Parameters for simple empirical catchment water quality models for simulating *Escherichia coli* in New Zealand rivers

June 2024

Prepared by:

Ton Snelder

Sandy Elliott[1]

Richard Muirhead[2]

Caroline Fraser

1. National Institute of Water and Atmosphere (NIWA)
2. AgResearch Ltd

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
New Zealand

LWP Client Report Number:	2024-09
Report Date:	June 2024

Quality Assurance Statement

Version	Reviewed By	
Final	Olivier Ausseil	

Table of Contents

Figures

Tables

Glossary

Term	Definition
Attenuation coefficient	The proportion of the export load that is lost between the sources and the instream observation point.
Concentration	Concentration of *E. coli* in water measured as colony forming units (cfu) or most probable number (MPN) per 100 ml. The different units reflect different test methods, but they both represent the number of viable bacteria in a 100 millilitre sample of water. In this study, concentration is reported as cfu 100 ml^{-1}.
DN2.4	Digital river network, version 2.4
Empirical catchment water quality model	Shortened to 'empirical model' in the report. These models use a purely empirical (statistical) approach to predict instream load or concentration at a point in the drainage network based on the proportion of upstream catchment occupied by various Types defined by a typology. There is no attempt to represent either contaminant loss from land or attenuation so the predictions are purely data driven.
ETC	Empirical Type Concentration. The expected instream concentration (cfu 100 ml^{-1}), realised at a point in the drainage network (post attenuation), that is generated by diffuse losses associated with a specific land-type that is defined by a typology. These values were estimated by fitting quantile regression models to median concentrations calculated for water quality stations. An ETC value can be interpreted as the expected proportional contribution of a land-type to concentration at a downstream point in the drainage network. Alternatively, it can be interpreted as the expected concentration for a catchment comprised of only that land-type.
ETY	Empirical Type Yield. The expected annual load per unit area (cfu ha^{-1} yr^{-1}), realised at a point in the drainage network (post attenuation), that is generated by diffuse losses associated with a specific land-type that is defined by a typology. These values were estimated by fitting quantile regression models to annual loads calculated for water quality stations. An ETY value can be interpreted as the expected proportional contribution of a land-type to yield at a downstream point in the drainage network. Alternatively, it can be interpreted as the expected yield for a catchment comprised of only that land-type.

Term	Definition
Export coefficient	Rates of diffuse *E.coli* loss from land (cfu ha^{-1} yr^{-1}) to streams.
Factor	Environmental variables that are used to define land-types in typologies including land use/cover, elevation and soil drainage.
Instream load	In this study, the in-river load calculated for a water quality monitoring station from infrequent (e.g., monthly) observations of *E. coli* concentration and daily flow.
Instream yield	Instream load standardised by (divided by) catchment area. In this study the units are giga cfu ha^{-1} yr^{-1}, where giga is 10^9.
Land use	Categorical description of land use. In this study a total of nine categories were used that include categories that, strictly speaking, are descriptions land cover (i.e., “Natural”, “Bare” and “Water”).
OLS	Ordinary Least Squares regression.
Process-based catchment water quality models	Models that represent separate processes such as contaminant loss from land and attenuation to produce a prediction of the instream load or concentration at downstream points in the drainage network. An example of a process-based model that is discussed in this report is CLUES.
Land-type	A type (i.e., a class of land) defined by a typology. In this report land-types are defined by land use categories with, optionally, one or more additional environmental categories describing variation in elevation, temperature, moisture and soil drainage. In places in this report, land-type is shortened simply to ‘type’.
Typology	A system of land types that is used to classify and group land areas that are alike in terms of their contribution to *E. coli* concentration and loads in the downstream drainage network. In this study, several typologies were defined and tested. These typologies were defined by subdivision of several factors into categories. The factors include land use or cover (sub-divided into categories such as Dairy, Sheep & Beef, Native, Exotic Forest, and Urban) and the environmental factors: soil drainage (sub-divided into categories of poorly drained and well drained), land elevation (sub-divided into categories of low and high), temperature and moisture.

Executive Summary

Faecal Indicator Bacteria (FIB) are a key freshwater contaminant that the National Policy Statement – Freshwater Management (NPS-FM) requires regional councils to manage. FIB that is discharged into freshwater indicates that pathogens that are harmful to humans may be present. The concentration of a specific FIB, the bacterium *Escherichia coli* (*E. coli*), is used as an indicator of human or animal faecal contamination and the risk of infectious human disease from waterborne pathogens in water used for contact recreation and drinking. The NPS-FM includes *E. coli* attributes that must be used by regional councils to set target states in rivers. In addition, the NPS-FM requires that regional councils define limits or management plans to resource use to achieve these targets. Typically, models are used to predict *E. coli* loads or concentration in freshwater receiving environments under both existing catchment conditions and some possible future set of conditions, termed scenarios, associated with change land use or management. Regional councils use the results of scenario analysis to formulate achievable targets and associated limits or management plans to achieve these.

E. coli is sampled at over 1,000 long term river water quality monitoring stations across New Zealand every month. The concentration of *E. coli* in each sample is measured in the laboratory and reported as colony forming units (cfu) or most probable number (MPN) per 100 ml. The different units reflect different test methods, but they both represent the number of viable organisms in a 100 millilitre sample of water. For the purposes of this study the two measurement units were assumed to be equivalent.

Analyses associated with NPS-FM implementation commonly use a class of catchment model that we call process-based models. Process-based models explicitly represent processes including contaminant loss in the catchment (source losses), transport to downstream receiving environments, and attenuation (i.e., reduction in loads between the point of discharge and downstream receiving environments by natural processes). Setting up process-based catchment *E. coli* models involves quantifying at least two types of parameters: (1) representing rates of *E. coli* loss in the catchment and (2) representing rates of attenuation. The parameters representing both diffuse losses and attenuation are typically calibrated by to instream loads observed at water quality monitoring stations. Calibration of process-based catchment *E. coli* models is challenging due to the complexity of the processes involved and data constraints. A significant challenge for process-based models is robust accounting and reporting of uncertainty. Combining all sources of uncertainty to fully characterise catchment model uncertainty is difficult and rarely undertaken. However, failing to quantify and report uncertainties can lead to overconfidence in the evidence produced by catchment modelling and limits the ability to make risk management-based decisions.

This study aimed to investigate the feasibility of fully empirical catchment *E. coli* models as an alternative to process-based models. This class of model is extremely simple and represents all processes leading to *E. coli* concentrations and loads in a receiving environment as the function of type of land (hereafter land-types) in the upstream catchment. While this approach is an extremely simplified representation of reality, it offers some advantages in terms of transparency, ease of implementation, and defensibility as well as more easily estimated model uncertainty.

We attempted to derive empirical models that can be used to predict *E. coli* concentrations (cfu 100 ml^{-1}) and loads as yields (cfu ha^{-1} yr^{-1}) as a function of the proportions of the upstream catchment occupied by land (see Table A), respectively. These models are expressed mathematically as:

$$Y = ETY_1P_1 + ETY_2P_2 + ETY_3P_3 + \cdots ETY_mP_m + PS_Y$$

$$C = ETC_1P_1 + ETC_2P_2 + ETC_3P_3 + \cdots ETC_mP_m + PS_C$$

where, Y is the yield (cfu ha^{-1} yr^{-1}) and C the median concentration (cfu 100 ml^{-1}) at the evaluation point, and PS_Y and PS_C are the yields or concentrations associated with upstream point sources, $P_1, P_2, P_3, \dots P_m$ are the proportions of catchment area occupied by each land-type in the upstream catchment, and $ETY_1, ETY_2, ETY_3 \dots ETY_m$ and $ETC_1, ETC_2, ETC_3 \dots ETC_m$ are the empirically derived parameters for yields and concentrations, respectively. The parameters are derived from statistical regression models fitted to observations of *E. coli* at all available river water quality monitoring stations nationally. The parameters can be interpreted as the expected proportional contribution of each land-type to concentration or yield. Alternatively, the parameters can be interpreted as the expected concentration or yield for a catchment comprised of only that land-type.

We failed to define a satisfactory empirical model for *E. coil* yields but did derive a satisfactory model for median concentrations. The model allows predictions of site-median *E. coli* concentrations, and the lower and upper bounds of the 90% prediction interval of this value, to be made for any catchment in New Zealand using the above equation and parameters (ETC) shown in Table A. The empirical model can also be used to simulate effects of land use change or mitigation actions on median *E. coli* concentrations by changing the proportion of catchment occupied by particular land-types and by applying appropriate changes to the model parameters.

The empirical models presented in this report provide simple and easily used tools that can be applied at any location within New Zealand. We have developed a dataset that provides proportions of catchment area occupied by each land-type defined by the *E. coli* concentration model for all segments of national digital river network[1] (catchment area >10 km^2). These data could also be used for national- and regional-scale assessments that aim to rapidly assess impacts of land use and land management scenarios at any location in New Zealand.

An important caveat that applies to the empirical models is associated with the national extent of the dataset that was used to derive the parameters. Because the water quality station data were limited, we were only able to derive robust ETCs for a limited number of land-types. This means the models only coarsely resolve landscape-scale variation in the contribution of land to *E. coli* concentrations in downstream receiving environments. An additional caveat that applies to the empirical model is that as the spatial extent of a modelled domain reduces (e.g., they are used at the scale of individual catchments), there is a reduction in applicability of the model parameters (ETCs). This is because the empirical models were fitted to a national dataset and the ETC values are therefore national in scope. As the model domain decreases, the parameters will be potentially biased (not represent the conditions associated with the smaller model domain).

We note that catchment models are used in scenario analyses. In this type of application, the objective is generally to evaluate relative differences in concentrations and loads between scenarios. It is reasonable to assume that the uncertainty of relative differences will be less than the uncertainty of predictions of absolute quantities. However, this study did not test the validity of this assumption, and this would be a useful direction for future research.

[1] The digital network associated with the River Environment Classification (version 2.4) described by Snelder and Biggs (2002).

Table A. ETC parameters for the empirical catchment E. coli concentration model for each of the 17 Types represented by the model. The values can be interpreted as the contribution of each land-type to the best estimate, upper and lower bounds for the 90% prediction intervals of E. coli *concentration (cfu 100 ml^{-1}). Note that the best estimate and the prediction interval limits are defined by three separate regression models that are independent of each other. This means that the parameter values are not necessarily expected to follow an order that is consistent with the point on the probability distribution that each model pertains to.*

Land-type	Best estimate	Prediction interval lower bound	Prediction Interval upper bound
Bare	0	0	0
Cropland	159	113	477
Dairy_PoorlyDrained	381	118	1987
Dairy_WellDrained	391	104	987
ExoticForest_HighElev_PoorlyDrained	25	21	-46
ExoticForest_HighElev_WellDrained	16	-38	139
ExoticForest_LowElev_PoorlyDrained	128	37	205
ExoticForest_LowElev_WellDrained	8	41	-104
Natural_HighElev	-1	3	17
Natural_LowElev	147	11	475
Orchard&Vineyard	310	61	240
Sheep&Beef_HighElev_PoorlyDrained	100	13	138
Sheep&Beef_HighElev_WellDrained	43	2	363
Sheep&Beef_LowElev_PoorlyDrained	349	48	865
Sheep&Beef_LowElev_WellDrained	183	89	724
Urban	754	115	2815
Water	0	0	0

1 Introduction

Faecal indicator bacteria (FIB) are a key freshwater contaminant that the National Policy Statement – Freshwater Management (NPS-FM, NZ Government 2023) requires regional councils to manage. FIB are discharged into freshwater by animals, effluent and waste water discharges, and stormwater run-off. The presence of FIB in freshwater indicates that other pathogens that are harmful to humans may be present. The concentration of a specific FIB, the bacterium *Escherichia coli* (*E. coli*), is used as an indicator of human or animal faecal contamination and the risk of infectious human disease from waterborne pathogens in water used for contact recreation and drinking. *E. coli* is sampled at over 1,000 long term river water quality monitoring stations across New Zealand every month. The concentration of *E. coli* in each sample is measured in the laboratory and reported as colony forming units (cfu) or most probable number (MPN) per 100 ml. The different units reflect different test methods, but they both represent the number of viable bacteria in a 100 millilitre sample of water.

The NPS-FM includes two *E. coli* attributes that must be used by regional councils to set target states in all rivers and at specified primary contact sites. The first attribute (NPS-FM Appendix 2A) is applicable to all rivers and comprises four statistics that are derived from monthly observations, including the median, the percentage of samples exceeding 260 and 540 *E. coli* 100 ml^{-1} and the 95^{th} percentile. Monitoring stations are graded into bands between A and E based on these statistics and the bands represent differing levels of infection risk associated with exposure to the site. The second attribute (NPS-FM Appendix 2B) is applicable to primary contact sites only. The 95^{th} percentile of *E. coli* concentrations of samples collected at primary contact sites during the bathing season are graded Excellent to Poor. Regional councils are required to use observation data to establish the baseline state of the *E. coli* attributes at monitoring stations and primary contact sites and to set target states for all rivers and lakes in their regions. If current states are not compliant with the target state (i.e., *E. coli* concentrations exceed that allowed by the chosen band), councils are required to set limits and/or develop management plans to achieve the target state.

Because there are potentially many limits and actions that can be used to achieve target states, finding the most acceptable solution involves exploration of options. Integral to this is the use of catchment models that provide a basis for simulating the impacts of land use and management on contaminant levels in freshwater receiving environments. Catchment models can be used to link *E. coli* losses from multiple sources in a catchment, including diffuse losses from land and point source losses, to concentrations and loads in downstream receiving environments (e.g., Semadeni-Davies and Elliott 2016; Muirhead 2019). Once set up and calibrated, catchment models can be used to explore the potential impact of alternative policy options, or planned mitigation actions, relating to changes in land management, land use and point sources.

Process-based catchment models are based on a mass balance, in which it is assumed that observed instream loads are the sum of the upstream source contributions, less any net loss of mass during transport down the drainage network. The net loss is referred to as "attenuation". Attenuation of *E. coli* occurs due to natural losses or decay of organisms along transport pathways, which may be due to a number of processes including die-off (Wilkinson et al. 2011). Setting up process-based catchment *E. coli* models involves defining parameters, which at least quantify rates of *E. coli* loss from multiple sources (including land) and attenuation rates. Depending on the level of detail of the processes a model represents, there are significant difficulties associated with calibration of process-based *E. coli* models in New

Zealand due to inadequate data and a lack of knowledge of the mechanisms that influence concentrations and fluxes of organisms through time. For example, Wilkinson et al. (2011) showed that pulses of *E. coli* in the Motueka and Sherry rivers were associated with storm-event flows. In addition, this study showed that *E. coli* was transferred to, and from, the riverbed in association with different stages of storm hydrographs. To adequately parameterise these processes in a catchment model would require more frequent *E. coli* observations than are generally available.

We propose a new class of purely empirical catchment *E. coli* models (hereafter 'empirical models'). Like process-based catchment models, the empirical models can be used to make "what if" simulations of changes in catchment *E. coli* loads or concentrations under changes in land use, land management and point sources. The primary difference between these empirical models and process-based water quality catchment models (hereafter 'process-based' models) is how the models are parameterised. Process-based models explicitly represent processes of contaminant loss from land (source losses) and attenuation to produce a prediction of the instream load or concentration at downstream points in the drainage network (Elliott et al. 2016). We acknowledge that process-based models may represent the processes in very lumped forms, for example loss from land may be represented by a single parameter whereas a more detailed representation might include different pathways such as overland and sub-surface flow.

In contrast, our empirical models represent loads and concentrations of *E. coli* at a point in the drainage network as a function of the sum of the proportions of catchment area occupied by different land-types weighted (i.e., multiplied) by constants that represent the contribution from those types. The land-types are defined by a typology that represents spatial variation in the *E. coli* contributions that arise due to differences in land use or cover (e.g., Dairy, Sheep & Beef, Exotic Forest, Natural) and environmental factors that further control *E. coli* contribution rates from land (e.g., elevation, soil drainage). The weights that are applied to each land-type therefore represent the diffuse source loss rate from that land-type and the attenuation of that loss, which means that empirical models represent these two processes with a single parameter.

For this study, we define an Empirical Type Yield (ETY) to be the expected annual load per unit area (units of cfu ha^{-1} yr^{-1}), realised at a point in the drainage network (post attenuation), generated by diffuse losses associated with a specific land-type that is defined by a typology. Tables of ETYs (and their uncertainties) for typologies that represent all land in New Zealand can be estimated using statistical models from the available monitoring data. These values can then be used to construct catchment models for any stream or river location in New Zealand without the need to calibrate parameters representing loss and attenuation. We also define the Empirical Type Concentration (ETC) to be the *E. coli* concentration realised at a point in the drainage network (post attenuation, units of cfu 100 ml^{-1}), generated by diffuse losses associated with a specific land-type that is defined by a typology. The available monitoring data can be used to derive tables of ETCs (and their uncertainties) for typologies that represent all land in New Zealand in the same manner as ETYs. An advantage of the empirical concentration models over the yield models is that there are significantly more sites available across New Zealand for which concentration data are available compared to load estimates. This means that, compared to ETYs, there is more statistical power and therefore the possibility of deriving robust ETCs for a more detailed typology (i.e., more land-types and greater environmental specificity).

The empirical models offer advantages in terms of transparency and ease of implementation. In addition, the uncertainties of the empirical models are easily estimated. The main disadvantage with the empirical approach is the number of land-types, and therefore the spatial detail that can be discriminated by the models and the accuracy of the predictions are limited by the availability and distribution of water quality data. It is also important to acknowledge that the empirical approach greatly simplifies the underlying processes. The simplifications include the lack of spatial detail associated with the distribution of the sources within catchments and the lumping of loss and attenuation in a single parameter. This means that empirical models are cruder than more detailed mechanistic models and cannot simulate certain types of interventions such as the impact of changes to point source discharges.

This study used the available data to attempt to estimate ETYs and ETCs for empirical *E. coli* models for New Zealand. We aimed to produce look up tables of these parameters for each land-type in simple typologies that could be used to predict *E. coli* loads and concentrations in rivers anywhere in New Zealand. These empirically based model parameters can provide alternative, or complementary, approaches to process-based catchment modelling that may be appropriate in some circumstances and that can be used to simulate the impact of management actions on *E. coli* loads and concentrations. A second aim of the study, therefore, was to show how such simulations could be made and describe the changes in *E. coli* for some example scenarios.

2 Overview

The analyses undertaken in this study are shown schematically in Figure 1. The first step involved assembling water quality monitoring data (including flows), point source discharge data and spatial environmental data, including data describing New Zealand's drainage network, which is described in Section 3.

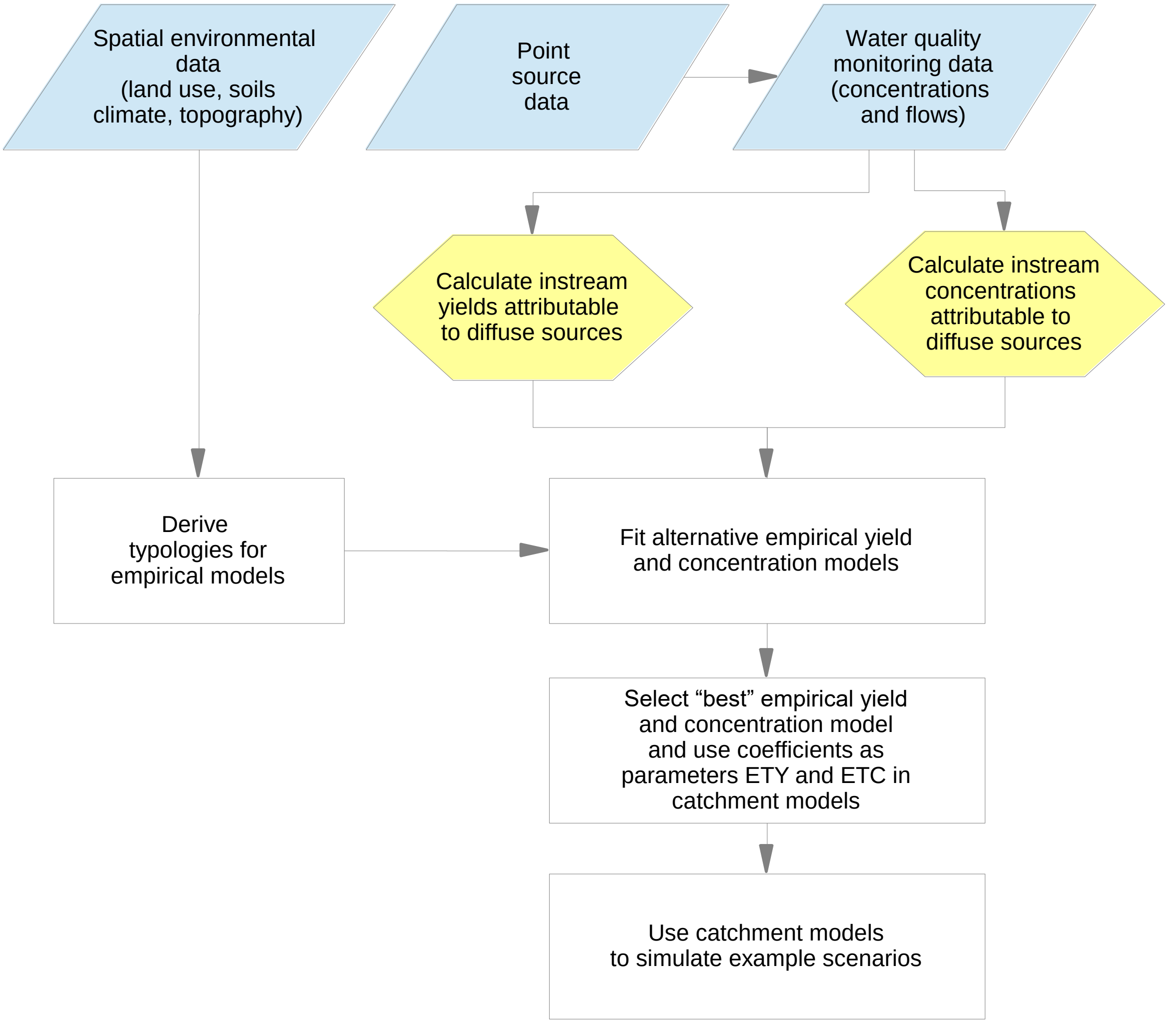

Figure 1. Schematic diagram of the input data and analyses undertaken by this study. The blue parallelograms indicate existing input data. The yellow hexagons indicate preparation of specific input data to subsequent analysis steps. The white rectangles indicate analyses and associated outputs from the study.

