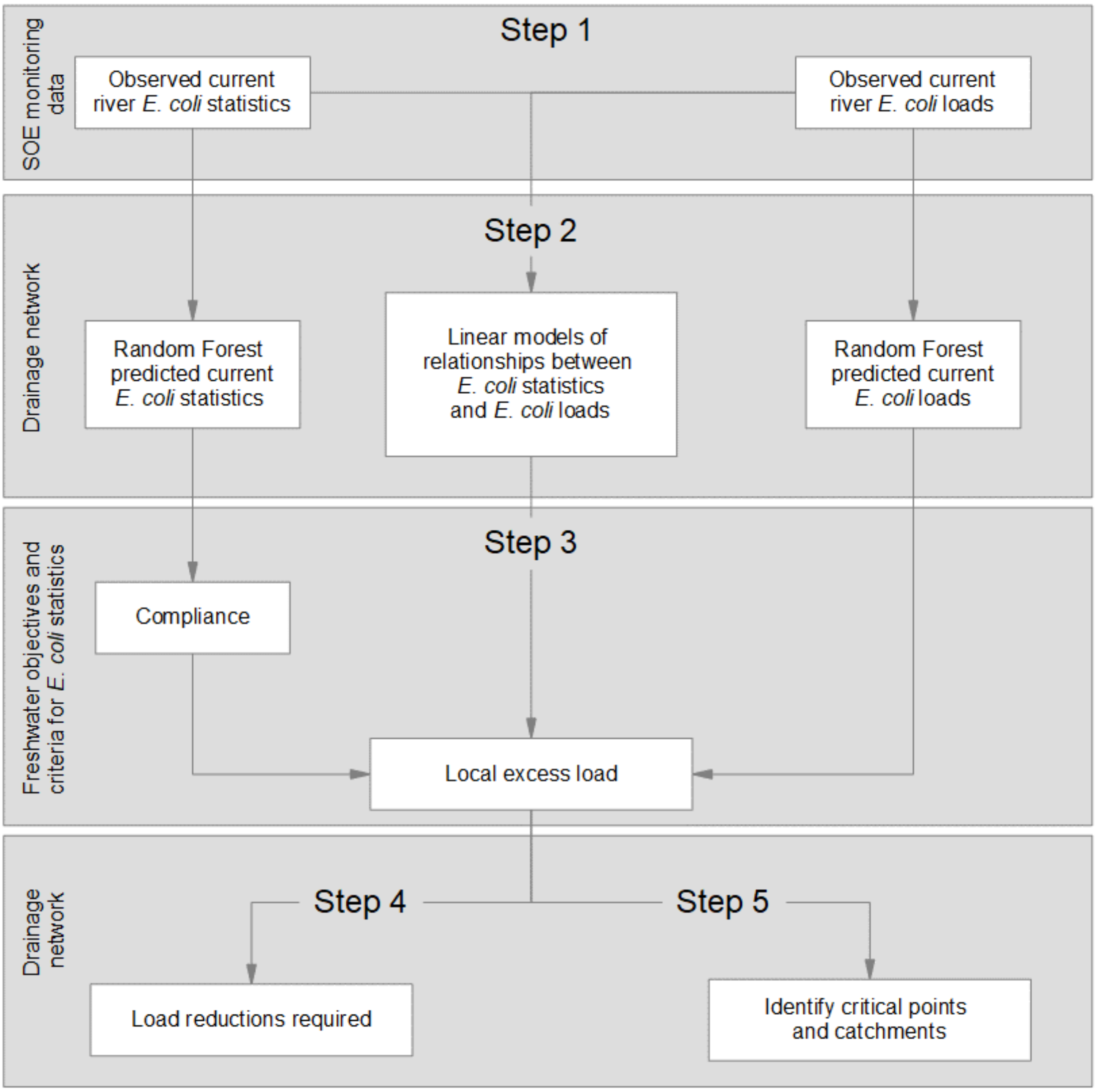

Figure 1. Schematic diagram of the assessment of E. coli load reductions required to achieve freshwater objectives.

The following sections describe the various components of the analysis shown in Figure 1 in more detail.

2.2 Spatial framework

The study area comprised the Southland region excluding Fiordland and Offshore Islands (Figure 2). The drainage network and river receiving environments were represented by the GIS-based digital drainage network, which underlies the River Environment Classification (REC version 2.4; Snelder and Biggs, 2002). This is the same drainage network that was the spatial framework used by Snelder (2020). The digital network was derived from 1:50,000 scale contour maps and represented the rivers within the study area as 44,000 segments bounded by upstream and downstream confluences, each of which is associated with a sub-catchment (Figure 2). The terminal segments of the river network (i.e., the most downstream points in each drainage network that discharges to the ocean) were identified.

Each segment in the network has been allocated to a reporting catchment that is an individual sea-draining catchment (Figure 3). The results of the load reductions required analyses can be reported at any spatial scale from individual receiving environments (i.e., river segments, Figure 2), to reporting catchments (Figure 3) and the whole study area.

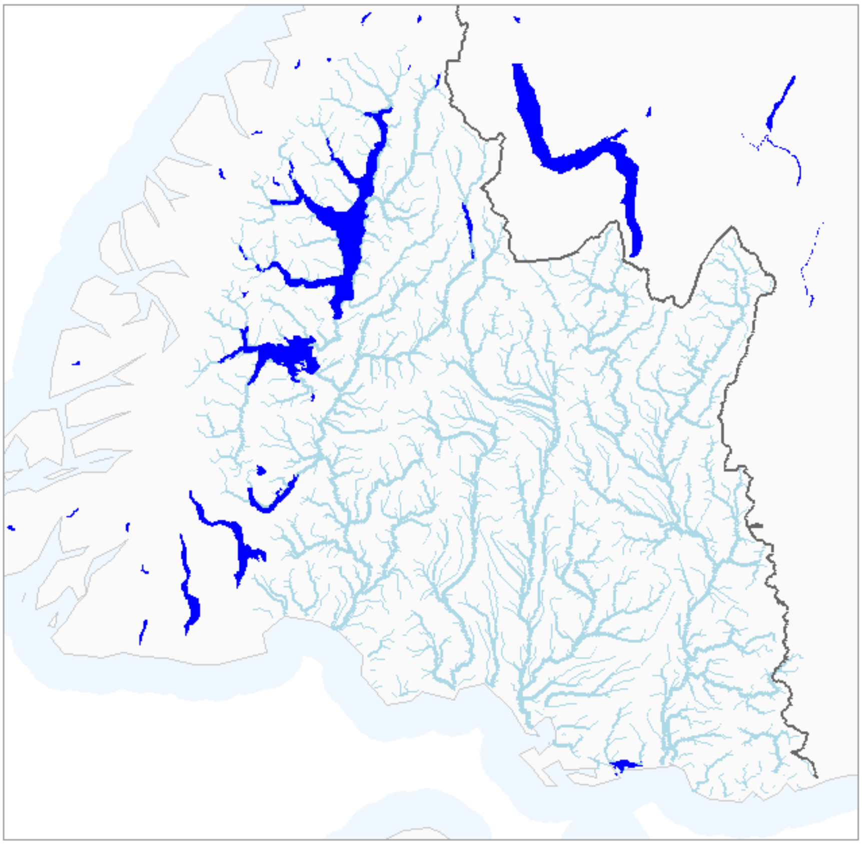

Figure 2. The digital network within the study area that provided the spatial framework for the analysis.

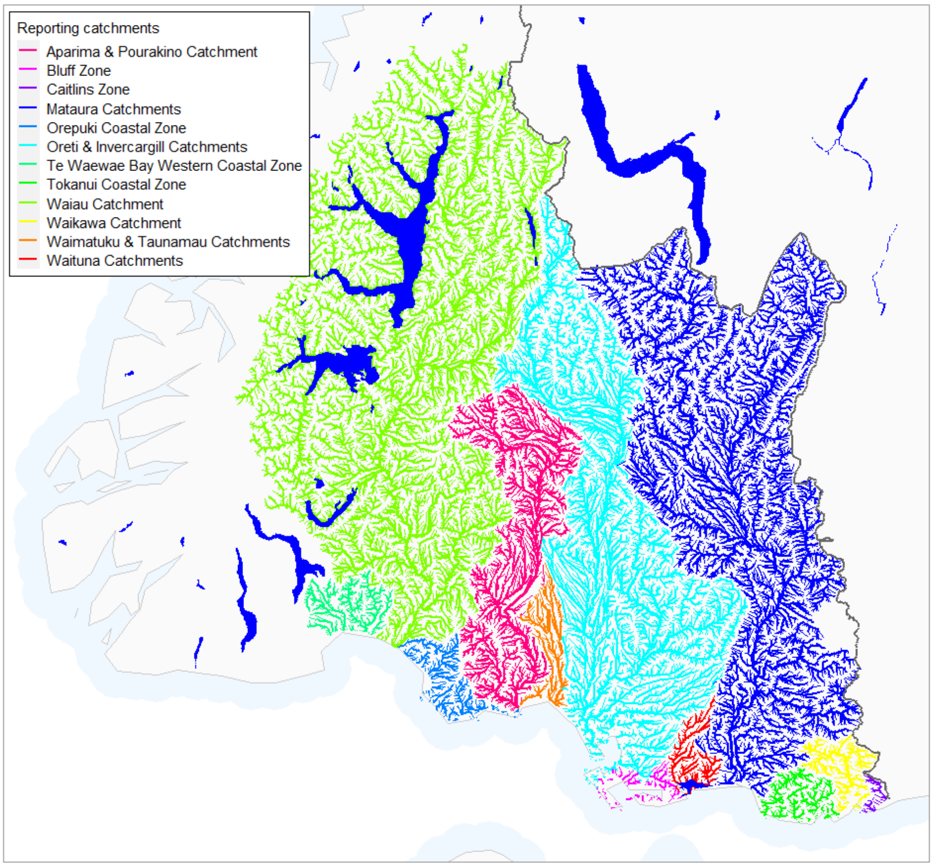

Figure 3. Reporting catchments used for summarising the results of the analysis.

2.3 *E. coli* criteria and predicted current *E. coli* statistics

The NOF defines five levels of environmental quality (termed "attribute states" in the NPS-FM) denoted A, B, C, D and E. The five attribute states are linked to threshold criteria for the four *E. coli* statistics shown in Table 1. The attribute states are associated with low (A) to high (E) concentrations of *E. coli*, which are linked to low to high risk of infection by microbiological pathogens for humans contacting the water. Each of the four criteria defined by Table 1 must be satisfied (i.e., the value of each statistic representing the state of a river receiving environment must be lower than the criteria) for that receiving environment to be in that attribute state. Thus, if one or more criteria cannot be satisfied in an attribute state, a lower attribute state applies.

The national targets to increase the proportions of large rivers that are suitable for primary contact are linked to NOF attribute states. The NPS-FM defines rivers as suitable for primary contact if they are in NOF attributes states A, B or C and large rivers are defined by network segments of a stream order of ≥4 as defined by the REC. In Southland, river segments of stream order of ≥4 have a minimum catchment area of 5.5 km^2 and approximately 75% of them have mean flows >1 m^3 s^{-1}.

Table 1. Criteria used to define the E. coli freshwater attribute states.

Criteria	**Attribute state**				
	A	**B**	**C**	**D**	**E**
Median *E. coli*/ 100ml^{-1} (Q50)	<130	130	130	260	>260
95th Percentile *E. coli* 100ml^{-1} (Q95)	<540	1000	1200	1200	>1200
Proportion of exceedances over 260 *E. coli* 100ml^{-1} (G260)	<0.2	0.3	0.34	0.5	>0.5
Proportion of exceedances over 540 *E. coli* 100ml^{-1} (G540)	<0.05	0.1	0.2	0.3	>0.3

The analysis described below was based on values of four NOF *E. coli* statistics (Table 1) that were predicted for all segments of the drainage network using spatial statistical regression modelling. The statistical modelling to predict the values of the four *E. coli* statistics for every network segment commenced by calculating each of the four statistics shown in Table 1 for SOE monitoring sites located in the Southland region. *E. coli* had been measured at each site on a monthly basis for the five-year period ending 2017 (Figure 4). The statistic values were calculated from the monitoring data for each site. The site values of each statistic were used as response variables in four regression models (one for each statistic) that were based on several similar national and regional studies (e.g., Whitehead, 2018) and the studies on which the current analysis was based (MFE, 2019; Snelder, 2020; Snelder *et al.*, 2020).

For each *E.coli* statistic (i.e., Median, Q95 G260, G540), a random forest (RF) regression model was fitted to the observed monitoring site values using predictor variables that describe various aspects of each site's catchment including the climate, geology and land cover. In addition, this study included five predictors that quantified the density of pastoral livestock in 2017 to indicate land use intensity. These predictors were based on publicly available information describing the density of pastoral livestock (https://statisticsnz.shinyapps.io/livestock_numbers/). These predictors improve the discrimination of catchment land use intensity compared to previous studies that have only had access to descriptions of the proportion of catchment occupied by different land cover categories (e.g., Whitehead, 2018). The densities of four livestock types (dairy, beef, sheep and deer) in each catchment were standardised using 'stock unit (SU) equivalents', which is a commonly used measure of metabolic demand by New Zealand's livestock (Parker, 1998). Stock unit equivalents that were applied to dairy, beef, sheep and deer were 8, 6.9, 1.35, and 2.3, respectively. These values represent adjustments to the original equivalents of Parker (1998) to account for increasing animal size and productivity since 1998 (Ross Monaghan, AgResearch *pers comm*). These five predictors express land use intensity as the total stock units and the stock units by each of the four livestock types divided by catchment area (i.e., SU ha^{-1}).

The site Median and Q95 values were $\log_{10}$ transformed to improve model performance (Whitehead, 2018). A logit transformation was applied before fitting the model for G260 and G540 values. A logit transformation is defined as:

$$logit = log\left(\frac{x}{1-x}\right) \qquad \text{Equation 1}$$

where x are the site G260 and G540 values, which are in the range 0 to 1. The logit transformed values range between $-\infty$ and $+\infty$.

In a previous study, Snelder (2018) showed that transformation of the G260 and G540 statistics did not improve the performance of the RF models but did improve their ability to discriminate variation in small values of the statistics. A total of 57 river water quality monitoring sites were used to fit the models for the four statistics (Figure 4). Two sites were removed from the regional data set (Waiau River at Sunnyside and Waiau River at Tuatapere) because the catchment characteristics described by the digital river network represent the unmodified catchment and are not representative of the source of water given diversion of the majority of the flow to Doubtful/Patea Sound by the Manapōuri Power Scheme.

The fitted models were combined with a database of predictor variables for every network segment in the region and used to predict the current (i.e., 2017) values of the four statistics for all segments. Model predictions were back-transformed and, in the case of the log transformed statistics (Median and Q95) corrected for re-transformation bias as described by Duan (1983).

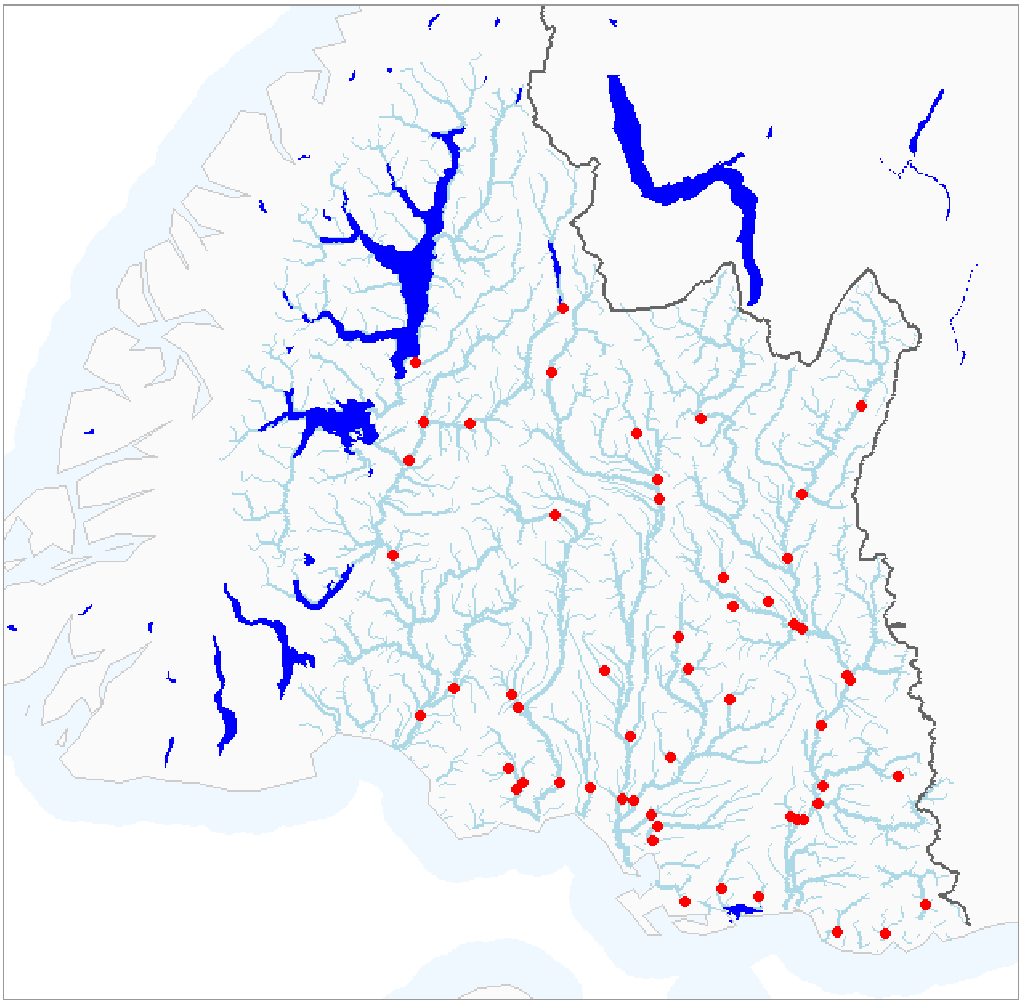

Figure 4. Locations of the 57 river SOE monitoring sites used to fit the E. coli concentration models.

The performance of the models and the uncertainty of the predictions were evaluated using three measures: regression R^2, Nash-Sutcliffe efficiency (NSE), bias. The regression R^2 value is the coefficient of determination derived from a regression of the observations against the predictions. The R^2 value indicates the proportion of the total variance explained by the model but is not a complete description of model performance (Piñeiro *et al.*, 2008). NSE indicates how closely the observations coincide with predictions (Nash and Sutcliffe, 1970). NSE values range from $-\infty$ to 1. An NSE of 1 corresponds to a perfect match between predictions and the observations. An NSE of 0 indicates the model is only as accurate as the mean of the observed data, and values less than 0 indicate the model predictions are less accurate than using the mean of the observed data. Bias measures the average tendency of the predicted values to be larger or smaller than the observed values. Optimal bias is zero, positive values indicate underestimation bias and negative values indicate overestimation bias (Piñeiro *et al.*, 2008). The normalization associated with R^2 and NSE allows the performance of the models of the four *E. coli* statistics to be directly compared. Model predictions were evaluated against two performance measures (R^2 and NSE) following the criteria proposed by Moriasi et al. (2015), outlined in Table 2.