The second step was to use the data to derive instream *E. coli* yields and concentrations at water quality monitoring stations that are attributable to diffuse sources (i.e., having removed the contribution from major point sources). In addition, the environmental data were used to construct typologies that were used to describe the composition of land-types in the catchments of all water quality monitoring stations. These steps are described in Sections 4.1 and 4.2.

The third step was to derive parameters for empirical models, an overview of which follows. The empirical models represent the yield or concentration of *E. coli* attributable to diffuse sources at a location in the drainage network as the weighted sum of the proportion of catchment land area occupied by a series of land-types. This is expressed mathematically as follows:

$$Z = \beta_1 P_1 + \beta_2 P_2 + \beta_3 P_3 + \cdots \beta_m P_m \qquad \textit{Equation 1}$$

where Z represents the concentration or yield of *E. coli* at a location in the drainage network, $P_1, P_2, P_{3,} \dots P_m$ are the proportions of catchment area occupied by each land-type in the upstream catchment, and $\beta_1, \beta_2, \beta_3 \dots \beta_m$ are coefficients derived from statistical regression models. The coefficients can be interpreted as the expected proportional contribution of each land-type to concentration or yield. Alternatively, the coefficients can be interpreted as the expected concentration or yield for a catchment comprised of only that type.

The coefficients $\beta_1, \beta_2, \beta_3 \dots \beta_m$ are derived by fitting linear regression models to the available water quality station data. These regression models have the same form as Equation 1, but Z represents data describing the observed yields or concentration at water quality stations after adjustment for point source contributions in the catchment. The regression model predictors are the proportion of the catchments of each water quality station that are occupied by each land-type. This is expressed mathematically as:

$$\begin{bmatrix} Z_1 \\ \vdots \\ Z_n \end{bmatrix} = \begin{bmatrix} P_{1,1} & \cdots & P_{1,m} \\ \vdots & \ddots & \vdots \\ P_{n,1} & \cdots & P_{n,m} \end{bmatrix} \times \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_m \end{bmatrix} \qquad \textit{Equation 2}$$

where *Z* is a 1 x *n* vector of the observed concentrations or yields at the *n* water quality stations after adjusting for any point source discharges in the catchment upstream, the *n* x *m* matrix represents the proportion of the catchment of each of *n* water quality stations (rows) in each of *m* land-types (columns), and *β* is 1 x *m* vector of the fitted regression coefficients for each of the *m* land-types. Note that there is one fitted regression coefficient for each land-type. Note also that the fitted model has no intercept term, which is consistent with concentration or yield being zero if there is no land.

To be clear that the derived coefficients (i.e., $\beta_1, \beta_2, \beta_3 \dots \beta_m$) are used as parameters in empirical concentration and yield models, we refer to them hereafter as empirical type yields (ETY) and empirical type concentrations (ETC). The ETY is the expected annual load per unit area realised at a point in the drainage network (i.e., having been attenuated) that is generated by a specific land-type that is defined by a typology. The units of ETYs are cfu ha^{-1} yr^{-1}. The ETC is the expected concentration, realised at a point in the drainage network (i.e., having been attenuated), that is generated by a specific land-type that is defined by a typology. The units of ETCs are cfu 100 ml^{-1}.

The general form of the empirical catchment water quality models for yield and concentration are given by:

$$Y = ETY_1 P_1 + ETY_2 P_2 + ETY_3 P_3 + \cdots ETY_m P_m + PS_Y \qquad \textit{Equation 3}$$

$$C = ETC_1 P_1 + ETC_2 P_2 + ETC_3 P_3 + \cdots ETC_m P_m + PS_C \qquad \textit{Equation 4}$$

where, Y is the yield and C the concentration at a point in the drainage network, PS_Y and PS_C are the yield or concentration forms of the catchment point source contributions (as described in 4.1.3), $P_1,\ P_2,\ P_{3,} \dots P_m$ are the proportions of area occupied by each land-type in the upstream catchment, and ETY_1, ETY_2, $ETY_3 \ldots ETY_m$ and ETC_1, ETC_2, $ETC_3 \ldots ETC_m$ are the empirically derived parameters for yields and concentrations, respectively.

The fourth set of analyses undertaken by this study were using the derived models to simulate some simple scenarios that concern the impact of mitigations and land use changes on *E. coli* concentrations at the water quality monitoring stations. The analyses and the results are described in Section 6.

3 Data

3.1 River data

3.1.1 Water quality data

River water quality data was obtained for long-term state of environment monitoring from regional council and NIWA records as part of the most recent national state of the environment assessment (Whitehead et al. 2021a). Methods describing the acquisition of river water quality monitoring data and processing are set out in Whitehead et al. (2021a). The dataset included 1030 water quality monitoring stations with *E. coli* observations up to the end of 2020.

3.1.2 Flow data

For the water quality monitoring stations that could be associated with river flow gauging stations, we obtained the entire time series of available mean daily flow data from regional councils and NIWA databases. River flow gauging stations were only reliably identified, and flow records obtained, for a subset (~330) of the water quality monitoring stations.

3.1.3 Point source data

Point source contributions of *E. coli* in the catchment of each monitoring station were obtained so that the total catchment mass loss from point sources on an annual basis could be estimated (see Methods). Point sources were based on a preliminary catalogue of annual point source loads (cfu y^{-1}) collated by NIWA (Annette Semadeni-Davies, *pers comm*.) from previous projects (e.g. Semadeni-Davies et al., 2018).

3.2 Drainage network

The hydrological connectivity for the analysis was defined by a GIS-based digital drainage network comprising rivers and catchment boundaries that is the basis for the River Environment Classification (REC; Snelder and Biggs, 2002). The digital network was derived from 1:50,000 scale contour maps; in version 2 (herein referred to as DN2.4) it represents New Zealand's rivers as 590,000 segments (delineated by upstream and downstream confluences), each of which is associated with a sub-catchment.

3.3 Land use

Land use nationally was defined based on land cover information from the Land Cover Database (LCDB v5.0[2]), and a further separation of pastoral land cover into Dairy and Sheep and Beef land uses based on information about the extent of Sheep and Beef and Dairy farm units obtained from Monaghan et al. (2021). We aggregated LCDB classes into nine simplified land use categories, where the aggregation included judgements about expectations of similarity in *E. coli* loss rates between land cover categories. The reclassified categories are

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

Natural, Exotic Forest, Sheep & Beef, Dairy, Cropland, Orchard & Vineyard, Urban, Bare, and Water and are further described in Table 1.

Table 1. Re-classification of LCDBV5 Name_2018 classes in land use categories used in this study.

Class_2018	LCDBV5 Name_2018	Reclassified category
0	Not land	Bare
1	Built up	Urban
2	Urban Park	Urban
5	Transport Inf	Urban
6	Mines&Dumps	Urban
10	Sand&Gravel	Bare
12	Landslide	Bare
14	Snow&Ice	Bare
15	Alpine Grass	Natural/ Sheep & Beef [2]
16	Gravel&Rock	Bare
20	Lake&Pond	Water
21	River	Water
22	Estuarine	Water
30	Cropland	Cropland
33	Orchard&Vineyard	Orchard&Vineyard
40	High Producing Grass	Dairy/Sheep&Beef [2]
41	Low Producing Grass	Dairy/Sheep&Beef [2]
43	Tussock Grassland	Natural/ Sheep & Beef [1]
44	Depleted Grassland	Natural/ Sheep & Beef [1]
45	Herbaceous Freshwater	Water
46	Herbaceous Saline	Water
47	Flaxland	Natural
50	Fernland	Natural
51	Gorse&Broom	Natural
52	Manuka&Kanuka	Natural
54	Broadleaved Indigenous hardwoods	Natural
55	Sub Alpine Shrubland	Natural
56	Mixed Exotic Shrubland	Natural
58	Grey Scrub	Natural
64	Forest Harvested	Exotic Forest
68	Deciduous Hardwood	Natural
69	Indigenous Forest	Natural
70	Mangrove	Water
71	Exotic Forest	Exotic Forest

Notes:

1. Depleted and Tussock grassland areas that were coincident with Sheep & Beef land use (as defined by Monaghan et al. 2021) were specified as Sheep & Beef; remaining areas were defined as natural.
2. Productive grassland areas that were coincident with Dairy land use (as defined by Monaghan et al. 2021) were specified as Dairy; remaining areas were defined as Sheep & Beef.

3.4 Environmental factors

Muirhead et al. (2023) used three environmental factors, in addition to land use, to describe the risk of *E. coli* loss from land: elevation, drainage and wetness. Muirhead et al. (2023) subdivided the three environmental factors into categories based on nominally defined thresholds. We adopted Muirhead et al.'s (2023) factors and categories as a starting point for defining the typologies used in this study and included an additional factor: temperature.

We obtained spatial data layers that covered all New Zealand and represented each of the four environmental factors. The elevation layer was obtained from the Land Environments of New Zealand dataset (Leathwick et al. 2003). We subdivided elevation into two categories, High and Low, based on the 350 m asl threshold proposed by Muirhead et al. (2023). The temperature layer was represented by mean annual temperature, which was also was obtained from the Land Environments of New Zealand dataset. We subdivided temperature into two categories, Cool and Warm, based on a 10 degrees Celsius threshold. We obtained the drainage layer from the Land Resources Information (LRI) spatial data layers (Newsome et al. 2008). We subdivided the LRIS drainage ordinal scale into two categories. Poorly Drained was defined by drainage classes 1 to 4 and Well Drained was defined by drainage class 5. We note that LRIS drainage ordinal category 4 is described as moderately well drained. We found that aggregating this ordinal category into our Well Drained category degraded the performance of our models and therefore allocated only the LRIS drainage class 5 to our Well Drained category. Finally, we followed Muirhead et al. (2023) and based moisture on the moisture classification of Srinivasan et al. (2021). Srinivasan et al. (2021) subdivide the moisture factor into four categories, dry, moist, wet and irrigated. However, we followed Muirhead et al. (2023) and aggregated irrigated into the moist category based on the assumption that *E. coli* loss would be similar under moist and irrigated conditions. For some typologies, we reduced this to two moisture categories by reclassifying the Srinivasan et al. (2021) Moist category as Wet. Various combinations of the land use categories and the categories associated with the four environmental factors were used to define typologies as described in Section 4.2.

4 Methods

4.1 Preparation of water quality data

4.1.1 Calculating instream loads

We calculated the annual instream loads of *E. coli* (i.e., cfu passing a specific location in a river over a year) at each water quality monitoring station that complied with the following data requirement criteria:

- Observations in at least 8 years in the 10 years up to the end of December 2020
- At least 60 total concurrent observations of flow and concentrations
- At least 80% of all quarters (defined as January-March, April-June, July-September, October-December) in the most recent 10 years.

Rating curve methods were used to calculate the instream loads at sites that had concurrent *E. coli* concentration observations and river daily mean flow records by (1) identifying the best rating curve method (out of four possible alternatives) for each site (through manual inspection of all possible rating curves for each site), and then (2) calculating loads by combining the best rating curve with the daily flow time series. A full description of the load calculation methodology is provided in Snelder et al. (2023).

We used all available flow-concentration observations at each site to characterise the rating curves and set temporal trend terms in the underlying rating curve models so that the load calculations represented the expected mean annual load for 2020. Setting temporal trend terms to a fixed year (for those models that use time-variable components) means that trends were accounted for in the calculation of loads. We also estimated 95% confidence intervals for the estimated instream loads, following a bootstrapping procedure (described in Snelder et al. 2023). For the following analysis, instream loads are generally reported as instream yields, which are the instream load divided by the upstream catchment area, with units of cfu ha^{-1} yr^{-1}.

4.1.2 Calculating median concentrations

We characterised *E. coli* concentrations at each water quality monitoring station as the median of all monthly observations for the 5-year period ending December 2020. The statistical precision of the median depends on the variability in the water quality observations and the number of observations. We therefore used filtering rules to restrict the sites that were used in our analysis to those for which the median could be calculated with reasonable precision. For a given level of variability, the precision of the median increases with the number of observations. As a general rule, the rate of increase in the precision of compliance statistics reduces for sample sizes greater than 30 (i.e., there are diminishing returns on increasing sample size with respect to precision above this number of observations; McBride 2005). In addition, because water quality observations tend to fluctuate seasonally, the precision of the calculated median is affected by how well each season is represented over the period of record. Our filtering rules therefore restricted site × variable combinations that were used in the analyses to those with measurements for at least 90% of the sampling intervals in that period (at least 56 of 60 months). Site by variable combinations that did not comply with these rules were excluded from the subsequent analysis. The time period and filtering rules are consistent with those used by Whitehead et al. (2021a).

4.1.3 Calculating point source contributions to instream yields and concentrations

Each point source in the dataset described in section 3.1.3 included location information in the form of a unique segment identifier (nzsegment). Point sources were assigned to the digital network based on the segment identifier, and load contributions were accumulated in the downstream direction of the network. Point source yield contributions at all segments of the digital network were estimated from the accumulated point source loads divided by upstream catchment area. Point source concentration contributions at all segments of the digital network were estimated by dividing point source loads by estimates of segment site mean flows (sourced from Woods et al. 2006), with appropriate conversion of units. We then compared this value to the observed median concentration of *E. coli*. Where the estimate of the point source contribution exceeded 5% of the observed median concentration of *E. coli* we removed the site from the analysis. Our rationale was that the exact contribution of point sources to the observed median concentration of *E. coli* cannot be accurately estimated from the load and therefore it was better to discard sites where point sources represented an appreciable contribution to concentration in general.

4.2 Definition of typologies

We primarily defined land-types to be used as predictors (i.e., $P_1, P_2, P_{3,} \dots P_m$ in Equation 1 and 2) using the land use categories described in Section 3.3 (Urban, Exotic Forest, Dairy, Orchard & Vineyard, Natural, Sheep & Beef, Cropland, Water, Bare). When fitting the models, we excluded the Water and Bare land use categories based on the assumption that these categories make a negligible contribution to catchment *E. coli* loss (i.e., we expected ETY and ETC for these land-types to be zero). The credibility of these assumptions was tested as part of the model fitting process by plotting *E. coli* loads and concentrations as a function of the proportion of catchment area occupied by land categorised as Water and Bare.

We defined additional typologies based on further subdivision of some of the land use categories by categorisation of the environmental factors described in Section 3.4, i.e., two elevation categories (Low and High), two drainage categories (Well Drained, Poorly Drained), two temperature categories (Cool and Warm) and three moisture categories (Dry, Moist and Wet). Different combinations of the land use categories with the environmental factors were used to produce typologies with differing numbers of land-types.

The combination of all possible factors and their categories produces a typology with a total of 9 x 2 x 3 x 3 x 2 = 324 potential land-types. We could not reliably estimate regression coefficients for this number (324) of land-types because it is large compared to the size of the fitting dataset (i.e., ~320 sites for yields and ~900 for concentrations). However, all other things being equal, the utility and credibility of a catchment model that includes many land-types is higher than the converse because it accounts for spatial variation in *E. coli* diffuse sources and allows for simulation of more nuanced management actions. We therefore derived a further 16 typologies (in addition to the typology based only on land use categories) that comprised differing numbers of types (i.e., each typology had a differing value of *m* in Equation 4).

The additional typologies were defined by successively subdividing some of the land use categories based on various combinations of the elevation, drainage, temperature and moisture categories. We did not know in advance how many regression coefficients could be reliably estimated (by the statistical modelling process) for each response variable (i.e., concentrations and yields). We therefore defined the typologies and fitted models to them (referred to as model 1, 2, 3 etc) and inspected the fitted coefficients to determine the “best” model (see section 4.3.3).

The definition of the typologies was subjective but was guided by the risk matrix of Muirhead et al. (2023) and expert opinion. Expertise was used to consider how the information available in the fitting dataset would produce fitted regression coefficients that reliably represent differences in *E. coli* contributions to yields and concentrations under differing land use and environmental conditions. Regression coefficients are most likely to be reliably estimated for land-types that are consistently occurring (i.e., are represented in the catchments of many sites) and have wide variation in occupancy across the fitting datasets (i.e., for which the predictor P_m covers a wide range of non-zero values). We expected that the land use categories Sheep & Beef, Exotic Forest and Natural would the most consistently occurring non-zero and widely variable land use categories because these land uses can be dominant at the scale of catchments. We therefore expected that further subdivision of these land use categories by the environmental factors were most likely to produce reliable regression coefficients. We expected that Dairy would have narrower ranges of occupancy given that this land use is rarely dominant at the scale of catchments of the monitoring stations and therefore expected that further subdivision of this land use category was less likely to produce reliable regression coefficients. The land use categories Cropland, Orchard & Vineyard and Urban tend to have low occupancy and we therefore did not expect that further subdivision of these land use categories would produce reliable regression coefficients. The 17 typologies (labelled 1 to 17) therefore included differing subdivisions of each of Sheep & Beef, Exotic Forest, Natural and Dairy land use categories by Elevation, Drainage, Moisture and Temperature. The number of land-types and a description of how these were defined is provided in Table 2.

Each typology comprised a series of land-types that were defined by the combination of factor categories, for example Sheep&Beef_LowElevation_PoorlyDrained_Wet. A complete description of the land-types defined by each of the typologies shown in Table 2 is contained in Appendix A (Table 5).

The next step was to evaluate the proportion of land in each of the land-types that occur the catchment of all monitoring stations for each of the 17 typologies. To achieve this, we converted all spatial layers into coincident raster layers with 200m x 200m cells. The raster resolution was a practical decision made for processing efficiency and took into consideration the requirement for national coverage, differences in the source data precision as well as a spatial scale that was commensurate with the typical smallest productive land use entities. As a test of the imprecision introduced by this choice, we compared estimates of rasterised catchment areas against catchment areas defined for the DN2.4. We found that for catchment areas greater than approximately 2 km^2 differences in the estimates were small (<5%).

The raster layers were overlaid to generate maps of each typology and further overlaid with a coincident raster layer of the DN2.4 sub-catchments. The area associated with each land-type within each sub-catchment was evaluated. Sub-catchment land-type areas were then accumulated in the downstream direction of the DN2.4 network to derive the total upstream areas associated with each land-type for each network segment. The accumulated type areas for each network segment were then normalised by the network segment upstream catchment area to provide estimates of the proportion of upstream area occupied by each type for each network segment.

Table 2. The 17 typologies that were tested, including the number of land-types in each and a description of how the land-types were defined.