Model uncertainty was quantified by the root mean square deviation (RMSD). RMSD is the mean deviation of the predicted values from their corresponding observations and is therefore a measure of the characteristic model uncertainty (Piñeiro *et al.*, 2008).

Table 2: Performance ratings for statistics used in this study. The performance ratings are from Moriasi et al. *(2015).*

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

For the statistics Median and Q95, model predictions require back transformation from the original $\log_{10}$ space to the original units (*E. coli* 100 mL^{-1}) using Equation 2.

$$Prediction = CF \times 10^{[log_{10}(x) - bias]} \qquad \text{Equation 2}$$

where x represents the untransformed prediction (in $\log_{10}$ space) from the model and CF is a factor to correct for retransformation bias (Duan, 1983).

For the statistics G260 and G540 the model predictions require back transformation from the original logit space to the using Equation 3.

$$Prediction = \frac{e^{x-bias}}{1+e^{x-bias}} \qquad \text{Equation 3}$$

where x represents the untransformed prediction (in logit space) from the model.

2.4 Estimated current river *E. coli* loads

Estimates of current loads of *E. coli* for all segments of the drainage network were made using river water quality monitoring data from river water quality SOE monitoring sites in the Southland region and statistical regression modelling in two steps. The first step calculated loads of *E. coli* for each SOE monitoring site using the methods described in Appendix A.

Loads were calculated for 54 sites that had at least 96 monthly concentration observations (80% of months) over the 10 years up to the end of 2017. Load calculations were based on mean daily flows that were provided for each monitoring site by Environment Southland. The load calculation method estimated the mean annual load and accounted for trends in the concentration data so that the final load estimates pertain to 2017[1]. The loads were expressed as yields by dividing by the area of the catchment upstream of each monitoring site (*E. coli* ha^{-1} yr^{-1}).

The second step used the same statistical regression modelling approach as for concentrations to fit random forest models to calculated monitoring site yields. The site yield values were $\log_{10}$ transformed to improve model performance (Snelder *et al.,* 2018). A total of 52 river water quality monitoring sites were used to fit the models (Figure 4). Two sites were removed from the regional data set (Waiau River at Sunnyside and Waiau River at Tuatapere) because the flows at these locations are regulated by Manapōuri Power Scheme and therefore, the catchment characteristics described by the digital river network are not representative of the source of water.

The fitted models were combined with a database of predictor variables for every network segment in the region and used to predict current yields of *E. coli* for all segments. Model predictions were back-transformed using Equation 2, which includes correcting for re-transformation bias based on the method of Duan (1983). The load model predictions were evaluated following the same criteria used for the concentration predictions (Table 2).

2.4.1 Loads and concentrations in the Waiau River main stem

Load predictions made for the Waiau River main stem by the RF models were not representative of actual loads because the predictor variables represent the characteristics of the unmodified catchment whereas a large proportion of the natural flow is diverted to Doubtful/Patea Sound by the Manapōuri Power Scheme. The predicted loads for the main stem of the Waiau River downstream of the Mararoa Weir control structure were discarded and replaced as follows.

The natural mean flow at the Mararoa Weir control structure is estimated to be 455 m^3 s^{-1} (Booker and Woods, 2014), whereas the measured mean flow at this location is 67 m^3 s^{-1} (i.e., 15% of the natural flow). The *E. coli* load for the segment representing the location of the Mararoa Weir control structure was therefore set to 15% of that predicted by the RF model. The *E. coli* load for each segment downstream of the Mararoa Weir was calculated by adding the load estimated for the upstream segment (starting with the segment representing the location of the Mararoa Weir location) to the predicted load for the local tributary joining the main stem of the Waiau River at that segment. The Waiau River main stem loads that were estimated as described were verified at two locations (Waiau River at Sunnyside, Waiau River at Tuatapere) by comparing the estimates with the loads calculated for these sites from actual concentration and flow records.

The *E. coli* concentration predictions made for the Waiau River main stem by the random forest models were also not representative of actual concentrations due to the modification to the natural flow regime by the Manapōuri Power Scheme. The predicted values of the four *E. coli* concentration statistics for the main stem of the Waiau River were discarded and replaced as follows. First, the measured values of the four statistics made at Sunnyside were used to replace the modelled values downstream of the Mararoa Weir control structure to midway

[1] Note that this report refers to ‘current loads and concentrations’ because the loads and concentrations estimated for 2017 are unlikely to be appreciably or statistically significantly different to loads at the time this study was conducted (2020).

between Sunnyside and Tuatapere. Second, the measured values of the four statistics made at Tuatapere were used to replace the modelled values from midway between Sunnyside and Tuatapere to the coast.

2.5 Linear models describing *E. coli* yield as function of attribute statistics

For the 52 water quality monitoring sites that were used to model the *E. coli* loads, we fitted models describing the relationship between the *E. coli* yield (i.e., the load divided by the catchment area) and each of the four attribute statistics. These models were linear regressions with appropriate transformations applied to linearise the modelled relationships. We log (base 10) transformed both the yield and the statistic values prior to fitting the models pertaining to the Median and Q95 statistics. We log (base 10) transformed just the yield values for the models pertaining to the G260 and G540 statistics.

The uncertainties of these models were evaluated using a leave-one-out cross validation process to obtain a set of independent predictions of the yields at each site. These independent predictions were then combined with the observed yields and the model statistics shown in Table 2 were used to describe the performance of the four separate models. Note that because the linear models were fitted to the $\log_{10}$ transformed values of the yield, the outputs obtained from the models were back-transformed (by raising to the power of 10) and corrected for re-transformation bias as described by Duan (1983).

2.6 Current state and compliance

The definition of FWOs for protection for human health in rivers across the Southland region is described by Norton and Wilson (2019). The *E. coli* FWOs represent levels of protection for human health in rivers across the Southland region that vary by management class as discussed in Section 2.8. The FWOs are expressed using the A to E grading system shown in Table 1 for each *E. coli* statistic and for every segment of the digital river network. This section describes how the predicted current values of the four *E. coli* statistics were used at step 3 of the analysis (Figure 1) to assess the current attribute state and compliance with FWOs, for each segment of the river network.

The current attribute state and compliance were assessed for each river segment in three steps. First, based on its FWO, each segment was assigned a criterion for each statistic based on the criteria in Table 1. Second, each segment was assigned a current attribute state based on the statistic that produced the lowest attribute state. For example, if the Median, Q95 and G540 were assigned to the B state but the G260 was assigned to the C state, the attribute state of the segment was assessed as C. Third, the predicted current values of each of the four *E. coli* statistics were compared to their corresponding criteria. If all four current *E. coli* statistics were less than their corresponding target values, the segment was compliant, otherwise it was considered noncompliant.

The predicted current values of the four *E. coli* statistics were also used to determine the proportion of large rivers that are currently suitable for primary contact. This assessment was performed by calculating the proportion of segments with stream order of ≥4 for which the predicted current attribute state was A, B or C. It is noted that the proportion of segments that are unsuitable is the complement of the proportion that are suitable (i.e., 1 – proportion suitable). The proportion suitable for primary contact was calculated for both the study area of focus for this report (i.e., all of Southland except Fiordland and the Offshore Islands), but also for the whole Southland region. This was done to enable comparison of the results for the whole Southland region with a previously reported estimate of the current proportion of large rivers suitable for primary contact in Southland (i.e., 62%; MFE 2018) and also to compare with the national targets laid out in the NPSFM (i.e., 80% by 2030 and 90% no later than 2040).

2.7 Load reductions required and critical catchments

The *E. coli* load reduction required to bring all segments into a compliant state was calculated in three steps (steps 3, 4 and 5; Figure 1). At step 3, for all noncompliant segments and each *E. coli* statistic, the *E. coli* yield corresponding to the criteria was estimated using the linear models describing the *E. coli* yield as a function of the four attribute statistics (section 2.5). Then, for each segment, the largest percentage reduction across all non-compliant *E. coli* statistics was found. The maximum allowable load (MAL) was evaluated as this largest percentage reduction applied to the predicted current *E. coli* load (i.e., predicted using the random forest model). The local excess load is then evaluated as the current load minus the MAL. For example, if the segment FWO were the A attribute state, the Q50 criteria would be 130 *E. coli* 100ml^{-1} (Table 1). The *E. coli* yield corresponding to a Q50 of 130 *E. coli* 100ml^{-1} would be estimated using the linear model to be 93 giga *E. coli* ha^{-1} yr^{-1} (see Figure 8). Then, the *E. coli* yield corresponding to the predicted current value of the statistic would be estimated from the linear models. If the predicted current Q50 value was 250 *E. coli* 100ml^{-1}, the corresponding *E. coli* yield would be estimated from the linear model as 163 giga *E. coli* ha^{-1} yr^{-1} (see Figure 8). The difference between the estimated current yield and the estimated yield to achieve the criteria can then be expressed as a percentage reduction, i.e., 42% ([163-93]/163). If the predicted current *E. coli* load were 190 giga *E. coli* ha^{-1} yr^{-1} the MAL would be evaluated as 110 giga *E. coli* ha^{-1} yr^{-1} (i.e., [100 – 42%] x 190). The local excess load would be evaluated as 190-110 = 80 giga *E. coli* ha^{-1} yr^{-1}.

At step 4 (Figure 1), the load reduction required was calculated by traversing the digital drainage network in the downstream direction. At all upstream-most segments, the load reduction required was defined to be the local excess load. Then, beginning at the upstream-most segments, the load reduction required was compared to the local excess load of the next segment downstream. If the local excess load at the next downstream segment was less than the load reduction required of the upstream segment, the load reduction required of the downstream segment was updated to be the load reduction required of the upstream segment. If the reverse applied, the local excess load of the downstream segment was updated to be its load reduction required. The load reduction required therefore took a positive value (*E. coli* yr^{-1}) at any segment in the catchment for which there was a local excess load at that, or any upstream, segment. Summaries of the load reductions required as mass per year (*E. coli* yr^{-1}) were produced for reporting catchments (Figure 3) and the whole study area. These summaries were evaluated by summing excess loads over all terminal segments (i.e., network of segments intersecting the coastline) of the summary area.

Finally, at step 5 (Figure 1), critical points and catchments were identified as follows. The terminal segment of every sea-draining catchment (the river mouth) was defined as a critical point and the ratio of the current load to MAL at that point was noted. A ratio of the current load to MAL greater than one indicates non-compliance and the larger this value is, the greater the extent to which the current load exceeds the MAL. From the terminal segment, the ratio of the current load to MAL at successive upstream river segments were obtained. At each segment, the ratio was compared with the ratio for the downstream critical point. If the ratio of the current load to MAL at the segment was greater than that of the downstream critical point, the segment was defined as a critical point and the load reduction required for the catchment upstream of this point is the local excess load of this segment. If the ratio of the current load to MAL at the segment is less than that of the downstream critical point the critical point is unchanged. The process continues upstream to the catchment headwaters. Maps indicating the load reductions required were produced for critical point catchments as yields by dividing by the upstream catchment area (*E. coli* ha^{-1} yr^{-1}) or as proportions of the current load (%). More details of the process of defining critical points and catchments are provided by Snelder *et al.* (2020)[2]. To map the critical catchment load reduction required, the quantity is expressed as a percentage of current *E. coli* load or as a yield (i.e., number of *E. coli* organisms per catchment area; *E. coli* ha^{-1} yr^{-1}).

2.8 Estimation of uncertainties

Our analysis was based on nine statistical models (i.e., random forest models to predict current values of the four *E. coli* statistics, random forest models to predict the current *E. coli* yield, and four linear regression models describing *E. coli* yield as function of four *E. coli* statistics). These models were all associated with uncertainties. The uncertainty of each model was quantified by its RMSD values (Table 2). These random errors propagate to all the assessments produced in this study including the assessments of current state and compliance, and the assessment of the load reduction required.

We inspected the residual errors for each of the models. There was no apparent geographic pattern in these errors and the pattern of errors was not explained by catchment characteristics. Because all models were derived from data pertaining to the same 52 sites that were common to all models, we expected that the residual errors from each model would be correlated to a degree with the errors of the other six models. We used the correlation matrix derived from the nine sets of model errors to describe the relationship between all pairs of model errors. We assumed that this correlation structure represents the correlation in the uncertainties when the models were combined in the assessment process.

We applied the same simple Monte Carlo analysis approach as Snelder *et al.* (2020) to estimate uncertainties in our assessments based on 100 'realisations' of our calculations in four steps. First, for a realisation (r), predictions made with all models were perturbed by a random error. Random errors were obtained by generating random normal deviates (ε_r) and applying these to predictions made using the models. When the response variables in the models were log (base 10) transformed the perturbed prediction for a realisation was derived as follows.

$$Prediction_r = CF \times 10^{[log_{10}(x) - bias + \varepsilon_r \times RMSD]} \qquad \text{Equation 4}$$

When the response variables in the models were logit transformed (i.e., the models of current values of G260 and G540) the perturbed prediction for a realisation was derived as follows.

$$Prediction_r = \frac{e^{x - bias + \varepsilon_r \times RMSD}}{(1 + e^{x - bias + \varepsilon_r \times RMSD})} \qquad \text{Equation 5}$$

Random normal deviates representing errors for each model (ε_r) were drawn from a multi-variate distribution with the same correlation structure as that between the observed errors. Because a concentration or load at any point in a catchment is spatially dependent on corresponding values at all other points in the catchment's drainage network, the values of the random normal deviates were held constant for each realisation within the river network

[2] Snelder et al. (2020) based the identification of critical points on excess loads, which were expressed as the ratio of the current load to the maximum allowable load.

representing a sea-draining catchment but differed randomly between sea-draining catchments.

At the second step, for each realisation we stored the perturbed predicted values of the four *E. coli* statistics, current load and load reduction required. At the third step, we repeated the procedure described above for each realisation. At the fourth step, we used the distribution of values of the four *E. coli* statistics, current load and load reduction required obtained from the 100 realisations to provide a best estimate and the uncertainty of the assessments. The uncertainty of the assessments of compliance and whether the segment was suitable for contact recreation were quantified by estimating the probability that each segment was compliant or suitable across the 100 realisations. Segment compliance and suitability for contact recreation was therefore assessed as a value between one (100% confident the segment is compliant or suitable) to zero (100% confident the segment is non-compliant or not suitable). For the current state and load reduction required assessments, the best estimate was represented by the mean value from the distribution of values. Where the mean value was negative, the load reduction requirement was taken to be zero. The uncertainty of these two assessments was quantified by their 90% confidence intervals. For the load reduction required assessment, the best estimates and the uncertainties were estimated from the 100 realisations for the reporting catchments and the entire region.

2.9 Freshwater objectives settings

2.9.1 Classification systems

The individual segments of the river network were assigned to one of Environment Southland's seven river management classes (Figure 5). The analysis allowed FWOs to be independently set for each river management class.

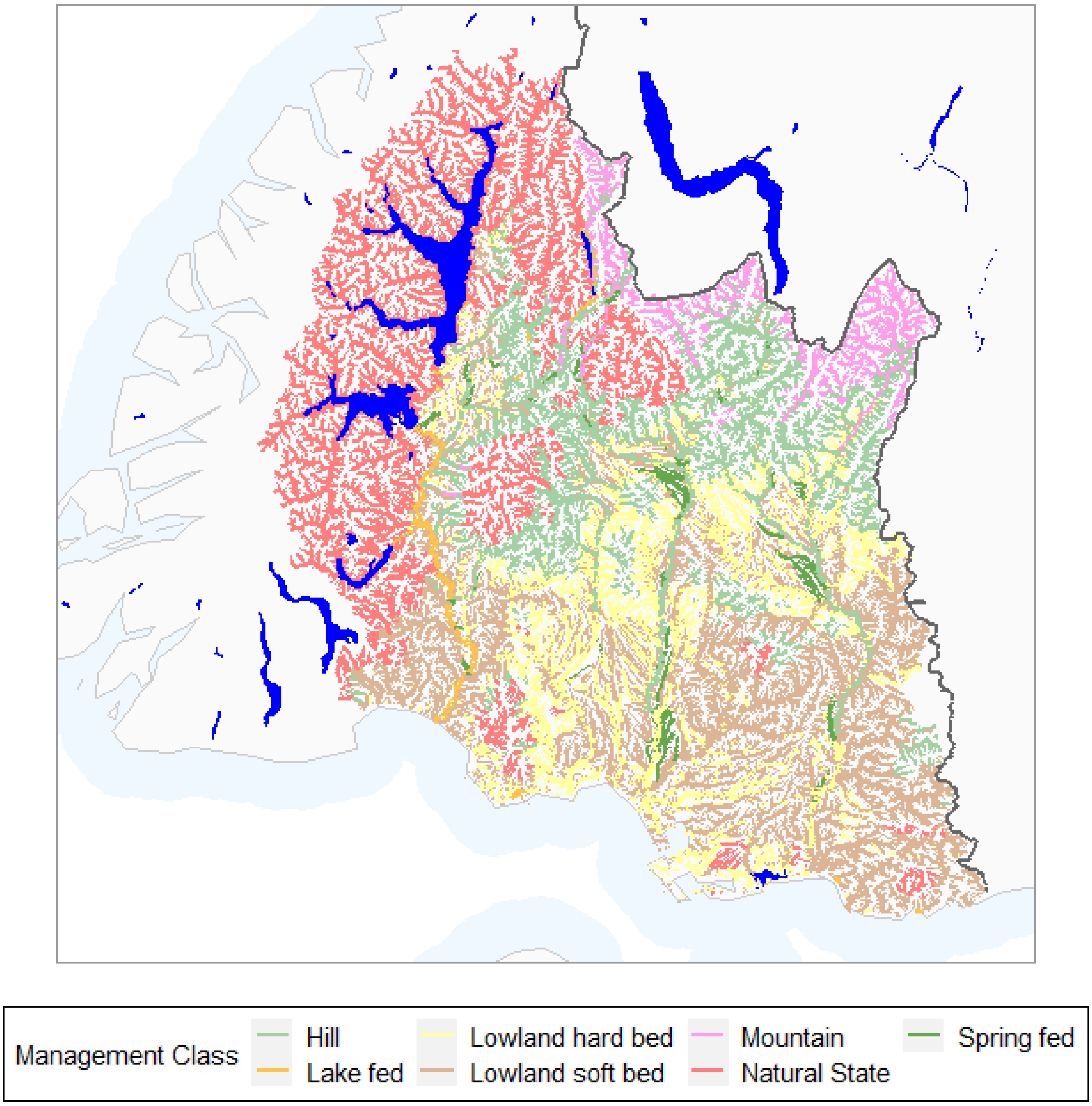

Figure 5. Map of the distribution of Environment Southland's river management classes.