Typology	Number of land-types	Description of typology
1	9	Subdivision of land use into nine categories
2	12	Subdivision of land use into nine categories. Further subdivision of Sheep & Beef, Exotic Forest and Natural into two drainage categories.
3	13	Subdivision of land use into nine categories. Further subdivision of Dairy, Sheep & Beef, Exotic Forest and Natural into two drainage categories.
4	12	Subdivision of land use into nine categories. Further subdivision of Sheep & Beef, Exotic Forest and Natural into two elevation categories.
5	13	Subdivision of land use into nine categories. Further subdivision of Dairy, Sheep & Beef, Exotic Forest and Natural into two elevation categories.
6	12	Subdivision of land use into nine categories. Further subdivision of Sheep & Beef, Exotic Forest and Natural into two temperature categories.
7	13	Subdivision of land use into nine categories. Further subdivision of Dairy, Sheep & Beef, Exotic Forest and Natural into two temperature categories.
8	15	Subdivision of land use into nine categories. Further subdivision of Sheep & Beef, Exotic Forest and Natural into three moisture categories.
9	17	Subdivision of land use into nine categories. Further subdivision of Dairy Sheep & Beef, Exotic Forest and Natural into three moisture categories.
10	17	Subdivision of land use into nine categories. Further subdivision of Sheep & Beef, and Exotic Forest into two elevation and two drainage categories. Subdivision of Natural into two elevation categories. Subdivision of Dairy into two moisture* categories.
11	17	Subdivision of land use into nine categories. Further subdivision of Sheep & Beef, and Exotic Forest into two elevation and two moisture categories. Subdivision of Natural into two elevation categories. Subdivision of Dairy into two moisture categories.
12	21	Subdivision of land use into nine categories. Further subdivision of Dairy, Sheep & Beef, and Exotic Forest into two elevation and two drainage categories. Subdivision of Natural into two elevation categories.
13	21	Subdivision of land use into nine categories. Further subdivision of Dairy, Sheep & Beef, Natural and Exotic Forest into two temperature and two drainage categories.
14	23	Subdivision of land use into nine categories. Further subdivision of Sheep & Beef into two elevation, two drainage categories, and two moisture* categories. Subdivision of Natural and Exotic Forest into two elevation and two drainage categories. Subdivision of Dairy into two drainage categories.
15	25	Subdivision of land use into nine categories. Further subdivision of Sheep & Beef into two elevation, two drainage categories, and two moisture* categories. Subdivision of Natural and Exotic Forest into two elevation and two drainage categories. Subdivision of Dairy into two elevation and drainage categories.
16	29	Subdivision of land use into nine categories. Further subdivision of Natural and Sheep & Beef into two elevation, two drainage categories, and two moisture categories. Further subdivision of Exotic Forest into two elevation and two drainage categories. Further subdivision of Dairy into two elevation and two drainage categories.
17	33	Subdivision of land use into nine categories. Further subdivision of Natural, Exotic Forest and Sheep & Beef into two elevation, two drainage categories, and two moisture categories. Further subdivision of Dairy into two elevation and two drainage categories.

* Two moisture categories were defined by reclassifying the Srinivasan *et al.* (2021) Moist category as Wet.

4.3 Derivation of empirical model parameters

4.3.1 Statistical modelling

We attempted to derive the regression coefficients shown in Equation 2 for each response variable (*E. coli* yields and concentrations). Prior to fitting both the yield and concentration models, we subtracted the estimate of the point source contribution from each of the water quality station yields so that the response (i.e., $Y_{1,n}$) was only representing the attenuated diffuse sources of *E. coli*.

There are two considerations with the process of fitting the statistical model expressed in Equation 2. First, the distribution of site concentrations and yields at the water quality stations will generally not be normally distributed. Normally distributed data (more specifically, regression residuals) is a requirement of ordinary least squares regression (OLS). It is therefore common to apply transformations, such as a logarithmic transformation, to normalise the response when fitting OLS models. However, transformation of the concentration or yields would mean that the fitted regression coefficients (i.e., $\beta_1, \beta_2, \beta_3 \dots \beta_n$) could not be interpreted as ETCs or ETYs for each land-type. We therefore used quantile regression instead of OLS to estimate the regression coefficients. Quantile regression is often used when the conditions of OLS are not met (Cade and Noon 2003). Whereas OLS estimates the conditional mean[3] of the response variable given some predictor variables, quantile regression estimates a specified quantile of the data. We fitted the model represented by Equation 2 to the median (i.e., the 0.5 quantile) value using quantile regression. The prediction from the model should be considered as an estimate of the median, conditional on the predictors (i.e., 50% of cases can be expected to be greater than or less than the prediction).

Because quantile regression is non-parametric, the fitted model does not describe the probability distribution within which prediction will lie. However, quantile regression models can be fitted to any quantile of the data. Therefore, in addition to fitting a model to the median (0.5 quantile), we also fitted models to the 0.05 and 0.95 quantiles (of the site-median concentrations and yields) to provide the lower and upper bounds of the 90% prediction interval[4]. Note that the median and the prediction interval limits are defined by three separate regression models that are independent of each other. This means that the coefficient values of the three models are not necessarily expected to follow an order that is consistent with the point on the probability distribution that each model pertains to. Quantile regression models were fitted using the quantreg package of the R Statistical Software (R Core Team 2023).

The second complication is that the predictors ($P_1, P_2, P_{3,} \dots P_m$) are what is referred to as compositional data. That is, $P_1, P_2, P_{3,} \dots P_n$ represent the composition of the catchment land as the proportions occupied by each type. Because the types are exhaustive and the predictors represent proportions, they sum to one and, therefore, the set of all proportions includes redundant information (e.g., $P_n = 1 - \sum_{i=1}^{i=n-1} P_i$). This means that another condition of multivariable regression, that the predictors are independent, is violated. Non-independence of predictors is referred to as multicollinearity because the implication is that there is correlation between the predictors.

When there is multicollinearity in the predictors of a regression model, the estimated coefficients (i.e., values of $\beta_1, \beta_2, \beta_3 \dots \beta_m$) can become sensitive to small changes in the model.

[3] The conditional mean of a random variable is its expected value – the value it would take "on average" over an arbitrarily large number of occurrences – given a certain set of "conditions". In a multiple linear regression model, these conditions are defined by the values of the independent (i.e., predictor) variables.
[4] The prediction interval indicates the range a future individual observation will fall.

For example, small changes in the predictors or cases that are included in the model can dramatically change the coefficient values or even their signs. This means that multicollinearity reduces the precision of the estimated coefficients and increases their *p*-values (i.e., decreasing their statistical significance and reducing confidence in the estimated values). Multicollinearity can therefore make it difficult to justify the model, and this increases as the severity of the multicollinearity increases. It is noted that multicollinearity is a problem for interpretation of the estimated coefficients but does not affect the predictions or the goodness-of-fit (performance) statistics of the model (Neter et al. 2004).

An option to avoid the problem of collinearity is to remove some of the strongly correlated predictors. In this analysis, we were wanting to evaluate the coefficients for all predictors, to provide parameter values for all land-types and, therefore, removing some of the predictors was not an option. However, the problems caused by multicollinearity reduce with increasing dataset size because sampling error reduces and precision increases as sample size increases (Mason and Perreault 1991). Because the datasets in this project were reasonably large, we adopted the approach of retaining all predictors and carefully inspecting the fitted coefficients and their standard errors to ensure that they were generally reasonable (i.e., were not so large as to render the coefficient unreliable). We also used cross validation (see Section 4.3.2) to generate multiple instances of the fitted coefficients and used these to evaluate the sensitivity of the coefficients to the fitting data.

4.3.2 Objective evaluation of the models

We fitted the 17 models described by Table 2 to each response variable (*E. coli* concentrations and yields) and then evaluated these to determine the “best” model based on four aspects: (1) the predictive performance, (2) the predictive performance compared to alternative frequently used models, (3) the ability to estimate the 95% confidence interval, and (4) the stability of the fitted coefficients. These evaluations were carried out based on independent predictions of the response variables made for each water quality station by cross validation. Cross validation was carried out by first subdividing the dataset (representing the concentrations and yields at each water quality station) randomly into 10 equally sized subsets that are hereafter referred to as “folds”. We fitted 10 “realisations” of each model (i.e., of the 0.05, 0.5 and 0.95 quantiles) by excluding one of the folds each time (the held-out fold). We used each of the 10 fitted models to predict the response for the associated held-out fold to obtain objective predictions (i.e., predictions for water quality stations that were not used in fitting the model) for each quantile and each water quality station.

We evaluated the predictive performance of the models using two statistics: Nash-Sutcliffe efficiency (NSE) and percent bias (PBIAS). NSE indicates how closely the observations coincide with predictions (Nash and Sutcliffe 1970). NSE values range from $-\infty$ to 1. A NSE of 1 corresponds to a perfect match between predictions and the observations. A 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). PBIAS is computed as the sum of the differences between the observations and predictions divided by the sum of the observations (Moriasi et al. 2007). The normalisation associated with NSE and PBIAS allows the performance of the models to be compared to criteria proposed by Moriasi et al. (2015), outlined in Table 3.

Table 3. Performance ratings for the measures of model performance used in this study. The performance ratings are from Moriasi et al. (2015).

Performance Rating	NSE	PBIAS
Very good	NSE > 0.65	|PBIAS| <15
Good	0.50 < NSE ≤ 0.65	15 ≤ |PBIAS| < 20
Satisfactory	0.35 < NSE ≤ 0.50	20 ≤ |PBIAS| < 30
Unsatisfactory	NSE ≤ 0.35	|PBIAS| ≥ 30

The second evaluation was a comparison of the NSE and PBIAS for the 0.5 quantile models with the same performance statistics achieved for equivalent random forest (RF) models. RF is a machine-learning method based on an ensemble of regression trees (Breiman 2001; Cutler et al. 2007). Because RF models can include many predictor variables and automatically fit non-linear relationships and high-order interactions, they achieve high accuracy. This means that RF models are an accepted method of making model-based predictions of current river concentrations and yields based on data obtained for water quality stations (e.g., Snelder et al. 2020; Whitehead et al. 2021b). RF-based models were fitted to the same response variable data as used in this study using a large set of predictors representing various aspects of the climate, topography, geology, land cover, and land use of the catchments of the water quality stations (see Snelder et al. 2023 for details). We expected that the RF models would perform better than our 0.5 quantile models but note that RF models do not produce interpretable coefficients that can be used as parameters in catchment nutrient water quality models. The purpose of the RF models, therefore, is to provide a fair benchmark against which to compare model performance.

The third evaluation was of the estimation of the 90% prediction interval by the 0.05 and 0.95 quantile models. From the cross-validation outputs we evaluated the proportion of the predictions of the median that were less than or greater than the predicted 0.05 and 0.95 quantiles, respectively (i.e., the proportion of the predictions of the median that fell outside the 90% prediction interval). We expected that on average (over the 10 cross validation realisations) 10% of the estimates of the median would lie outside of the 90% prediction interval.

The fourth evaluation was of the stability of the fitted coefficients. From the cross-validation outputs, we retained the fitted coefficients (i.e., values of $\beta_1, \beta_2, \beta_3 \dots \beta_m$) for each realisation. We compared the mean values and the standard deviation of the fitted coefficients over the 10 realisations to the coefficients and their standard errors estimated for the full models (i.e., the models fitted to the entire dataset). We interpreted agreement of the mean and standard deviation of the coefficients estimated from the cross validation with their counterparts estimated from the full dataset to indicate that collinearity in the predictors was not causing sensitivity in the estimated coefficients (i.e., they were stable and reliable).

We undertook a final evaluation of the plausibility of the fitted model coefficients by comparing them to the response variable (*E. coli* yields and concentrations) at the water quality stations. For each land-type we plotted the response variable against the proportion of catchment occupied, with the fitted model coefficient plotted at the position indicating a proportion occupancy of 1. We expected that, for each type, the fitted model coefficient would tend to be consistent with the observed response variables at sites having high occupancy (i.e., the fitted coefficients for each type would be similar to the observed response variables in catchments that are dominated by that type). This expectation is consistent with the physical meaning of the fitted model coefficients as the expected proportional contribution of each type to concentration or yield.

4.3.3 Determination of the best model

For each response variable (*E. coli* yields and concentrations), we inspected the fitted models and, based on several considerations, selected a single “best” model. First, we used the NSE values as a measure of predictive performance of the models and assumed that, all other things being equal, better models have higher NSE values. Second, we considered that, all things being equal, better models would be derived from more detailed typologies (i.e., typologies with more land-types and therefore fitted models with more predictors) because this would better discriminate variation in land use and environmental factors.

Third, we considered coefficient values of credible models would be positive (i.e., we expected all land-types to contribute *E. coli* apart from Bare and Water). We also considered that a ranking of the fitted coefficient values of the most credible models would be consistent with the *E. coli* runoff risk ranking matrix of Muirhead et al. (2023). Muirhead et al.'s (2023) risk matrix proposes that risk decreases with land use categories in the following order: Urban, Pastoral (i.e., Dairy and Sheep & Beef), Horticulture, Arable, Exotic Forest and other non-productive land uses. The risk matrix proposes that for pastoral land uses, the *E. coli* runoff risk is higher for low elevation locations than high elevation due to farm intensity, with stocking rates reducing as farm elevation increases. The risk matrix also proposes that risk increases with decreasing drainage and increases with increasing wetness primarily due to increasing risk of overland flow, this being the dominant pathway for *E. coli* losses from land.

Our fourth consideration was the significance of the fitted coefficients. We interpreted these as measures of confidence in their representation of the true value of the ETC or ETY. We considered that, all other things being equal, significant fitted coefficients were preferable to non-significant coefficients and models with greater numbers of significant coefficients were preferable to the converse. We also expected that the number of significant coefficients would decrease (i.e., p-values would increase) with increasing numbers of predictors (i.e., land-types) due to decreasing statistical power.

For each response variable, we considered that the “best” model represents a trade-off between the number of land-types, the NSE value, the consistency of the coefficient values with our prior understanding, and the proportion of the coefficients that were <0 and were significant ($p < 0.05$). This means that the best model is a judgement that does not mean that other models and their associated typologies are not useful or better in some circumstances. In addition, we note that given different or updated datasets, different models would be derived. Therefore, we declare a “best” model in this study but do not regard this as the only possible model.

Finally, we quantified the overall uncertainty of the best model as 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). We calculated the RMSD after log (base 10) transformation of the predicted and observed values to achieve approximately normal error distributions.

5 Results

5.1 Water quality monitoring station concentrations and instream yields

The total number of water quality monitoring stations used in this study, i.e., that met the minimum data requirements and had a catchment area >2 km^2, for concentrations and yields was 869 and 320, respectively (Figure 2 and Figure 3). The median *E. coli* concentrations varied over three orders of magnitude (Figure 4), and instream yields varied over two orders of magnitude (Figure 5).

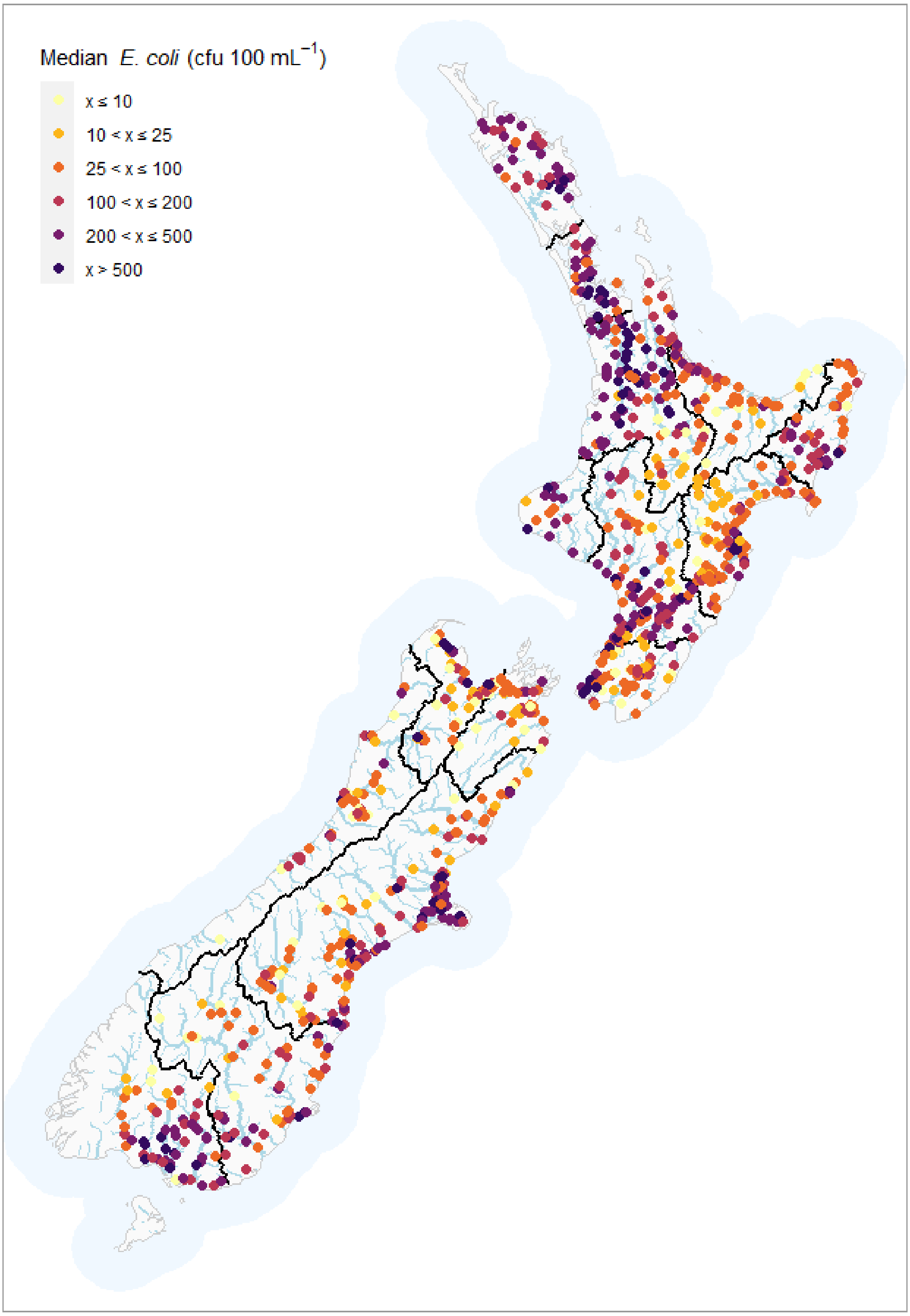

Figure 2: Locations of water quality monitoring stations with median E. coli *concentration data. The sites are coloured to indicate the evaluated site median concentrations (cfu 100 ml^{-1}).*

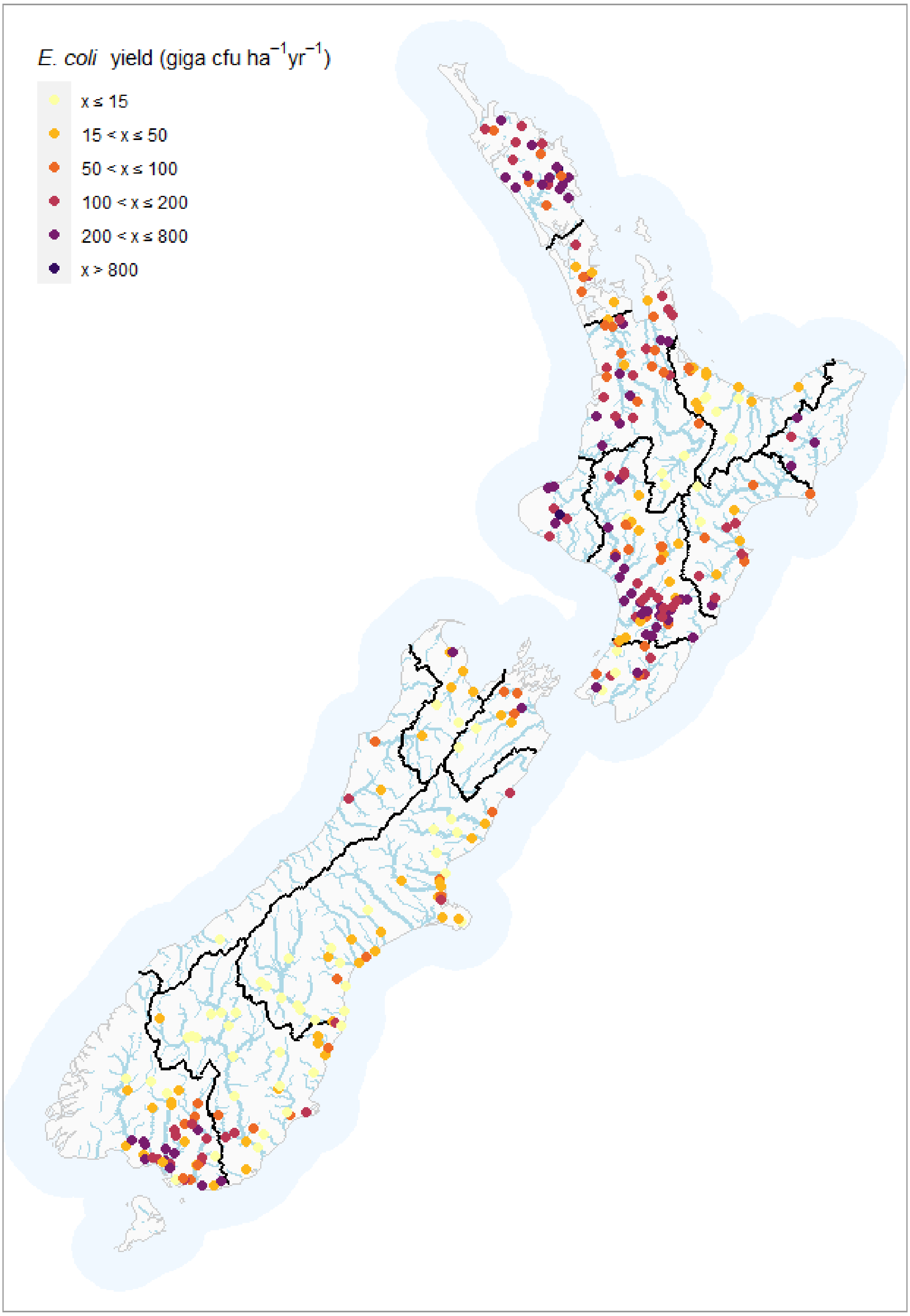

Figure 3: Locations of water quality monitoring stations with E. coli *yield data. The sites are coloured to indicate the evaluated site* E. coli *yields (giga cfu ha^{-1} year^{-1}).*