2.9.2 Objectives

To proceed with the analysis, it is necessary to nominate FWOs in terms of a band (A, B, C, D or E) for all river receiving environments (represented by network segments) in the study area. Because FWOs could be set for each river management class individually, there were many potential combinations of FWO by management class. For this reason, and to make the analysis and presentation of results manageable, it is useful to nominate the specific set, or sets, of FWOs for which load reduction estimates are desired.

In November 2020, Environment Southland and the Te Ao Marama Inc (TAMI) board approved a directive to the Regional Forum to test the implications of achieving "a state of hauora (well-being) for water across the region within a generation (25-30 years)"[3]. The 'approval in principle' sets the minimum level of hauora for the draft FWOs for the region. However, because the setting of FWOs is an iterative process that must include consideration of implications, the analysis here took an 'envelope' approach by generating load reduction results for two FWO options. The definition of the minimum level of hauora was based on hauora principles as reported in Bartlett *et al.* (2020) and is represented by the minimum state (in A, B or C terms) necessary to support hauora for human health in rivers. This is referred to hereafter as the Hauora option (Table 3). Another option was defined using a method based primarily on the Proposed Southland Water and Land Plan water quality standards, while also ensuring the NPS-FM minimum requirement of "maintain or improve" was satisfied, hereafter referred to as the "pSWLP option" (Table 3). The derivation of both the Hauora and the pSWLP options are summarised in Bartlett *et al.* (2020) and further detail is provided in Norton and Wilson (2019). These two sets of objectives represent scenarios (i.e., possible future states). It is possible to re-run the analyses and produce load reduction estimates for scenarios comprising any combination of FWOs.

For the analyses that follow, the FWO band and the linked criteria for the four *E. coli* statistics in Table 1 are the basis for the analysis of compliance and load reductions required. The FWOs for rivers belonging to the Natural State management class were assumed to be no change in current state for both the Hauora and pSWLP options. This means there are no bands associated with FWOs for rivers in the Natural State management class. Therefore, there is no assessment of compliance and no load reduction required for any river segments belonging to the Natural State management class. This is appropriate from a practical perspective because there is little to no resource use, and therefore no potential for load reductions, in the catchments of rivers belonging to the Natural State management class.

Table 3. Nominated bands defining the 'Hauora' and 'pSWLP' options for all river management classes. The NA bands for the river Natural State management classes indicate the objective is no change in current state.

Management class	**FWO options**	
	Hauora	**pSWLP**
Natural State	NA	NA
Lowland soft bed	A	B
Lowland hard bed	A	B
Hill	A	B
Mountain	A	A
Lake fed	A	A
Spring fed	A	B

[3] https://www.es.govt.nz/repository/libraries/id:26gi9ayo517q9stt81sd/hierarchy/about-us/meetings/2020/November/Strategy%20and%20Policy%20Committee%20-%2025%20November%202020/Strategy%20and%20Policy%20Committee%20Agenda%20-%202020%20November%2025.pdf

3 Results

3.1 Performance of random forest *E. coli* statistics models

The random forest models of the four *E. coli* statistics had at least satisfactory performance (Table 4), as indicated by the criteria of Moriasi et al. (2015; Table 2). The mapped predictions for all four statistics had similar coarse-scale spatial patterns, with relatively high values in low-elevation areas on the Southland plains and low values in high elevation areas (Figure 6). These patterns were consistent with expectations and reflect the influence of increasing proportions of catchments occupied by agricultural and other land uses.

Table 4. Performance of random forest models of the four E. coli statistics; Median, Q95, G260 and G540. Median and Q95E. coli ha^{-1} yr^{-1}. G260 and G540The overall performance rating is based on the criteria of Moriasi et al. (2015) shown in Table 2.

Statistic	N	R^2	NSE	BIAS	RMSD	Performance rating
Median	57	0.55	0.54	-0.04	0.37	Satisfactory
Q95	57	0.61	0.60	-0.01	0.37	Good
G260	57	0.56	0.55	-0.09	0.91	Satisfactory
G540	57	0.60	0.59	-0.07	0.74	Good

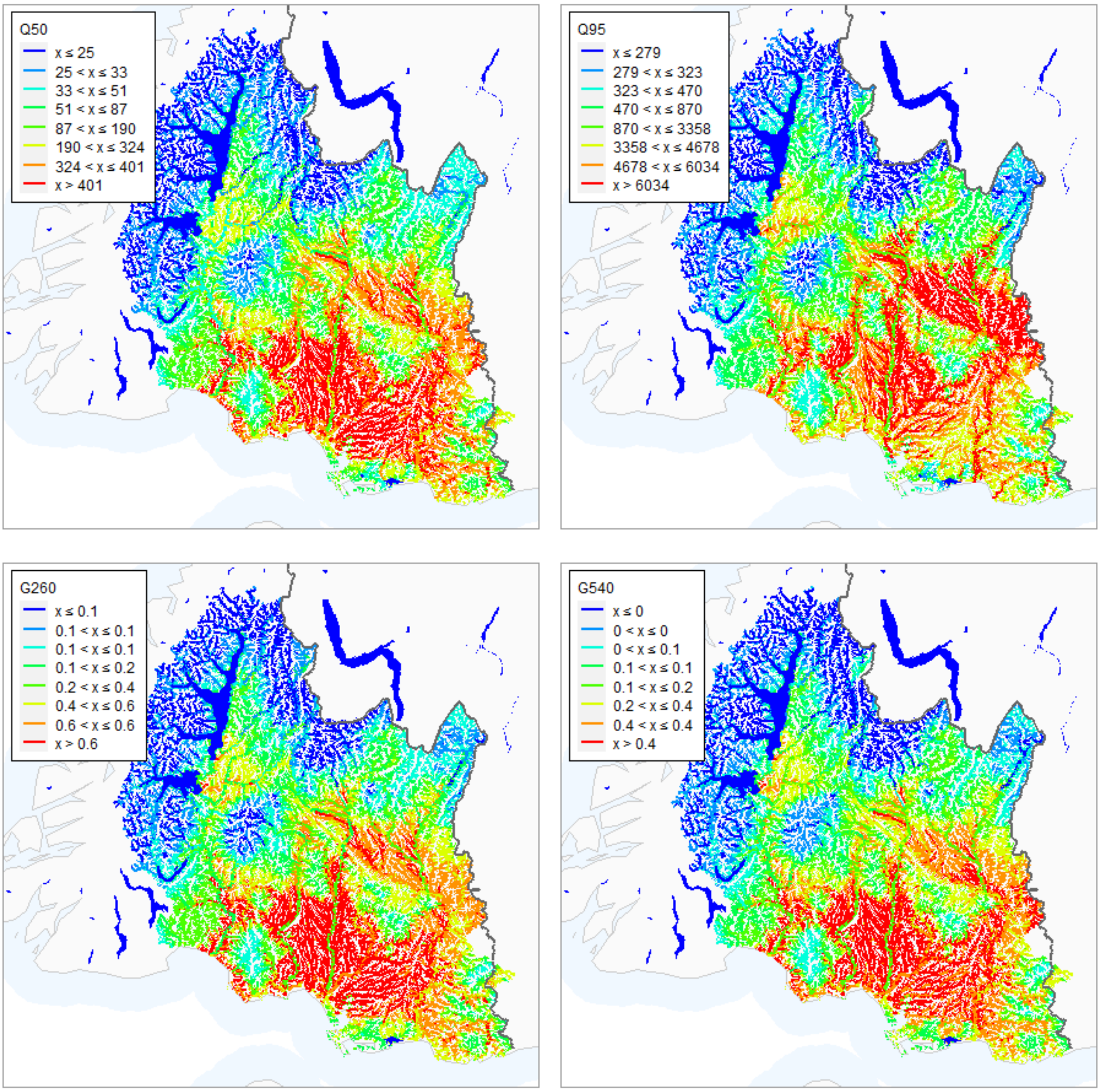

Figure 6. Predicted patterns of the current value of the four E. coli statistics. Note that the breakpoints shown in the map legend are nominal and have no special significance (i.e. are not guidelines or standards).

3.2 Performance of the random forest *E. coli* current yield model

The random forest models of *E. coli* annual yield had satisfactory performance (Table 5), as indicated by the criteria of Moriasi et al. (2015; Table 2). The mapped predictions of annual yield of *E. coli* had similar coarse-scale spatial patterns as the *E. coli* statistics with relatively high values in low-elevation areas on the Southland plains and low values in high elevation areas (Figure 7). These patterns were consistent with expectations and reflect the increasing concentration of *E. coli* in association with increasing proportions of catchments occupied by agricultural and other land uses.

Table 5. Performance of the random forest models of E. coli annual yield.E. coli ha^{-1} yr^{-1}. The overall performance rating is based on the criteria of Moriasi et al. (2015) shown in Table 2.

Variable	N	R^2	NSE	BIAS	RMSD	Performance rating
E. coli yield	52	0.48	0.46	-0.03	0.43	Satisfactory

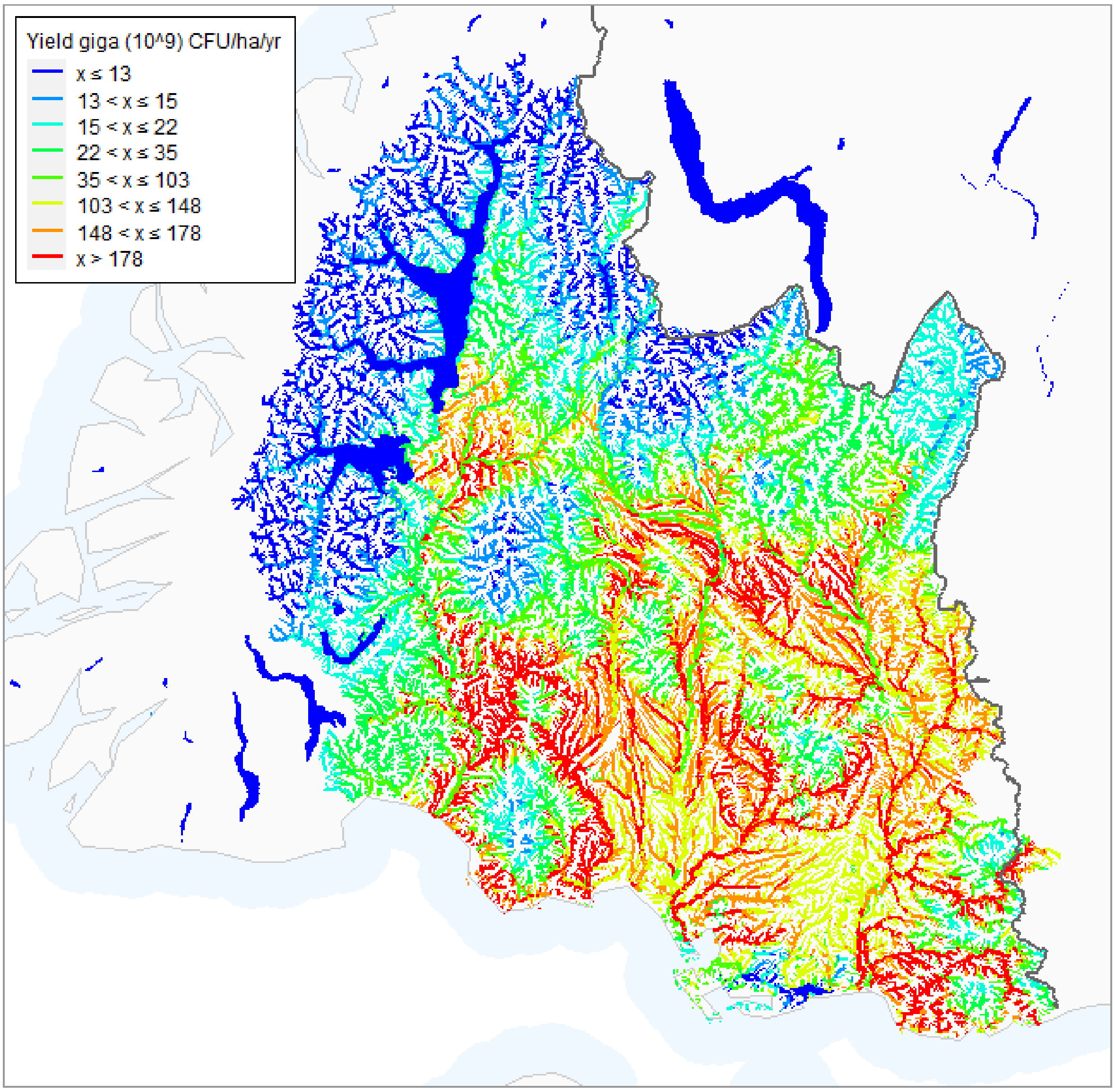

Figure 7. Predicted patterns of the current E. coli loads (as yields giga (10^9) E. coli ha^{-1} yr^{-1}). Note that the breakpoints shown in the map legend are nominal and have no special significance (i.e., are not guidelines or standards).

The estimated loads for two monitoring stations on the main stem of the Waiau River (Waiau River at Sunnyside, Waiau River at Tuatapere) were close to the 'observed loads' (i.e., loads calculated from the observed concentrations and flows at the sites). The observed yield at Sunnyside was 10.4 giga *E. coli* ha^{-1} yr^{-1} compared and estimated yield of 10.2 giga *E. coli* ha^{-1} yr^{-1}. The observed yield at Tuatapere was 28.6 giga *E. coli* ha^{-1} yr^{-1} compared and estimated yield of 43.3 giga *E. coli* ha^{-1} yr^{-1}. It is noted that there is considerable uncertainty associated with the estimated observed loads.

3.3 Performance of the linear models of *E. coli* yield as function of attribute statistics

With appropriate transformation, *E. coli* yield was linearly related to the four *E. coli* attribute statistics (Figure 8). The linear models had at least satisfactory performance as indicated by the criteria of Moriasi *et al.* (2015; Table 2) and low bias (Table 6).

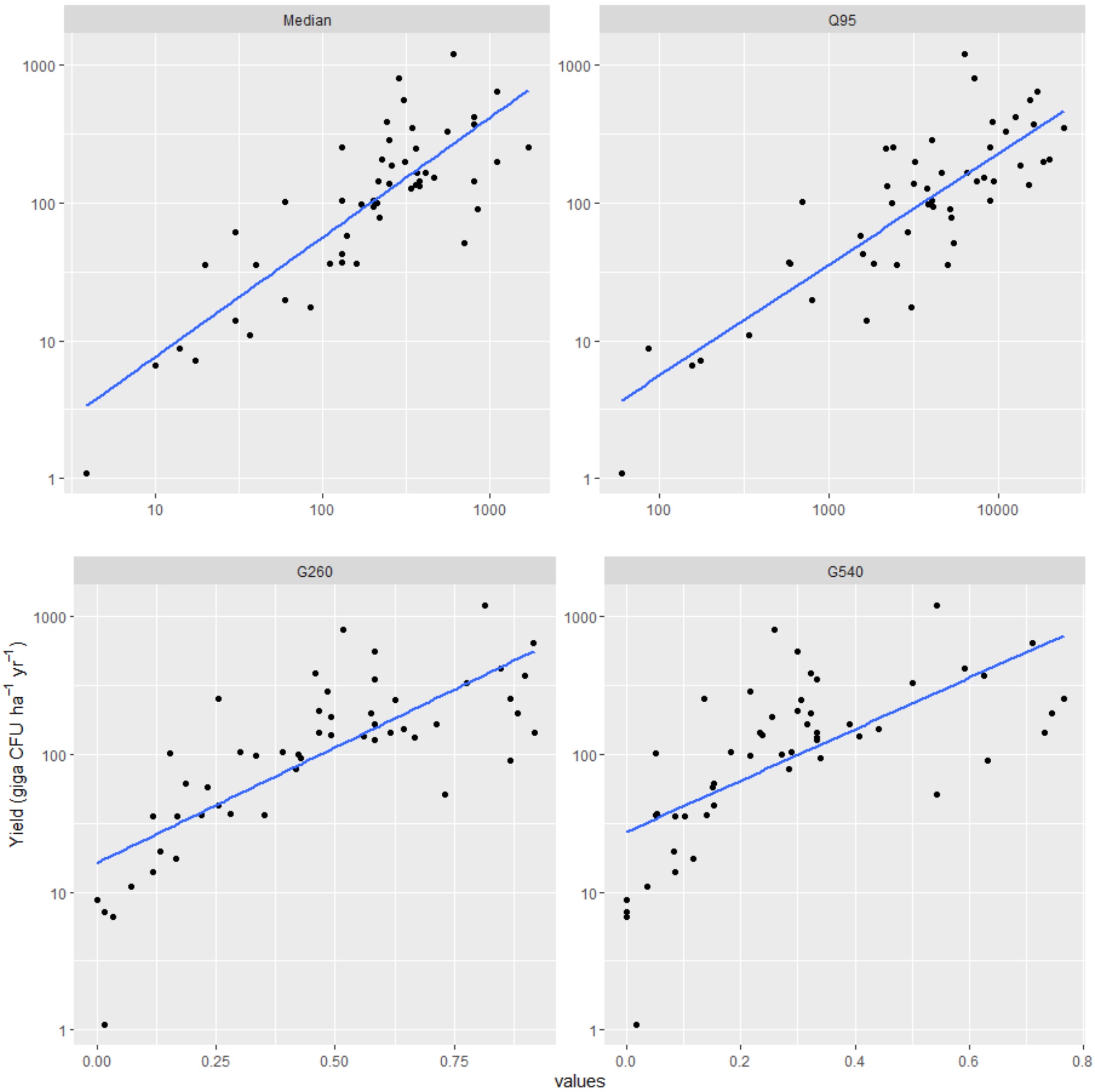

Figure 8. Linear relationships between E. coli yield and the four E. coli statistics. The black points represent the yield and E. coli statistic for the 52 sites and the blue line indicates the fitted linear regression. Note that yield was log (base 10) transformed in all models and the Median and Q95 values were log (base 10) transformed in these models, respectively.

Table 6. Performance of the linear models describing E. coli yield as function of the four E. coli attribute statistics; Median, Q95, G260 and G540. The overall performance rating is based on the criteria of Moriasi et al. (2015) shown in Table 2.