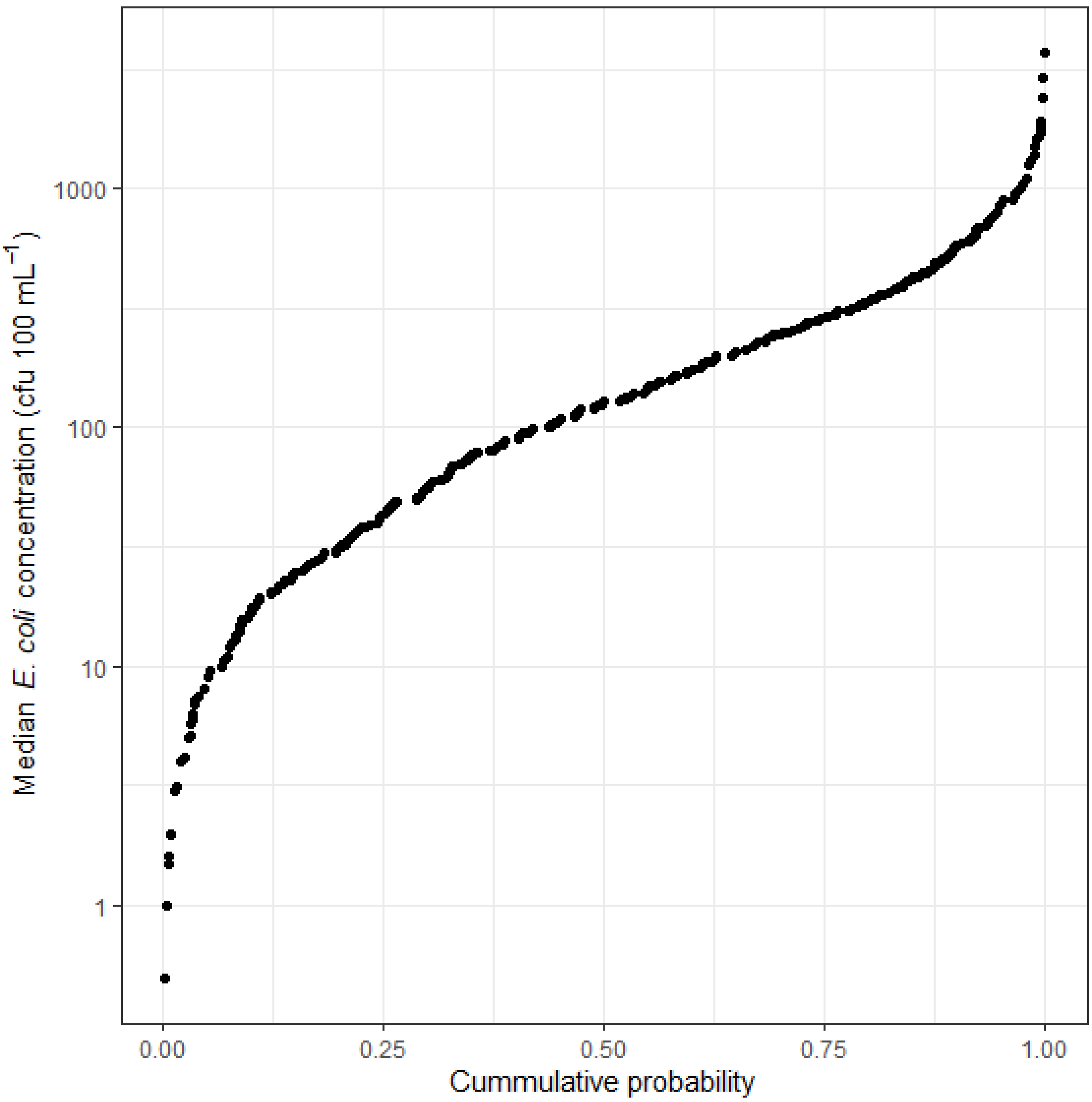

Figure 4: Cumulative distribution of median E. coli *concentrations at water quality monitoring stations. See Figure 2 for site locations. Note that the y-axis has a log-scale.*

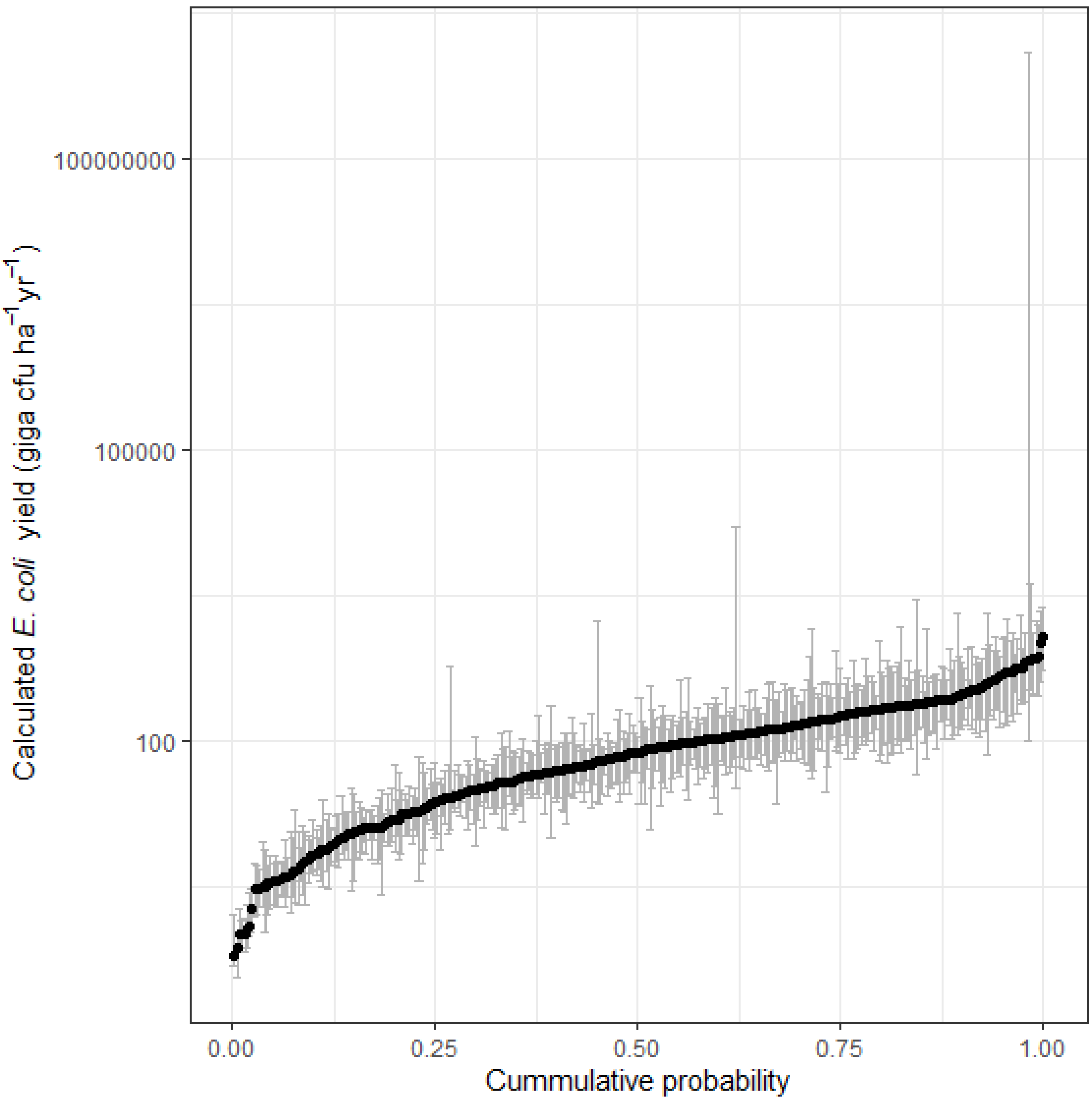

Figure 5: Cumulative distribution of estimated instream E. coli *yields at water quality monitoring stations. See Figure 3 for site locations. The error bars indicate the 95% confidence interval for the instream yields. Note that the y-axis has a log-scale.*

5.2 Empirical concentration model

For sites with concentration data, the Sheep & Beef, Exotic Forest and Natural land use types were the most consistently occurring (i.e., are represented in the catchments of many sites) and had the widest variation in occupancy (i.e., occupancy ranged from zero to 100% of catchment area, Figure 6). The land use categories Bare, Cropland, Orchard & Vineyard and Water tended to have low occupancy (i.e., the proportion of catchment area occupied by these categories was most commonly zero) and variation across sites was restricted (i.e., very few sites had >50% occupancy by any of these categories). Visual inspection of Figure 6 supported our assumption that contributions to *E. coli* by land categorised as Bare or Water is negligible.

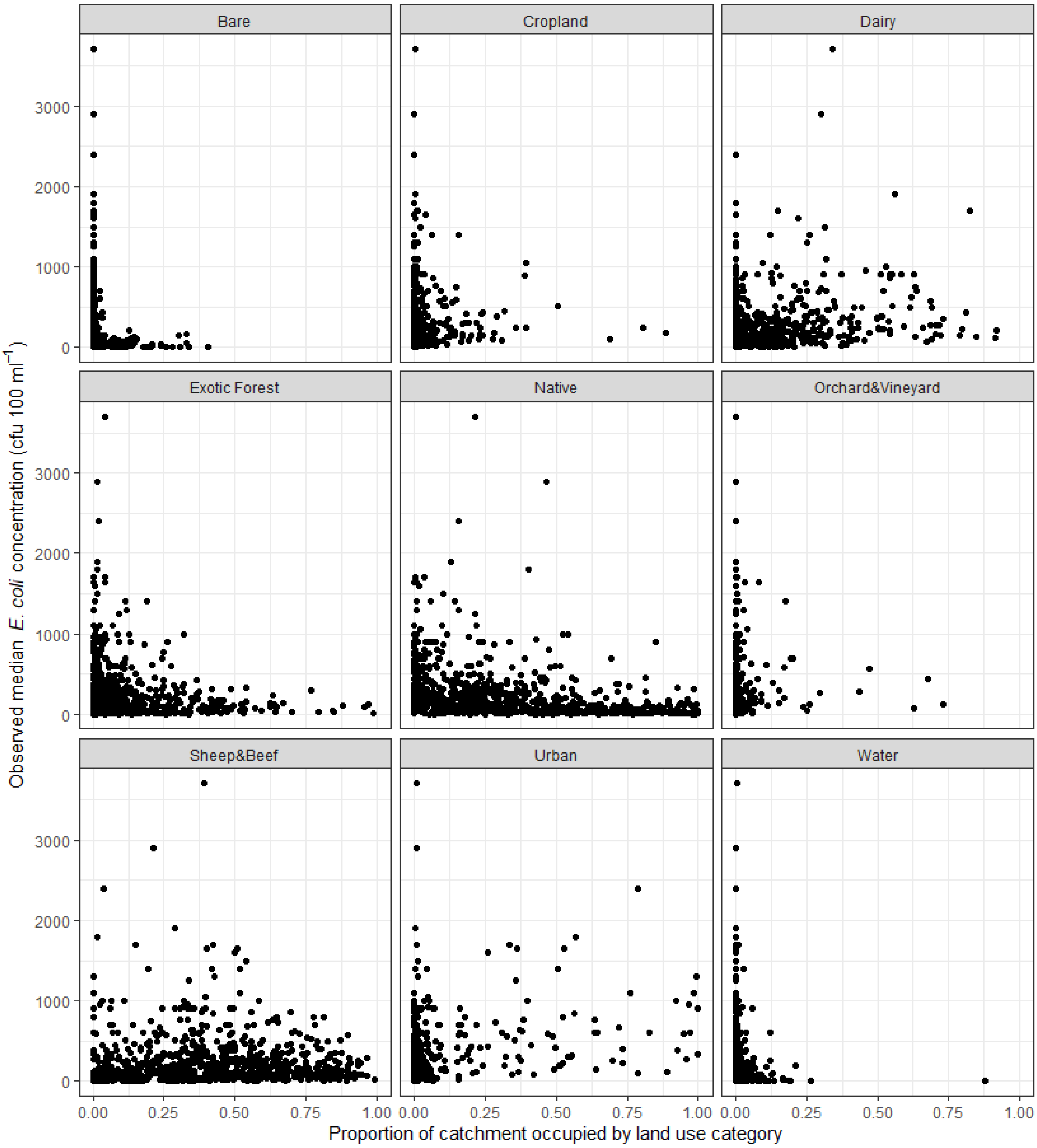

Figure 6. Observed site median E. coli *concentrations as a function of the proportion of catchment area occupied by the nine land use categories.*

The median *E. coli* concentrations (model response) for the 869 sites were not normally distributed (they were right skewed), justifying the use of quantile regression (Figure 4). The proportion of significant fitted coefficient values decreased with increasing numbers of predictors for models pertaining to all three quantiles (i.e., median (0.5 quantile), 0.05 and 0.95 quantiles; Figure 7). There was also an increasing number of negative fitted coefficient values as the number of predictors included in the models increased (Figure 7).

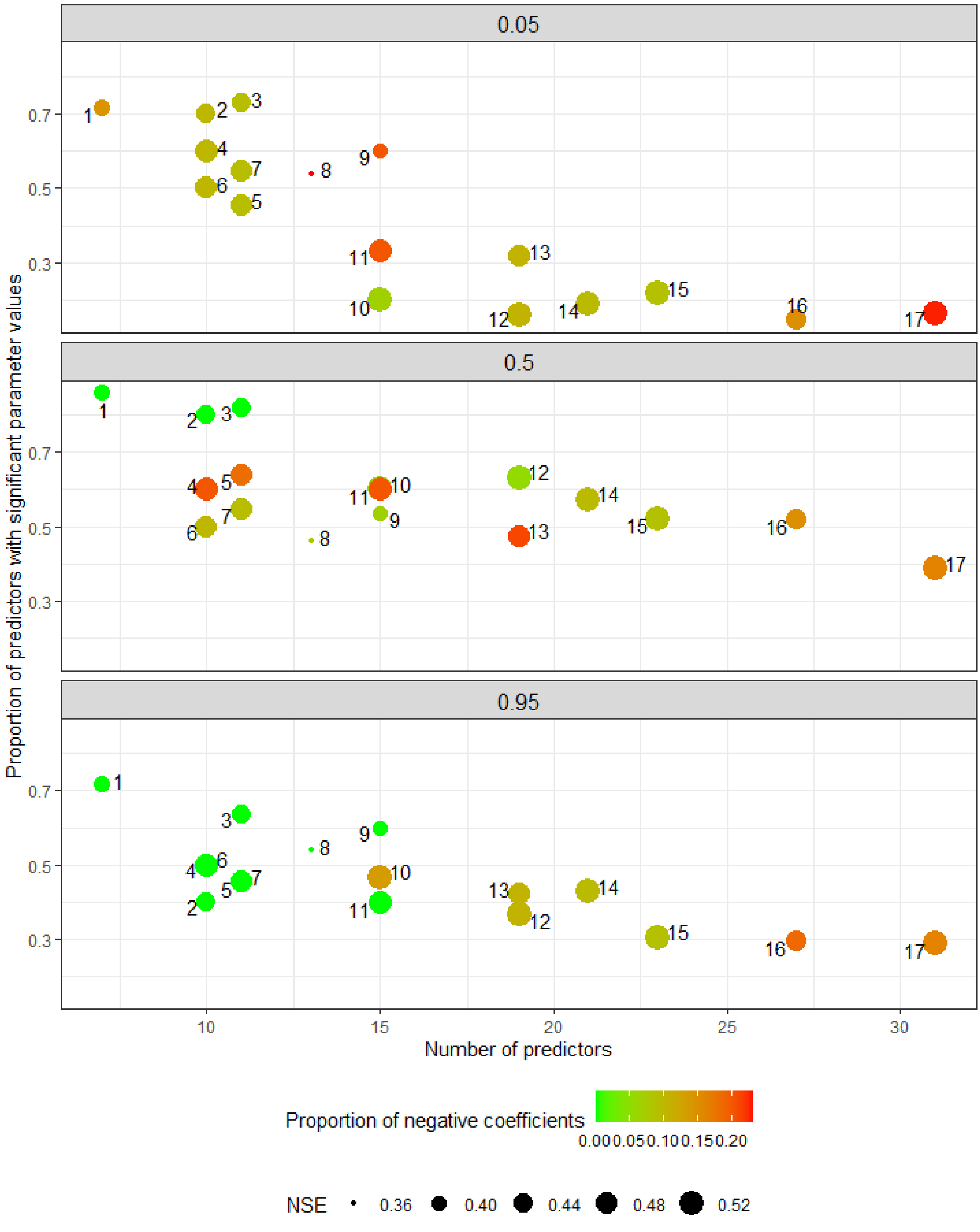

Figure 7. Proportion of significant coefficients versus number of predictors for quantile models for E. coli *concentration fitted to 14 sets of predictors (different typologies) for models pertaining to the 0.05, 0.5, and 0.95 quantiles. The numbers beside each point indicate the model number (1 to 14). The colour indicates the proportion of fitted coefficients that were negative.*

From 17 potential models (Table 2) we judged Model 10 to be the best. Model 10 included 15 predictors (i.e., land-types) excluding the Bare and Water and was based on subdivision of the land use categories by elevation and drainage (see Table 2 for details). Our rationale for the choice of Model 10 is set out below but we emphasise that this choice is a judgement and that other models and their associated typologies may be better in some circumstances.

For Model 10, 20%, 60% and 47% of the fitted coefficients were significant for the 0.05, 0.5, and 0.95 quantile models. The model had a cross validated NSE of 0.51 and a PBIAS of -1.6%, which indicates good performance (NSE > 0.5, |PBIAS|<15%; Table 3) based on the criteria of Moriasi et al. (2015; Table 3, Figure 7).

The performance of Model 10 compares favourably with the performance of a RF model fitted to the same dataset (NSE 0.71, Snelder et al. 2023). We expected the RF model to significantly out-perform the models fitted by this study because the RF model included a larger number of predictors and fitted non-linear relationships and high-order interactions (Cutler et al. 2007).

For Model 10, there was one negative coefficient for the 0.5 and 0.05 quantile models and the 0.95 quantile models had two negative coefficients (Figure 8). The negative coefficient for the 0.5 quantile model was associated with the Natural_HighElev land-type with a value of -1, which was not statistically significant (Figure 8). We judged that this very small negative value could either be ignored or set to zero and in either case was a credible ETC parameter for a catchment model for this land-type (i.e., we expected this land-type makes a small to negligible contribution to *E. coli* concentrations).