Statistic	N	R^2	NSE	BIAS	RMSD	Performance rating
Median	52	0.55	0.55	0.00	0.39	Satisfactory
Q95	52	0.38	0.38	-0.01	0.46	Satisfactory
G260	52	0.66	0.66	0.00	0.34	Good
G540	52	0.67	0.66	0.00	0.34	Good

3.4 Correlation of model errors

The model errors were strongly correlated (Pearson correlation coefficient (r) > 0.6) between all pairs of models that were used to predict current values of the *E. coli* statistics and loads (Table 7). The model errors were strongly correlated (r > 0.6) between all pairs of the linear regression models that were used to describe the relationships between the *E. coli* statistics and loads (as yields). Correlations between the RF and linear regression models were low (r < 0.4). The correlation structure shown in Table 7 was used to generate random normal deviates (ε_r) for each model in the Monte Carlo analysis.

Table 7. Correlation of errors between all pairs of models used in the analysis. The table is a lower triangular matrix showing the correlations of model errors between all pairs of models. RF indicates random forest models and LM indicates linear regression models.

Model	RF Median	RF Q95	RF G260	RF G540	RF Load	LM Median	LM Q95	LM G260
RF Q95	0.62							
RF G260	0.92	0.61						
RF G540	0.85	0.61	0.92					
RF Load	0.67	0.64	0.62	0.46				
LM Median	-0.0	0.27	-0.0	-0.2	0.67			
LM Q95	0.25	-0.1	0.2	0.03	0.64	0.57		
LM G260	0.18	0.37	0.04	-0.1	0.74	0.89	0.58	
LM G540	0.25	0.38	0.14	-0.1	0.79	0.85	0.66	0.95

3.5 Current state

The measured current state for the 59 SOE monitoring sites is shown in Table 8 in terms of attribute bands. These results indicate that poor attribute grades (i.e., D and E) occur in rivers that are in the Lowland, Spring fed and Hill management classes and that there is a dominance of SOE sites in the D and E state. In addition, Table 8 indicates that the SOE sites provide some representation of rivers in all management classes.

Table 8. Measured current state as numbers of SOE monitoring sites in each attribute band by management class.

Attribute band	Natural state	Lowland hard bed	Lowland soft bed	Hill	Lake fed	Mountain	Spring fed
A	1	0	0	2	0	2	0
B	1	1	0	2	0	0	0
C	0	0	0	0	1	0	0
D	0	5	7	8	1	0	2
E	0	11	13	2	0	0	0

The best estimate of the proportions of the study area river network segments in the five NOF attribute states are shown in Table 9. The 90% confidence intervals for these estimates are very wide. For example, the best estimate of the proportion of segments that are in the A band is 17% but the 90% confidence intervals extend from 2% to 40% (Table 9). This large uncertainty reflects the, at best good, performance of the models used to predict the *E. coli* statistics (Table 4). This uncertainty is also indicated by the maps of the estimated probability of network segments belonging to the five NOF *E. coli* attribute states (Figure 9). Over much of the network, the probability of segments being in a specific NOF attribute band was less than 50% indicating low certainty about the true state. The exceptions were segments in lowland parts of the region, which generally had >80% probability of being in the NOF E band.

Table 9. Proportion of all segments (%) predicted to be in each attribute band. Note that this information by reporting catchment and management class are tabulated in Appendix B.

Attribute band	Best estimate	5% confidence limit	95% confidence limit
A	19	2	40
B	18	6	33
C	8	3	23
D	21	13	35
E	34	19	51

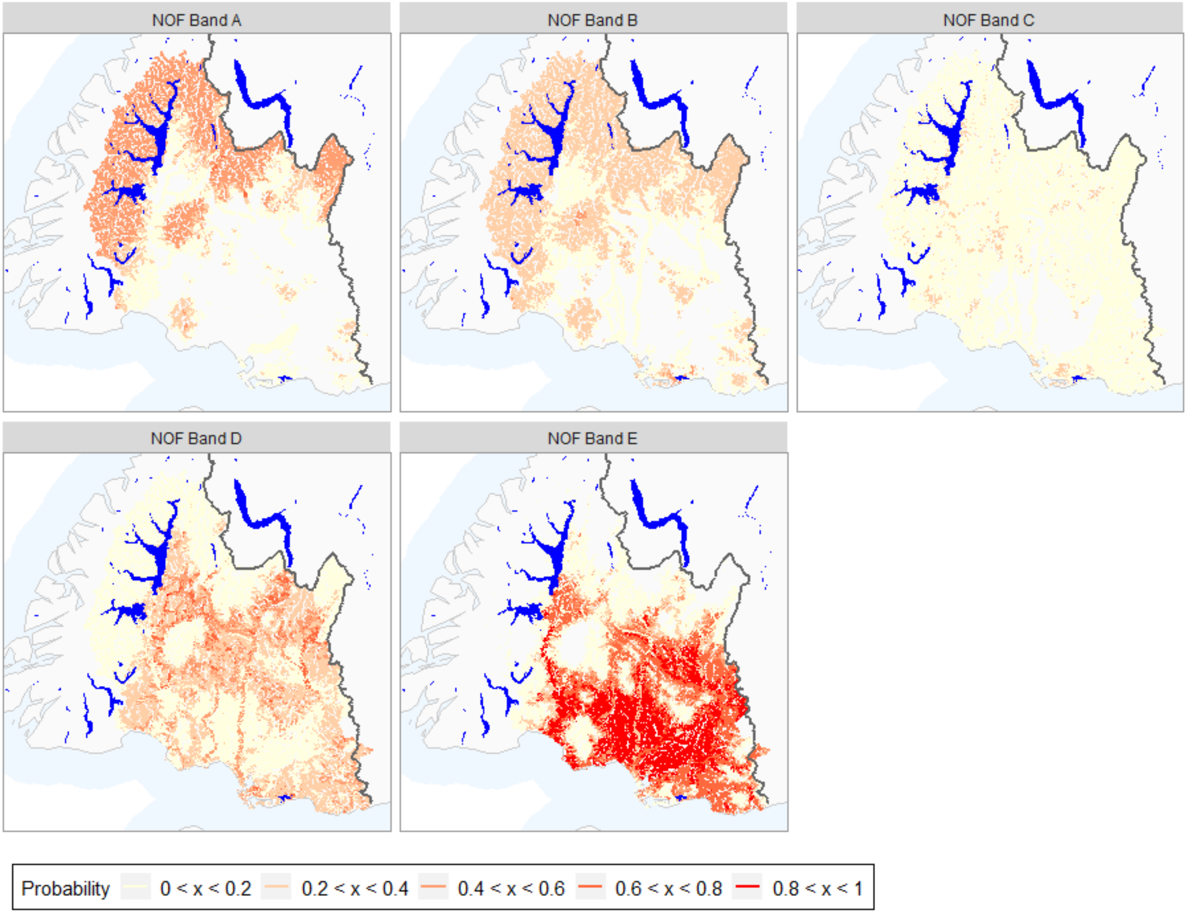

Figure 9. Estimated probability of network segments belonging to the five NOF E. coli attribute states.

The best estimate of the proportion of river segments in the study area (i.e., excluding Fiordland and the Offshore Islands) of stream order ≥4 that are currently suitable for primary contact (i.e., NOF attributes states A, B or C) is 45%. The 90% confidence interval for this estimate is wide with the lower and upper bounds defined by 31% and 54%, respectively. This uncertainty is also indicated by a map of the estimated probability that network segments are currently suitable for primary contact (Figure 10). The map indicates that segments that have high probability of being suitable for primary contact are smaller streams located in headwater catchments and the larger main-stem rivers traversing much of the lowland areas have low probability (i.e., <20%) of being suitable, which is equivalent to high probability (i.e., >80%) of being unsuitable.

For the whole Southland region (i.e., including Fiordland and the Offshore Islands) the best estimate of the proportion of river segments of stream order ≥4 that are currently suitable for primary contact (i.e., NOF attributes states A, B or C) is 61%. This is comparable to a previous estimate of 62% of these large rivers being suitable for primary contact in Southland reported in MFE (2018). These estimates of current state reported estimate of the current proportion of large rivers suitable for primary contact in Southland (i.e., 62%; MFE 2018) and also to compare with the national targets laid out in the NPSFM (i.e., 80% by 2030 and 90% no later than 2040).

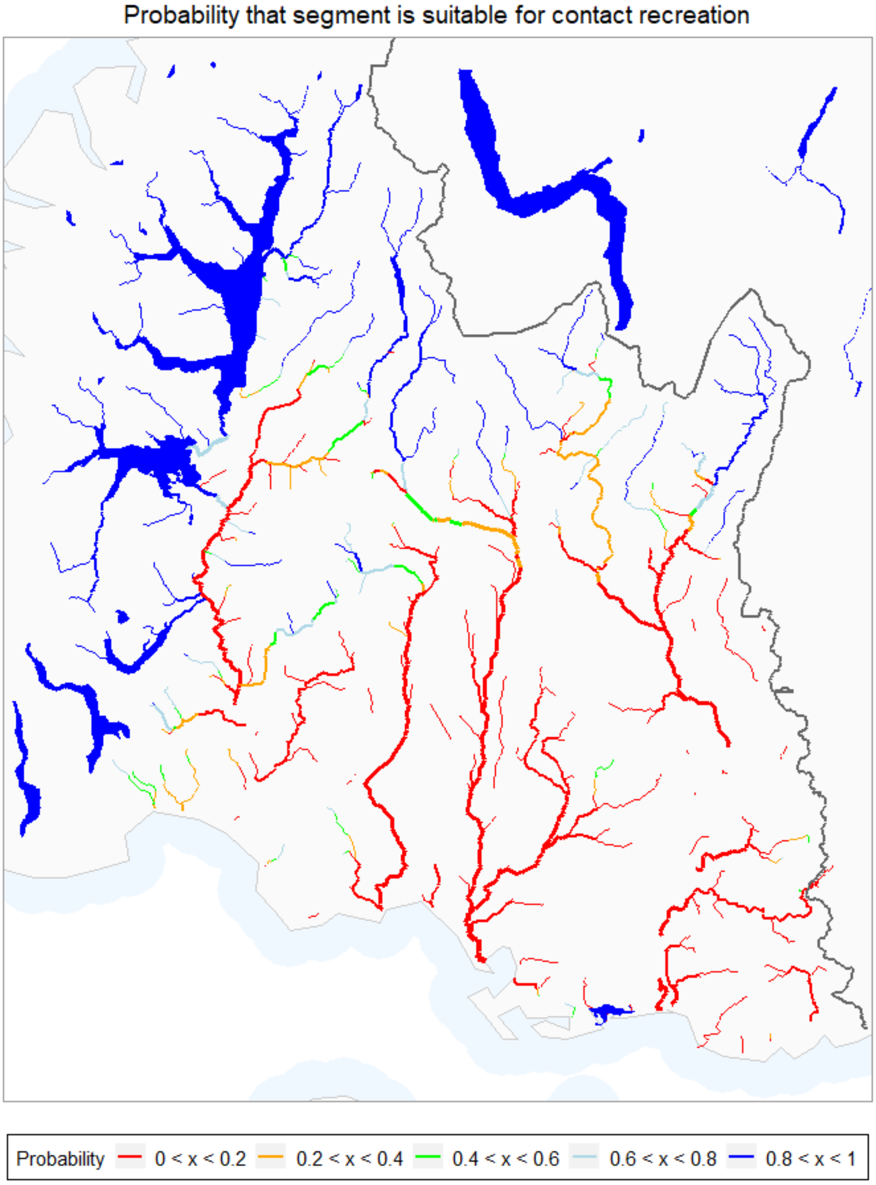

Figure 10. Estimated probability that network segments with stream order ≥ 4 are suitable for primary contact recreation.

3.6 Hauora option

3.6.1 Compliance

The estimated probability that values of the four *E. coli* statistics are compliant with the criteria defined for the Hauora FWOs settings was greater than 0.6 for 54%, 41%, 46% and 34% of segments for the Median, Q95, G260 and G540, respectively (Figure 11). The estimated probability that all statistics complied with the Hauora FWOs criteria was greater than 0.6 for 33% of segments (Figure 12). The probability of compliance was greatest for segments in the headwater areas of the individual catchments, and particularly in the higher elevation parts of the study area. The probability of compliance was lowest for segments in the low elevation parts of the region. This was consistent with the predicted pattern in the current values of all four *E. coli* statistics shown in Figure 6 and Figure 7.

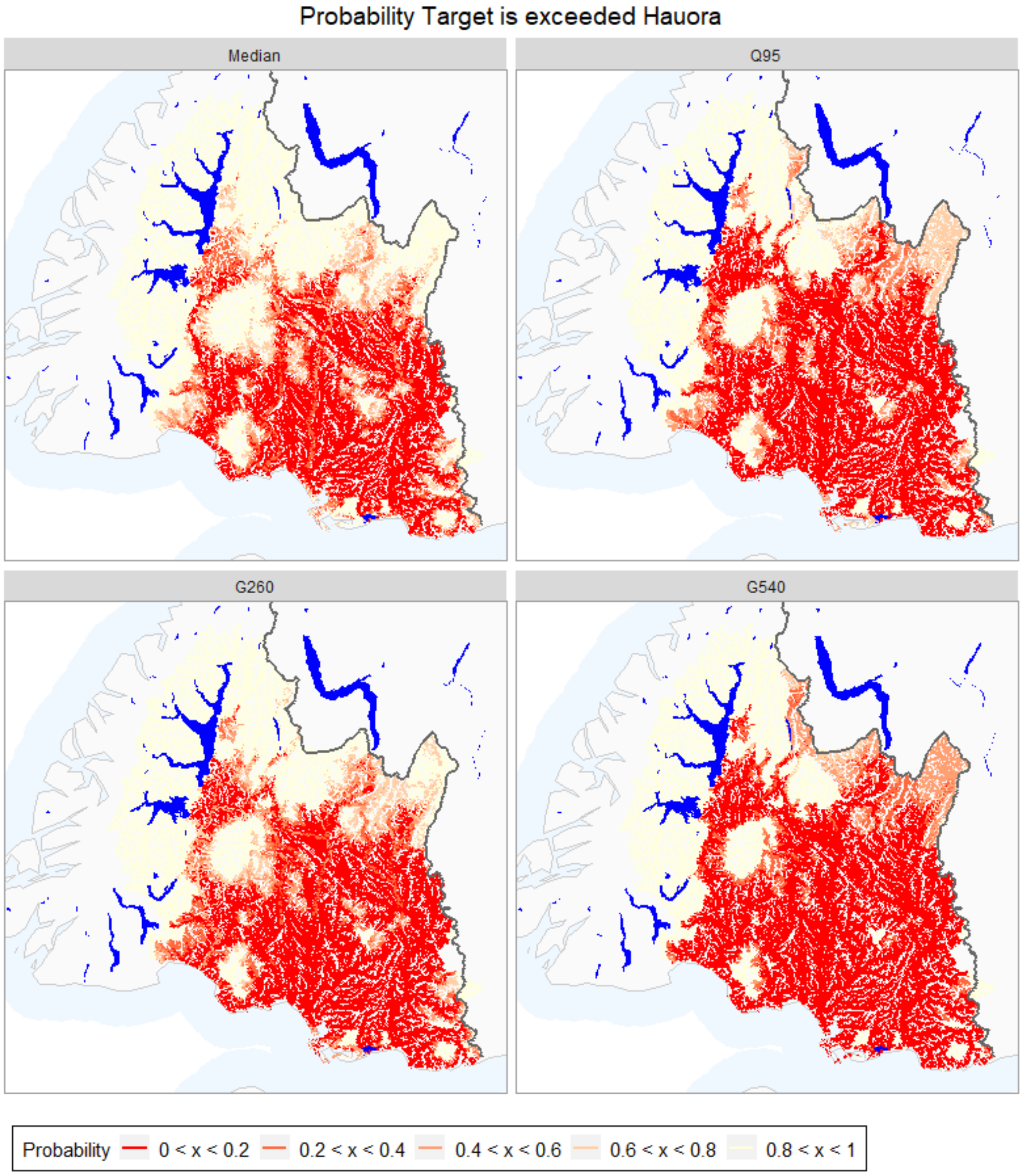

Figure 11. Probability of compliance with the criteria for each of the four E. coli statistics pertaining to the Hauora FWOs settings. Each map represents the probability that segments achieve the criteria for the E. coli statistic that is associated with the Hauora FWOs settings.

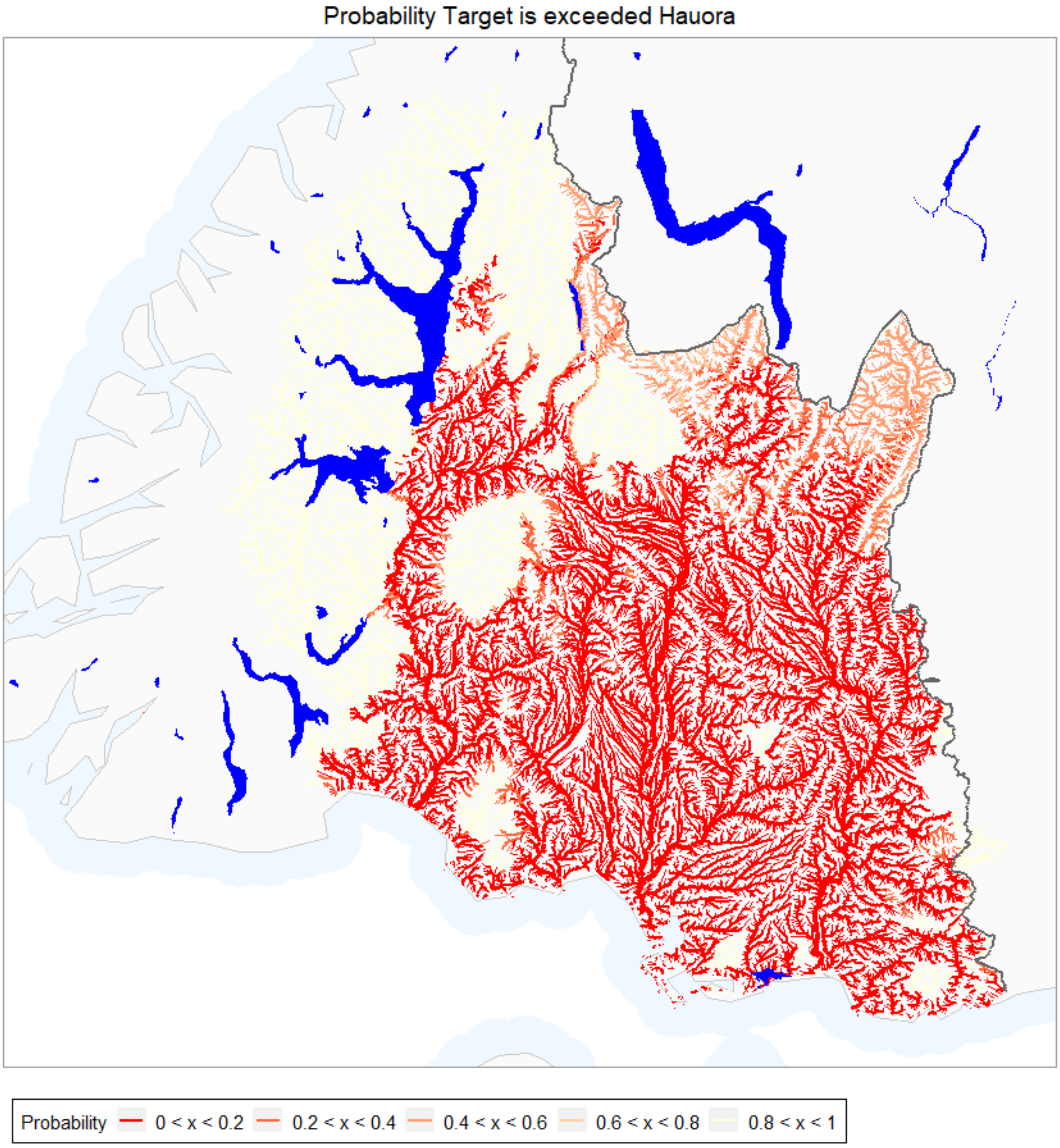

Figure 12. Probability of compliance with any of the criteria pertaining to the Hauora FWOs settings. This map represents the overall probability that segments achieve the Hauora FWOs settings.

3.6.2 Local excess loads

The local excess load is the amount by which the current *E. coli* load at a river segment would need to be reduced to achieve the FWO for that receiving environment. For the Hauora settings, the best estimate[4] of the local excess *E. coli* yield for rivers exceeded 10 giga *E. coli* ha^{-1} yr^{-1} for 93% of river segments, exceeded 50 giga *E. coli* ha^{-1} yr^{-1} for 73% of river segments and exceeded 180 giga *E. coli* ha^{-1} yr^{-1} for 49% of river segments (Figure 13). Note that the 10, 50 and 180 giga *E. coli* ha^{-1} yr^{-1} are nominal breakpoints for communication purposes and correspond to the legend thresholds on Figure 13. These values have no special significance

[4] The best estimate for each segment is the median of the estimates of the local excess loads over the 100 realisations.

(i.e., are not guidelines or standards). There were no segments that were classified other than Natural State river management class for which the local excess *E. coli* loads were zero; it is noted that 32% of river segments are in the Natural State management class and therefore have no assessment of compliance and zero load reduction required.

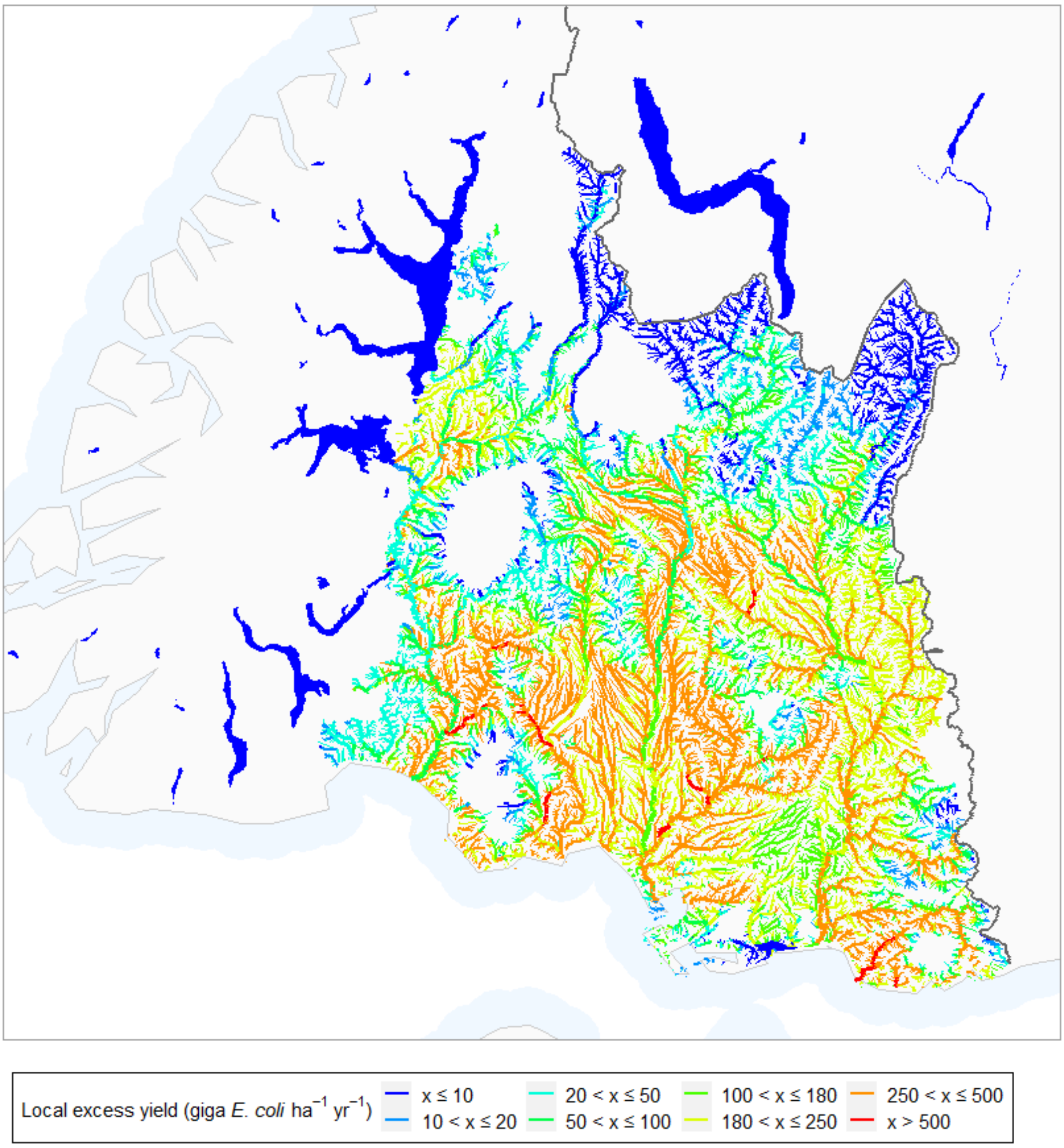

Figure 13. Local excess E. coli loads for the Hauora settings. Note that the breakpoints for the local excess yield in the map legend are nominal and have no special significance (i.e., are not guidelines or standards).

3.6.3 Critical point catchment load reductions required

The load reduction required for critical point catchments is the minimum load reduction that ensures the loads for all receiving environments in the critical catchment do not exceed the MAL (and therefore all FWOs in the catchment to be achieved). The load reductions required therefore differ from the local excess loads in that they consider all river segments in a critical point catchment. The load reduction required for the Hauora settings is expressed below as yields (i.e., *E. coli* ha^{-1} yr^{-1}) and as a percentage of the current load.

The load reductions required by the Hauora settings for critical point catchments are shown on Figure 14 and Figure 15. When critical point catchments *E. coli* load reductions required were expressed as yields, critical point catchments with reduction requirements of greater 50 giga *E. coli* ha^{-1} yr^{-1} occupied 99% of the study area (Figure 14). Critical point catchments with reduction requirements of greater 150 giga *E. coli* ha^{-1} yr^{-1} occupied 61% of the study area (Figure 14). The comparison of load reductions expressed as yields (*E. coli* ha^{-1} yr^{-1}) with those expressed as proportion of current load (%) indicates that reduction requirements in catchments with low yield reductions (e.g., much of the Waiau River catchment) are nevertheless large in relative terms. Critical point catchments with *E. coli* load reductions of greater than 50% occupied 99% of the study area and critical point catchments with *E. coli* load reductions of greater than 90% occupied 51% of the study area (Figure 15).

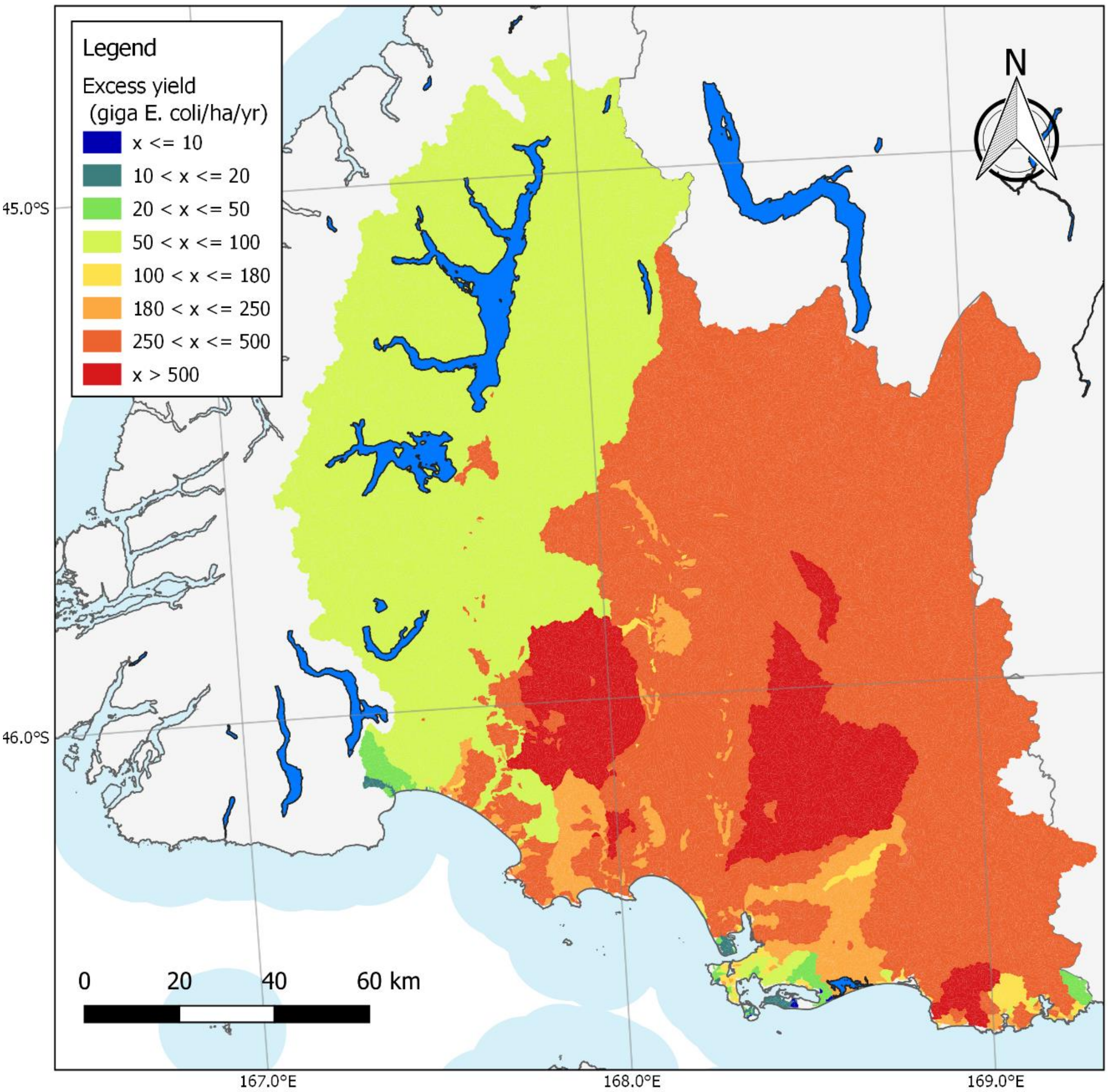

Figure 14. The E. coli load reduction required, expressed as yields, for critical point catchments and the Hauora settings. The critical point catchment colours indicate the median E. coli load reductions required to allow all FWOs be achieved in the critical point catchment (including the critical point at the bottom of the catchment).

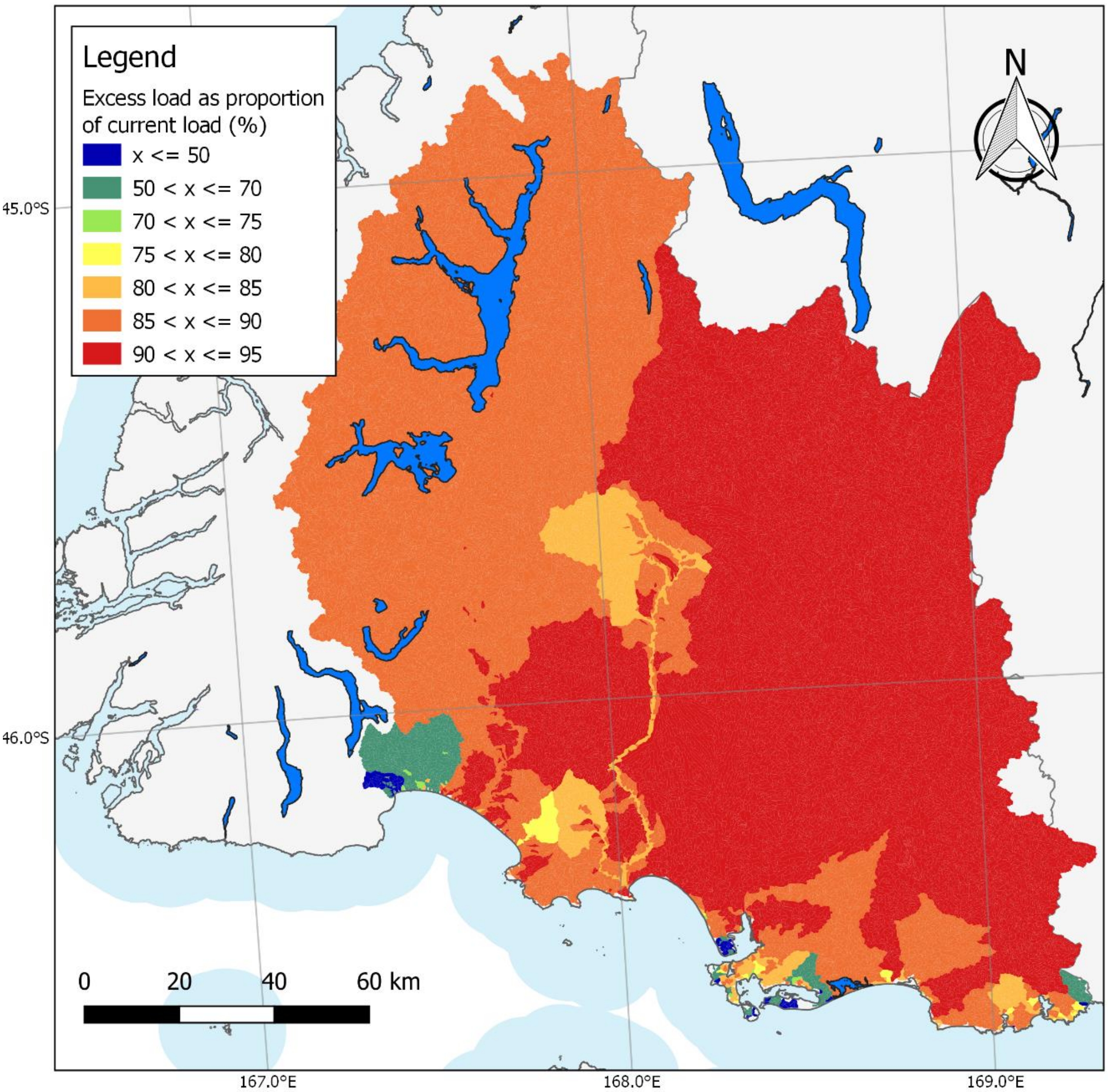

Figure 15. The E. coli load reduction required, expressed as proportion of the current load (%), for critical point catchments and the Hauora settings. The critical point catchment colours indicate the E. coli load reductions required to allow all FWOs be achieved in the critical point catchment (including the critical point at the bottom of the catchment).

3.6.4 Reporting catchments and regional load reductions required

The load reduction required by the Hauora settings for each reporting catchment and for whole study area are shown in Table 10. For the whole study area, the best estimate of *E. coli* load reductions required was 347 peta *E. coli* yr^{-1}, which represents 90% of the current load. The *E. coli* load reductions required were highest (>85%) in the Matāura, Ōreti & Invercargill Catchments, Tokanui catchments, Waiau Catchment, Waikawa Catchment and Waimatuku & Taunamau Catchments. The *E. coli* load reductions required were lowest (<70%) in the Bluff Zone, Catlins Zone, and Te Waewae Bay Western Coastal Zone.

Table 10. Current load and load reduction required for E. coli by reporting catchments and for the study area for the Hauora settings. Note that loads are expressed in absolute terms in units of E. coli organisms per year (peta E. coli yr^{-1}) and as a proportion of current load (%). The first value in each column is the best estimate, which is the median value over the 100 Monte Carlo realisations. The values in parentheses are the lower and upper bounds of the 90% confidence interval.

Reporting catchment	Total load (peta *E. coli* yr^{-1})	Load reduction required (peta *E. coli* yr^{-1})	Load reduction required (%)
Aparima & Pourakino Catchments	62 (6.6 - 195)	52 (3 - 163)	74 (38 - 93)
Bluff Zone	0.4 (0.2 - 0.8)	0.3 (0.1 - 0.6)	65 (49 - 78)
Catlins Zone	0.5 (0.2 - 1.1)	0.3 (0.1 - 0.9)	69 (46 - 86)
Matāura Catchments	164 (27.3 - 509)	152 (22.3 - 468.3)	89 (76 - 98)
Orepuki Coastal Zone	6 (3.4 - 11)	5.4 (2.6 - 9.9)	85 (77 - 90)
Ōreti & Invercargill Catchments	122 (26 - 326)	115 (19.9 - 323.9)	89 (73 - 100)
Te Waewae Bay Western Coastal Zone	2.9 (1 - 7)	2 (0.5 - 5.5)	64 (40 - 85)
Tokanui Coastal Zone	11 (3.4 - 26)	10 (2.7 - 26)	91 (79 - 102)
Waiau Catchment	60 (11 - 166)	54 (8.6 - 156)	86 (64 - 96)
Waikawa Catchment	10 (1.9 - 26)	9 (1.5 - 24.2)	86 (72 - 95)
Waimatuku & Taunamau Catchments	8.3 (1.5 - 22)	7.5 (1.1 - 21)	86 (71 - 95)
Waituna Catchments	5 (0.4 - 12)	4.3 (0.2 - 10.9)	78 (50 - 93)
Total (whole study area)	379 (209 - 928)	347 (168 - 879)	90 (80 - 96)

3.7 pSWLP option

3.7.1 Compliance

The estimated probability that values of the four *E. coli* statistics are compliant with the criteria defined by the pSWLP option was greater than 0.6 for 55%, 49%,53% and 41% of segments for the Median, Q95, G260 and G540, respectively (Figure 16). The estimated probability that any statistic exceeded the pSWLP option criteria was greater than 0.6 for 37% of segments (Figure 17). The probability of compliance was highest for segments in the headwater areas of the individual catchments, and particularly in the higher elevation parts of the study area. The probability of compliance was lowest for segments in the low elevation parts of the region. This was consistent with the predicted pattern in the current values of all four *E. coli* statistics shown in Figure 6 and Figure 7.