Models 1 to 7 were judged to be less appropriate than Model 10 because these had fewer predictors and were therefore less able to discriminate spatial variation in *E. coli* contributions. We note that Models 4 and 5 had NSE values of approximately 0.5 whereas Models 6 and 7 had NSE values of approximately 0.46. The typologies underlying Models 4 and 5 were based on subdivision of the land use categories by the two elevation categories whereas Models 6 and 7 were based on subdivision by temperature. We interpreted this as evidence that elevation was a better predictor of contribution of land to *E. coli* concentrations than temperature. Models 8 and 9 had the same and two more predictors than Model 10, respectively. However, these models had significantly lower performance than Model 10 (NSE was 0.36 and 0.39, respectively). The typologies underlying Models 8 and 9 were based on subdivision of the land use categories by the three moisture categories. We interpreted the lower NSE values for Models 8 and 9 compared to Models 4 and 5 as evidence that elevation was a better predictor of contribution of land to *E. coli* concentrations than moisture. In addition, Model 11 was similar to Model 10 with the same number of predictors but was based on subdivision of Sheep & Beef, Dairy and Exotic Forest by two moisture categories rather than two drainage categories. Model 11 had similar performance to Model 10 (NSE=0.5, PBAIS = -1.6%) but Model 11 had three negative coefficients.

Model 12 had 21 predictors (including Bare and Water), the same performance (NSE=0.51, PBIAS=-1.5%), more significant coefficients than Model 10 and only one negative coefficient for the 0.5 quantile model (Figure 8). However, the further subdivision of Dairy into two elevation and two drainage categories (for Model 12) compared to just into two drainage categories (Model 10) was not well supported by the data (there was generally low occupancy for the four Model 12 Dairy categories) and the coefficient for the Dairy_HighElev_PoorlyDrained category, in particular, was unrealistic. We therefore judged that Model 12 was a less appropriate model than Model 10. Models 13 to 17 had a greater number of predictors than Model 10 but had similar performance (NSE ~ 0.5, PBBIAs ~ -1.6%)

and had at least two negative coefficients. Because we expected all land-types to contribute *E. coli* apart from Bare and Water, we judged that these models could not provide credible ETC parameters for a catchment model. Taken together, these results help to justify selecting a typology based on subdivision using elevation and drainage rather than temperature or wetness, and to choose Model 10 as the best model.

The fitted coefficients for Model 10 were generally consistent with expectations (see Section 4.3.3, Figure 8 and Table 4). For example, the highest values were associated with Urban, and Dairy land uses, and the lowest values were associated with Natural land cover. Within a land use, the coefficients were consistently lower for the High elevation category and Well Drained drainage category, which is consistent with the expectations set out in Muirhead et al. (2023). The standard errors for the fitted coefficients were largest for the Cropland and Orchard & Vineyard land-types, which is consistent with their low occupancy in our dataset (Figure 6).

Table 4. ETC parameters for each of the 17 types derived from the best E. coli *concentration model (Model 10). The values can be interpreted as the contribution of each type to the median* E. coli *concentration (cfu 100 ml^{-1}). Note that the types Bare and Water were excluded from the regression model and are assumed to have ETC values of zero. Note that the best estimate and the prediction interval limits are defined by three separate regression models that are independent of each other. This means that the parameter values are not necessarily expected to follow an order that is consistent with the point on the probability distribution that each model pertains to.*

Land-type	Best estimate	Prediction interval lower bound	Prediction interval
Cropland	159	113	477
Dairy_PoorlyDrained	381	118	1987
Dairy_WellDrained	391	104	987
ExoticForest_HighElev_PoorlyDrained	25	21	-46
ExoticForest_HighElev_WellDrained	16	-38	139
ExoticForest_LowElev_PoorlyDrained	128	37	205
ExoticForest_LowElev_WellDrained	8	41	-104
Natural_HighElev	-1	3	17
Natural_LowElev	147	11	475
Orchard&Vineyard	310	61	240
Sheep&Beef_HighElev_PoorlyDrained	100	13	138
Sheep&Beef_HighElev_WellDrained	43	2	363
Sheep&Beef_LowElev_PoorlyDrained	349	48	865
Sheep&Beef_LowElev_WellDrained	183	89	724
Urban	754	115	2815

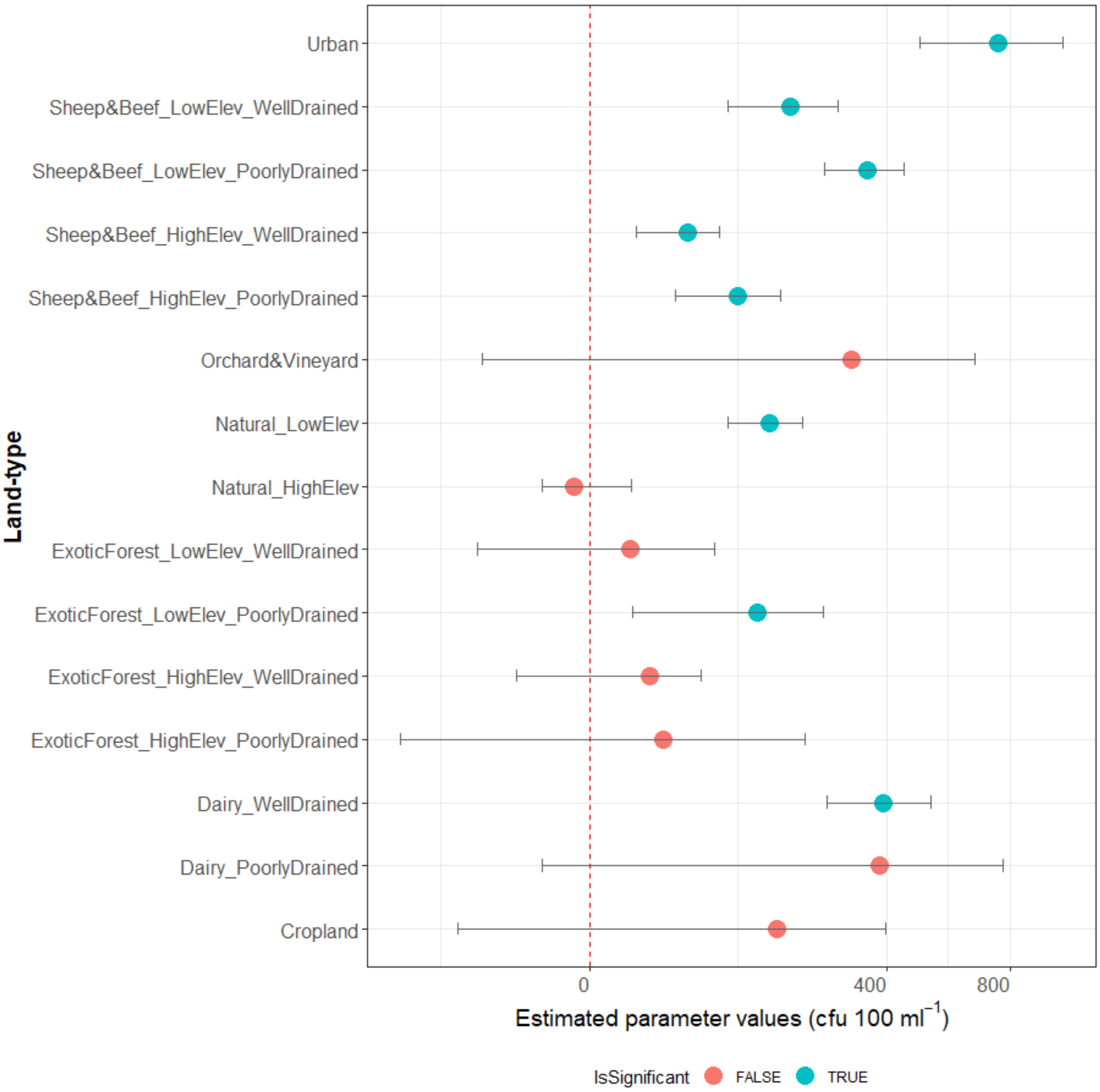

Figure 8. Fitted coefficients for the 0.5 quantile of the best E. coli *concentration model (Model 10). Note that the x-axis is transformed to provide greater resolution of values with low magnitudes compared to higher magnitudes. The error bars indicate the standard errors for the fitted coefficients. Note that the types Bare and Water were excluded from the regression model and are assumed to have ETC values of zero.*

For each land-type, the fitted coefficients for Model 10 were reasonably consistent with the observations of *E. coli* concentrations at the water quality stations having high occupancy by that type (Figure 9). For some types, the data did not include many or any water quality stations with high (e.g., >0.7) occupancy (e.g., Dairy_PoorlyDrained). However, for some types there was good representation by sites with high occupancy (e.g., Natural_HighElev_WellDrained, Urban, Sheep&Beef_LowElev_PoorlyDrained). In these cases, as the proportion occupancy increased, the central tendency of the observed *E. coli* concentrations converged on the fitted model coefficients (Figure 9). In addition, where there was not good representation by sites with high occupancy, the fitted coefficients were generally consistent with the extrapolated value of trend lines (red lines, Figure 9) fitted to the data.

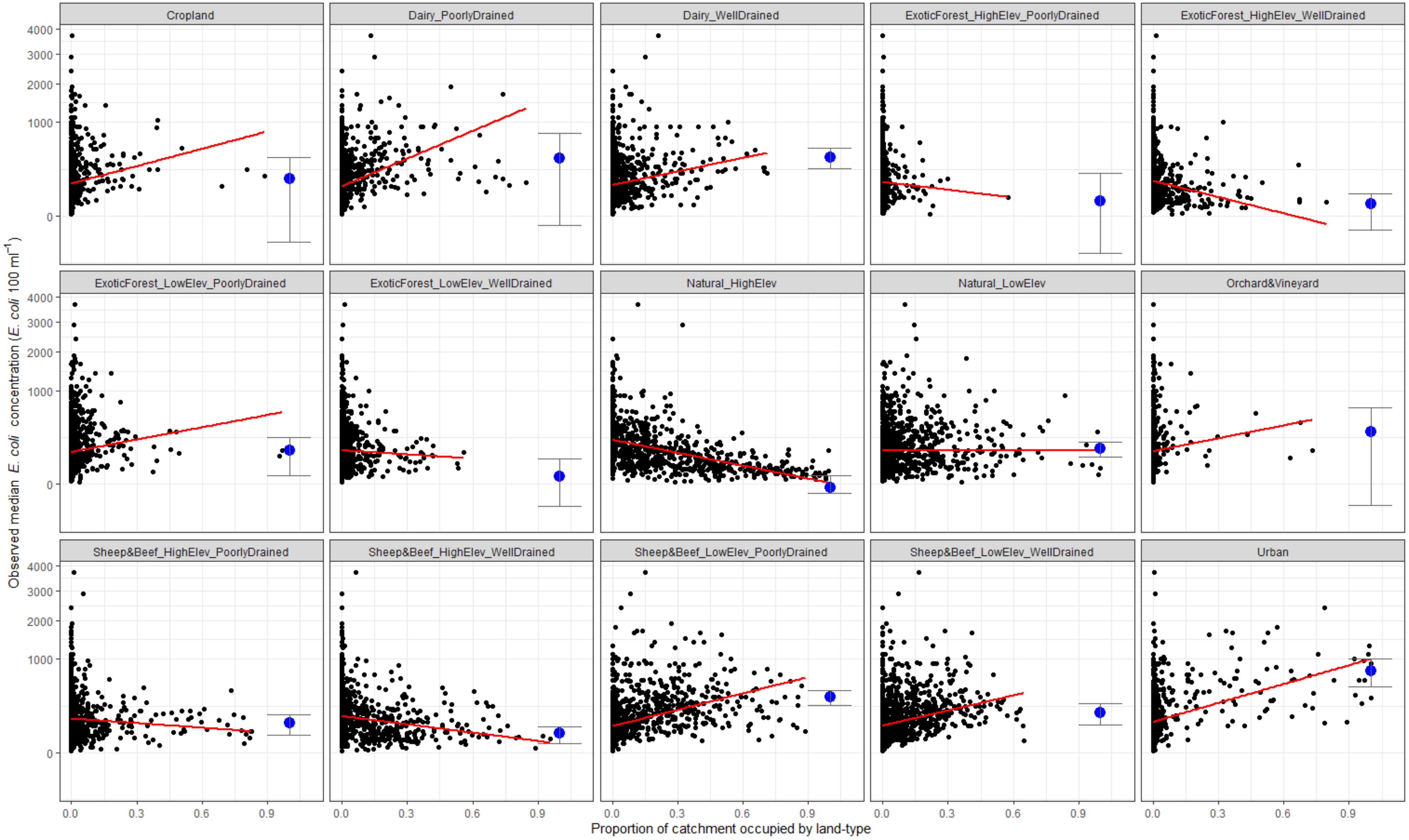

Figure 9. Observed E. coli *concentrations versus the proportion of catchment occupied by each land-type (panels). For each type, the ETC values for the best model are indicated as a blue dot at the position on the x-axis indicating a proportion occupancy of 1. The red line is a linear regression indicating the expected value of* E. coli *concentration conditional on the proportion of catchment occupied by the type. The purpose of this line is show that as occupancy approaches 1, the central tendency of the observations tend to converge to the fitted coefficient. Note that the types Bare and Water were excluded from the regression model and are assumed to have ETC values of zero.*

The mean of the proportion of predictions that were within the 90% prediction interval over the 10 cross validation folds was 88% (range 83% to 94%, Figure 10). This indicates that the 90% prediction interval is a reliable measure of the uncertainty of the empirical concentration model predictions.

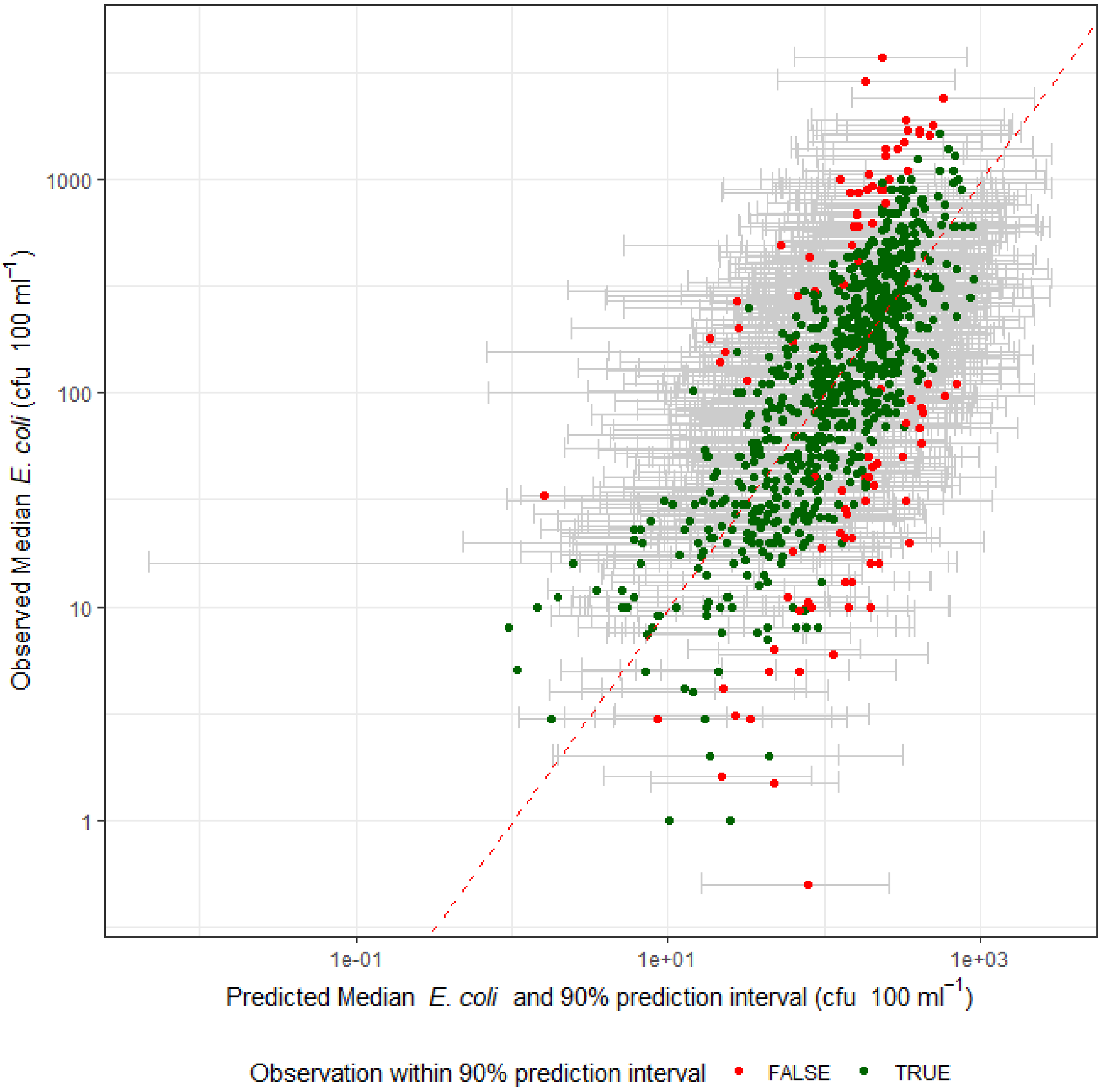

Figure 10. Observed versus predicted site median E. coli *concentrations (points) and 90% prediction interval (grey error bars). The predictions and the estimated 95% prediction interval are independently derived for each water quality station by the cross validation. The green points indicate the observations that are within the 90% prediction interval and the red points indicate the observations that are outside of the 90% prediction interval. Note that the observed values are plotted on the Y-axis and predicted values on the X-axis, following Piñeiro* et al. *(2008). Red line: one-to-one line.*

The RMSD of Model 10 (in log base 10 space) was 0.41. The 90% prediction interval for an estimated value can be evaluated as:

$$90\%\ prediction\ interval = 10^{[log_{10}(x)\ \pm\ 1.65\ \times\ RMSD]} \qquad \textit{Equation 5}$$

where x is the estimated value in the original units, and RMSD is the reported model error (0.41). For example, the 90% prediction interval for an estimated value of 100 cfu 100 ml^{-1} is the range 21 to 475 cfu 100 ml^{-1}. This large uncertainty in the model prediction is evident in the predictions shown in Figure 10 (noting that the axes are log_{10} transformed).

The mean of each of the coefficient values for each quantile (i.e., 5^{th}, 50^{th} and 95^{th} quantiles) over the 10 versions of the best model fitted by cross validation were generally consistent with the fitted coefficients for the full models as indicated by points lying close to the one-to-one line in Figure 11. In addition, the standard deviations of the coefficient values over the 10 realisations of the models fitted by cross validation were approximately equal to the corresponding standard errors (see also Figure 11) for the fitted coefficients for the full model. This is an objective indication of the stability and reliability of the fitted parameters.

Based on the observations above, the fitted coefficients for Model 10 (Table 4) were judged to be a credible set of ETC parameters for an empirical catchment water quality model for *E. coli* concentrations of the form indicated by Equation 4.

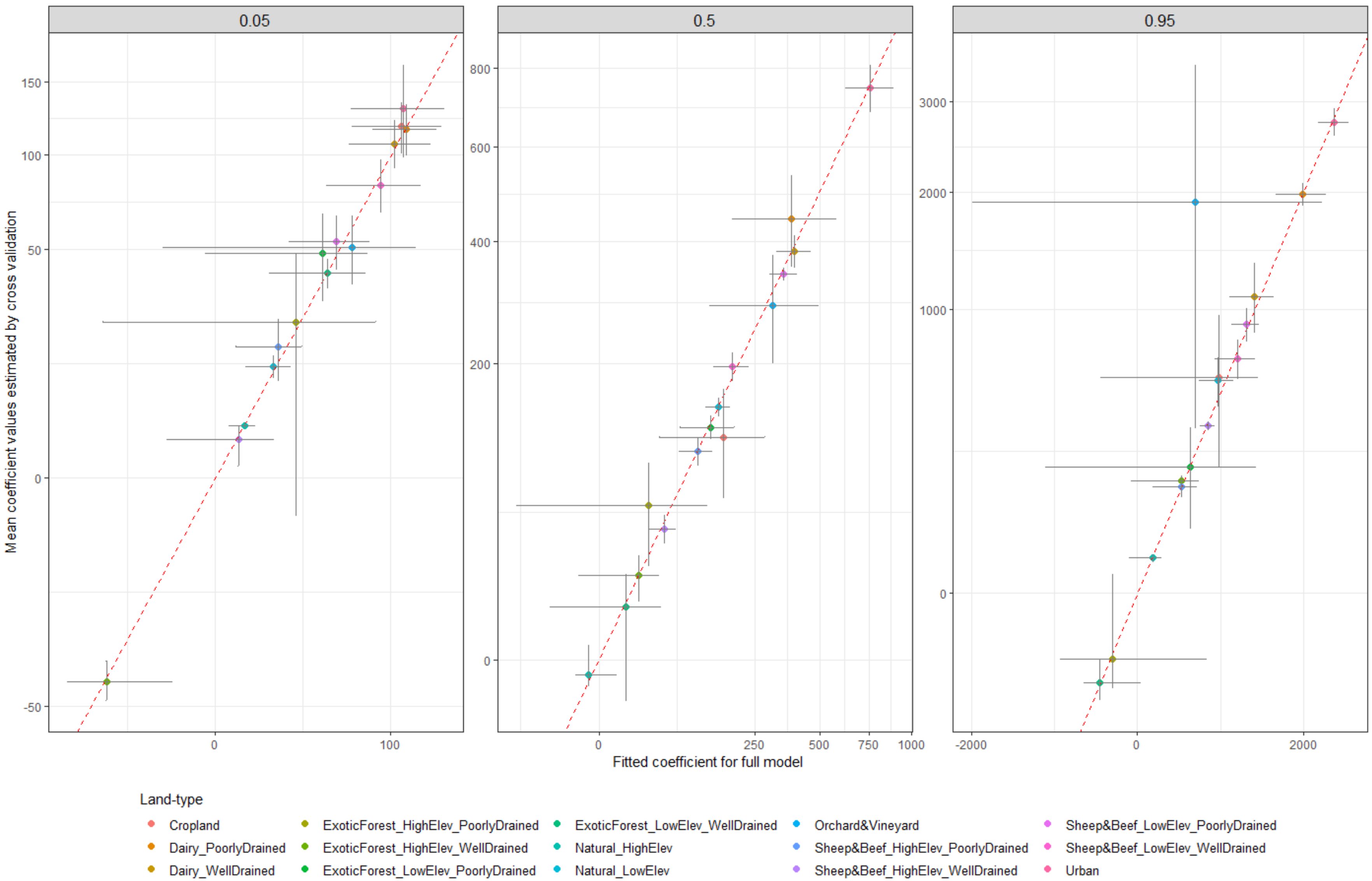

Figure 11. Comparison of coefficients fitted to each type in the full E. coli *concentration model with the mean of 10 realisations of the same coefficients fitted by cross validation. The vertical error bars indicate the standard deviation of the coefficients over the 10 cross validation folds. The horizontal error bars are the standard errors of the coefficients fitted in the full models. The red dashed line is one-to-one and indicates perfect agreement.*

5.3 Empirical yield model

For sites with yield data, the Sheep & Beef and Natural land use types were the most consistently occurring (i.e., are represented in the catchments of many sites) and had the widest variation in occupancy (i.e., occupancy ranged from zero to 100%, Figure 12). The land use categories Bare, Cropland, Orchard & Vineyard and Water tended to have low occupancy (i.e., the proportion of catchment area occupied by these categories was most commonly zero) and variation across sites was restricted (i.e., very few sites had >50% occupancy by any of these categories). Visual inspection of Figure 12 supported our assumption that contributions to *E. coli* by land categorised as Bare or Water is negligible. The *E. coli* yields (model response) for the 320 sites were not normally distributed (right skewed), justifying the use of quantile regression (Figure 5).

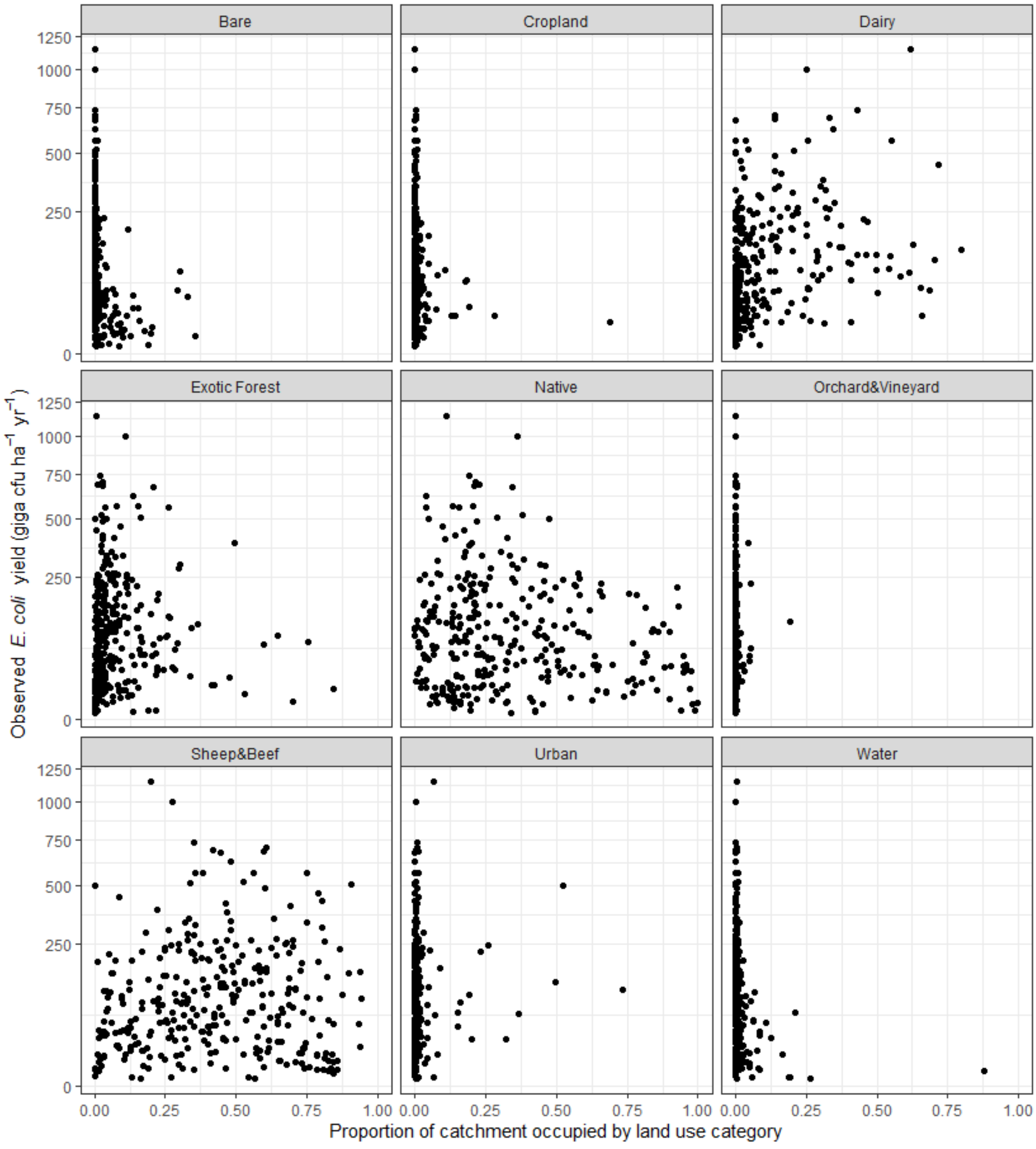

Figure 12. Observed site median E. coli *yields as a function of the proportion of catchment area occupied by the nine land use categories.*

The proportion of significant fitted coefficient values decreased with increasing numbers of predictors for models pertaining to all three quantiles (i.e., median (0.5 quantile), 0.05 and 0.95 quantiles; Figure 13). There was also a trend in negative fitted coefficient values as the number of predictors included in the models increased (Figure 12). The performance of the fitted models, as measured by NSE, varied between 0.12 (Model 1) and 0.42 (Model 11, Figure 12). All models had absolute PBIAS values no greater than 6%.

Models 1, 2, 3, 8 and 9 did not achieve satisfactory performance (NSE > 0.35, |PBIAS|<20%; Table 3) based on the criteria of Moriasi et al. (2015, Figure 12). Although the remaining models did achieve satisfactory performance, they had at least three negative coefficients and the proportion of significant fitted model coefficients was not greater than 30% for any of the 0.5 quantile models. We therefore judged that no model would provide a satisfactory basis for defining empirical catchment model parameters (i.e., ETY values).

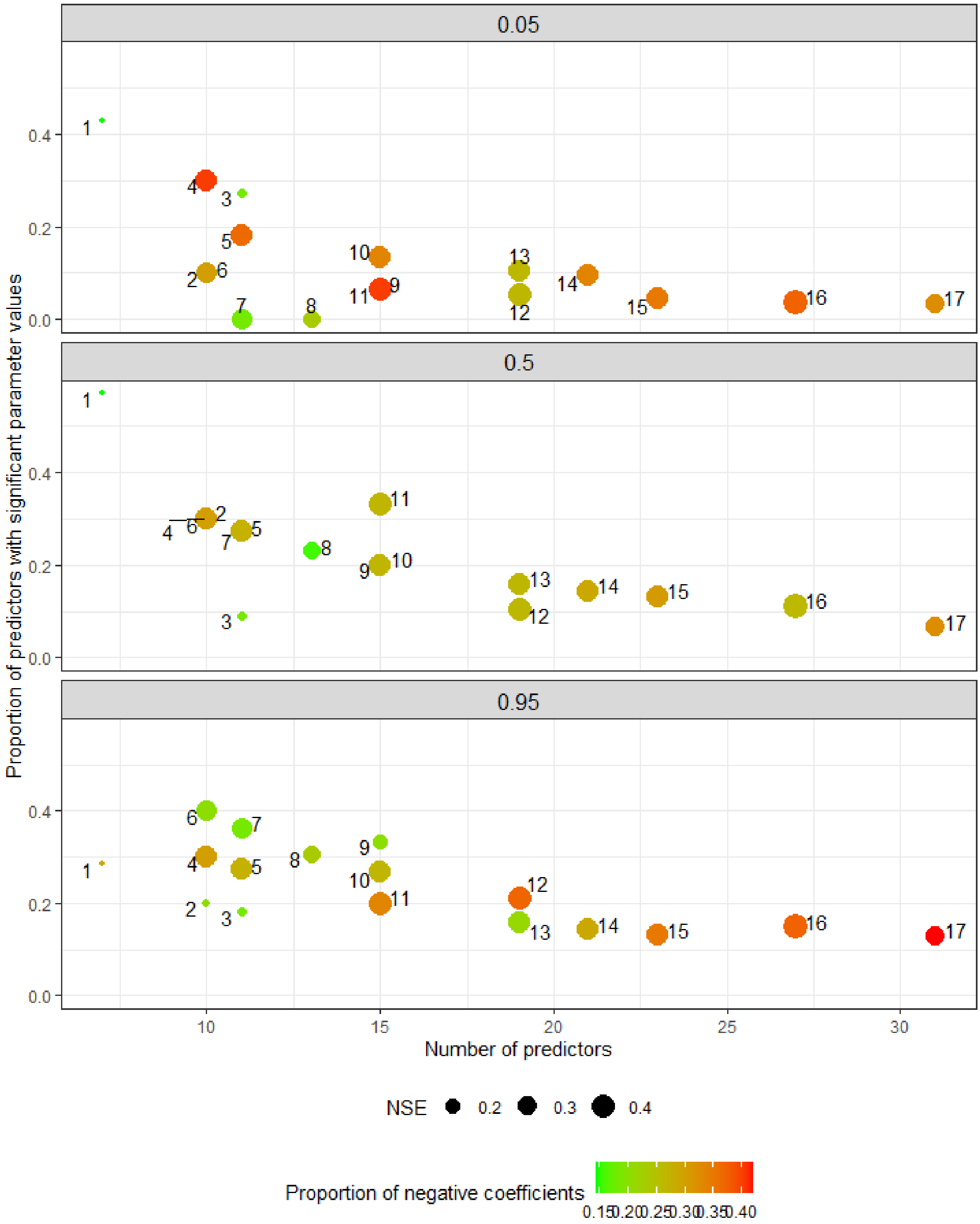

Figure 13. Proportion of significant coefficients versus number of predictors for quantile models for E. coli *yield fitted to 17 sets of predictors (different typologies) for models pertaining to the 0.05, 0.5, and 0.95 quantiles. The numbers beside each point indicate the model number (1 to 17).*

6 Example simulations

6.1 Scenarios

We used the empirical *E. coli* catchment concentration model derived above to simulate three catchment management scenarios. Simulations were made for the catchments of the 869 water quality monitoring stations (Figure 2) as follows:

- Scenario 1: apply default mitigation reductions of 50% for Sheep & Beef and Dairy land in each catchment.
- Scenario 2: Scenario 1 plus convert one half of all Sheep & Beef land in each catchment to Exotic Forest
- Scenario 3: Scenario 1 plus convert one quarter of both Sheep & Beef and Dairy land in each catchment to Natural

We used the land-types and fitted coefficients of Model 10 (Table 4) to provide the ETC parameters for the catchment model (Equation 4). We made one modification to the fitted coefficients and changed the ETC parameter for the Natural_HighElev land-type to zero. The outputs are reported below as both reductions from the current concentration (%) and as absolute concentrations (cfu 100 ml^{-1}).

6.2 Calculations

The first step in the calculation was to use the model to predict this study's baseline (i.e., for the 5-year period to the end of 2020) *E. coli* concentration at all sites as:

$$Predicted\ baseline\ concentration = \sum_{i=1}^{m} P_i \times ETC_i \qquad \textit{Equation 6}$$

where P_i is the proportion of catchment area occupied by the ith land-type, ETC_i is the empirical type concentration parameter for the type, *m* is the number of land-types represented by the model, which in this case is 17 (i.e., the 15 defined in Table 4 and ETC parameters equal to zero for the Bare and Water land-types). Note that Equation 5 differs from Equation 4 by not including the term representing the concentration form of the catchment point source contributions (i.e., PS_C). This means that the analyses performed here ignored any point source contribution, but these were expected to be small for the reasons set out in Section 4.1.3.

We simulated the *E. coli* concentration for Scenario 1 by applying a default mitigation rate of 50% to all pastoral land (i.e., to all land-types that had either a Sheep & Beef or Dairy land use category) and setting the mitigation rate for all other land uses to zero. The 50% mitigation rate is the most likely effectiveness of stream fencing as a mitigation option for reducing *E. coli* concentrations in streams (Muirhead 2019). We note that we have not allowed here for any fencing that is already in place during the baseline period. The use of 50% as an estimate of effectiveness is therefore probably optimistic relative to the change that is achievable. The *E. coli* concentration for all sites under Scenario 1 was predicted as:

$$Predicted\ scenario\ concentration = \sum_{i=1}^{m} P_i \times ETC_i \times (1 - Mitigation\ rate_i) \qquad \textit{Equation 7}$$

where $Mitigation\ rate_i$ is the assumed mitigable proportion of the concentration for the ith land-type.

For Scenario 2, we converted 50% of all land in the Sheep & Beef land use category to the Exotic Forest land use category. Both the Sheep & Beef and Exotic Forest land use categories are split into four land-types, based on combinations of elevation and drainage (Table 4). We therefore reassigned half the proportion of catchment area in each of the four Sheep & Beef land-types to an Exotic Forest land-type with the same elevation and soil categories and reduced the proportion of catchment area in each Sheep & Beef land use type by 50%.

For Scenario 3, we converted 25% of all Sheep & Beef and Dairy land use categories into the Natural category. The Natural land use category is split into two land-types, based on elevation alone (Table 4). We therefore reassigned one quarter of the proportion of catchment area in the two Sheep & Beef High Elevation land-types (i.e., two drainage categories; Sheep&Beef_HighElev_PoorlyDrained and Sheep&Beef_HighElev_WellDrained) to Natural-High Elevation. Similarly, we reassigned one quarter of the proportion of catchment area in the two Sheep & Beef Low Elevation land-types (i.e., two drainage categories; Sheep&Beef_LowElev_PoorlyDrained and Sheep&Beef_LowElev_WellDrained) to Natural-Low Elevation. We reassigned one quarter of the proportion of catchment area in all Dairy types (only differentiated by Well Drained and Pooly Drained) to Natural-low elevation[5]. The proportion of catchment area in each of the Sheep & Beef and Dairy land-types were reduced by 25%. Note that a more general description of how land use changes should be simulated when using the empirical catchment *E. coli* concentration model is provided in Appendix C.

We predicted the *E. coli* concentrations for scenario 2 and 3 by changing the proportions of land area in the catchment of each water quality station as described above and applying Equation 7. Note that this means that Scenario 2 and 3 represent the combination of land use changes plus mitigation measures on Sheep & Beef and Dairy land.

We estimated the change in concentration for each scenario as:

$$Change\ in\ concentration = \frac{Predicted\ baseline\ concentration - Scenario\ concentration}{Predicted\ baseline\ concentration} \qquad \text{Equation 8}$$

Because we knew the actual (i.e., observed concentration) at each LAWA site, we estimated the new predicted concentration for each scenario as:

$$Estimated\ concentration = Observed\ concentation\ \times Change\ in\ concentration \qquad \text{Equation 9}$$

6.3 Results

Figure 14 shows the results for all three scenarios as predicted reductions of *E. coli* (from current concentration) plotted against the proportion of the catchment occupied by pastoral land use (i.e., Sheep & Beef or Dairy). For Scenario 1 the reduction from the baseline *E. coli* concentration increased with increasing proportion of the catchment occupied by pastoral land use to reach an approximate mean reduction of around 50% for catchments having high occupancy by pastoral land use. This simply reflects the “default” 50% reduction in *E. coli* losses from all pastoral land uses assumed in this scenario. There are sites that have very low proportion of catchment area occupied by pastoral land use, but which have reductions of 50%. This occurs in catchments with very small pastoral area and the remainder being occupied by the Natural_HighElev land-type, which has an ETC parameter value of zero. In these catchments, the reduction from the pastoral land is 50% and the contribution of the remaining land is zero and, therefore, the reduction is 50%. The between-site variation in

[5] Note that this makes the assumption that all land in the Dairy land use category is below 350 m ASL, which is generally, but not always true.

percentage reduction that is achievable under Scenario 1 reflects the variable composition of non-pastoral land uses in individual catchments.

Figure 14 shows that Scenario 2 generally achieves greater reductions than Scenario 1 and Scenario 3. The reductions achieved under Scenarios 2 and 3 increase with increasing proportion of the catchment occupied by pastoral land use. This is to be expected because Scenarios 2 and 3 involve changing fixed proportions of current pastoral land use to Exotic Forest and Natural land use categories, respectively, and therefore the overall reductions in *E. coli* concentrations increase with the amount of current pastoral land use. Figure 14 shows that although there are generally larger reductions for Scenario 2 compared to Scenario 3, there is considerable variation and for some sites this pattern is reversed. This is because different catchments comprise different amounts of Sheep & Beef and Dairy and different combinations of the underlying ETC parameter values and therefore, the reductions vary.

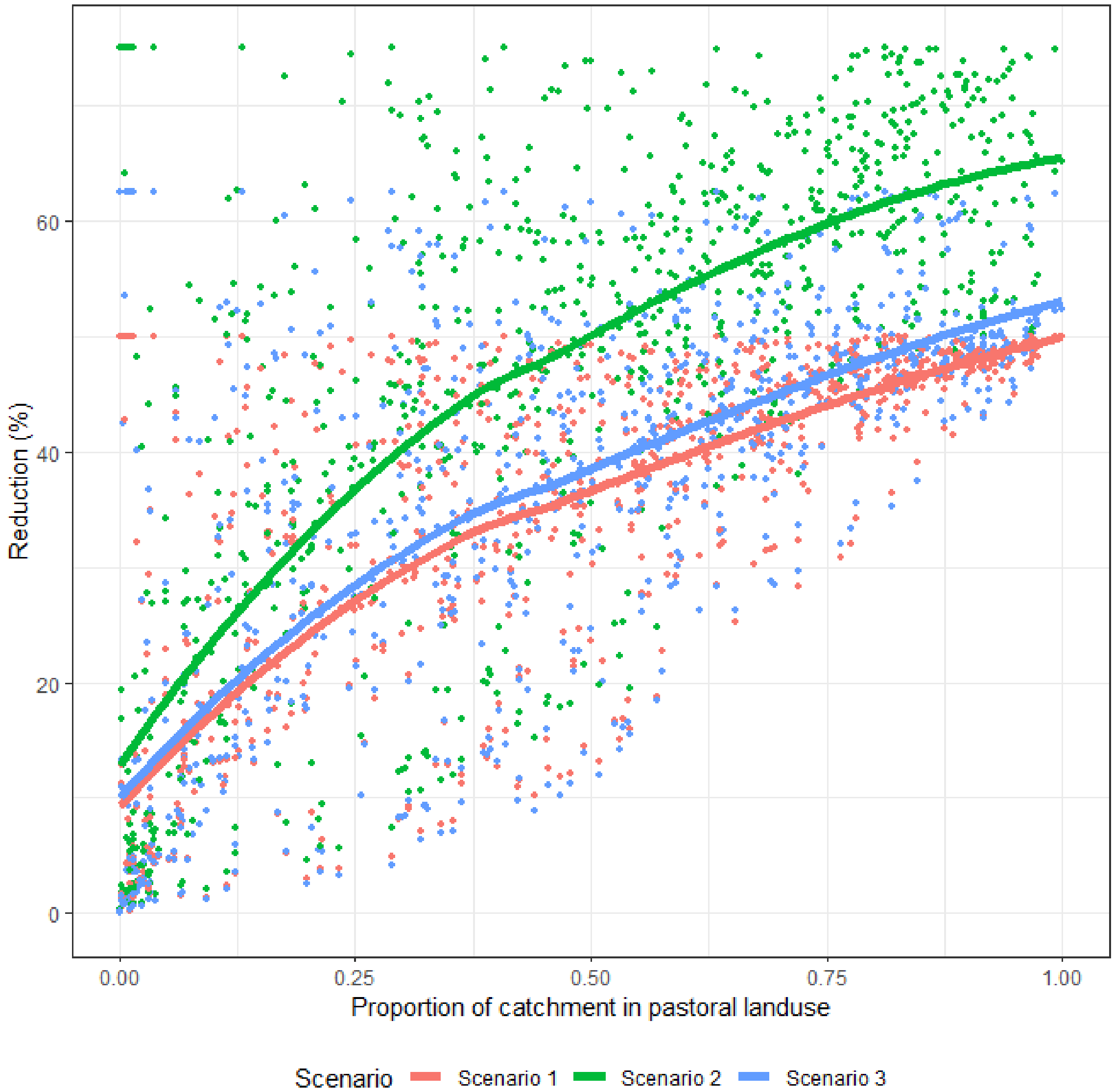

Figure 14: Outputs of the simulations of the three scenarios. The plot shows the predicted reductions for E. coli *against proportion of the catchment that is currently in pastoral land use. The solid lines are smoothed representations of the mean response (i.e., reduction) versus proportion of the catchment in pastoral land use.*

Figure 15 shows maps of the sites coloured by the predicted *E. coli* reductions under the three scenarios. These maps indicate geographic variation in the extent to which mitigation is predicted to reduce *E. coli* concentrations. For example, for Scenario 1, the reductions in *E. coli* concentration are generally large in the parts of Waikato, Manawatu-Whanganui, Taranaki, Canterbury and Southland regions. This is because catchments in these regions generally have higher proportions of pastoral land use and therefore have larger areas of mitigable land. In contrast, catchments in the West Coast, Tasman and Marlborough regions and in parts of the Bay of Plenty and Auckland had lower reductions because their catchments tend to have lower proportions of pastoral land use.