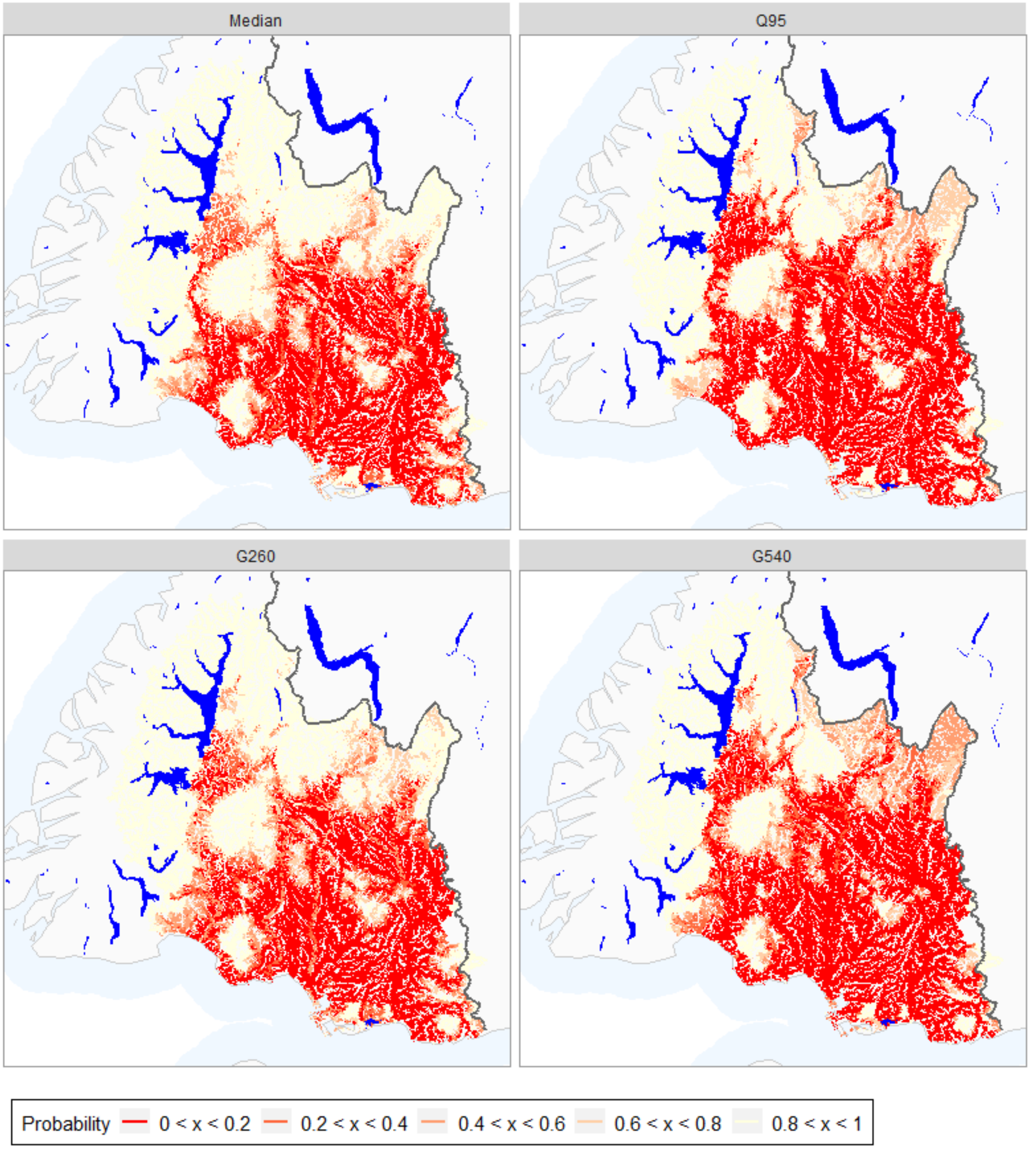

Figure 16. Probability of compliance with the criteria for each of the four E. coli statistics pertaining to the pSWLP option. Each map represents the probability that segments achieve the criteria for the E. coli statistic that is associated with the pSWLP option.

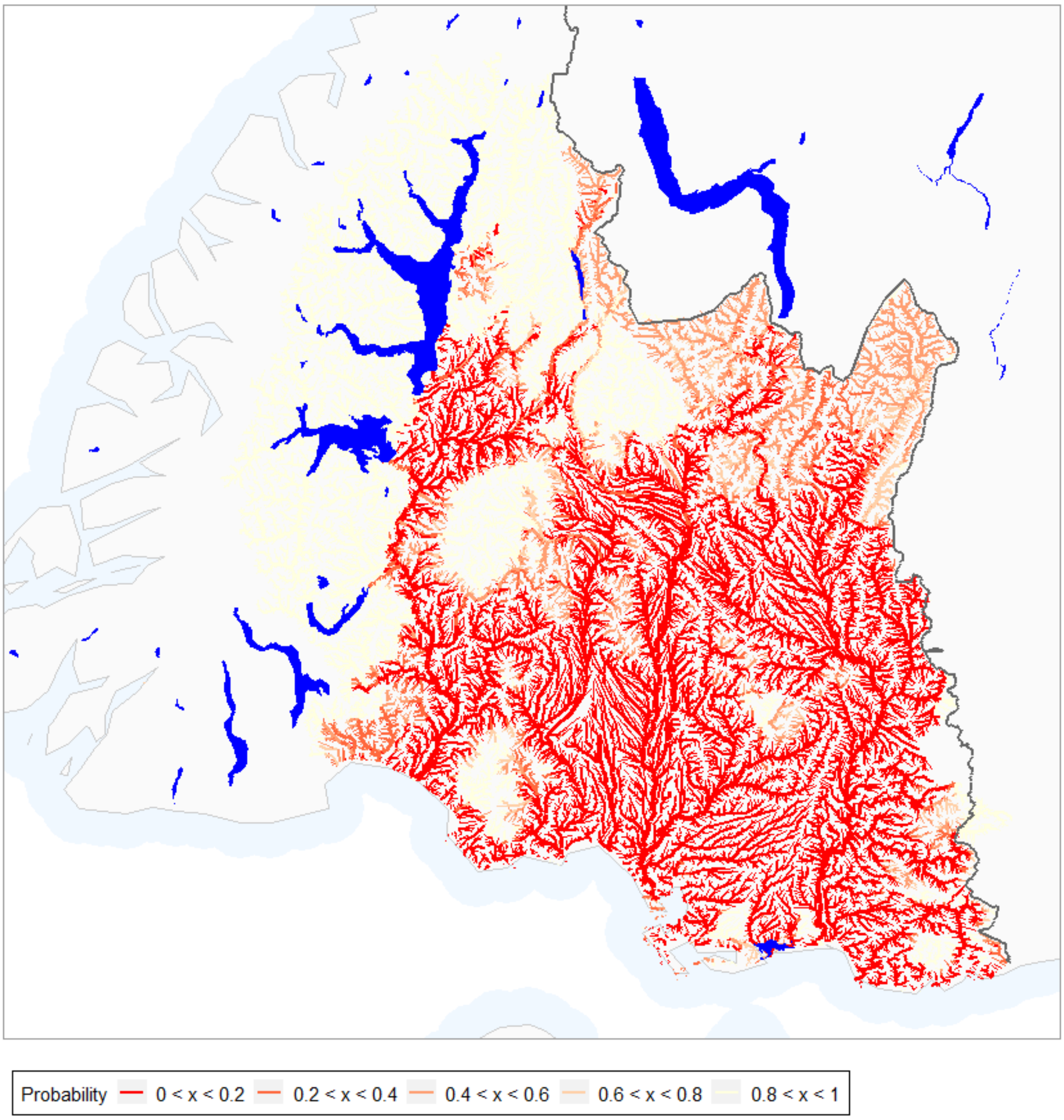

Figure 17. Probability of compliance with any of the criteria pertaining to the pSWLP option. This map represents the overall probability that segments achieve the pSWLP option.

3.7.2 Local excess loads

The local excess load is the amount by which the current *E. coli* load at a river segment would need to be reduced to achieve the FWO for that receiving environment. For the pSWLP option, the best estimate[5] of the local excess *E. coli* yield for rivers exceeded 10 giga *E. coli* ha^{-1} yr^{-1} for 91% of river segments, exceeded 50 giga *E. coli* ha^{-1} yr^{-1} for 77% of river segments and exceeded 180 giga *E. coli* ha^{-1} yr^{-1} for 60% of river segments (Figure 18). Note that the 10, 50 and 180 giga *E. coli* ha^{-1} yr^{-1} are nominal breakpoints for communication purposes and correspond to the legend thresholds on Figure 18. These values have no special significance (i.e., are not guidelines or standards). There were no segments that were classified other than Natural State river management class for which the local excess *E. coli* loads were zero; it is

[5] The best estimate for each segment is the median of the estimates of the local excess loads over the 100 realisations.

noted that 32% of river segments are in the Natural State management class and therefore have no assessment of compliance and no load reduction required.

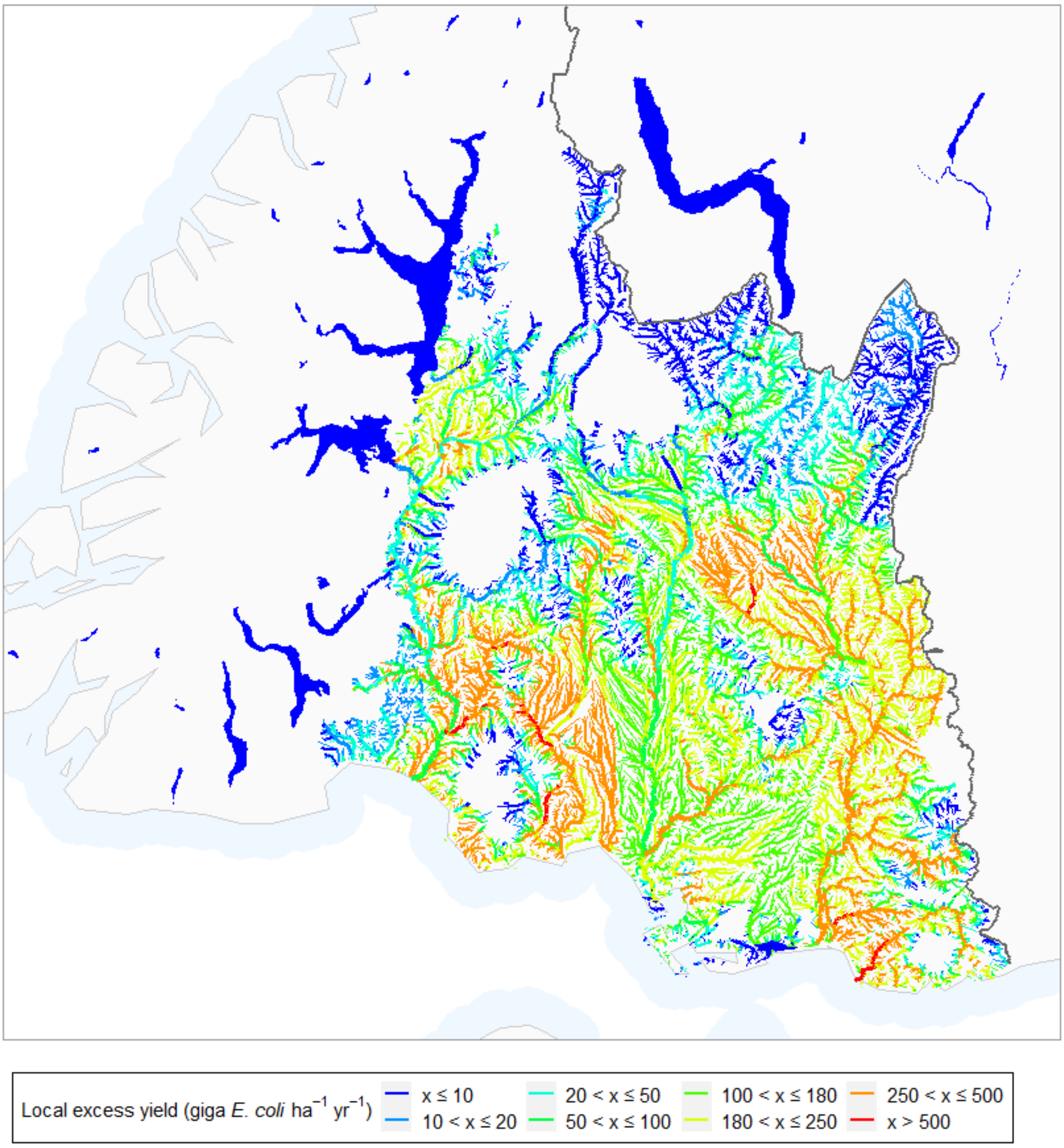

Figure 18. Local excess E. coli loads for the pSWLP option. Note that the breakpoints for the local excess yield in the map legend are nominal and have no special significance (i.e., are not guidelines or standards).

3.7.3 Critical point catchment load reductions required

The load reductions required by the pSWLP option for critical point catchments are shown on Figure 19 and Figure 20. When critical point catchments *E. coli* load reductions required were expressed as yields, critical point catchments with reduction requirements of greater 50 giga *E. coli* ha^{-1} yr^{-1} occupied 98% of the study area (Figure 19). Critical point catchments with reduction requirements of greater 150 giga *E. coli* ha^{-1} yr^{-1} occupied 58% of the study area (Figure 19). The comparison of load reductions expressed as yields (*E. coli* ha^{-1} yr^{-1}) with those expressed as proportion of current load (%) indicates that reduction requirements in catchments with low yield reductions (e.g., much of the Waiau River catchment) are nevertheless large in relative terms. Critical point catchments with *E. coli* load reductions of greater than 50% occupied 99% of the study area and critical point catchments with *E. coli* load reductions of greater than 90% occupied 24% of the study area (Figure 20).

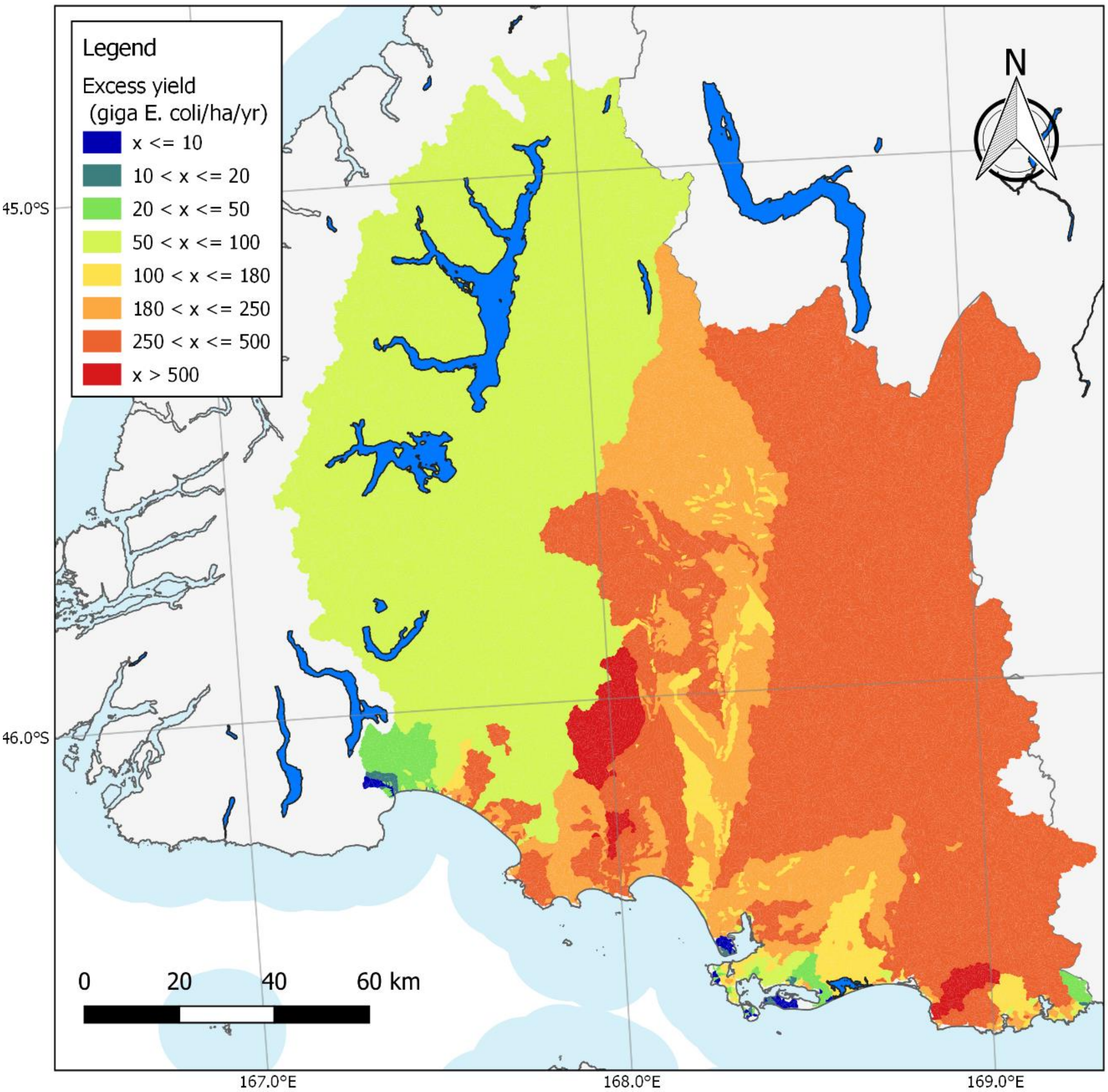

Figure 19. The E. coli load reduction required, expressed as yields, for critical point catchments and the pSWLP option. The critical point catchment colours indicate the mean TN load reductions required to allow all FWOs be achieved in the critical point catchment (including the critical point at the bottom of the catchment).

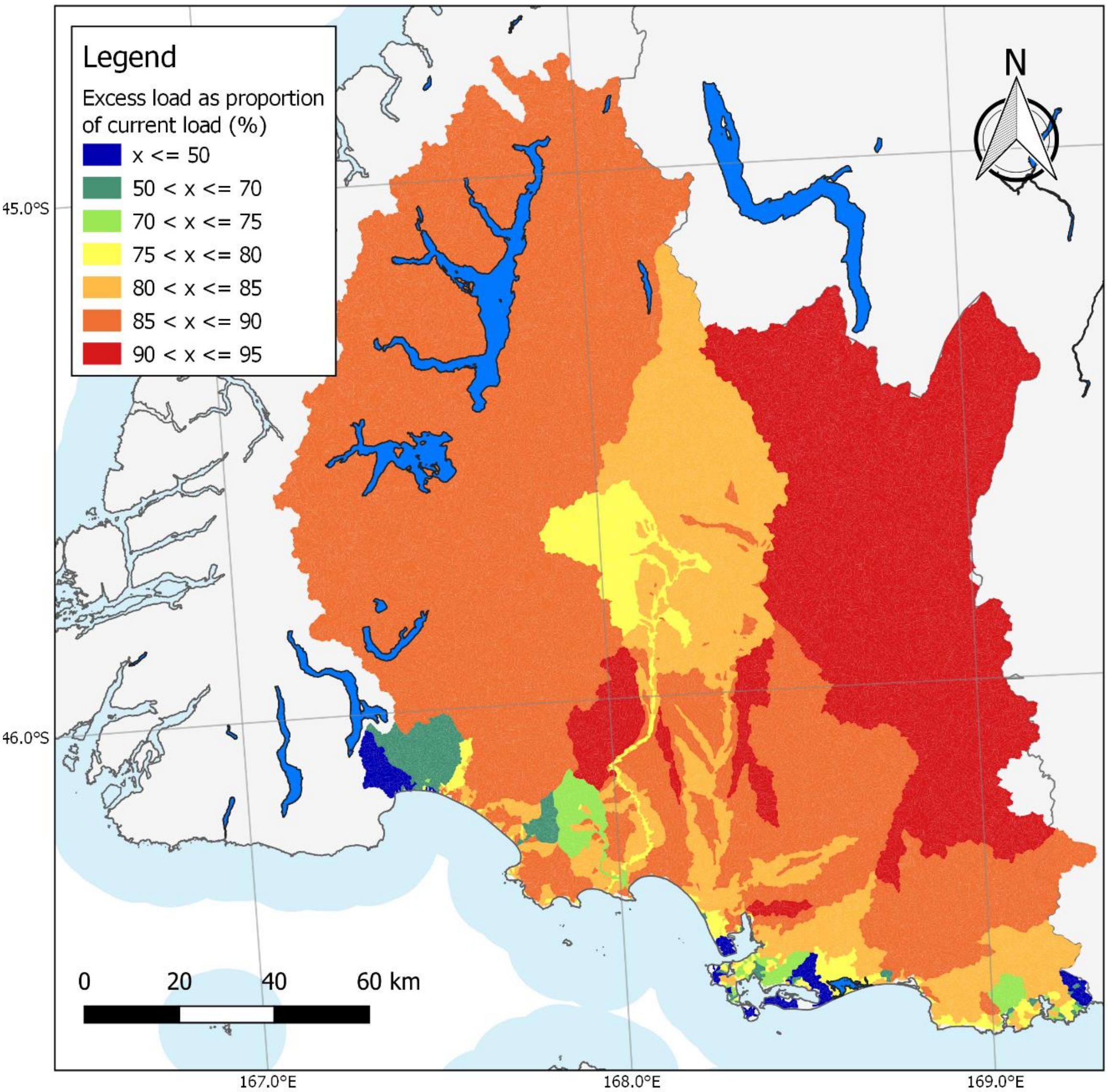

Figure 20. The E. coli load reduction required, expressed as proportion of the current load (%), for critical point catchments and the pSWLP option. The critical point catchment colours indicate the E. coli load reductions required to allow all FWOs be achieved in the critical point catchment (including the critical point at the bottom of the catchment).

3.7.4 Reporting catchments and regional load reductions required

The load reduction required by the pSWLP option for each reporting catchment and for whole study area are shown in Table 11. For the whole study area, the best estimate of *E. coli* load reductions required was 311 peta *E. coli* yr^{-1}, which represents 82% of the current load. The *E. coli* load reductions required were highest (>80%) in the Matāura, Orepuki Coastal Zone, Tokanui catchments, Waiau Catchment and Waimatuku & Taunamau Catchments. The *E. coli* load reductions required were lowest (<65%) in the Aparima & Pourakino Catchments, Bluff Zone, Catlins Zone, and Te Waewae Bay Western Coastal Zone.

Table 11. Current load and load reduction required for E. coli by reporting catchments and for the whole study area for the pSWLP option. Note that loads are expressed in absolute terms in units of E. coli organisms per year (peta E. coli yr^{-1}) and as a proportion of current load (%). The first value in each column is the best estimate, which is the median value over the 100 Monte Carlo realisations. The values in parentheses are the lower and upper bounds of the 90% confidence interval.

Reporting catchment	Total load (peta *E. coli* yr^{-1})	Load reduction required (peta *E. coli* yr^{-1})	Load reduction required (%)
Aparima & Pourakino Catchments	67 (9.7 - 251)	49 (3.8 - 228)	63 (23 - 88)
Bluff Zone	0.4 (0.2 - 0.7)	0.2 (0.1 - 0.4)	50 (30 - 68)
Catlins Zone	0.5 (0.2 - 1.2)	0.3 (0.1 - 0.9)	59 (31 - 83)
Matāura Catchments	181 (32 - 556)	161 (19 - 530)	82 (63 - 96)
Orepuki Coastal Zone	6.4 (3.5 - 11)	5.3 (2.5 - 9.2)	80 (71 - 88)
Ōreti & Invercargill Catchments	88 (16 - 240)	73 (7.6 - 222)	78 (50 - 95)
Te Waewae Bay Western Coastal Zone	2.8 (0.9 - 6.6)	1.6 (0.2 - 4.2)	51 (20 - 79)
Tokanui Coastal Zone	9.5 (3.5 - 23)	8.3 (2.4 - 21)	82 (63 - 98)
Waiau Catchment	60 (9.5 - 202)	54 (6.6 - 199)	84 (62 - 96)
Waikawa Catchment	8.4 (2.2 - 24)	6.9 (1.1 - 20)	76 (55 - 92)
Waimatuku & Taunamau Catchments	9.4 (1.4 - 22)	8.3 (0.9 - 20)	80 (57 - 93)
Waituna Catchments	3.6 (0.5 - 12)	2.8 (0.2 - 10.9)	67 (28 - 91)
Total (whole study area)	373 (171 - 875)	311 (135 - 783)	82 (72 - 94)

3.8 Comparison between FWO settings

A comparison of the *E. coli* load reductions required by the Hauora and pSWLP options for the reporting catchments is shown in Table 12. The best estimate for the load reductions is always less for the pSWLP option compared to the Hauora option. However, the 90% confidence intervals for the two FWO options overlap in all cases. This indicates that from a practical perspective the amount of effort (i.e., the reduction in *E. coli* loads required) to achieve the two sets of FWOs are indistinguishable. This is because the models have considerable uncertainty and the concentrations and corresponding loads that separate the two sets of FWOs are similar, relative to this uncertainty.

Table 12. Comparison of the load reductions required for individual reporting catchments for the Hauora and pSWLP options. The load reductions are shown as proportion of current load (%). The first value in each column is the best estimate, which is the median value over the 100 Monte Carlo realisations. The values in parentheses are the lower and upper bounds of the 90% confidence interval.

Reporting catchment	**Load reduction required for Hauora option (%)**	**Load reduction required for pSWLP option (%)**
Aparima & Pourakino Catchments	74 (38 - 93)	63 (23 - 88)
Bluff Zone	65 (49 - 78)	50 (30 - 68)
Catlins Zone	69 (46 - 86)	59 (31 - 83)
Matāura Catchments	89 (76 - 98)	82 (63 - 96)
Orepuki Coastal Zone	85 (77 - 90)	80 (71 - 88)
Ōreti & Invercargill Catchments	89 (73 - 100)	78 (50 - 95)
Te Waewae Bay Western Coastal Zone	64 (40 - 85)	51 (20 - 79)
Tokanui Coastal Zone	91 (79 - 102)	82 (63 - 98)
Waiau Catchment	86 (64 - 96)	84 (62 - 96)
Waikawa Catchment	86 (72 - 95)	76 (55 - 92)
Waimatuku & Taunamau Catchments	86 (71 - 95)	80 (57 - 93)
Waituna Catchments	78 (50 - 93)	67 (28 - 91)
Total (whole study area)	90 (80 - 96)	82 (72 - 94)

4 Discussion and conclusions

4.1 Load reductions required

This report has predicted *E. coli* load reductions needed to achieve options for freshwater objectives (FWO) for human health in rivers in Southland. The options for FWOs are defined in terms of NOF attribute bands for all river receiving environments (represented by network segments) in the study area. Two sets of FWO scenarios were nominated and are referred to as the 'Hauora' option and the 'pSWLP' option. The Hauora option is defined based on Hauroa Principle D (Bartlett et al, 2020) and represent the most aspirational (highest quality) objectives (i.e., A band state) while the criteria for the pSWLP option were defined based on the draft minimum FWOs developed by Environment Southland and represent existing water quality standards in regional plans (i.e., A band state for Mountain and Lake fed classes; B band state for Lowland soft bed, Lowland hard bed, Spring fed and Hill classes).

The study area includes all of Southland except Fiordland and Stewart Island. The load reductions required for each scenario were estimated for all individual river segments and these individual results were also aggregated to report on individual 'reporting catchments' and the whole study area.

For the whole study area and the Hauora option, the best estimate of the *E. coli* load reductions required were 347 peta *E. coli* yr^{-1}, which represents a best estimate of 90% of the current loads. For the whole study area and the pSWLP option, the *E. coli* load reductions required were estimated to be 311 peta *E. coli* yr^{-1}, which represent 82% of the current loads. These large load reductions reflect the stringency of the FWOs for the Hauora and pSWLP options, which are all A band and a mix of A and B bands, respectively. We note however that an analysis based on meeting the NOF C band state (i.e., the minimum state deemed suitable for primary contact in the national targets laid out in Appendix 3 of the NPSFM) indicates a regional load reduction of 77% of the current *E. coli* load. This result indicates that *E. coli* loads in the Region's rivers far exceed even that minimal accepted state. It is noted that the 90% confidence intervals for the two sets of load reduction estimates are strongly overlapping. This indicates that, from a practical perspective, the reduction in *E. coli* loads required to achieve the two sets of FWOs are indistinguishable. The difference in the percentage load reduction for both the whole study area and the reporting catchments were always consistent with expectations, that is, reductions were always less for the pSWLP option than the Hauora option (Table 12).

4.2 Comparison with previous studies and national targets

This study also provided an estimate that 61% of the large rivers (i.e., those with stream order ≥4) across the whole region (i.e., including Fiordland and the Offshore Islands) are currently suitable for primary contact according to the NPSFM Appendix 3 criteria (i.e., in attribute state C or better). This estimate is comparable to a previous estimate of 62% of these large rivers in Southland being suitable for primary contact as reported in MFE (2018). It is noteworthy that the current draft FWOs approved in principle by Environment Southland and the Te Ao Marama Inc (TAMI) board in November 2020, and indeed both sets of FWO options assessed in this report, would see 100% of rivers being suitable for primary contact (i.e., better than C band) within a generation (25-30 years). This is a step further than the regional primary contact targets set by Environment Southland in December 2018 for 66% of rivers to be suitable by 2030 and 80% by 2040, with continued improvement beyond 2040. Those regional targets were Environment Southland's intended contribution to meeting the national targets laid out in the 2017 NPSFM (unchanged in the NPSFM 2020) of 80% being suitable by 2030 and 90% being suitable no later than 2040.

4.3 Uncertainties

Uncertainty is an unavoidable aspect of this study because it is based on simplifications of reality and because it has been informed by limited data. The study estimated the statistical uncertainty of the *E. coli* load reduction estimates that are associated with two key components of the analyses: the modelled regional river *E. coli* statistics and loads (see Sections 3.1 and 3.2). The statistical uncertainty of these models is associated with their inability to perfectly predict the statistics and load observed at water quality monitoring sites; the error associated with these predictions is quantified by the model RMSD values (Table 4 and Table 5).

A simple Monte Carlo analysis was used to combine the above model uncertainties and to make assessments of the uncertainty of several characteristics and quantities. The Monte Carlo analysis recognises and is based on the uncertainties associated with the nine individual models that are used in the assessment. We have presented the results of the analyses differently depending on the assessed characteristics. In general, the mean of results obtained from 100 Monte Carlo realisations has been used to represent the best estimate of any quantity. For example, we provide a best estimate of the proportion of network segments in each of the five NOF attribute bands (Table 9). We also use the 90% confidence interval for our estimates of these proportions to indicate the uncertainty of these estimates. The lower and upper confidence limits can be interpreted as the values for which we are 95% confident the proportions are not lower than or greater than. The estimated load reductions required for the region and the reporting catchments have followed this same approach with a best estimate and 90% confidence intervals given (Table 10 and Table 11). We have presented maps showing compliance with criteria associated with the FWOs (e.g., Figure 12 and Figure 17). These maps show the estimated probability that segments comply with the criteria. We note that the distributions of load reductions over the 100 realisations (and mean and 90% confidence intervals) derived by the analyses can be obtained for each of the 44,000 segments represented in the analysis. These data are not presented in this report but are available as supplementary files.

An important conclusion from the analysis of uncertainty is that we are 95% confident that *E. coli* load reductions are required to achieve both sets of FWOs (i.e., the Hauora and the pSWLP options) for all reporting catchments. This is because the lower bound of the 90% confidence interval for load reduction required is greater than zero for both sets of FWOs (Table 10 and Table 11).

There are sources of uncertainty that this study has not accounted for. A key uncertainty is the source of *E. coli* at any point in the river network. An underlying implicit assumption in this study is that *E. coli* concentrations at any point are the outcome of load and that this load is attributable to contributions from all land in the upstream catchment. This is not necessarily true, concentrations at a location may be more strongly influenced by immediate local sources than contributions from upstream. Local sources may be from local land areas or may be associated with transfer of *E. coli* from the river bed, particularly during high flow events (Wilkinson *et al.*, 2011). The assumption that *E. coli* loads at any point are the outcome contributions from all land in the upstream catchment is manifested in our analysis by the additive reconciliation of local load reductions in the downstream direction at step 4 of the analysis to obtain the load reduction required at every point in the drainage network. This analysis is based on an assumption that any reduction upstream of a location contributes to the load reduction necessary at that location. This assumption would be violated if local contributions were important determinants of concentrations and loads at a point. The existence of these processes is not well understood or represented by our analysis and are therefore sources of additional uncertainty associated with the estimation of load reduction required.

The uncertainties described in this study indicate that the best estimates and maps are appropriately considered as indicative of the regional-scale patterns of compliance and *E. coli* load reductions required. The broad scale patterns provide a reliable indication of the relative differences in compliance and load reductions required between locations. However, there is considerable uncertainty associated with the absolute values of the *E. coli* load reductions required and these become larger as the spatial scale over which the reductions are evaluated is reduced. It is unlikely that these uncertainties can be significantly reduced in the short to medium term (i.e., in less than 5 to 10 years) because, among other factors, the modelling is dependent on the collection of long-term water quality monitoring data.

4.4 Informing decision-making on limits

The NPS-FM requires regional councils to set limits on resource use to achieve environmental outcomes (e.g., FWOs). This report helps inform Environment Southland's process of setting limits by assessing the approximate magnitude of *E. coli* load reductions needed to achieve several options for FWOs, with a quantified level of confidence and risk associated with each option. However, this report does not consider what kinds of limits on resource might be used to achieve any load reductions, how such limits might be implemented, over what timeframes and with what implications for other values. The NPS-FM requires regional councils to have regard to these and other things when making decisions on setting limits. This report shows that these decisions will ultimately need to be made in the face of uncertainty about the magnitude of load reductions needed.

Acknowledgements

Thanks is due to Roger Hodson for assistance with water quality data and load calculations. Early versions of the report were reviewed by Ned Norton, and Karen Wilson, Roger Hodson and Nick Ward. Their comments and corrections greatly improved this report.

References

Bartlett, M., J. Kitson, N. Norton, and K. Wilson, 2020. Draft Murihiku Southland Freshwater Objectives: Providing for Hauora, the Health and Well-Being of Waterbodies in Murihiku Southland. Environment Southland and Te Ao Marama Inc, Invercargill, New Zealand.

Booker, D.J. and R.A. Woods, 2014. Comparing and Combining Physically-Based and Empirically-Based Approaches for Estimating the Hydrology of Ungauged Catchments. Journal of Hydrology 508:227–239.

Cohn, T.A., 2005. Estimating Contaminant Loads in Rivers: An Application of Adjusted Maximum Likelihood to Type 1 Censored Data. Water Resources Research 41. http://onlinelibrary.wiley.com/doi/10.1029/2004WR003833/full. Accessed 21 Jan 2016.

Cohn, T.A., D.L. Caulder, E.J. Gilroy, L.D. Zynjuk, and R.M. Summers, 1992. The Validity of a Simple Statistical Model for Estimating Fluvial Constituent Loads: An Empirical Study Involving Nutrient Loads Entering Chesapeake Bay. Water Resources Research 28:2353–2363.

Cohn, T.A., L.L. Delong, E.J. Gilroy, R.M. Hirsch, and D.K. Wells, 1989. Estimating Constituent Loads. Water Resources Research 25:937–942.

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

Efron, B., 1981. Nonparametric Estimates of Standard Error: The Jackknife, the Bootstrap and Other Methods. Biometrika 68:589–599.

MFE, 2019. Essential Freshwater: Impact of Existing Periphyton and Proposed Dissolved Inorganic Nitrogen Bottom Lines. Ministry for the Environment & Statistics NZ, Wellington, New Zealand.

MFE and MoH, 2003. Microbiological Water Quality Guidelines for Marine and Freshwater Recreational Areas. Ministry for the Environment. https://www.mfe.govt.nz/sites/default/files/microbiological-quality-jun03.pdf.

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

Nash, J.E. and J.V. Sutcliffe, 1970. River Flow Forecasting through Conceptual Models Part I—A Discussion of Principles. Journal of Hydrology 10:282–290.

Norton, N. and K. Wilson, 2019. Developing Draft Freshwater Objectives for Southland. Environment Southland publication, Environment Southland, Invercargill, New Zealand.

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

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

Parker, W.J., 1998. Standardisation between Livestock Classes: The Use and Misuse of the Stock Unit System. Proceedings of the Conference New Zealand Grassland Association., pp. 243–248.

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

Roygard, J.K.F., K.J. McArthur, and M.E. Clark, 2012. Diffuse Contributions Dominate over Point Sources of Soluble Nutrients in Two Sub-Catchments of the Manawatu River, New Zealand. New Zealand Journal of Marine and Freshwater Research 46:219–241.

Snelder, T., 2018. Assessment of Recent Reductions in E. Coli and Sediment in Rivers of the Manawatū-Whanganui Region. LWP Ltd, Christchurch, New Zealand.