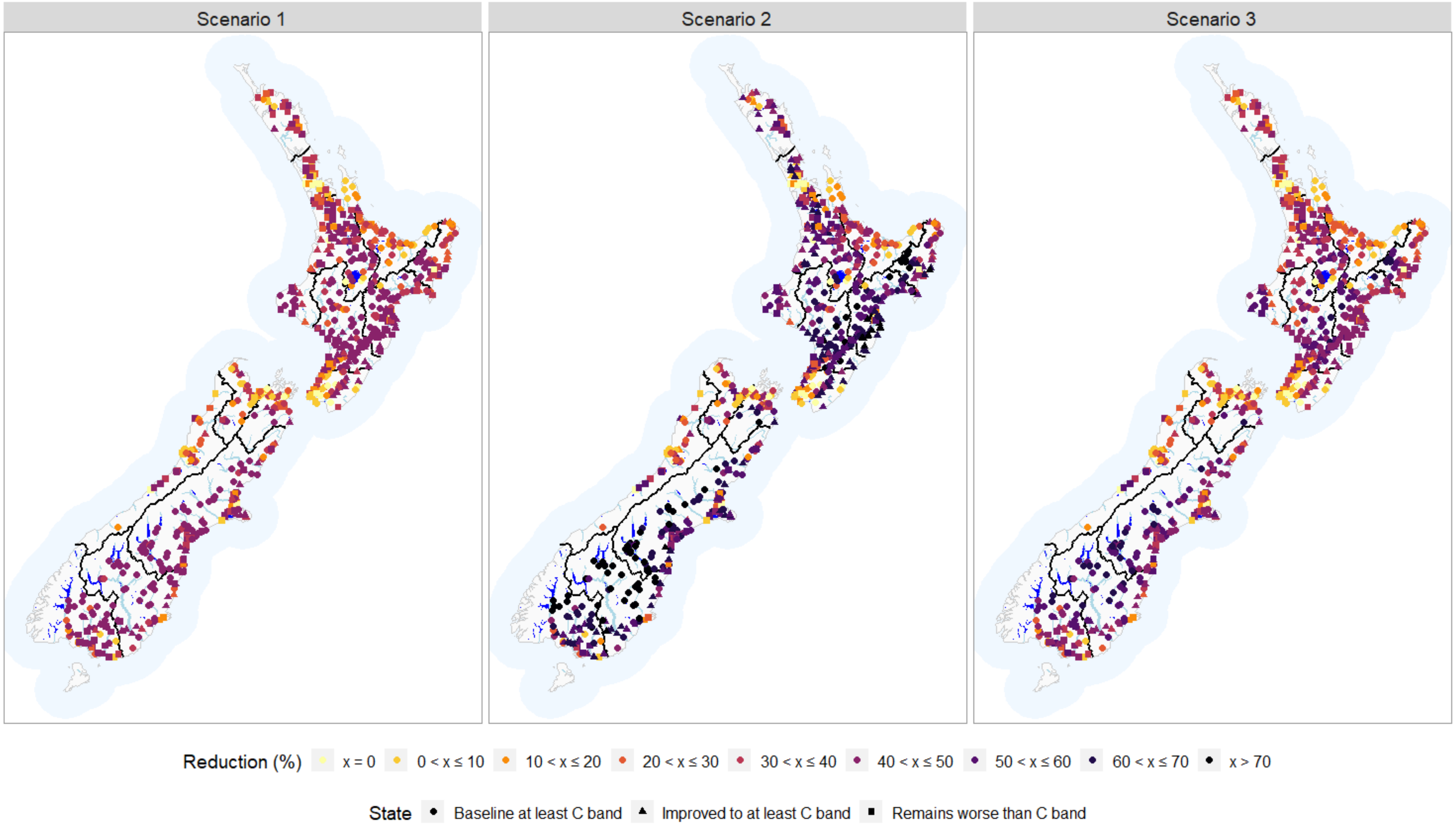

Figure 15: Maps showing outputs of the simulations of the three scenarios. The maps show water quality monitoring stations coloured by the predicted reductions for median E. coli *concentration. Stations represented by round points have baseline median* E. coli *concentrations <130 cfu 100 ml^{-1} and those represented by triangles and squares are improved to be <130 cfu 100 ml^{-1} or remain >130 cfu 100 ml^{-1} under the scenario, respectively.*

Figure 16 compares current and estimated median concentrations of *E. coli* for the three scenarios. This plot indicates that there are some water quality stations that have very little or no change in concentrations under the scenarios. This occurs when stations have little or no pastoral land use in their catchments. The plot also indicates that the change in absolute concentrations generally increases from scenario 1 to scenario 3.

Under the scenarios, some water quality stations with current median *E. coli* concentrations exceeding 130 cfu 100 ml^{-1} will have levels decreased to less than or equal to 130 cfu 100 ml^{-1}. These stations are below the horizontal and to the right of the vertical grey lines in Figure 16. The value of 130 cfu 100 ml^{-1} is the upper threshold for suitability for primary contact (C band) as defined by the NPS-FM. For Scenarios 1, 2 and 3, the proportion of stations that are predicted to be moved from unsuitable to suitable under the scenarios are 38%, 61% and 39% of stations with baseline median *E. coli* concentrations greater than 130 cfu 100 ml^{-1}, respectively. The location of these stations is shown in

Figure 15. Note that 55% of stations have baseline median *E. coli* concentrations greater than 130 cfu 100 ml^{-1}.

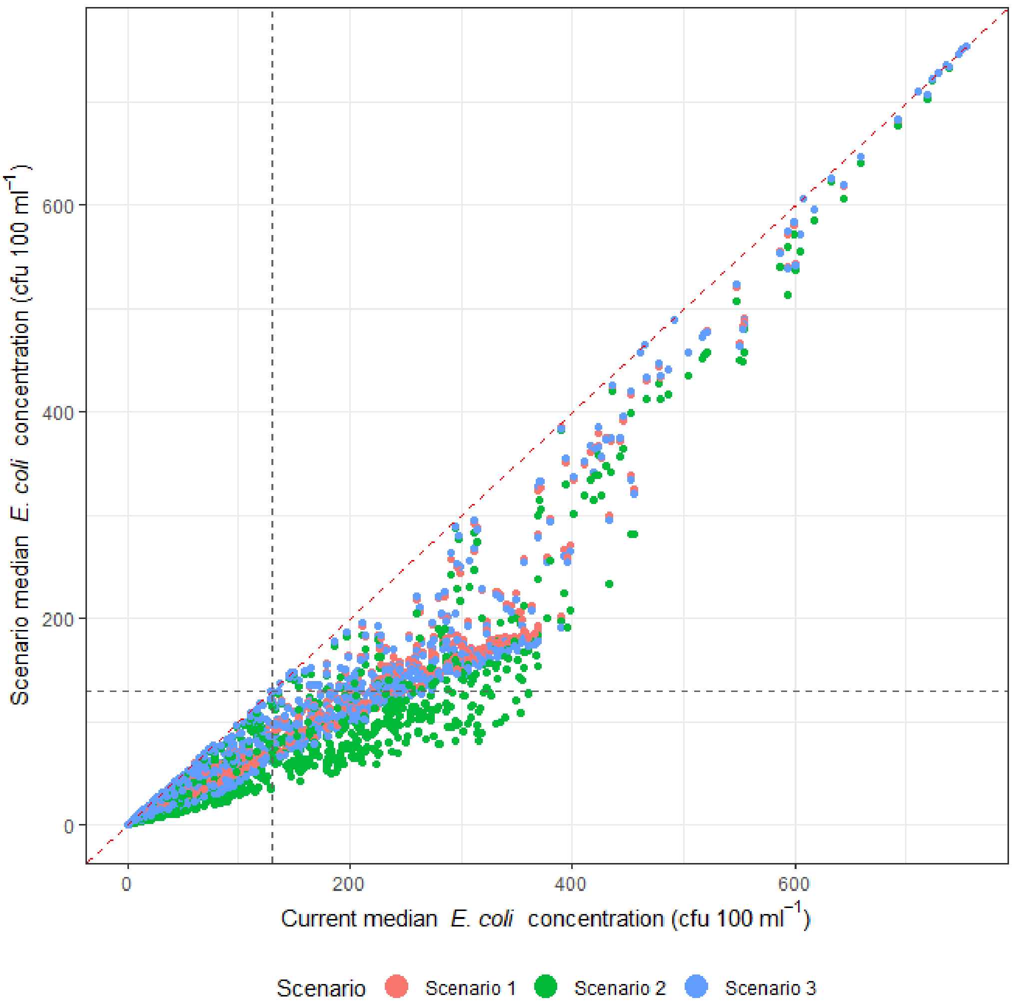

Figure 16: Comparison between current and estimated concentrations for the three scenarios. The red dashed line is one to one. Site and scenario combinations lying on this line are predicted to have no change in E. coli *concentrations for the indicated scenario.*

7 Discussion

7.1 Empirical catchment models for *E. coli*

In this study, we developed an alternative class of empirical catchment model for predicting median *E. coli* concentrations. We were not able to develop a satisfactory empirical *E. coli* yield model.

The empirical *E. coli* concentration model has calibrated lumped parameters that, for each land-type, represent loss of *E. coli* from land, transport and attenuation in the drainage network, and dilution in the water column. The model provides a simple and easily used tool that can be applied at any location within New Zealand. In other words, the empirical model is a simple alternative to setting up process-based catchment *E. coli* models. In addition, the empirical model allows the user to estimate the 90% prediction interval as an estimate of the imprecision of its predictions. The approach provides a simpler and more transparent method for simulating the impacts of land management on median *E. coli* concentrations than process-based models, which may be appropriate for at least some applications.

We have developed a dataset that provides proportions of catchment area occupied by each land-type used by the *E. coli* concentration model for all segments of the DN2.4 (>10 km^2). These data allow estimates of *E. coli* concentration to be made at any location in New Zealand very easily. In addition, the data can be used to rapidly assess impacts of land use and land management scenarios on *E. coli* concentrations at any location in New Zealand.

Defining the land-types that were included in the empirical model involved expert judgement and was strongly influenced by the work of Muirhead et al. (2023). The definition of the 'best' empirical model was also based on expert judgement that involved making trade-offs between model performance, the proportion of significant coefficients, and optimising for the total number of land-types. Future research or applications could refine the approach, potentially with more exhaustive exploration of the sets of land-types (typologies, their factors and the threshold values used to define the categories), incorporating the observation uncertainties into the fitting process, using updated observed water quality datasets, and improving the criteria used to define the "best" model based on specific model purpose.

The main reason we were unable to derive a satisfactory empirical catchment model for predicting *E. coli* loads (as yields) was that there were fewer sites in our fitting datasets. Another reason may be that the *E. coli* yield estimates are themselves very uncertain (Figure 5) which adds noise to the model response. This outcome emphasises the importance of having continuous flow measurement at water quality monitoring stations and more frequent or more targeted sampling over the full range of flows (to reduce load estimate uncertainty). We note that these requirements are important for the calibration of process-based models, as well as for empirical models.

7.2 Application of the empirical models and limitations

This study demonstrates the potential application of the empirical *E. coli* concentration model to assessing the potential benefits of land use mitigations, and of land use change on median *E. coli* concentrations. As such, the model is potentially useful to support engagement and decision-making processes relating to *E. coli* target attribute states. However, users need to understand the limitations of the empirical model and recognise that the very simplified representation of processes means that the model cannot answer some types of questions and its predictions are approximations.

An important limitation that applies to the empirical model is associated with the national scale of the quantile regression models that were used to derive its parameters (i.e., the ETCs). Because the water quality station data were limited, we were only able to derive robust ETCs for a small number of (17) land-types. This limits the spatial resolution of the empirical model. In addition, as the spatial extent of a modelled domain reduces, the specificity of the ETC values will diminish because the calibration target was the central tendency (i.e., median value[6]) of all the national water quality stations.

The ETC parameter for each land-type is a lumped value that represents the loss of *E. coli* from land, transport and attenuation in the drainage network and dilution in the water column. The use of lumped parameters to represent all these processes has limitations. First, the lumped parameters mean that the empirical model has no spatial discretisation of *E. coli* sources. Therefore, the estimation of the contribution of a land-type to *E. coli* concentration at a location has no consideration of the distance between a source and the location of interest. This is a simplification of the real-world situation because it is likely that sources that are close to a location of interest are a more important determinant of concentration than those that are further away due to die-off (i.e., attenuation) of organisms during transport in the drainage system. Second, because attenuation is not explicitly represented and is lumped, the empirical model has no spatial discretisation of *E. coli* attenuation. This is an approximation because attenuation is likely to vary spatially. For example, if a location of interest is immediately downstream of a lake or reservoir, it is likely that there will be significant attenuation of all *E. coli* lost in the upstream catchment as it passed through the lake. In contrast, if there is significant input of *E. coli* between a lake or reservoir and the location of interest, attenuation of this contribution will be less. This is important because, if the proportions of catchment area of land-types were the same in the previous two cases, the empirical model would predict the same concentration.

A limitation of using observations of *E. coli* concentrations from many catchments across New Zealand is that there will be differences in the degree to which mitigation measures, such as stream fencing, have already been deployed across the fitting dataset. This means that the derived ETC parameters represent the 'average' of the varying effects of existing mitigation across catchments. It also means that the changes in concentrations associated with user-defined scenarios involve implicit assumptions about the existing level of implementation at the site of interest. These limitations contribute to the uncertainty at the site level of both the predicted current *E. coli* concentrations and the changes to those concentrations under land use change scenarios.

We note also that it is understood that river faecal microbial dynamics are strongly influenced by both mobilisation of *E. coli* from land and remobilisation of channel store during storm flows (Wilkinson et al. 2011). Our simple empirical model has no representation of the remobilisation of *E. coli* from channel stores and has no representation of the frequency of these remobilisation events. For example, two catchments could have the same proportions of catchment area in various land-types but have contrasting flow regimes (e.g., high base flows and infrequent high flows compared to low base flows and frequent high flows). The difference in flow regimes is likely to contribute to differences in median *E. coli* concentrations between the two catchments but this will not be represented by our model. This and the limitations set out above are partly why predictions of absolute values made with the model have large uncertainties (Figure 10).

[6] Note that this is the median of the site median values because the quantile regression model was fitted to the median of the site median *E. coli* concentrations (as well as the 0.05 and 0.95 quantiles).

Given the above limitations, the most appropriate application of this class of model is in scenario analysis over broad spatial areas (e.g., multiple catchments to regions). In addition, appropriate applications are where the objective is not to evaluate absolute concentrations resulting from a set of actions, but rather to evaluate relative differences between a baseline and scenario. There is greater confidence in the relative difference between a baseline and scenario, and between scenarios, than in the absolute values of the predictions themselves. We note that the accuracy of these relative differences relies on the assumption that the relative differences in the ETC values for different land-types are applicable to the catchment of interest. We note that testing the validity of these assumptions was beyond the scope of this study but would be a useful direction for future research.

7.3 Catchment water quality models and simulations are uncertain

Because models are dependent on the long-term collection of data, the uncertainties associated with water quality models in general, and their use to make simulations of the impact of land management actions on water quality, cannot be reduced appreciably in the short to medium term. However, catchment water quality models will generally need to be used to inform decision makers about appropriate responses to water quality issues including actions such as limiting resource use, requiring mitigations and land use changes. These decisions will ultimately need to be made in the face of considerable uncertainty.

To some extent the large uncertainties associated with estimates of absolute *E. coli* concentrations are less important when the model is used to assess relative differences between two simulations (e.g., between a baseline and a mitigation scenario). In other words, when using the model to make simulations, users should focus on the predicted relative change in concentration between scenarios than the absolute values of the predictions. This is advantageous because there is likely some commonality in the sources of uncertainty between scenarios (e.g., because some uncertainty is due to within-land-type variability or the lack of representation of spatial variation in attenuation) and this means that the uncertainty in the relative change will be less than the absolute values. However, methods for understanding and quantifying the uncertainty of these relative differences for both our empirically based approach and for process-based catchment models have not been developed. Defining and quantifying uncertainties in relative differences between scenarios presents a considerable technical challenge that needs further research.

8 Conclusions

This study derived a credible set of ETC parameters for an empirical catchment *E. coli* concentration model of the form indicated by Equation 4. We also demonstrated the potential application of this model to assessing the potential benefits of land use mitigations, and of the impact of land use change on median *E. coli* concentrations. The model is potentially useful to support engagement and decision-making processes relating to *E. coli* target attribute states.

This model's ETC parameters were derived using the best currently available information. Nevertheless, there is considerable uncertainty associated with the model's predictions of median *E. coli* concentrations. A strength of the model is that its uncertainty is quantified for any prediction as the 90% prediction interval. In contrast, quantification of uncertainty is a significant challenge for process-based models and is rarely undertaken. The uncertainty reflects the complexity and variability of the processes that control *E. coli* concentrations in catchments. Catchment *E. coli* modelling would significantly benefit from more collection of data from different land-uses and trials of mitigations, and from research on instream attenuation and hydrological processes that impact on the microbial water quality metrics.

The example simulations carried out using the derived model show that even under the substantial actions envisaged by scenarios 2 and 3 (i.e., converting significant areas of pastoral land use to Exotic Forest or Natural land use), a large proportion of the water quality stations with current median *E. coli* concentrations exceeding 130 cfu 100 ml^{-1} were not predicted to have concentrations less than 130 cfu 100 ml^{-1} under the scenarios. This indicates that improving water quality to achieve suitability for primary contact (C band) as defined by the NPS-FM is a significant challenge.

9 Acknowledgements

We thank Amy Whitehead (NIWA) for assembly of water quality and flow data from regional councils and the National River Quality Monitoring Network. Thanks also to Olivier Ausseil and Laura Keenan whose reviews greatly improved early drafts of this report.