Snelder, T., 2020. Assessment of Nutrient Load Reductions to Achieve Freshwater Objectives in the Rivers, Lakes and Estuaries of Southland. To Inform the Southland Regional Forum Process. LWP Client Report, LWP Ltd, Christchurch, New Zealand.

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., S.T. Larned, and R.W. McDowell, 2018. Anthropogenic Increases of Catchment Nitrogen and Phosphorus Loads in New Zealand. New Zealand Journal of Marine and Freshwater Research 52:336–361.

Snelder, T.H., A.L. Whitehead, C. Fraser, S.T. Larned, and M. Schallenberg, 2020. Nitrogen Loads to New Zealand Aquatic Receiving Environments: Comparison with Regulatory Criteria. New Zealand Journal of Marine and Freshwater Research:1–24.

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

Wilkinson, R.J., L.A. McKergow, R.J. Davies-Colley, D.J. Ballantine, and R.G. Young, 2011. Modelling Storm-Event E. Coli Pulses from the Motueka and Sherry Rivers in the South Island, New Zealand. New Zealand Journal of Marine and Freshwater Research 45:369–393.

Appendix A Calculation of E. coli loads at monitoring sites

A1 Water quality data

We obtained E. coli monitoring data for 59 river SOE monitoring sites from the ES database. E. coli was generally observed at the river sites on a monthly basis. These sites had variable start and end dates and total numbers of observations. Most sites had greater than 200 observations for over 15 years. Five sites were excluded from the subsequent load calculation as they had fewer than 96 observations (80% of monthly observations over a ten-year period).

A2 Flow Data

We obtained daily timeseries of flow for all 59 E. coli monitoring sites from ES. Of these 59 sites, 17 were observed timeseries and the remaining 42 sites were simulated timeseries of flow. Start years for the flow records ranged from 1955 to 2002 (one site started in 2011 but was already excluded due to short E. coli record). All but 4 sites had more than 20 years of daily flow observations, with a median of 43 years of flow data across all sites.

A3 Load calculations

Calculation of E. coli loads at monitoring sites generally comprise two steps: (1) the generation of a series of flow and concentration pairs representing 'unit loads' and (2) the summation of the unit loads over time to obtain the total load. In practice step 1 precedes step two but in the explanation that follows, we describe step 2 first.

If flow and concentration observations were available for each day, the export coefficient, (the mean annual load, standardised by the upstream catchment area) would be the summation of the daily flows multiplied by their corresponding concentrations:

$$L = \frac{K}{A_c N} \sum_{j=1}^{N} C_j Q_j \qquad \text{(Equation A1)}$$

where L: mean annual export coefficient (giga *E. coli* yr^{-1} ha^{-1}), A_c: catchment area, ha, K: units conversion factor, C_j: contaminant concentration for each day in period of record (mg m^{-3}), Q_j: daily mean flow for each day in period of record (m^3 s^{-1}), and N: number of days in period of record.

In this summation, the individual products represent unit loads. Because concentration data are generally only available for infrequent days (i.e., generally in this study, monthly observations), unit loads can only be calculated for these days. However, flow is generally observed continuously, or the distribution of flows can be estimated for locations without continuous flow data, and there are often relationships between concentration and flow, time and/or season. Rating curves exploit these relationships by deriving a relationship between the sampled nutrient concentrations (c_i) and simultaneous observations of flow (q_i). Depending on the approach, relationships between concentration and time and season may be included in the rating curve. This rating curve is then used to generate a series of flow and concentration pairs (i.e., to represent Q_j *and* C_j in Equation A1) for each day of the entire sampling period (i.e., step 1 of the calculation method; Cohn *et al.*, 1989). The estimated flow and concentration pairs are then multiplied to estimate unit loads, and these are then summed and transformed by K, N and A_c to estimate mean annual export coefficients (i.e., step 2 of the calculation method; Equation A1).

There are a variety of approaches to defining rating curves. Identifying the most appropriate approach to defining the rating curve requires careful inspection of the available data for each site and contaminant. The details of the approaches and the examination of the data are described below in Section A3.3.

For each site, we calculated the load for each contaminant using three commonly used and recommended methods that are based on different types of rating curves, which we refer to as the the flow stratification method, the seven-parameter (L7) rating method and the five-parameter (L5) rating method. We expressed all contaminant loads as annual export coefficients (i.e., for giga *E. coli* yr^{-1} ha^{-1}) by dividing the annual load (kg yr^{-1}) by the catchment area (ha). Loads were estimated for an evaluation date of 31/12/2019 (rather than a long term mean load).

A3.1 Methods for defining rating curves

A3.1.1 Flow stratification

Roygard *et al.* (2012) employed a flow stratification approach to defining rating curves. This approach is based on a non-parametric rating curve, which is defined by evaluating the average concentration within equal increments of the flow probability distribution (flow 'bins'). In their application, Roygard *et al.* (2012) employed ten equal time-based categories (flow decile bins), defined using flow distribution statistics and then calculated mean concentrations within each bin. This non-parametric rating curve can then be used to estimate nutrient concentrations, $\hat{C}$, for all days with flow observations. At step 2, the load is calculated following Equation A1a, providing an estimate of average annual load over the observation time period.

$$L = \frac{K}{A_c N} \sum_{j=1}^{N} \hat{C}_j Q_j \qquad \text{Equation A1a}$$

where $\widehat{C}_j$ is calculated mean concentration associated with the flow quantile bin of the flow Q_j., and all other variables are as per Equation A1.

A3.1.2 L7 model

Two regression model approaches to defining rating curves of Cohn *et al.*(1989, 1992) and Cohn (2005) are commonly used to calculated loads. The regression models relate the log of concentration to the sum of three explanatory variables: discharge, time, and season. The L7 model is based on seven fitted parameters given by:

$$ln(\widehat{C}_i) = \beta_1 + \beta_2\left[ln(q_i) - \overline{(ln(q))}\right] + \beta_3\left[ln(q_i) - \overline{(ln(q))}\right]^2 + \beta_4(t_i - \overline{T}) + \beta_5(t_i - \overline{T})^2 + \beta_6 sin(2\pi t_i) + \beta_7 cos(2\pi t_i) \qquad \text{Equation A2}$$

where, *i* is the index for the concentration observations, $\beta_{1,2,..7}$: regression coefficients, t_i: time in decimal years, $\overline{T}$: mean value of time in decimal years, $\overline{(ln(q))}$ mean of the natural log of discharge on the sampled days, and $\widehat{C}_i$: is the estimated i^{th} concentration.

The coefficients are estimated from the sample data by linear regression, and when the resulting fitted model is significant ($p < 0.05$), it is then used to estimate the concentration on each day in the sample period, $ln(\widehat{C}_j)$. The resulting estimates of $ln(\widehat{C}_j)$ are back-transformed (by exponentiation) to concentration units. Because the models are fitted to the log transformed concentrations the back-transformed predictions were corrected for retransformation bias. We used the smearing estimate of Duan (1983) as a correction factor (S):

$$S = \frac{1}{n}\sum_{i=1}^{n} e^{\hat{\varepsilon}_i} \qquad \text{Equation A3}$$

where, $\hat{\varepsilon}$ are the residuals of the regression models, and n is the number of flow-concentration observations. The smearing estimate assumes that the residuals are homoscedastic and therefore the correction factor is applicable over the full range of the predictions.

The average annual load is then calculated by combining the flow and estimated concentration time series:

$$L = \frac{KS}{A_c N}\sum_{j=1}^{N} \hat{C}_j Q_j \qquad \text{Equation A1b}$$

If the fitted model is not significant, $\widehat{C}_j$ is replaced by the mean concentration and S is unity.

To provide an estimate of the load at a specific date, (i.e. t^{est} = 1/3/2004) a transformation is performed so that the year components of all dates (t_j) are shifted such that all transformed dates lie within a one-year period centred on the proposed observation date (i.e. Y=1/9/2003 to 31/8/2004). For example, flow at time t=13/6/2007 would have a new date of Y =13/6/2004, and a flow at time t=12/11/1998 would have a new date of Y=12/11/2003.

$$ln\left(\widehat{C_j^Y}\right) = \beta_1 + \beta_2\left[ln(q_j) - \overline{(ln(q))}\right] + \beta_3\left[ln(q_j) - \overline{(ln(q))}\right]^2 + \beta_4(Y_j - \overline{T}) + \beta_5(Y_j - \overline{T})^2 + \beta_6 sin(2\pi Y_j) + \beta_7 cos(2\pi Y_j) \qquad \text{Equation A2a}$$

where $\widehat{C_j^Y}$ is the estimated j^{th} concentration for the estimation year, and Y_j is the transformed date of the i^{th} observation, and all other variables are as per Equation A3. The regression coefficients ($\beta_{1,2,..7}$) are those derived from fitting Equation A2 to the observation dataset. It follows that the estimated load for the year of interest can be calculated by:

$$L^Y = \frac{KS}{A_c N}\sum_{j=1}^{N} \hat{C}_j^Y Q_j \qquad \text{Equation A1c}$$

A3.1.3 L5 Model

The L5 model is the same as L7 model except that two quadratic terms are eliminated:

$$ln(\widehat{C}_i) = \beta_1 + \beta_2(ln(q_i)) + \beta_3(t_i) + \beta_4 sin(2\pi t_i) + \beta_5 cos(2\pi t_i) \qquad \text{Equation A4}$$

The five parameters are estimated, and loads are calculated in the same manner as the L7 model. Following the approach outlined for the L7 model, the L5 model can be adjusted when used for prediction to provide estimates for a selected load estimation date:

$$ln\left(\widehat{C_j^Y}\right) = \beta_1 + \beta_2[ln(q_j)] + +\beta_4(Y_j - \overline{T}) + +\beta_6 sin(2\pi Y_j) + \beta_7 cos(2\pi Y_j) \qquad \text{Equation A4a}$$

A3.2 Precision of load estimates

The statistical precision of a sample statistic, in this study the mean annual load, is the amount by which it can be expected to fluctuate from the population parameter it is estimating due to sample error. In this study, the precision represents the repeatability of the estimated load if it was re-estimated using the same method under the same conditions. Precision is characterised by the standard deviation of the sample statistic, commonly referred to as the standard error. We evaluated the standard error of each load estimate by bootstrap resampling (Efron, 1981). For each load estimate we constructed 100 resamples of the concentration data (of equal size to the observed dataset), each of which was obtained by random sampling with replacement from the original dataset. Using each of these datasets, we recalculated the site load and estimated the 95% confidence intervals, using the boot r package.

A3.3 Identifying a best load estimate

We developed an expert judgement-based methodology to evaluate the ‘best’ rating curve approach for each site and used this to make a ‘best’ load estimate. We did this by inspecting summaries of the flow-concentration-time (Q-C-T) data and model diagnostic information and performance measures pertaining to the rating curves. Data availability and sampling distribution with season, time and flow were also considered in this assessment. An example of the diagnostic plots that we used in this process is shown in Figure 21.

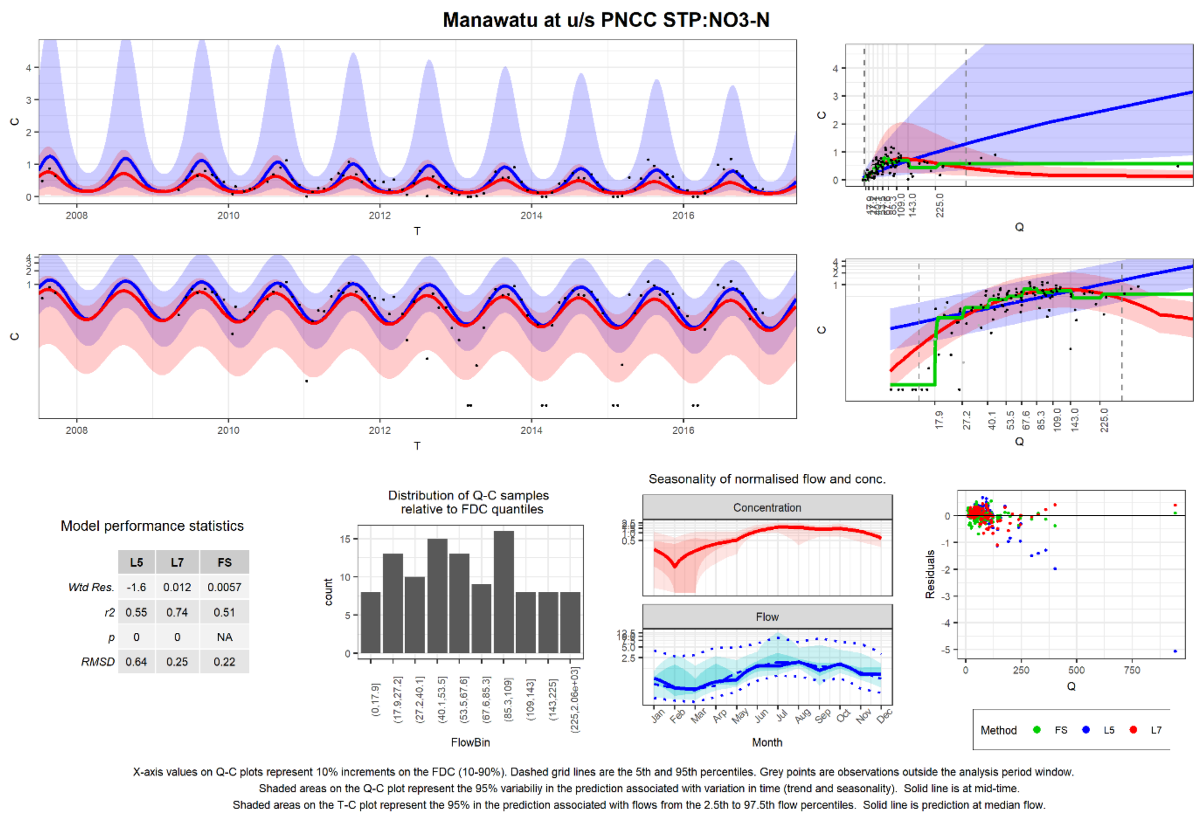

Figure 21: Example of a diagnostic information summary page that was used to examine the rating curves fitted to each site.

Appendix B Proportions of segments predicted to be in each attribute band by reporting catchment and management class.

Table 13. Proportion of segments (%) predicted to be in each attribute band by reporting catchment.

Reporting catchment	Attribute band	Best estimate	5% confidence limit	95% confidence limit
Aparima & Pourakino Catchment	A	8	2	33
Aparima & Pourakino Catchment	B	14	2	23
Aparima & Pourakino Catchment	C	7	0	19
Aparima & Pourakino Catchment	D	23	12	50
Aparima & Pourakino Catchment	E	49	13	83
Bluff Zone	A	9	1	29
Bluff Zone	B	23	8	42
Bluff Zone	C	21	8	39
Bluff Zone	D	28	18	41
Bluff Zone	E	19	11	30
Caitlins Zone	A	10	1	65
Caitlins Zone	B	24	1	63
Caitlins Zone	C	12	1	50
Caitlins Zone	D	29	5	71
Caitlins Zone	E	25	4	71
Mataura Catchments	A	14	0	39
Mataura Catchments	B	13	0	25
Mataura Catchments	C	8	1	24
Mataura Catchments	D	25	10	54
Mataura Catchments	E	40	4	64
Orepuki Coastal Zone	A	8	0	23
Orepuki Coastal Zone	B	12	3	23
Orepuki Coastal Zone	C	7	1	17
Orepuki Coastal Zone	D	24	14	37
Orepuki Coastal Zone	E	49	35	62
Oreti & Invercargill Catchments	A	10	0	26
Oreti & Invercargill Catchments	B	11	1	20
Oreti & Invercargill Catchments	C	5	0	11
Oreti & Invercargill Catchments	D	20	11	52
Oreti & Invercargill Catchments	E	54	10	77
Te Waewae Bay Western Coastal Zone	A	15	0	49
Te Waewae Bay Western Coastal Zone	B	24	2	55
Te Waewae Bay Western Coastal Zone	C	14	1	46
Te Waewae Bay Western Coastal Zone	D	27	5	68

Reporting catchment	Attribute band	Best estimate	5% confidence limit	95% confidence limit
Te Waewae Bay Western Coastal Zone	E	20	3	58
Tokanui Coastal Zone	A	4	0	20
Tokanui Coastal Zone	B	9	1	22
Tokanui Coastal Zone	C	6	1	19
Tokanui Coastal Zone	D	30	9	56
Tokanui Coastal Zone	E	51	19	72
Waiau Catchment	A	32	0	73
Waiau Catchment	B	27	1	63
Waiau Catchment	C	9	1	32
Waiau Catchment	D	17	6	45
Waiau Catchment	E	16	1	35
Waikawa Catchment	A	4	0	21
Waikawa Catchment	B	10	1	22
Waikawa Catchment	C	7	1	19
Waikawa Catchment	D	28	10	64
Waikawa Catchment	E	51	7	87
Waimatuku & Taunamau Catchments	B	0	1	1
Waimatuku & Taunamau Catchments	C	0	0	3
Waimatuku & Taunamau Catchments	D	17	0	87
Waimatuku & Taunamau Catchments	E	83	13	100
Waituna Catchments	A	3	0	19
Waituna Catchments	B	6	1	18
Waituna Catchments	C	9	1	20
Waituna Catchments	D	22	4	75
Waituna Catchments	E	59	18	96

Table 14. Proportion of segments (%) predicted to be in each attribute band by management class.

Management class	Attribute band	Best estimate	5% confidence limit	95% confidence limit
Natural State	A	42	1	87
Natural State	B	33	5	76
Natural State	C	10	1	42
Natural State	D	11	1	52
Natural State	E	3	0	11
Lowland hard bed	A	1	0	3
Lowland hard bed	B	4	2	7
Lowland hard bed	C	4	1	7
Lowland hard bed	D	27	13	50
Lowland hard bed	E	64	38	82
Lowland soft bed	A	1	0	4
Lowland soft bed	B	4	1	7
Lowland soft bed	C	4	1	7
Lowland soft bed	D	23	10	42
Lowland soft bed	E	68	45	86
Hill	A	14	1	35
Hill	B	22	9	34
Hill	C	13	5	26
Hill	D	34	19	48
Hill	E	17	3	37
Lake fed	A	15	1	35
Lake fed	B	13	1	29
Lake fed	C	5	0	15
Lake fed	D	14	3	39
Lake fed	E	53	23	64
Mountain	A	49	10	89
Mountain	B	31	1	75
Mountain	C	10	0	50
Mountain	D	9	0	58
Mountain	E	0	0	4
Spring fed	A	2	0	4
Spring fed	B	2	1	4
Spring fed	C	2	0	6
Spring fed	D	22	4	54
Spring fed	E	72	40	93