10 References

Breiman L (2001) Random Forests. Mach Learn 45:5–32

Cade BS, Noon BR (2003) A gentle introduction to quantile regression for ecologists. Front Ecol Environ 1:412–420

Cutler DR, Edwards JTC, Beard KH, et al (2007) Random forests for classification in ecology. Ecology 88:2783–2792

Elliott AH, Semadeni-Davies AF, Shankar U, et al (2016) A national-scale GIS-based system for modelling impacts of land use on water quality. Environmental Modelling & Software 86:131–144

Leathwick J, Overton J, McLeod M (2003) An environmental domain classification of New Zealand and its use as a tool for biodiversity management. Conserv Biol 17:1612–1623

Mason CH, Perreault Jr WD (1991) Collinearity, power, and interpretation of multiple regression analysis. J Mark Res 28:268–280

McBride GB (2005) Using statistical methods for water quality management: issues, problems and solutions. John Wiley & Sons, New Jersey

Monaghan R, Manderson A, Basher L, et al (2021) Quantifying contaminant losses to water from pastoral landuses in New Zealand I. Development of a spatial framework for assessing losses at a farm scale. N Z J Agric Res 64:344–364

Moriasi DN, Arnold JG, Van Liew MW, et al (2007) Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Trans ASABE 50:885–900

Moriasi DN, Gitau MW, Pai N, Daggupati P (2015) Hydrologic and water quality models: Performance measures and evaluation criteria. Trans ASABE 58:1763–1785

Muirhead R, Elliot S, Snelder T (2023) Development of an E. coli runoff risk matrix. AgResearch Ltd

Muirhead RW (2019) The effectiveness of streambank fencing to improve microbial water quality: a review. Agric Water Manag 223:105684

Nash JE, Sutcliffe JV (1970) River flow forecasting through conceptual models part I—A discussion of principles. J Hydrol 10:282–290

Neter J, Kutner MH, Nachtsheim CJ, Wasserman W (2004) Applied linear statistical models, 4th edn. McGraw-Hill, Chicago, IL

Newsome PFJ, Wilde RH, Willoughby EJ (2008) Land resource information system spatial data layers: data dictionary. Landcare Research New Zealand

NZ Government (2023) National Policy Statement for Freshwater Management 2014 (amended 2023)

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

R Core Team (2023) R: A language and environment for statistical computing.

Semadeni-Davies A, Elliott S (2016) Modelling the effect of stock exclusion on E. coli in rivers and streams: National application. Ministry for Primary Industries

Semadeni-Davies, A., Yalden, S., Sukias, J. and Elliott, S. (2018) National E. coli modelling: Supplementary material to support setting draft regional targets for swimmable rivers, NIWA client report prepared for Ministry for the Environment.

Snelder T, Smith H, Plew D, Fraser C (2023) Nitrogen, phosphorus, sediment and Escherichia coli in New Zealand's aquatic receiving environments: Comparison of current state to national bottom lines. LWP Ltd, Christchurch, New Zealand

Snelder TH, Biggs BJF (2002) Multi-scale river environment classification for water resources management. J Am Water Resour Assoc 38:1225–1240

Snelder TH, Whitehead AL, Fraser C, et al (2020) Nitrogen loads to New Zealand aquatic receiving environments: comparison with regulatory criteria. N Z J Mar Freshw Res 54:527–550

Srinivasan MS, Muirhead RW, Singh SK, et al (2021) Development of a national-scale framework to characterise transfers of N, P and Escherichia coli from land to water. N Z J Agric Res 64:286–313

Whitehead A, Fraser CE, Snelder TH, et al (2021a) Water quality state and trends in New Zealand Rivers. Analyses of national data ending in 2020. NIWA, Christchurch

Whitehead A, Fraser CE, Snelder TH (2021b) Spatial modelling of river water-quality state. Incorporating monitoring data from 2016 to 2020. NIWA, Christchurch

Wilkinson RJ, McKergow LA, Davies-Colley RJ, et al (2011) Modelling storm-event E. coli pulses from the Motueka and Sherry Rivers in the South Island, New Zealand. N Z J Mar Freshw Res 45:369–393

Woods RA, Hendrikx J, Henderson R, Tait A (2006) Estimating mean flow of New Zealand rivers. J Hydrol N Z 45:95–110

Appendix A Definition of potential typologies for the empirical catchment *E. coli* concentration models

Table 5. Typologies and associated sets of land-types included in this study. Land-types are defined by category names linked by an under-score (e.g. Dairy_WellDrained). Land use definitions are defined in Table 1. Note that the Bare and Water land use categories were not explicitly included in all models and are not listed in the land-types column of this table but are counted in the number of land-types.

Typology	Number of land-types	Land-types
1	9	Natural, Dairy, Sheep&Beef, Cropland, Water, Orchard&Vineyard, Urban, Bare, Exotic Forest
2	12	Natural_WellDrained, Dairy, Sheep&Beef_WellDrained, Cropland, Water, Orchard&Vineyard, Urban, Bare, ExoticForest_PoorlyDrained, Natural_PoorlyDrained, ExoticForest_WellDrained, Sheep&Beef_PoorlyDrained
3	13	Natural_WellDrained, Dairy_WellDrained, Sheep&Beef_WellDrained, Cropland, Dairy_PoorlyDrained, Water, Orchard&Vineyard, Urban, Bare, ExoticForest_PoorlyDrained, Natural_PoorlyDrained, ExoticForest_WellDrained, Sheep&Beef_PoorlyDrained
4	12	Natural_HighElev, Dairy, Sheep&Beef_LowElev, Cropland, Water, Orchard&Vineyard, Urban, Bare, Natural_LowElev, ExoticForest_LowElev, ExoticForest_HighElev, Sheep&Beef_HighElev
5	13	Natural_HighElev, Dairy_LowElev, Sheep&Beef_LowElev, Cropland, Dairy_HighElev, Water, Orchard&Vineyard, Urban, Bare, Natural_LowElev, ExoticForest_LowElev, ExoticForest_HighElev, Sheep&Beef_HighElev
6	12	Natural_Warm, Natural_Cold, Dairy, Sheep&Beef_Cold, Cropland, Water, Orchard&Vineyard, Urban, Bare, ExoticForest_Cold, ExoticForest_Warm, Sheep&Beef_Warm
7	13	Natural_Warm, Natural_Cold, Dairy_Cold, Sheep&Beef_Cold, Cropland, Dairy_Warm, Water, Orchard&Vineyard, Urban, Bare, ExoticForest_Cold, ExoticForest_Warm, Sheep&Beef_Warm
8	15	Natural_Dry, Natural_Moist, Dairy, Sheep&Beef_Wet, Cropland, Water, Orchard&Vineyard, Urban, Bare, ExoticForest_Moist, ExoticForest_Wet, ExoticForest_Dry, Sheep&Beef_Moist, Sheep&Beef_Dry, Natural_Wet
9	17	Natural_Dry, Natural_Moist, Dairy_Dry, Sheep&Beef_Wet, Cropland, Water, Orchard&Vineyard, Urban, Bare, ExoticForest_Moist, ExoticForest_Wet, ExoticForest_Dry, Dairy_Moist, Sheep&Beef_Moist, Sheep&Beef_Dry, Dairy_Wet, Natural_Wet
10	17	Natural_HighElev, Dairy_WellDrained, Sheep&Beef_LowElev_WellDrained, Cropland, Dairy_PoorlyDrained, Water, Orchard&Vineyard, Urban, Bare, Natural_LowElev, ExoticForest_LowElev_PoorlyDrained, ExoticForest_HighElev_PoorlyDrained, Sheep&Beef_HighElev_WellDrained, ExoticForest_HighElev_WellDrained, Sheep&Beef_LowElev_PoorlyDrained, Sheep&Beef_HighElev_PoorlyDrained, ExoticForest_LowElev_WellDrained
11	17	Natural_HighElev, Dairy_Dry, Sheep&Beef_LowElev_Wet, Cropland, Water, Orchard&Vineyard, Urban, Bare, Natural_LowElev, ExoticForest_LowElev_Wet, ExoticForest_HighElev_Dry, Dairy_Wet, Sheep&Beef_HighElev_Wet, Sheep&Beef_HighElev_Dry, ExoticForest_LowElev_Dry, Sheep&Beef_LowElev_Dry, ExoticForest_HighElev_Wet
12	21	Natural_HighElev_WellDrained, Dairy_LowElev_WellDrained, Sheep&Beef_LowElev_WellDrained, Cropland, Dairy_HighElev_PoorlyDrained, Water, Orchard&Vineyard, Urban, Bare, Natural_LowElev_WellDrained, ExoticForest_LowElev_PoorlyDrained, Dairy_LowElev_PoorlyDrained, ExoticForest_HighElev_PoorlyDrained, Sheep&Beef_HighElev_WellDrained, Natural_HighElev_PoorlyDrained, Dairy_HighElev_WellDrained, ExoticForest_HighElev_WellDrained, Natural_LowElev_PoorlyDrained, Sheep&Beef_LowElev_PoorlyDrained, Sheep&Beef_HighElev_PoorlyDrained, ExoticForest_LowElev_WellDrained

Typology	Number of land-types	Land-types
13	21	Natural_Warm_WellDrained, Natural_Cold_WellDrained, Dairy_Cold_WellDrained, Sheep&Beef_Cold_WellDrained, Cropland, Dairy_Warm_PoorlyDrained, Water, Orchard&Vineyard, Dairy_Warm_WellDrained, Urban, Bare, ExoticForest_Cold_PoorlyDrained, ExoticForest_Warm_PoorlyDrained, Sheep&Beef_Warm_WellDrained, Dairy_Cold_PoorlyDrained, Natural_Warm_PoorlyDrained, ExoticForest_Cold_WellDrained, Natural_Cold_PoorlyDrained, Sheep&Beef_Cold_PoorlyDrained, Sheep&Beef_Warm_PoorlyDrained, ExoticForest_Warm_WellDrained
14	23	Natural_HighElev_WellDrained, Dairy_WellDrained, Sheep&Beef_LowElev_WellDrained_Wet, Cropland, Dairy_PoorlyDrained, Water, Orchard&Vineyard, Urban, Bare, Natural_LowElev_WellDrained, ExoticForest_LowElev_PoorlyDrained, ExoticForest_HighElev_PoorlyDrained, Sheep&Beef_HighElev_WellDrained_Wet, Sheep&Beef_HighElev_WellDrained_Dry, Sheep&Beef_LowElev_WellDrained_Dry, Natural_HighElev_PoorlyDrained, ExoticForest_HighElev_WellDrained, Natural_LowElev_PoorlyDrained, Sheep&Beef_LowElev_PoorlyDrained_Dry, Sheep&Beef_HighElev_PoorlyDrained_Wet, Sheep&Beef_LowElev_PoorlyDrained_Wet, Sheep&Beef_HighElev_PoorlyDrained_Dry, ExoticForest_LowElev_WellDrained
15	25	Natural_HighElev_WellDrained, Dairy_LowElev_WellDrained, Sheep&Beef_LowElev_WellDrained_Wet, Cropland, Dairy_HighElev_PoorlyDrained, Water, Orchard&Vineyard, Urban, Bare, Natural_LowElev_WellDrained, ExoticForest_LowElev_PoorlyDrained, Dairy_LowElev_PoorlyDrained, ExoticForest_HighElev_PoorlyDrained, Sheep&Beef_HighElev_WellDrained_Wet, Sheep&Beef_HighElev_WellDrained_Dry, Sheep&Beef_LowElev_WellDrained_Dry, Natural_HighElev_PoorlyDrained, Dairy_HighElev_WellDrained, ExoticForest_HighElev_WellDrained, Natural_LowElev_PoorlyDrained, Sheep&Beef_LowElev_PoorlyDrained_Dry, Sheep&Beef_HighElev_PoorlyDrained_Wet, Sheep&Beef_LowElev_PoorlyDrained_Wet, Sheep&Beef_HighElev_PoorlyDrained_Dry, ExoticForest_LowElev_WellDrained
16	29	Natural_HighElev_WellDrained_Dry, Natural_HighElev_WellDrained_Wet, Dairy_LowElev_WellDrained, Sheep&Beef_LowElev_WellDrained_Wet, Cropland, Dairy_HighElev_PoorlyDrained, Water, Orchard&Vineyard, Urban, Bare, Natural_LowElev_WellDrained_Wet, ExoticForest_LowElev_PoorlyDrained, Dairy_LowElev_PoorlyDrained, ExoticForest_HighElev_PoorlyDrained, Sheep&Beef_HighElev_WellDrained_Wet, Sheep&Beef_HighElev_WellDrained_Dry, Sheep&Beef_LowElev_WellDrained_Dry, Natural_HighElev_PoorlyDrained_Wet, Dairy_HighElev_WellDrained, ExoticForest_HighElev_WellDrained, Natural_LowElev_PoorlyDrained_Wet, Sheep&Beef_LowElev_PoorlyDrained_Dry, Natural_LowElev_PoorlyDrained_Dry, Sheep&Beef_HighElev_PoorlyDrained_Wet, Natural_HighElev_PoorlyDrained_Dry, Natural_LowElev_WellDrained_Dry, Sheep&Beef_LowElev_PoorlyDrained_Wet, Sheep&Beef_HighElev_PoorlyDrained_Dry, ExoticForest_LowElev_WellDrained
17	33	Natural_HighElev_WellDrained_Dry, Natural_HighElev_WellDrained_Wet, Dairy_LowElev_WellDrained, Sheep&Beef_LowElev_WellDrained_Wet, Cropland, Dairy_HighElev_PoorlyDrained, Water, Orchard&Vineyard, Urban, Bare, Natural_LowElev_WellDrained_Wet, ExoticForest_LowElev_PoorlyDrained_Wet, Dairy_LowElev_PoorlyDrained, ExoticForest_HighElev_PoorlyDrained_Dry, Sheep&Beef_HighElev_WellDrained_Wet, Sheep&Beef_HighElev_WellDrained_Dry, ExoticForest_LowElev_PoorlyDrained_Dry, Sheep&Beef_LowElev_WellDrained_Dry, Natural_HighElev_PoorlyDrained_Wet, ExoticForest_HighElev_PoorlyDrained_Wet, Dairy_HighElev_WellDrained, ExoticForest_HighElev_WellDrained_Dry, Natural_LowElev_PoorlyDrained_Wet, Sheep&Beef_LowElev_PoorlyDrained_Dry, Natural_LowElev_PoorlyDrained_Dry, Sheep&Beef_HighElev_PoorlyDrained_Wet, ExoticForest_HighElev_WellDrained_Wet, Natural_HighElev_PoorlyDrained_Dry, Natural_LowElev_WellDrained_Dry, Sheep&Beef_LowElev_PoorlyDrained_Wet, Sheep&Beef_HighElev_PoorlyDrained_Dry, ExoticForest_LowElev_WellDrained_Wet, ExoticForest_LowElev_WellDrained_Dry

Appendix B Fitted coefficients for the best empirical *E. coli* concentration model

Table 6. Best E. coli *concentration model (Model 10) fitted coefficients. Coefficients for each quantile, their standard errors (St Error) and p-values.*

Land-type	Quantile	Coefficient	St Error	P value
Cropland	0.05	112.5	48.2	0.020
Dairy_PoorlyDrained	0.05	118.5	35.5	0.001
Dairy_WellDrained	0.05	104.1	48.5	0.032
ExoticForest_HighElev_PoorlyDrained	0.05	21.4	51.9	0.681
ExoticForest_HighElev_WellDrained	0.05	-38.1	33.8	0.259
ExoticForest_LowElev_PoorlyDrained	0.05	37.2	34.1	0.275
ExoticForest_LowElev_WellDrained	0.05	40.9	32.1	0.203
Natural_HighElev	0.05	2.8	2.5	0.255
Natural_LowElev	0.05	11.0	7.3	0.133
OrchardVineyard	0.05	60.9	70.5	0.387
Sheep&Beef_HighElev_PoorlyDrained	0.05	13.0	12.3	0.292
Sheep&Beef_HighElev_WellDrained	0.05	1.8	8.4	0.835
Sheep&Beef_LowElev_PoorlyDrained	0.05	47.6	27.4	0.082
Sheep&Beef_LowElev_WellDrained	0.05	88.8	50.4	0.078
Urban	0.05	114.8	63.1	0.069
Cropland	0.5	158.8	116.2	0.172
Dairy_PoorlyDrained	0.5	381.1	173.9	0.029
Dairy_WellDrained	0.5	391.1	67.3	0.000
ExoticForest_HighElev_PoorlyDrained	0.5	24.8	95.7	0.795
ExoticForest_HighElev_WellDrained	0.5	16.4	16.8	0.331
ExoticForest_LowElev_PoorlyDrained	0.5	127.8	54.2	0.019
ExoticForest_LowElev_WellDrained	0.5	7.6	30.8	0.805
Natural_HighElev	0.5	-1.0	4.9	0.831
Natural_LowElev	0.5	146.6	29.9	0.000
OrchardVineyard	0.5	310.2	194.1	0.110
Sheep&Beef_HighElev_PoorlyDrained	0.5	99.6	36.2	0.006
Sheep&Beef_HighElev_WellDrained	0.5	43.5	16.1	0.007
Sheep&Beef_LowElev_PoorlyDrained	0.5	349.2	52.5	0.000
Sheep&Beef_LowElev_WellDrained	0.5	182.8	49.2	0.000
Urban	0.5	754.4	149.3	0.000
Cropland	0.95	476.6	569.9	0.403
Dairy_PoorlyDrained	0.95	1986.8	553.5	0.000
Dairy_WellDrained	0.95	987.2	350.2	0.005
ExoticForest_HighElev_PoorlyDrained	0.95	-45.7	412.8	0.912
ExoticForest_HighElev_WellDrained	0.95	138.8	126.2	0.272
ExoticForest_LowElev_PoorlyDrained	0.95	204.9	826.2	0.804
ExoticForest_LowElev_WellDrained	0.95	-103.7	105.1	0.324
Natural_HighElev	0.95	17.1	19.5	0.382
Natural_LowElev	0.95	475.4	190.8	0.013
OrchardVineyard	0.95	240.1	2280.1	0.916
Sheep&Beef_HighElev_PoorlyDrained	0.95	138.1	104.5	0.187
Sheep&Beef_HighElev_WellDrained	0.95	363.1	80.1	0.000
Sheep&Beef_LowElev_PoorlyDrained	0.95	864.8	210.0	0.000
Sheep&Beef_LowElev_WellDrained	0.95	723.9	275.6	0.009
Urban	0.95	2815.4	428.1	0.000

Appendix C Simulating changes in land use using the model

The typology used for the best concentration model (Model 10, Table 2) defines land-types based on nine land use categories, and two environmental factors: elevation and drainage, which are both subdivided into two categories. The ETC parameters of the concentration model merge some of these land-types. For example, the land use categories Cropland, Orchard & Vineyard and Urban are not subdivided by the environmental factors (i.e., both the elevation and the drainage categories are merged) and Natural is only subdivided by elevation (i.e., the drainage categories are merged, see Table 4). However, we use the full typology to implement a scenario that involves simulating a land use change as described below.

Land use change is specified as a percentage change (dLU) from one land use to another. For example, 50% of Sheep & Beef land in each catchment is changed to Exotic Forest in scenario 2 in this study.

Land use change scenarios involve a reallocation of the proportion of area occupied by different land use categories (i.e., land under Sheep & Beef land use changed to land under Exotic Forest use). Because the drainage and elevation components of the typology are environmental factors, these do not change under the scenario. However, the transfer of land uses can only occur between land-types that have the same environmental categories. Therefore, the land use change scenario needs to be simulated in a way in which the unchanging nature of the environmental categories is maintained. This is achieved by subdividing the catchment by the “full typology”. For Model 10 the full typology is defined by nine land use categories all of which are further subdivided by the two environmental factors, resulting in four land-types per land use category. The change in land use is then specified for each land-type represented by the full typology in the catchment.

Consider the scenario 2 example where land use is changed from Sheep & Beef to Exotic Forest. The new proportion of catchment area (P) of Exotic Forest for the scenario, is given by:

$$P_{EF,i,j}^{scenario} = P_{EF,i,j}^{original} + P_{SB,i,j}^{original} \times dLU \qquad \text{Equation 10}$$

where $P_{EF,i,j}^{original}$ is the original proportion of the catchment occupied by Exotic Forest for the i^{th} and j^{th} elevation and drainage categories, $P_{SB,i,j}^{original}$ is the original proportion of the catchment occupied by Sheep & Beef for the i^{th} and j^{th} elevation and drainage categories and $P_{EF,i,j}^{Sc}$ is the new proportion of the catchment occupied by Exotic Forest for the i^{th} and j^{th} elevation and drainage categories for the scenario. The new proportion of catchment area (P) of Sheep & Beef for the scenario, is given by:

$$P_{SB,i,j}^{scenario} = P_{SB,i,j}^{original} \times (1 - dLU) \qquad \text{Equation 11}$$

For all unchanged land uses, we set $P_{x,i,j}^{Sc} = P_{x,i,j}^{original}$. The full typology proportions of area are then summed over each of the Model 10 land use categories where the Model 10 types are coarser than the full typology (e.g., the proportion of the Model 10 land-type “Natural Low Elevation” is $P_{N,low,well}^{scenario} + P_{N,low,poor}^{scenario}$). Finally, the scenario concentration can be calculated using Equation 7.